

Prepared in cooperation with the
North Dakota State Water Commission, Devils Lake Basin Joint Water Resource Board,
and Red River Joint Water Resource District

Water-Quality Trend Analysis and Sampling Design for the Devils Lake Basin, North Dakota, January 1965 Through September 2003

Scientific Investigations Report 2006-5238

Water-Quality Trend Analysis and Sampling Design for the Devils Lake Basin, North Dakota, January 1965 Through September 2003

By Karen R. Ryberg and Aldo V. Vecchia

Prepared in cooperation with the
North Dakota State Water Commission, Devils Lake Basin Joint Water Resource
Board, and Red River Joint Water Resource District

Scientific Investigations Report 2006–5238

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

U.S. Department of the Interior
DIRK KEMPTHORNE, Secretary

U.S. Geological Survey
Mark D. Myers, Director

U.S. Geological Survey, Reston, Virginia: 2006

For product and ordering information:
World Wide Web: <http://www.usgs.gov/pubprod>
Telephone: 1-888-ASK-USGS

For more information on the USGS--the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:
World Wide Web: <http://www.usgs.gov>
Telephone: 1-888-ASK-USGS

Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Although this report is in the public domain, permission must be secured from the individual copyright owners to reproduce any copyrighted materials contained within this report.

Suggested citation:

Ryberg, K.R., and Vecchia, A.V., 2006, Water-quality trend analysis and sampling design for the Devils Lake Basin, North Dakota, January 1965 through September 2003: U.S. Geological Survey Scientific Investigations Report 2006-5238, 64 p.

Contents

Abstract.....	1
Introduction.....	1
Description of Devils Lake Basin.....	5
Streamflow and Concentration Data Used for Water-Quality Trend Analysis	7
Time-Series Model Used for Water-Quality Trend Analysis.....	11
Analysis of Streamflow-Related Variability in Concentration Data.....	13
Water-Quality Trend Analysis	26
Dissolved Major Ions.....	26
Nutrients.....	29
Trace Element.....	29
Sampling Design	34
Summary.....	39
References Cited.....	43
Appendix.....	45
Parametric Flow-Adjustment Model	46
Survival Regression for Dissolved Nitrite Plus Nitrate as Nitrogen.....	47
Variogram Analysis of Flow-Adjusted Concentrations.....	58
Model Robustness	60

Figures

1–2. Maps showing:	
1. Location of the Devils Lake Basin in North Dakota.	2
2. Locations of stations used for water-quality trend analysis.	3
3–22. Graphs showing:	
3. Elevation of Devils Lake, North Dakota, June 30, 1867, through September 30, 2003.	4
4. Daily streamflow for 1960 through September 2003 for the Mauvais Coulee near Cando, North Dakota, and Edmore Coulee near Edmore, North Dakota, stations.	6
5. Calcium concentrations for 1965 through September 2003 for the Lake Irvine near Churchs Ferry, North Dakota, station.....	11
6. Streamflow anomalies for the Mauvais Coulee near Cando, North Dakota, and Edmore Coulee near Edmore, North Dakota, stations.....	12
7. Measured and fitted dissolved sulfate concentrations for 1970 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.....	14
8. Flow-adjusted dissolved sulfate concentrations for 1970 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.....	14
9. Partial residuals of dissolved sulfate concentrations for the Mauvais Coulee near Cando, North Dakota, station in relation to streamflow anomalies.	16
10. Measured and fitted dissolved ammonia concentrations for 1985 through September 2003 for the Edmore Coulee near Edmore, North Dakota, station.	17

11.	Flow-adjusted dissolved ammonia concentrations for 1985 through September 2003 for the Edmore Coulee near Edmore, North Dakota, station.	17
12.	Partial residuals of dissolved ammonia concentrations for the Edmore Coulee near Edmore, North Dakota, station in relation to streamflow anomalies.....	18
13.	Measured and fitted dissolved nitrite plus nitrate concentrations and censored values for 1980 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station	19
14.	Flow-adjusted dissolved nitrite plus nitrate concentrations and censored values for 1985 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station	19
15.	Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis	20
16.	Fitted trends for flow-adjusted dissolved calcium concentrations	30
17.	Fitted trends for flow-adjusted dissolved sulfate concentrations	31
18.	Fitted trends for flow-adjusted dissolved chloride concentrations	32
19.	Fitted trends for flow-adjusted dissolved nitrite plus nitrate concentrations	33
20.	Fitted trends for flow-adjusted dissolved ammonia concentrations	35
21.	Fitted trends for flow-adjusted total phosphorus concentrations	37
22.	Fitted trends for flow-adjusted dissolved strontium concentrations.....	38
A1.	Dissolved nitrite plus nitrate concentrations for the Lake Alice near Churchs Ferry, North Dakota, station	55
A2.	Dissolved nitrite plus nitrate concentrations for the Mauvais Coulee near Cando, North Dakota, station.....	55
A3.	Dissolved nitrite plus nitrate concentrations and censored values for the Mauvais Coulee near Cando, North Dakota, station	56
A4.	Fourth root of squared residuals for flow-adjusted dissolved sulfate concentrations for the Starkweather Coulee near Webster, North Dakota, station	59
A5.	Fourth root of squared residuals for flow-adjusted dissolved chloride concentrations for the Edmore Coulee near Edmore, North Dakota, station	59
A6.	Estimated variogram for flow-adjusted dissolved sulfate concentrations for the Big Coulee near Churchs Ferry, North Dakota, station for season 1 (March through May)	63
A7.	Estimated variogram for flow-adjusted dissolved sulfate concentrations for the Big Coulee near Churchs Ferry, North Dakota, station for season 2 (June through November)	63
A8.	Estimated variogram for flow-adjusted total phosphorus concentrations for the Channel A near Penn, North Dakota, station for season 1 (March through May)	64
A9.	Estimated variogram for flow-adjusted total phosphorus concentrations for the Channel A near Penn, North Dakota, station for season 2 (June through November)	64

Tables

1. Stations used for water-quality trend analysis.....	8
2. Constituents used for water-quality trend analysis.....	8
3. Stations and constituents used for water-quality trend analysis and number of samples for concentration data	9
4. Stations and constituents used to test significance of Channel A step trend and linear trend from 1980 through September 2003.....	27
5. Number of samples required to obtain specified error tolerance for estimated median concentrations	41
A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).....	49
A2. Fitted survival regression model coefficients for dissolved nitrite plus nitrate for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).....	57
A3. Estimated seasonal standard deviations and correlation ranges for flow-adjusted concentrations	61
A4. Results of test of significance for Channel A step trend and linear trend from 1980 through September 2003 for the Big Coulee near Churchs Ferry, North Dakota, station	62

Conversion Factors, Abbreviations, and Datum

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi ²)	259.0	hectare (ha)
square mile (mi ²)	2.590	square kilometer (km ²)
Flow rate		
cubic foot per second (ft ³ /s)	0.02832	cubic meter per second (m ³ /s)

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

$$^{\circ}\text{F} = (1.8 \times ^{\circ}\text{C}) + 32$$

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

$$^{\circ}\text{C} = (^{\circ}\text{F} - 32) / 1.8$$

Vertical coordinate information is referenced to the National Geodetic Vertical Datum of 1929 (NGVD 29).

Horizontal coordinate information is referenced to the North American Datum of 1927 (NAD 27).

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).

Water-Quality Trend Analysis and Sampling Design for the Devils Lake Basin, North Dakota, January 1965 Through September 2003

By Karen R. Ryberg and Aldo V. Vecchia

Abstract

This report presents the results of a study conducted by the U.S. Geological Survey, in cooperation with the North Dakota State Water Commission, the Devils Lake Basin Joint Water Resource Board, and the Red River Joint Water Resource District, to analyze historical water-quality trends in three dissolved major ions, three nutrients, and one dissolved trace element for eight stations in the Devils Lake Basin in North Dakota and to develop an efficient sampling design to monitor the future trends.

A multiple-regression model was used to detect and remove streamflow-related variability in constituent concentrations. To separate the natural variability in concentration as a result of variability in streamflow from the variability in concentration as a result of other factors, the base-10 logarithm of daily streamflow was divided into four components—a 5-year streamflow anomaly, an annual streamflow anomaly, a seasonal streamflow anomaly, and a daily streamflow anomaly. The constituent concentrations then were adjusted for streamflow-related variability by removing the 5-year, annual, seasonal, and daily variability. Constituents used for the water-quality trend analysis were evaluated for a step trend to examine the effect of Channel A on water quality in the basin and a linear trend to detect gradual changes with time from January 1980 through September 2003.

The fitted upward linear trends for dissolved calcium concentrations during 1980–2003 for two stations were significant. The fitted step trends for dissolved sulfate concentrations for three stations were positive and similar in magnitude. Of the three upward trends, one was significant. The fitted step trends for dissolved chloride concentrations were positive but insignificant. The fitted linear trends for the upstream stations were small and insignificant, but three of the downward trends that occurred during 1980–2003 for the remaining stations were significant. The fitted upward linear trends for dissolved nitrite plus nitrate as nitrogen concentrations during 1987–2003 for two stations were significant. However, concentrations during recent years appear to be lower than those for the 1970s and early 1980s but higher than those for the late 1980s and early 1990s. The fitted downward linear trend for dissolved

ammonia concentrations for one station was significant. The fitted linear trends for total phosphorus concentrations for two stations were significant. Upward trends for total phosphorus concentrations occurred from the late 1980s to 2003 for most stations, but a small and insignificant downward trend occurred for one station. Continued monitoring will be needed to determine if the recent trend toward higher dissolved nitrite plus nitrate as nitrogen and total phosphorus concentrations continues in the future.

For continued monitoring of water-quality trends in the upper Devils Lake Basin, an efficient sampling design consists of five major-ion, nutrient, and trace-element samples per year at three existing stream stations and at three existing lake stations. This sampling design requires the collection of 15 stream samples and 15 lake samples per year rather than 16 stream samples and 20 lake samples per year as in the 1992–2003 program. Thus, the design would result in a program that is less costly and more efficient than the 1992–2003 program but that still would provide the data needed to monitor water-quality trends in the Devils Lake Basin.

Introduction

The Devils Lake Basin in North Dakota is a 3,810-square-mile closed subbasin in the Red River of the North (Red River) Basin (fig. 1). About 3,320 square miles of the total 3,810 square miles is tributary to Devils Lake, and the remainder is tributary to East and West Stump Lakes (fig. 2). The Devils Lake Basin contributes naturally to the Red River Basin only when the elevation of Devils Lake is greater than 1,459 feet. At an elevation of about 1,447 feet, Devils Lake begins to spill into the Stump Lakes; and at an elevation of about 1,459 feet, the combined lakes begin to spill through Tolna Coulee into the Sheyenne River, a tributary to the Red River. Devils Lake is characterized by large fluctuations in elevation (fig. 3) and in concentrations of dissolved chemical constituents. From 1867 through September 2003, the elevation ranged from 1,400.9 feet in October 1940 to 1,448.3 feet in July 2001. The elevation recorded in July 2001 was about 24.7 feet higher than the elevation recorded in July 1993.

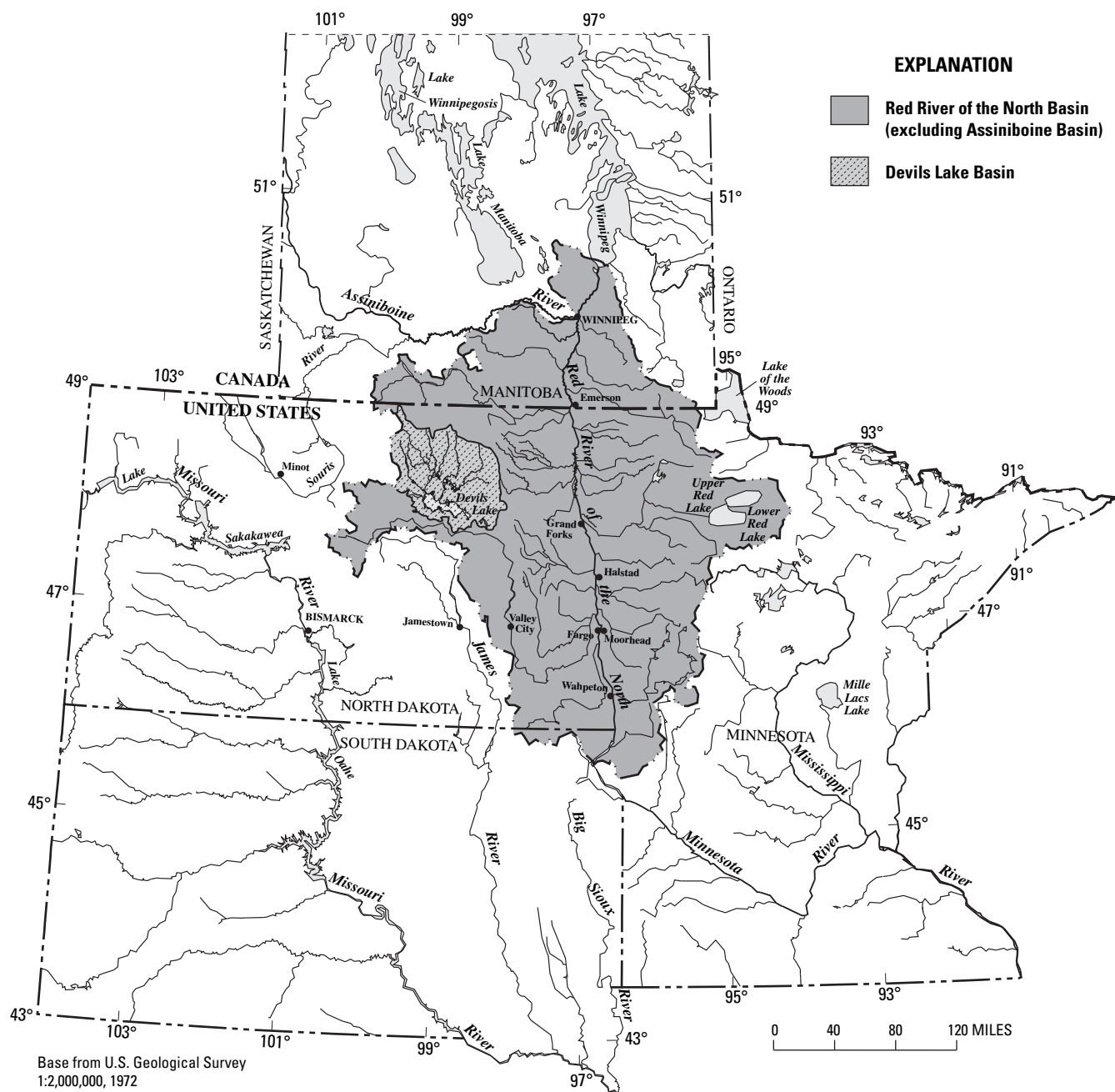


Figure 1. Location of the Devils Lake Basin in North Dakota.

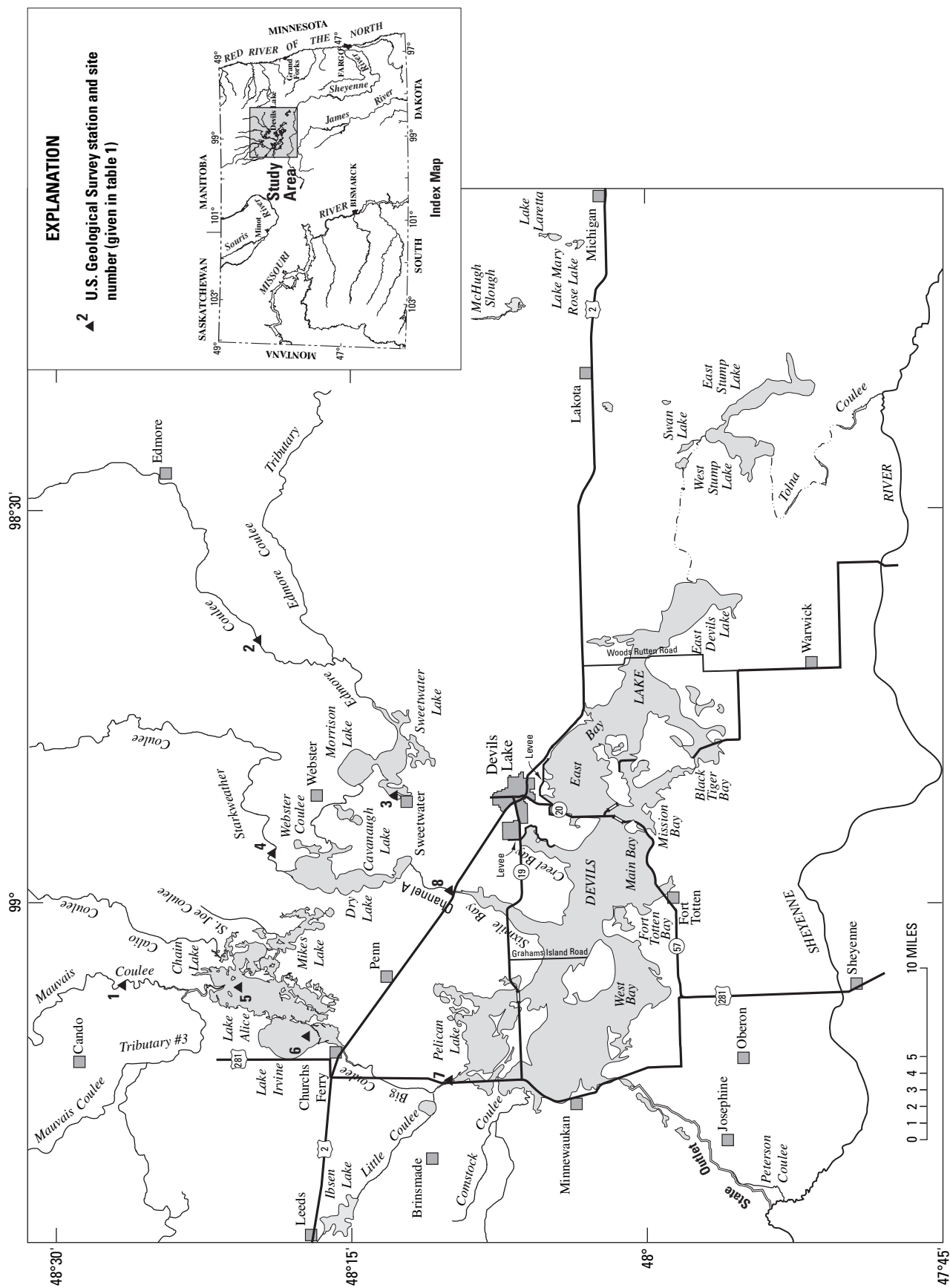


Figure 2. Locations of stations used for water-quality trend analysis.

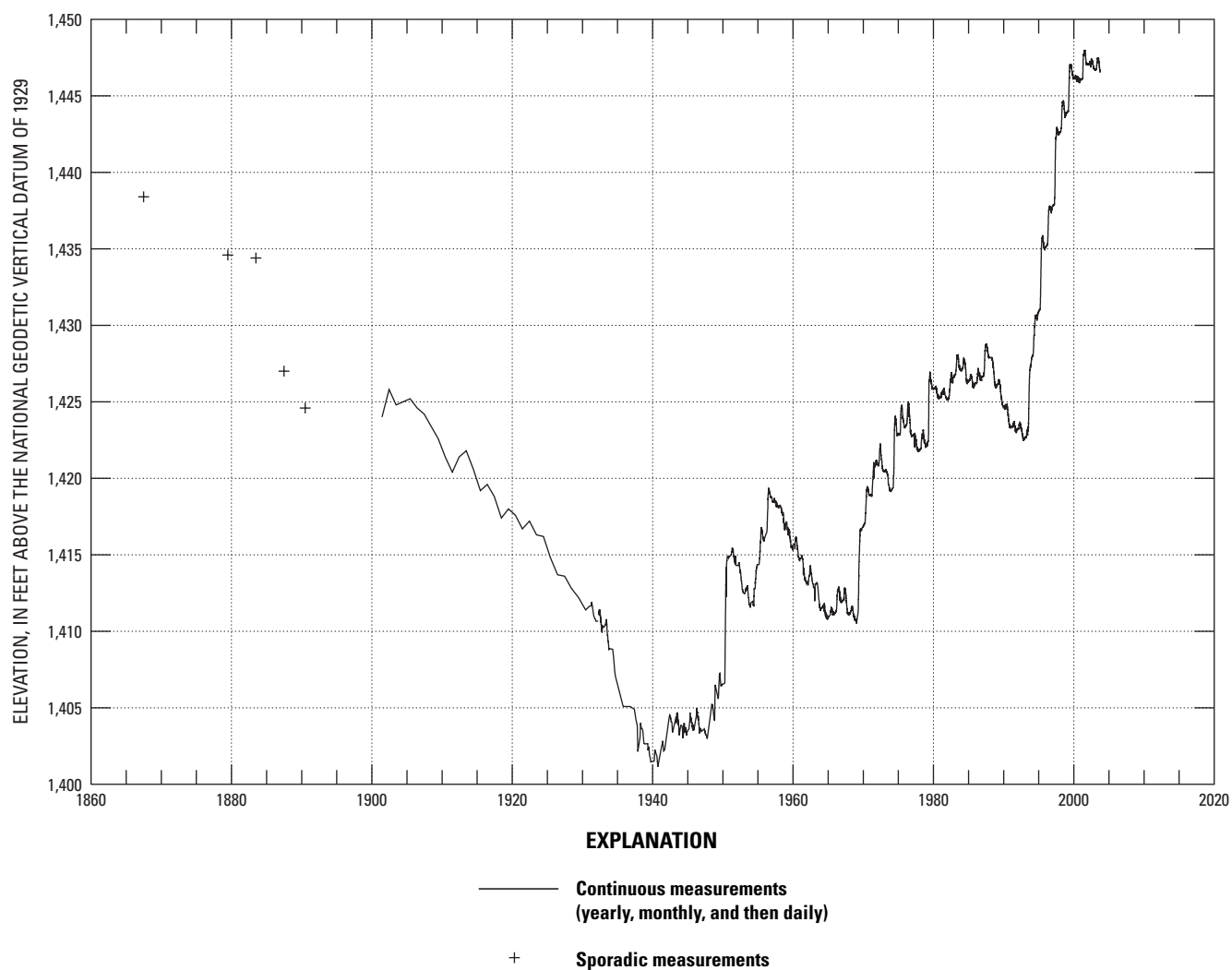


Figure 3. Elevation of Devils Lake, North Dakota, June 30, 1867, through September 30, 2003.

Maintaining the quality of surface waters in the Devils Lake Basin is important for the protection of agricultural resources, fisheries, waterfowl and wildlife habitat, and recreation in the basin and for the operation of the State of North Dakota Devils Lake outlet. To maintain the quality of surface waters in the Devils Lake Basin, information was needed on historical water-quality trends in the basin, and an efficient sampling design was needed to monitor the future trends. Therefore, the U.S. Geological Survey (USGS), in cooperation with the North Dakota State Water Commission, the Devils Lake Basin Joint Water Resource Board, and the Red River Joint Water Resource District, conducted this study to analyze historical water-quality trends in selected constituents and selected streams and lakes in the Devils Lake Basin and to develop an efficient sampling design to monitor the future trends. The Devils Lake outlet transfers water from the Devils Lake Basin to the Sheyenne River (fig. 2) to mitigate flooding in the Devils Lake Basin. The maximum outlet discharge is based, in part, on the sulfate concentration, in milligrams per liter, in the Sheyenne River downstream from the outlet insertion point (North Dakota State Water Commission, accessed October 17, 2005).

This report, which focuses on upper-basin streams and lakes that are tributary to Devils Lake, presents the results of the water-quality trend analysis and presents a sampling design to monitor future trends in the upper Devils Lake Basin. The results presented in the report are based on streamflow data for January 1960 through September 2003 and on concentration data for January 1965 through September 2003. The constituents evaluated for the report include three dissolved major ions (calcium, sulfate, and chloride), three nutrients (dissolved nitrite plus nitrate as nitrogen, dissolved ammonia, and total phosphorus), and one dissolved trace element (strontium). The constituents were evaluated for eight stations in the Devils Lake Basin, including three on the upstream tributaries, three on the chain of lakes, one on Big Coulee, and one on Channel A (fig. 2). All of the constituents were evaluated for the stream stations. However, for the lake stations, few data are available for some of the constituents. Therefore, for those stations, constituents were evaluated on the basis of data availability.

Description of Devils Lake Basin

The surface-water drainage system of the Devils Lake Basin is a complex system of interconnected tributaries and lakes (Wiche and others, 1986). Most surface runoff from the upper basin flows through several major tributaries (Mauvais, Calio, Starkweather, and Edmore Coulees; fig. 2) into the interconnected chain of lakes north of Devils Lake (Sweetwater, Morrison, Dry, Mikes, and Chain Lakes and Lakes Alice and Irvine; fig. 2), and outflow from the upstream chain of lakes flows into Devils Lake through either Big Coulee or Channel A. Before 1979, all outflow from the upstream chain

of lakes flowed through Big Coulee into Pelican Lake and eventually into Devils Lake. However, in 1979, the Ramsey County and Cavalier County Water Management Boards constructed Channel A, which connects Dry Lake to Sixmile Bay of Devils Lake (fig. 2), and a levee was constructed across the natural outlet of Dry Lake to Mikes Lake. Thus, since 1979, outflow from Sweetwater, Morrison, and Dry Lakes has flowed through Channel A into Devils Lake and outflow from the remaining lakes has flowed through Big Coulee into Devils Lake. During extremely large floods, Dry Lake has overtopped the levees across the natural outlet, causing some flow to enter the natural watercourse.

Since the end of glaciation about 10,000 years ago, Devils Lake has fluctuated between the extremes of spilling into the Sheyenne River and being dry. Research by the North Dakota Geological Survey indicates Devils Lake has overflowed into the Sheyenne River at least twice during the past 4,000 years and has spilled into the Stump Lakes several times. The North Dakota State Geologist concluded that the natural condition for Devils Lake is either rising or falling, and the lake should not be expected to remain at any elevation for a long period (Bluemle, 1991; Murphy and others, 1997). This natural condition is illustrated, in part, in figure 3. Because Devils Lake consists of a series of bays connected by bridge openings or culverts, the water quality between the bays can differ substantially (Sether and others, 1999).

The elevation of Devils Lake declines during dry conditions and rises during wet conditions because surface runoff from the Devils Lake Basin drains into Devils Lake. Except for two distinct periods of drought from the late 1950s to the late 1960s and from the late 1980s to the early 1990s, conditions from 1950 through September 2003 generally were wetter than conditions from 1870 through 1949. The rise in elevation during the 1970s and 1980s culminated in August 1987 when the elevation of Devils Lake reached 1,428.8 feet, the highest elevation since the 1880s (fig. 3).

Daily streamflow in Mauvais and Edmore Coulees was high during 1970-87 except for occasional dry years (fig. 4). A severe drought began late in the summer of 1987, and the elevation of Devils Lake declined to 1,422.6 feet by February 1993 (fig. 3). During the drought, the volume of water in Devils Lake decreased about 37 percent, from 884,000 acre-feet in August 1987 to 558,400 acre-feet in February 1993 (Wiche and others, 2000). Daily streamflow in Mauvais and Edmore Coulees was at or near zero many times during the drought (fig. 4).

The drought also caused elevation declines on other lakes in the Devils Lake Basin, including on the chain of lakes north of Devils Lake. According to Wiche and others (2000), "Dry Lake was below the outlet elevation (1,445 feet above sea level) to Channel A during most of 1988-92. Crops were planted in the lakebed of Lake Irvine during part of the drought.... Much-below-average streamflow into the chain of lakes and above-average evaporation from the lake surfaces during most of the drought resulted in little flow out of the chain of lakes into Devils Lake."

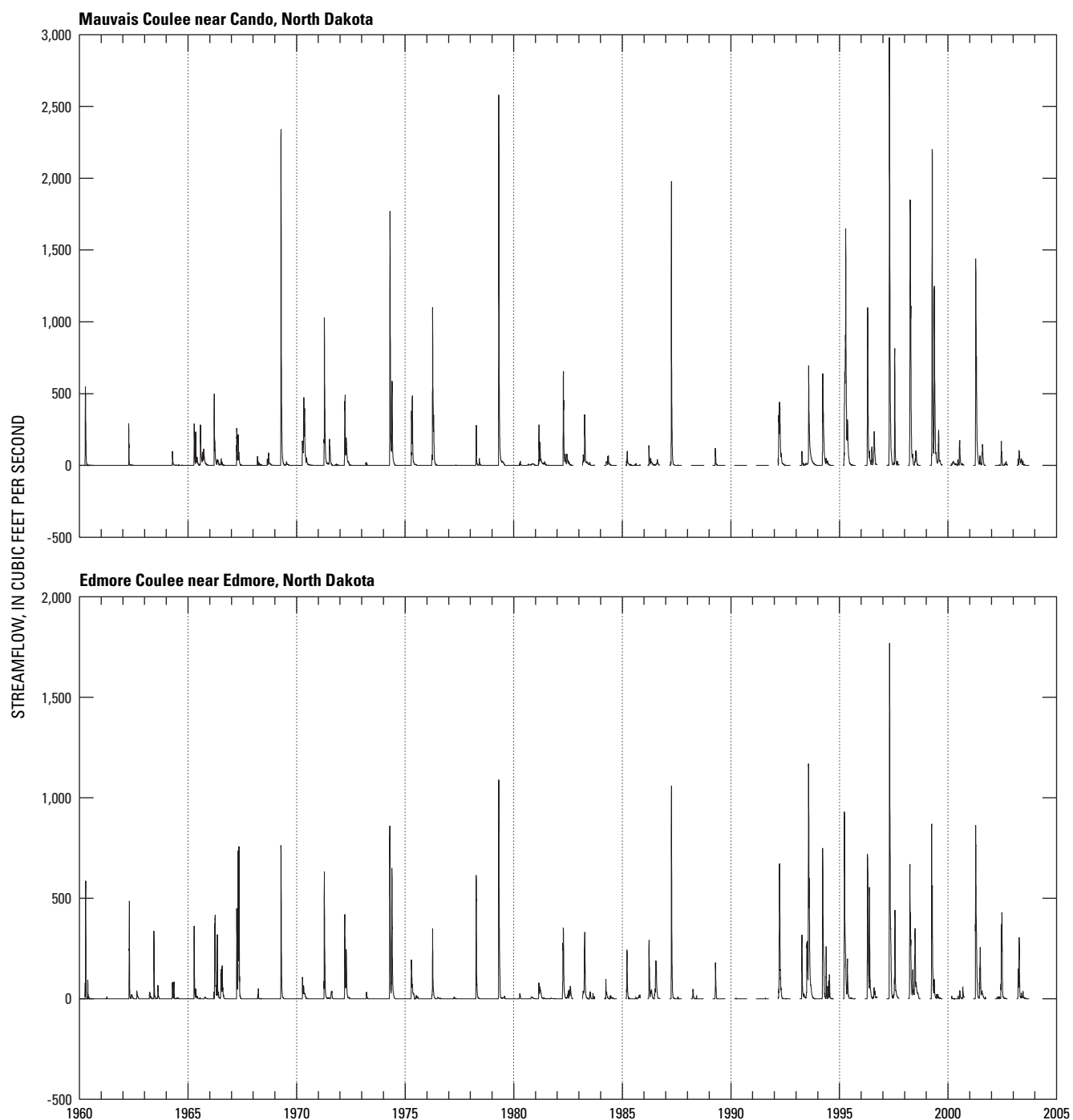


Figure 4. Daily streamflow for 1960 through September 2003 for the Mauvais Coulee near Cando, North Dakota, and Edmore Coulee near Edmore, North Dakota, stations.

Since 1993, severe flooding has occurred in the Devils Lake Basin. Devils Lake began to spill into the Stump Lakes in 2001 and continued to do so through 2003. The elevations of Devils Lake and the Stump Lakes were higher in the early 2000s than at any other time since the 1800s (Wiche and others, 2000).

Streamflow and Concentration Data Used for Water-Quality Trend Analysis

From 1957 through September 2003, the USGS collected water-quality samples from streams and lakes in the Devils Lake Basin in cooperation with the Devils Lake Basin Joint Water Resource Board, the North Dakota State Water Commission, the North Dakota Department of Health, the U.S. Army Corps of Engineers, and other local, State, and Federal agencies. Sample-collection techniques followed published USGS protocols (Wilde and others, various dates). The water-quality samples were analyzed by the USGS National Water Quality Laboratory, the North Dakota State Water Commission Laboratory, or the North Dakota Department of Health Laboratory. The USGS National Water Quality Laboratory followed strict, well-documented protocols (U.S. Geological Survey, various dates). The North Dakota State Water Commission Laboratory and the North Dakota Department of Health Laboratory followed U.S. Environmental Protection Agency approved methods and procedures. External agencies and customer organizations audit the State laboratories to assure the laboratory analytical methods and quality-assurance/quality-control procedures. The U.S. Environmental Protection Agency reviews the laboratory procedures about every third year, and the Board of Quality Systems of the USGS reviews the laboratory procedures periodically.

The streamflow and concentration data used for the water-quality trend analysis were obtained from the USGS National Water Information System (NWIS; <http://water-data.usgs.gov/nd/nwis/>). Ryberg and others (2005) described much of the data for the study area and presented a graphical user interface for graphical analysis of the data. During development of the graphical user interface, the large amount of historical data available from the USGS was reduced by selecting the stations and constituents that provided the most useful and reliable information for the assessment of water quality in the Devils Lake Basin. The most recent data included were for water year 2003¹. The amount of historical data then was reduced a second time by selecting the stations and constituents for which the period of record and the number of concentrations was sufficient for trend analysis. The eight stations used for the trend analysis are given in table 1, and the locations of the stations are shown in figure 2. The constitu-

ents used are given in table 2. The number of samples for each constituent at each site, the number of samples for specific periods during 1965 through September 2003, and the percentage of censored values (values that are known to be less than the laboratory reporting level but for which an exact value is not known) are given in table 3. Because of the drought of 1988-92, few or no data are available for many stations from 1987 through 1992. For example, the scatter plot of calcium concentrations for the Lake Irvine near Churchs Ferry station (site 6) (fig. 5) shows a gap in the data because no water-quality samples were collected from Lake Irvine between October 15, 1986, and March 16, 1993.

Streamflow is important in the analysis of water-quality trends because much of the variability in concentration is caused by variability in streamflow (Vecchia, 2003). Therefore, for this trend analysis, all available streamflow data for the Devils Lake Basin were examined for use as possible predictors of water quality. Daily mean streamflow data for the Mauvais Coulee near Cando station (site 1) were used for the Lake Alice near Churchs Ferry, Lake Irvine, and Big Coulee near Churchs Ferry stations (sites 5, 6, and 7, respectively) (table 1, fig. 2). Streamflow at those stations is contributed to Devils Lake through Big Coulee. Daily mean streamflow data for the Edmore Coulee near Edmore station (site 2) were used for the Sweetwater Lake at Sweetwater, Starkweather Coulee near Webster, and Channel A near Penn stations (sites 3, 4, and 8, respectively). Streamflow at those stations was contributed to Devils Lake through Big Coulee until 1979 and now is contributed to Devils Lake through Channel A. Streamflows for the Mauvais Coulee and Edmore Coulee stations (sites 1 and 2, respectively) were selected as being representative of streamflows in the Devils Lake Basin because of the long period of record and geographic location of each station.

The Mauvais Coulee station (site 1) data set contains streamflow values for June 1956 through September 2003 with seasonal records only since 1982 except for water year 1993 for which the record is complete. The seasonal records contain values for March 1 through September 30 except for 1985 when data were collected through October 30. The Edmore Coulee station (site 2) data set contains streamflow values for April through June 1956 and for July 1957 through September 2003 with seasonal records only since 1982 except for water year 1993 for which the record is complete. Again, the seasonal records contain values for March 1 through September 30 except for 1985 when data were collected through October 30. The streamflow data used for the trend analysis are for January 1, 1960, through September 30, 2003. The streamflow values used are shown in figure 4.

Strontium was the only trace element included in the water-quality trend analysis. Although many trace elements were considered for inclusion, most were not included because of a large percentage of censored values, a small number of samples, or both. Dissolved arsenic initially was included but, because most of the measured arsenic concentrations were reported by the analyzing laboratory as whole numbers, in micrograms per liter, the data were coarse and had few unique

¹A water year is defined as the 12-month period from October 1 through September 30 and is designated by the calendar year in which it ends. Thus, the year ending September 30, 2003, is called water year 2003.

values. Dissolved iron also was included initially but most of the measured iron concentrations were reported by the analyzing laboratory to two decimal places, in milligrams per liter, and then entered into the USGS water-quality database as micrograms per liter. The conversion from milligrams per liter to micrograms per liter moved the decimal point three places to the right and caused the iron data to be coarse and overly discrete (for example, 60, 70, 80, 110, and 140). Therefore, because use of continuous distribution methods is not appropriate for coarse or overly discrete data, dissolved arsenic and dissolved iron concentrations were not included in the trend analysis.

Table 1. Stations used for water-quality trend analysis.

Site number (figure 2)	U.S. Geological Survey station number	U.S. Geological Survey station name	Approximate contributing drainage area (square miles)	Latitude	Longitude
1	05056100	Mauvais Coulee near Cando, North Dakota	377	48°26'53"	99°06'08"
2	05056200	Edmore Coulee near Edmore, North Dakota	282	48°20'12"	98°39'36"
3	05056220	Sweetwater Lake at Sweetwater, North Dakota	380	48°12'37"	98°52'15"
4	05056239	Starkweather Coulee near Webster, North Dakota	210	48°19'14"	98°56'25"
5	05056250	Lake Alice near Churchs Ferry, North Dakota	1,600	48°19'33"	99°07'16"
6	05056260	Lake Irvine near Churchs Ferry, North Dakota	1,620	48°16'57"	99°10'25"
7	05056400	Big Coulee near Churchs Ferry, North Dakota	1,462	48°10'40"	99°13'15"
8	05056410	Channel A near Penn, North Dakota	790	48°10'00"	98°58'47"

Table 2. Constituents used for water-quality trend analysis.

Constituent	Unit	U.S. Geological Survey National Water Information System parameter code
Major ions		
Calcium, dissolved	Milligrams per liter	P00915
Sulfate, dissolved	Milligrams per liter	P00945
Chloride, dissolved	Milligrams per liter	P00940
Nutrients		
Nitrite plus nitrate as nitrogen, dissolved	Milligrams per liter	P00631
Ammonia, dissolved	Milligrams per liter	P00608
Phosphorus, total	Milligrams per liter	P00665
Trace element		
Strontium, dissolved	Micrograms per liter	P01080

Table 3. Stations and constituents used for water-quality trend analysis and number of samples for concentration data.

Site number (figure 2)	U.S. Geological Survey station name	Number of samples for				Percent of censored values for 1965 through September 2003
		1965 through September 2003	1965-79	1980-89	1990 through September 2003	
Calcium, dissolved						
1	Mauvais Coulee near Cando, North Dakota	106	17	29	60	0
2	Edmore Coulee near Edmore, North Dakota	96	12	23	61	0
3	Sweetwater Lake at Sweetwater, North Dakota	83	45	10	28	0
4	Starkweather Coulee near Webster, North Dakota	90	0	29	61	0
5	Lake Alice near Churchs Ferry, North Dakota	73	35	10	28	0
6	Lake Irvine near Churchs Ferry, North Dakota	75	35	10	30	0
7	Big Coulee near Churchs Ferry, North Dakota	198	128	24	46	0
8	Channel A near Penn, North Dakota	75	0	17	58	0
Sulfate, dissolved						
1	Mauvais Coulee near Cando, North Dakota	106	17	29	60	0
2	Edmore Coulee near Edmore, North Dakota	96	12	23	61	0
3	Sweetwater Lake at Sweetwater, North Dakota	83	45	10	28	0
4	Starkweather Coulee near Webster, North Dakota	90	0	29	61	0
5	Lake Alice near Churchs Ferry, North Dakota	73	35	10	28	0
6	Lake Irvine near Churchs Ferry, North Dakota	75	35	10	30	0
7	Big Coulee near Churchs Ferry, North Dakota	198	128	24	46	0
8	Channel A near Penn, North Dakota	77	0	17	60	0
Chloride, dissolved						
1	Mauvais Coulee near Cando, North Dakota	106	17	29	60	0
2	Edmore Coulee near Edmore, North Dakota	96	12	23	61	0
3	Sweetwater Lake at Sweetwater, North Dakota	83	45	10	28	0
4	Starkweather Coulee near Webster, North Dakota	90	0	29	61	0
5	Lake Alice near Churchs Ferry, North Dakota	73	35	10	28	0
6	Lake Irvine near Churchs Ferry, North Dakota	75	35	10	30	0
7	Big Coulee near Churchs Ferry, North Dakota	198	128	24	46	0
8	Channel A near Penn, North Dakota	77	0	17	60	0

Table 3. Stations and constituents used for water-quality trend analysis and number of samples for concentration data.—Continued

Site number (figure 2)	U.S. Geological Survey station name	Number of samples for				Percent of censored values for 1965 through September 2003
		1965 through September 2003	1965-79	1980-89	1990 through September 2003	
Nitrite plus nitrate as nitrogen, dissolved						
1	Mauvais Coulee near Cando, North Dakota	65	0	12	53	48
2	Edmore Coulee near Edmore, North Dakota	67	0	11	56	39
3	Sweetwater Lake at Sweetwater, North Dakota	67	30	10	27	40
4	Starkweather Coulee near Webster, North Dakota	61	0	12	49	39
5	Lake Alice near Churchs Ferry, North Dakota	57	21	10	26	26
6	Lake Irvine near Churchs Ferry, North Dakota	58	21	10	27	36
7	Big Coulee near Churchs Ferry, North Dakota	118	61	10	47	32
8	Channel A near Penn, North Dakota	67	0	12	55	37
Ammonia, dissolved						
1	Mauvais Coulee near Cando, North Dakota	65	0	12	53	12
2	Edmore Coulee near Edmore, North Dakota	66	0	10	56	12
4	Starkweather Coulee near Webster, North Dakota	61	0	12	49	13
7	Big Coulee near Churchs Ferry, North Dakota	57	0	10	47	7
8	Channel A near Penn, North Dakota	65	0	12	53	5
Phosphorus, total						
1	Mauvais Coulee near Cando, North Dakota	57	0	3	54	0
2	Edmore Coulee near Edmore, North Dakota	59	0	3	56	0
4	Starkweather Coulee near Webster, North Dakota	54	0	3	51	0
7	Big Coulee near Churchs Ferry, North Dakota	49	0	2	47	0
8	Channel A near Penn, North Dakota	57	0	3	54	0
Strontium, dissolved						
1	Mauvais Coulee near Cando, North Dakota	81	0	22	59	0
2	Edmore Coulee near Edmore, North Dakota	80	0	19	61	0
3	Sweetwater Lake at Sweetwater, North Dakota	69	36	5	28	0
4	Starkweather Coulee near Webster, North Dakota	79	0	20	59	0
5	Lake Alice near Churchs Ferry, North Dakota	58	25	5	28	0
6	Lake Irvine near Churchs Ferry, North Dakota	60	25	5	30	0
7	Big Coulee near Churchs Ferry, North Dakota	87	20	20	47	0
8	Channel A near Penn, North Dakota	78	0	17	61	0

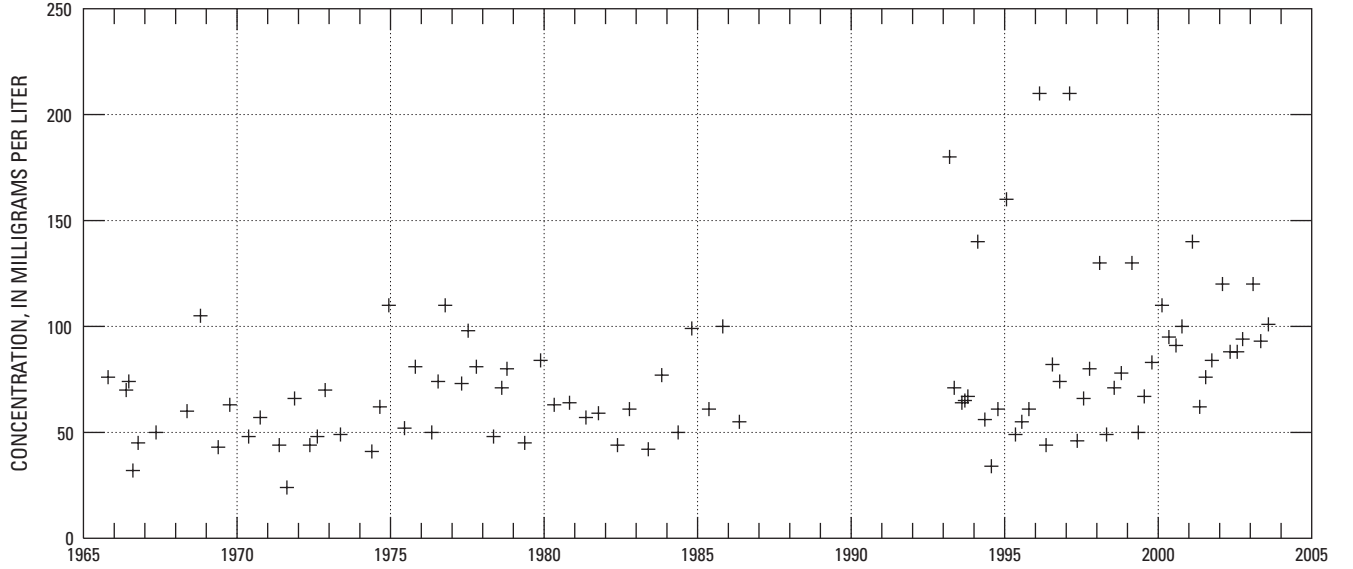


Figure 5. Calcium concentrations for 1965 through September 2003 for the Lake Irvine near Churchs Ferry, North Dakota, station.

Time-Series Model Used for Water-Quality Trend Analysis

For this report, a parametric multiple-regression model was used to detect and remove the streamflow-related variability in constituent concentrations. The approach used is similar to the approach used by Vecchia (2000, 2003, 2005). The methods used to fit the regression model are described in the appendix.

To separate the natural variability in concentration as a result of variability in streamflow from the variability in concentration as a result of other factors, the base-10 logarithm of daily streamflow for each station was divided into four components—a 5-year streamflow anomaly, an annual streamflow anomaly, a seasonal streamflow anomaly, and a daily streamflow anomaly. The four components were defined as follows:

$$X(t) = C_x + L_x(t) + A_x(t) + S_x(t) + D_x(t) \quad (1)$$

where

- $X(t)$ is the base-10 logarithm of the maximum of 0.1 or the daily streamflow, in cubic feet per second, for time t ;
- t is time, in decimal years;
- C_x is equal to $\text{mean}\{X(u), 1960 \leq u \leq 2003.75\}$;
- $L_x(t)$ is the 5-year streamflow anomaly (dimensionless) for time t and is equal to $\text{mean}\{X(u), t - 5 < u \leq t\} - C_x$;
- $A_x(t)$ is the annual streamflow anomaly (dimensionless) for time t and is equal to $\text{mean}\{X(u), t - 1 < u \leq t\} - \text{mean}\{X(u), t - 5 < u \leq t\}$;

$S_x(t)$ is the seasonal streamflow anomaly (dimensionless) for time t and is equal to $\text{mean}\{X(u), t - 0.25 < u \leq t\} - \text{mean}\{X(u), t - 1 < u \leq t\}$;

and

$D_x(t)$ is the daily streamflow anomaly (dimensionless) for time t and is equal to $X(t) - \text{mean}\{X(u), t - 0.25 < u \leq t\}$.

The 5-year anomaly, $L_x(t)$, represents the 5-year variability in streamflow from the long-term mean; the annual anomaly, $A_x(t)$, represents the annual variability in streamflow from the 5-year mean; the seasonal anomaly, $S_x(t)$, represents the seasonal variability in streamflow from the annual mean; and the daily anomaly, $D_x(t)$, represents the daily variability in streamflow from the seasonal mean. Vecchia (2003) used a similar equation to identify natural variability in constituent concentrations as a result of variability in streamflow.

To calculate the 5-year anomaly, 5 years of data, up to and including time t , were needed. Therefore, the complete set of anomalies was computed for January 1, 1965, through September 30, 2003. That time period also was used for the remainder of the water-quality trend analysis.

The 5-year, annual, seasonal, and daily anomalies for all stations used in the water-quality trend analysis are closely related. The anomalies for the Mauvais Coulee near Cando and Edmore Coulee near Edmore stations (sites 1 and 2, respectively) are shown in figure 6. Streamflows for those stations are subject to 5-year variability and a high degree of interannual variability. The seasonal anomalies indicate the highest streamflows generally occurred in June or July except during 1992 when the highest streamflows occurred in March or April.

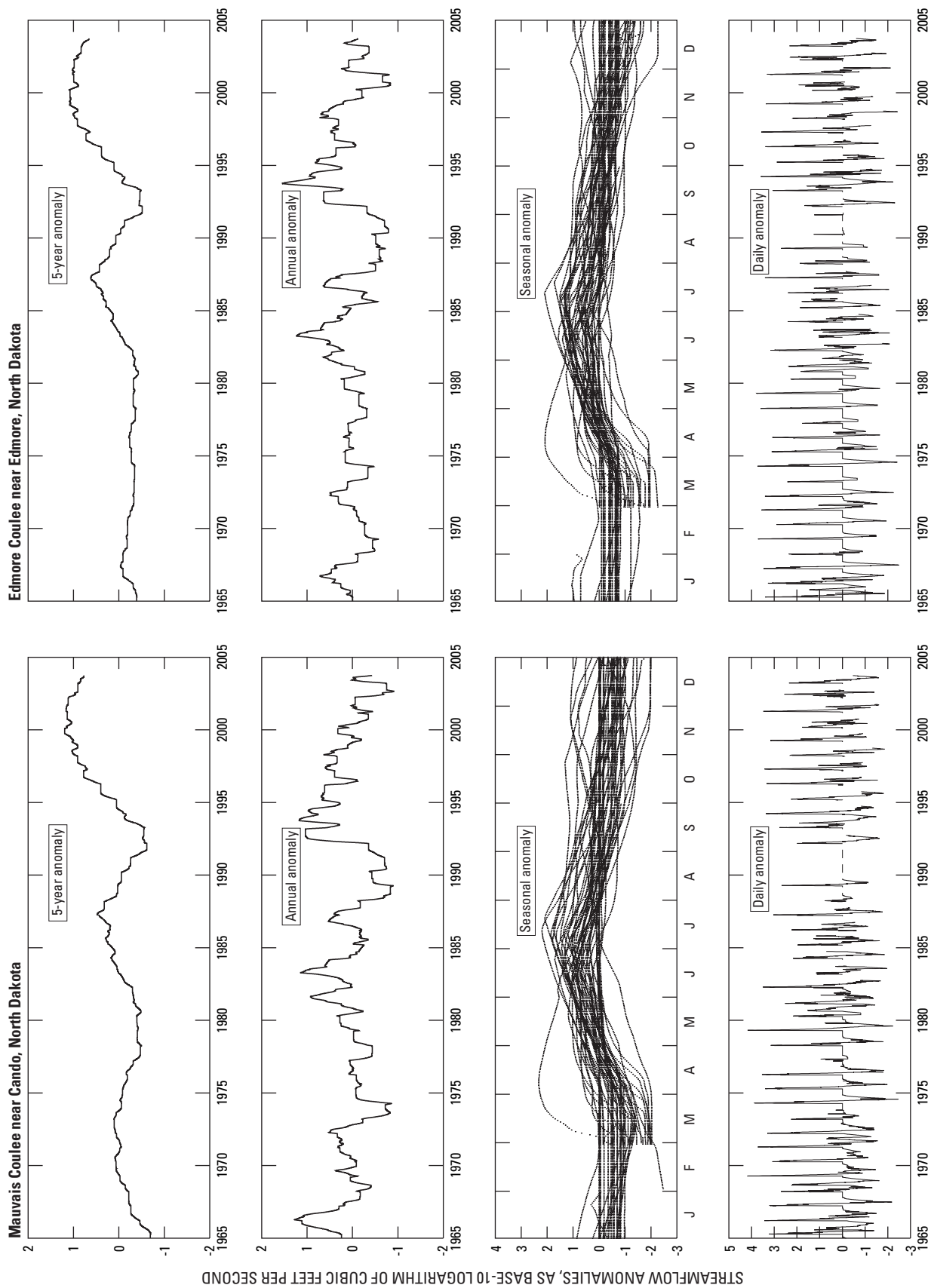


Figure 6. Streamflow anomalies for the Mauvais Coulee near Cando, North Dakota, and Edmore Coulee near Edmore, North Dakota, stations.

The seasonal anomalies indicated by the straight lines in January, February, October, November, and December are for years in which little or no streamflow occurred during the winter months. The seasonal anomalies indicated by the lines that begin March 1 are for years in which the station was operated seasonally and data collection began March 1.

The daily anomalies show a greater degree of high-frequency variability than the remaining anomalies. Individual meteorological events, temperature changes, and other factors contribute to this variability. The breaks in the daily anomaly plots in figure 6 indicate dates for which data are not available because of seasonal operation of the station.

Analysis of Streamflow-Related Variability in Concentration Data

For a water-quality trend analysis, streamflow-related variability generally is considered nuisance variability that needs to be removed before the trends are analyzed. If the streamflow-related variability is not properly fitted and removed, trends might be detected when no trends exist. Flow-adjusted concentrations, defined as the residuals from the flow-adjustment model plus the intercept (see appendix), are estimates of what the actual concentrations would have been if no streamflow-related variability occurred (that is, if streamflow conditions were constant) during the entire trend-analysis period.

The measured dissolved sulfate concentrations for the Mauvais Coulee near Cando station (site 1) and the fitted concentrations from the stream-station model for dissolved sulfate (table A1) are shown in figure 7. Most of the variability in the fitted concentrations ($R_a^2 = 76.3$ percent; table A1) can be attributed to streamflow-related variability. Both long-term (interannual) and short-term (seasonal and daily) variability are evident. However, as indicated by the flow-adjusted concentrations (fig. 8), no obvious trends exist. Therefore, the interannual variability shown in figure 7 for 1988–2003 probably is an artifact of the severe drought that occurred during 1988–92 and the extreme wet conditions that followed the drought.

The relation between concentration and each of the streamflow anomalies (eq. 1) can be analyzed by examining partial residual plots. The partial residual plot for a given streamflow anomaly shows the fitted relation between the log-transformed concentration and the given anomaly after the variability caused by the remaining anomalies is removed (see appendix). For example, to determine the partial residuals of dissolved sulfate concentrations for the Mauvais Coulee station (site 1) for the annual streamflow anomaly, the variability caused by the 5-year, seasonal, and daily streamflow anomalies was subtracted from the measured concentrations. The resulting residuals are shown in figure 9. The lines in the figure correspond to the fitted relation between the partial residuals and the given anomaly. For this example, annual

wet conditions (conditions for the present year are wetter than average conditions for the previous 5 years; indicated by a positive annual anomaly) tend to result in lower concentrations than do dry conditions (indicated by a negative annual anomaly). However, extended wet conditions (indicated by a positive 5-year anomaly) tend to have little effect on the concentrations. The slope for the 5-year anomaly is 0.028 (table A1), indicating only a slight increase for the high 5-year anomalies. However, the 95-percent confidence interval for the slope is (-0.030, 0.086). Therefore, because zero is contained in the confidence interval, the 5-year anomaly is not significant. Wet conditions on short time scales (less than 1 year) (indicated by positive seasonal and daily anomalies) tend to result in substantially lower concentrations than do dry conditions (indicated by negative seasonal and daily anomalies). The negative slopes indicate both the seasonal and daily anomalies are highly significant.

The measured dissolved ammonia concentrations for the Edmore Coulee near Edmore station (site 2) and the fitted concentrations from the stream-station model for dissolved ammonia (table A1) are shown in figure 10. Much of the variability in the fitted concentrations ($R_a^2 = 50.7$ percent; table A1) can be attributed to streamflow-related variability. However, the variability in the concentrations for this station is in sharp contrast to the variability in dissolved sulfate concentrations for the Mauvais Coulee station (site 1) (fig. 7). For this example, concentrations were high at the end of the severe drought (1992) and then decreased after the onset of the extreme wet conditions that followed the drought. However, as indicated by the flow-adjusted concentrations (fig. 11), no obvious trends exist. Therefore, the interannual variability shown in figure 11 for 1988–2003 probably is an artifact of the severe drought that occurred during 1988–92 and the extreme wet conditions that followed the drought.

The partial residuals of dissolved ammonia concentrations for the Edmore Coulee station (site 2) are shown in figure 12. Extended wet conditions tend to result in lower concentrations than do dry conditions. The slope for the 5-year anomaly is -0.322 (table A1), and the 95-percent confidence interval is (-0.506, -0.138). Annual wet conditions tend to have little effect on the concentrations. The zero slope indicates the annual anomaly is not significant. Seasonal wet conditions tend to result in substantially lower concentrations than do dry conditions. The slope for the seasonal anomaly is -0.295 (table A1), and the 95-percent confidence interval is (-0.443, -0.146). Daily wet conditions tend to have little effect on the concentrations. The slope for the daily anomaly is -0.026 (table A1), indicating only a slight decrease for the high daily anomalies. However, the 95-percent confidence interval for the slope is (-0.217, 0.165). Therefore, because zero is contained in the confidence interval, the daily anomaly is not significant.

The measured dissolved nitrite plus nitrate concentrations for the Mauvais Coulee station (site 1) and the fitted concentrations from the stream-station model for dissolved nitrite plus nitrate (table A2) are shown in figure 13. Much of the variability in the fitted concentrations probably can be

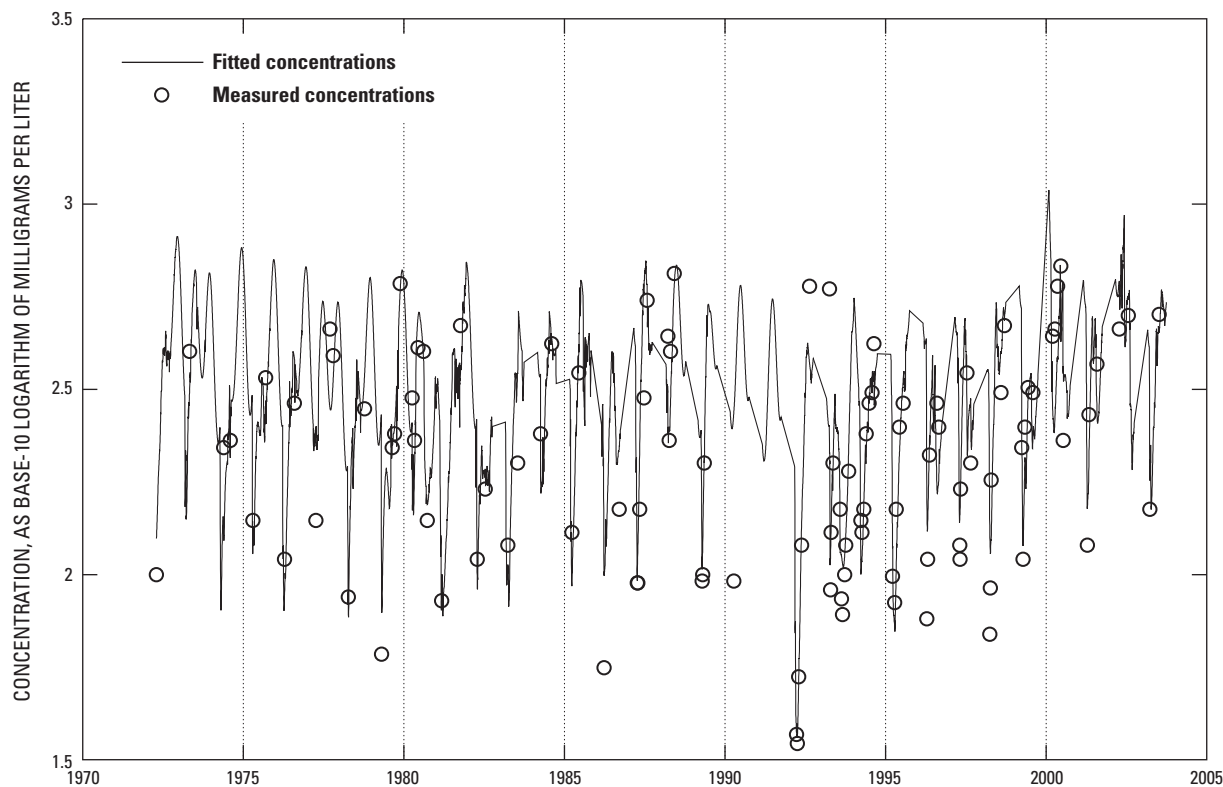


Figure 7. Measured and fitted dissolved sulfate concentrations for 1970 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.

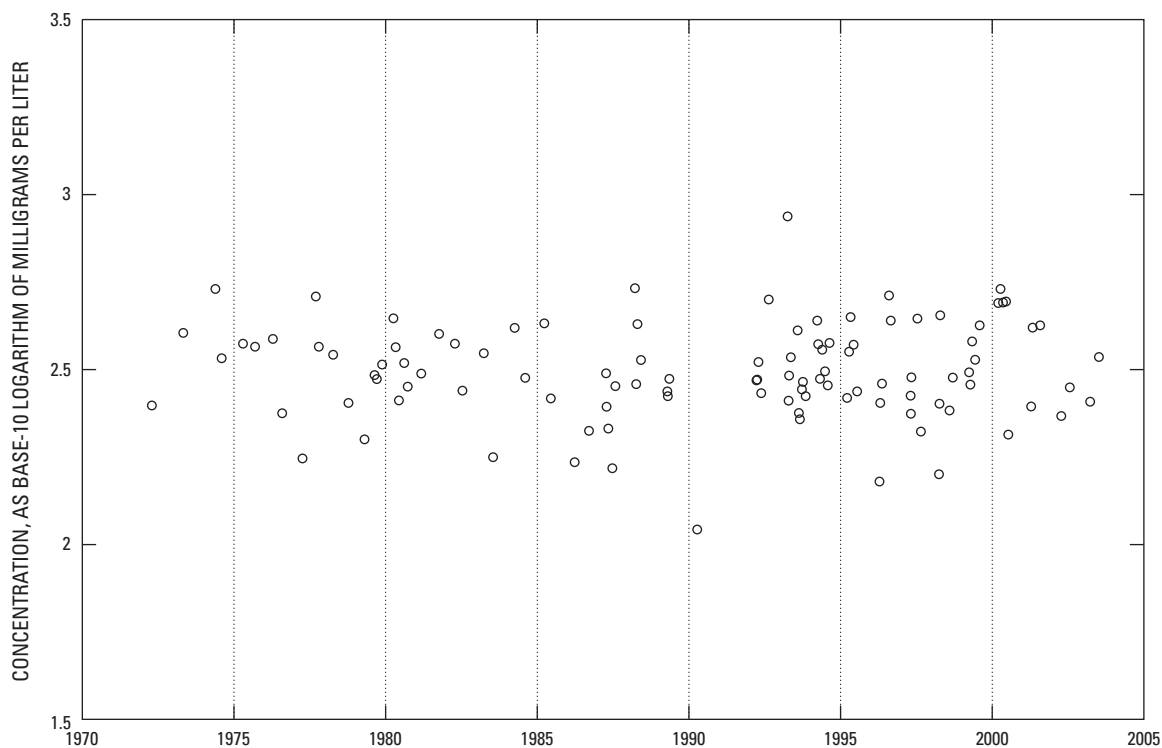


Figure 8. Flow-adjusted dissolved sulfate concentrations for 1970 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.

attributed to streamflow-related variability. However, because of the large percentage of censored values for this constituent (table 3), an alternative technique known as survival regression was used to fit the model (see appendix) and the fitted model was difficult to evaluate for low concentrations. The flow-adjusted concentrations, which in this case were defined as the residuals from the survival regression model plus the intercept, are shown in figure 14. Although these concentrations also were difficult to evaluate, no obvious trends exist.

The effects of adjusting concentrations for streamflow-related variability are shown in figure 15. The first set of box plots shows measured dissolved calcium concentrations for 1980 through September 2003 for the eight stations used for the water-quality trend analysis, and the second set of box plots shows flow-adjusted dissolved calcium concentrations. The second set of box plots indicates a decrease in streamflow-related variability for all of the stations. Because data are not available for some stations before 1980 (table 3), all data collected before that year were excluded to ensure a consistent period of record for the box plots. In some cases, the number of samples (fig. 15) is smaller for the flow-adjusted concentrations than for the measured concentrations because streamflow was not always measured when samples were collected.

In addition to reducing the streamflow-related variability in the constituent concentrations, the flow-adjustment process also may change the location of the distribution of the concentrations. For example, the median measured dissolved calcium concentrations for all of the stations are similar, but the median flow-adjusted concentrations for the Sweetwater Lake at Sweetwater and Starkweather Coulee stations (sites 3 and 4, respectively) are substantially lower than the median flow-adjusted concentrations for the remaining stations (fig. 15). The median measured concentrations may be unbiased estimates of the actual long-term median concentrations because sample-collection times do not represent random sampling dates during the year or during a full range of streamflow conditions. The flow-adjustment process accounts for the bias by adjusting the measured concentrations for seasonality and for streamflow-related variability.

The box plots for flow-adjusted dissolved sulfate concentrations (fig. 15) indicate large reductions in the variability of the distribution of the flow-adjusted concentrations in relation to the distribution of the measured concentrations for most stations. As for calcium, the median flow-adjusted sulfate concentrations were lower for the Starkweather Coulee station (site 4) than for the remaining stations.

The box plots for flow-adjusted dissolved chloride concentrations (fig. 15) indicate a decrease in streamflow-related variability and in bias in the measured concentrations for most stations. The median flow-adjusted concentrations for the Edmore Coulee and Starkweather Coulee stations (sites 2 and

4, respectively) were most affected by the flow-adjustment process. The median flow-adjusted concentrations for the Edmore Coulee station (site 2) were considerably higher than the median measured concentrations for that station. However, for the Starkweather Coulee station (site 4), the median flow-adjusted concentrations were considerably lower than the median measured concentrations. The low flow-adjusted concentrations for the Starkweather Coulee station (site 4) indicate that, for typical streamflow conditions, dissolved chloride concentrations for that station are much lower than concentrations for the remaining stations.

The box plots for flow-adjusted dissolved ammonia concentrations (fig. 15) indicate only a small decrease in streamflow-related variability. However, the flow-adjusted concentrations for the Big Coulee near Churchs Ferry and Channel A near Penn stations (sites 7 and 8, respectively), were much more symmetrically distributed than the measured concentrations for those stations. Thus, the flow-adjustment process also removes potential skewness in log-transformed concentrations. The concentrations for the Edmore Coulee station (site 2) are a particularly useful example of the effects of the flow-adjustment process. Most of the water-quality samples for that station were collected from May through October and the concentrations generally were low. However, the concentrations for February through April, when only a few water-quality samples were collected, generally were high. Weighting the sampling more heavily toward the May through October period caused a downward bias in the median measured concentrations, and the median flow-adjusted concentrations generally were higher than the median measured concentrations. The Sweetwater Lake, Lake Alice near Churchs Ferry, and Lake Irvine near Churchs Ferry stations (sites 3, 5, and 6, respectively) were not included in the evaluation of dissolved ammonia concentrations because too few data are available for those stations.

The box plots for flow-adjusted total phosphorus concentrations (fig. 15) indicate a decrease in streamflow-related variability and in bias in the measured concentrations for most stations included in the water-quality trend analysis. The low flow-adjusted concentrations for the Starkweather Coulee and Channel A stations (sites 4 and 8, respectively) indicate that, for typical streamflow conditions, concentrations for those stations are much lower than for the remaining stations. The Sweetwater Lake, Lake Alice, and Lake Irvine stations (sites 3, 5, and 6, respectively) were not included in the evaluation of total phosphorus concentrations because too few data are available for those stations.

The box plots for flow-adjusted dissolved strontium concentrations (fig. 15) indicate a decrease in streamflow-related variability for all stations.

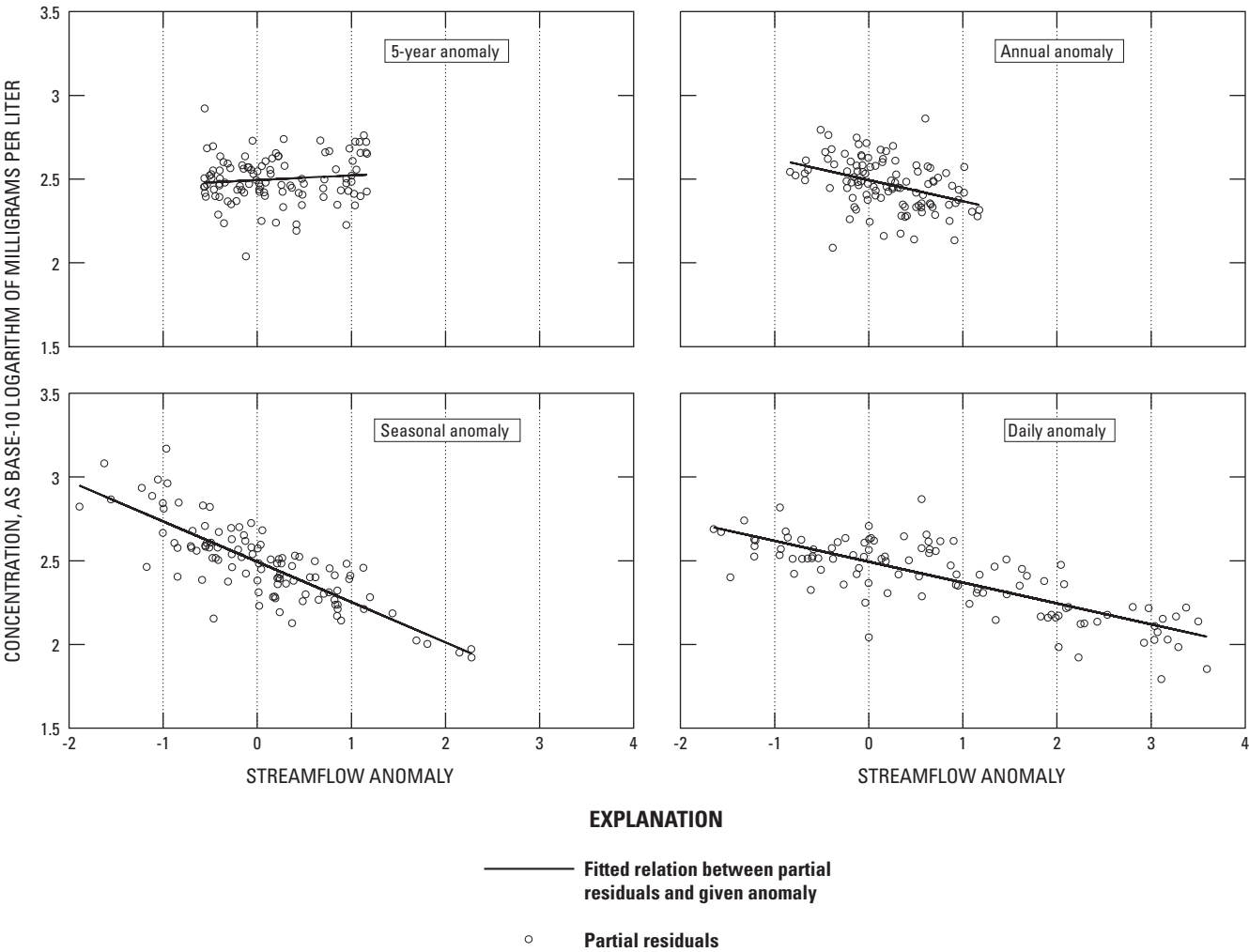


Figure 9. Partial residuals of dissolved sulfate concentrations for the Mauvais Coulee near Cando, North Dakota, station in relation to streamflow anomalies.

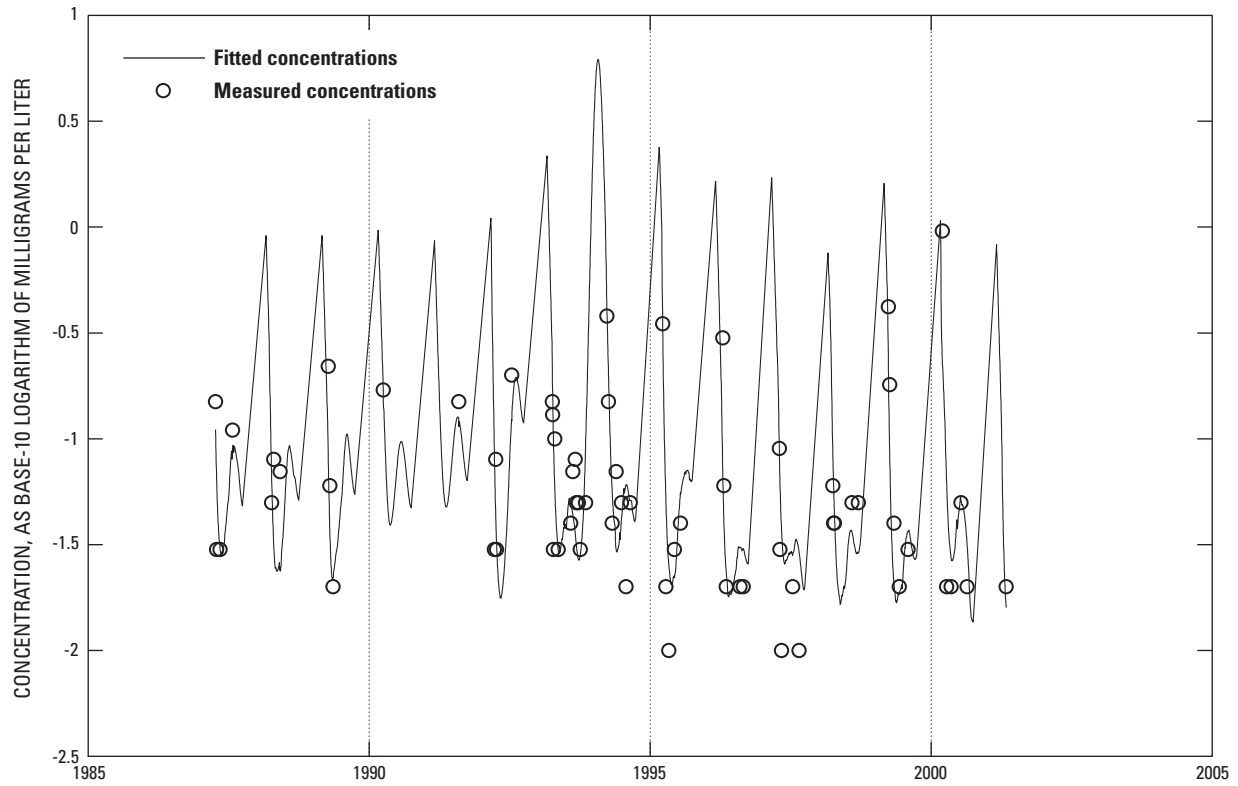


Figure 10. Measured and fitted dissolved ammonia concentrations for 1985 through September 2003 for the Edmore Coulee near Edmore, North Dakota, station.

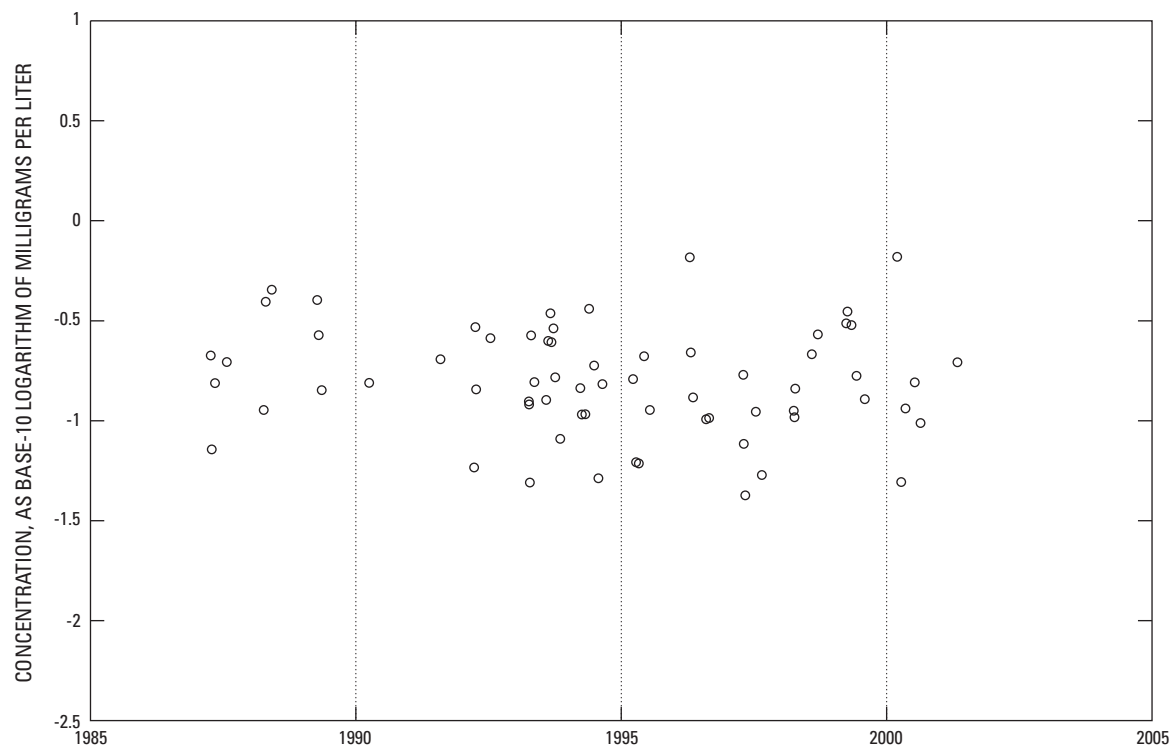


Figure 11. Flow-adjusted dissolved ammonia concentrations for 1985 through September 2003 for the Edmore Coulee near Edmore, North Dakota, station.

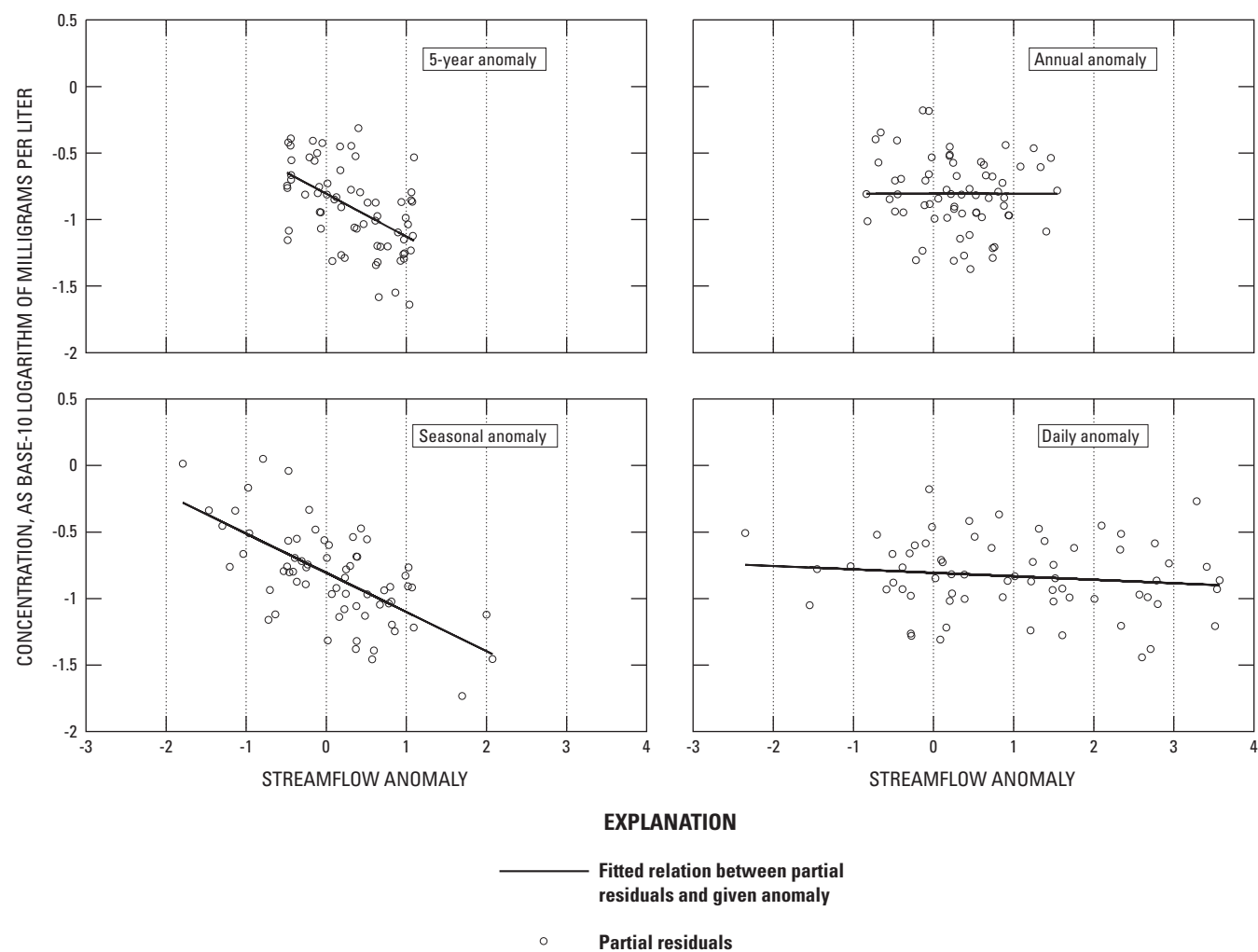


Figure 12. Partial residuals of dissolved ammonia concentrations for the Edmore Coulee near Edmore, North Dakota, station in relation to streamflow anomalies.

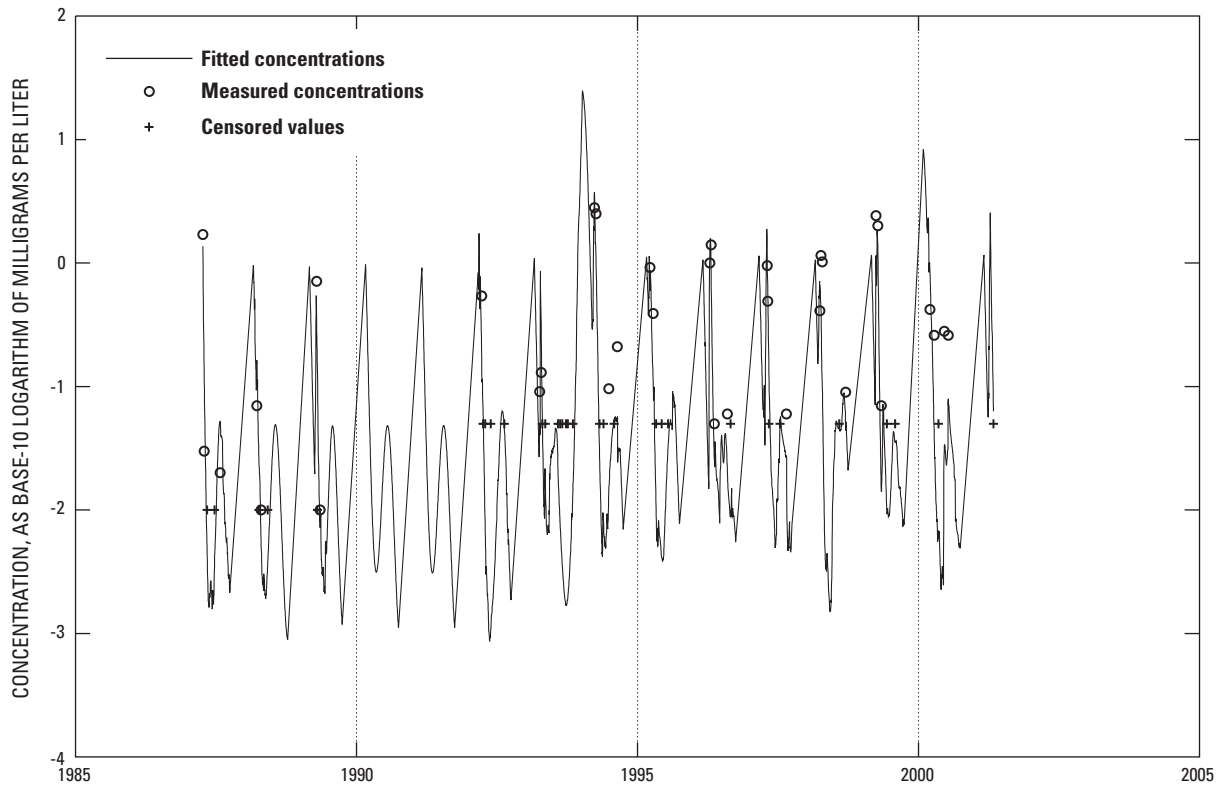


Figure 13. Measured and fitted dissolved nitrite plus nitrate concentrations and censored values for 1980 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.

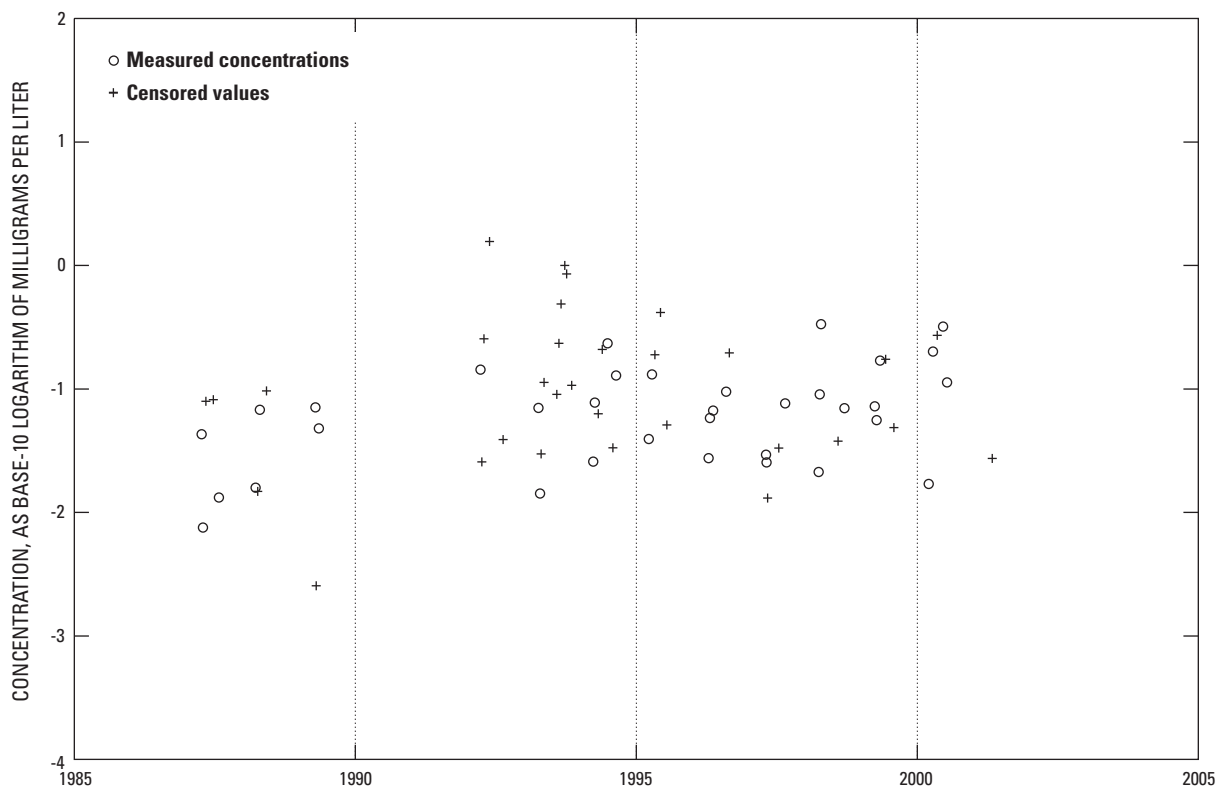


Figure 14. Flow-adjusted dissolved nitrite plus nitrate concentrations and censored values for 1985 through September 2003 for the Mauvais Coulee near Cando, North Dakota, station.

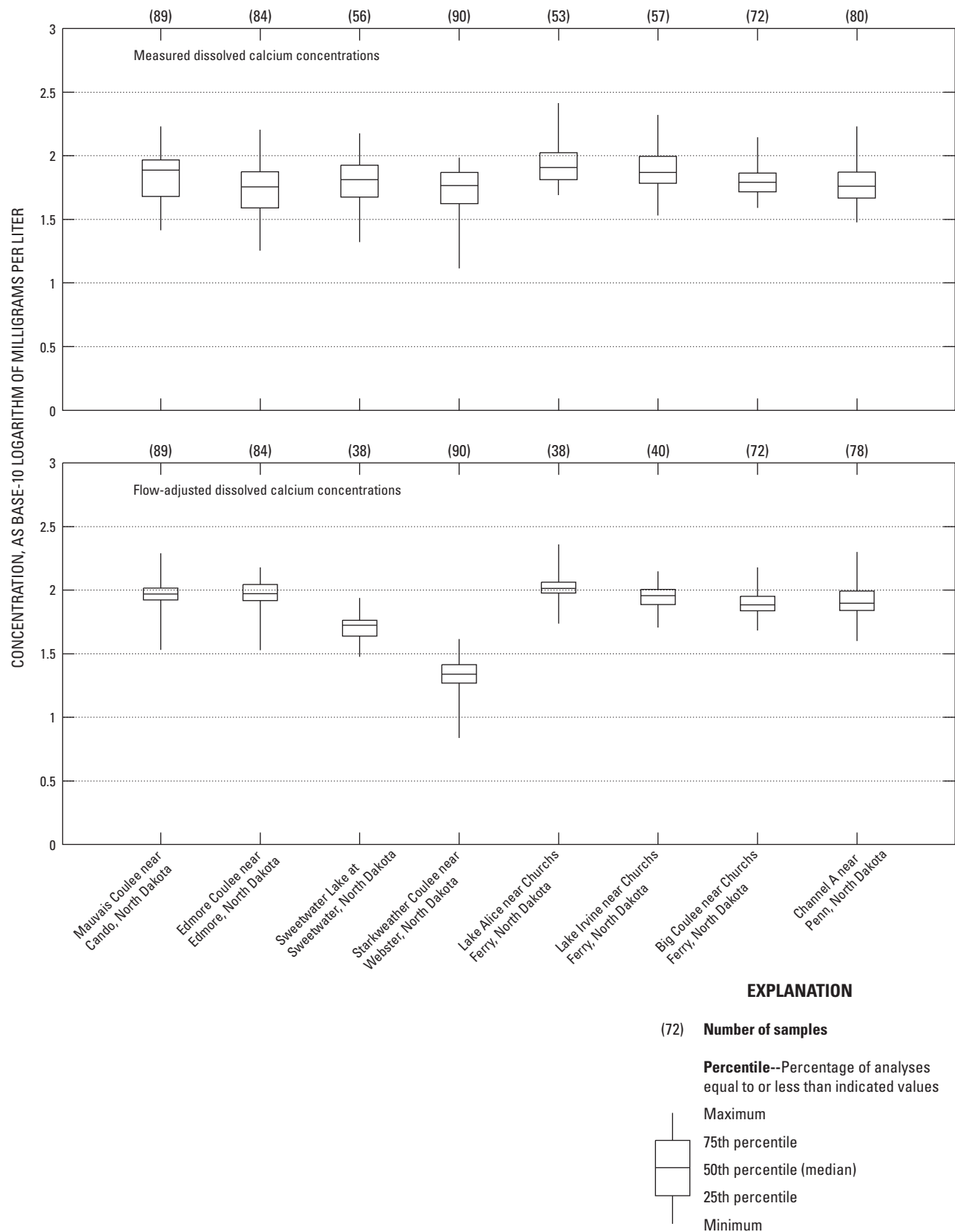


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.

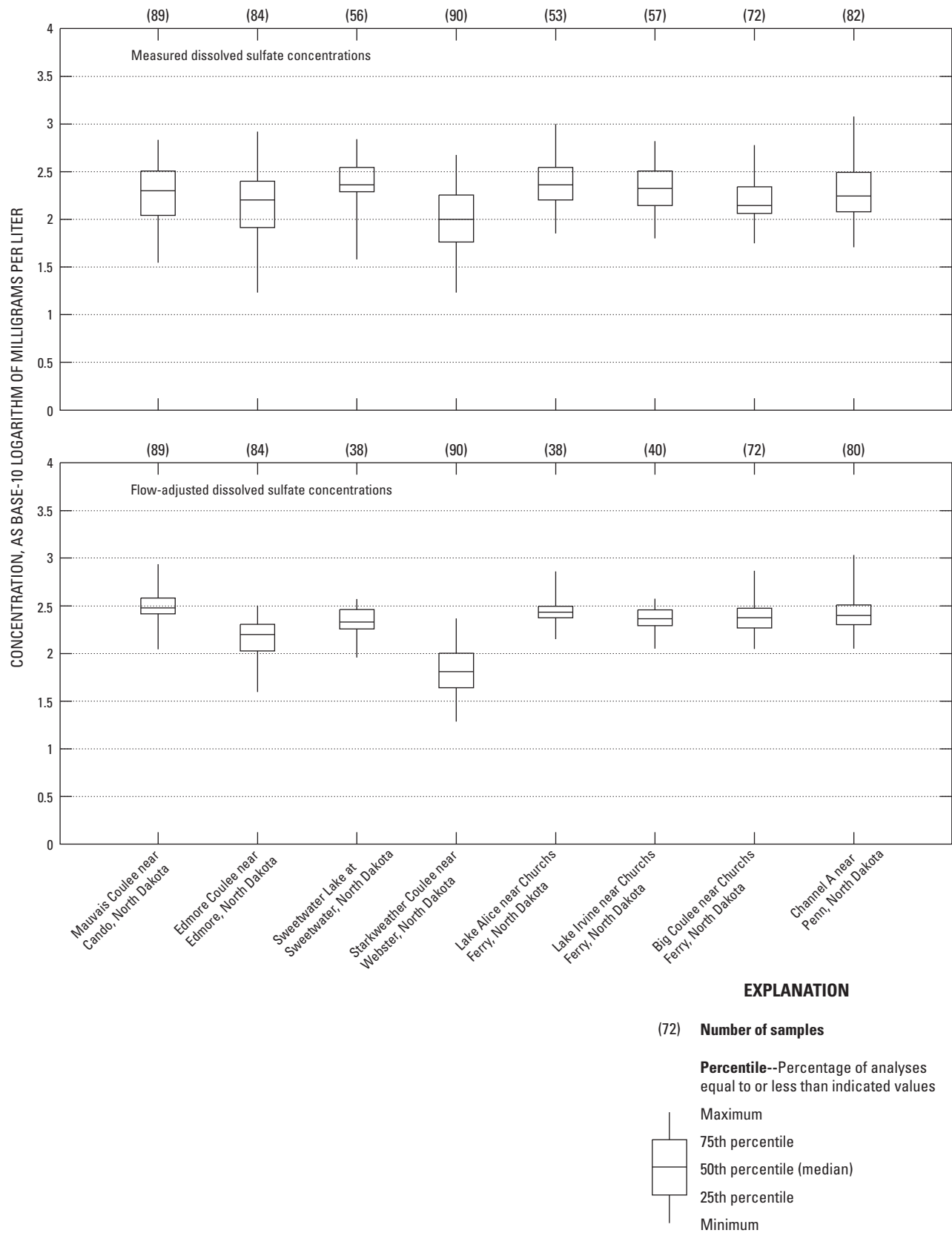


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.—Continued

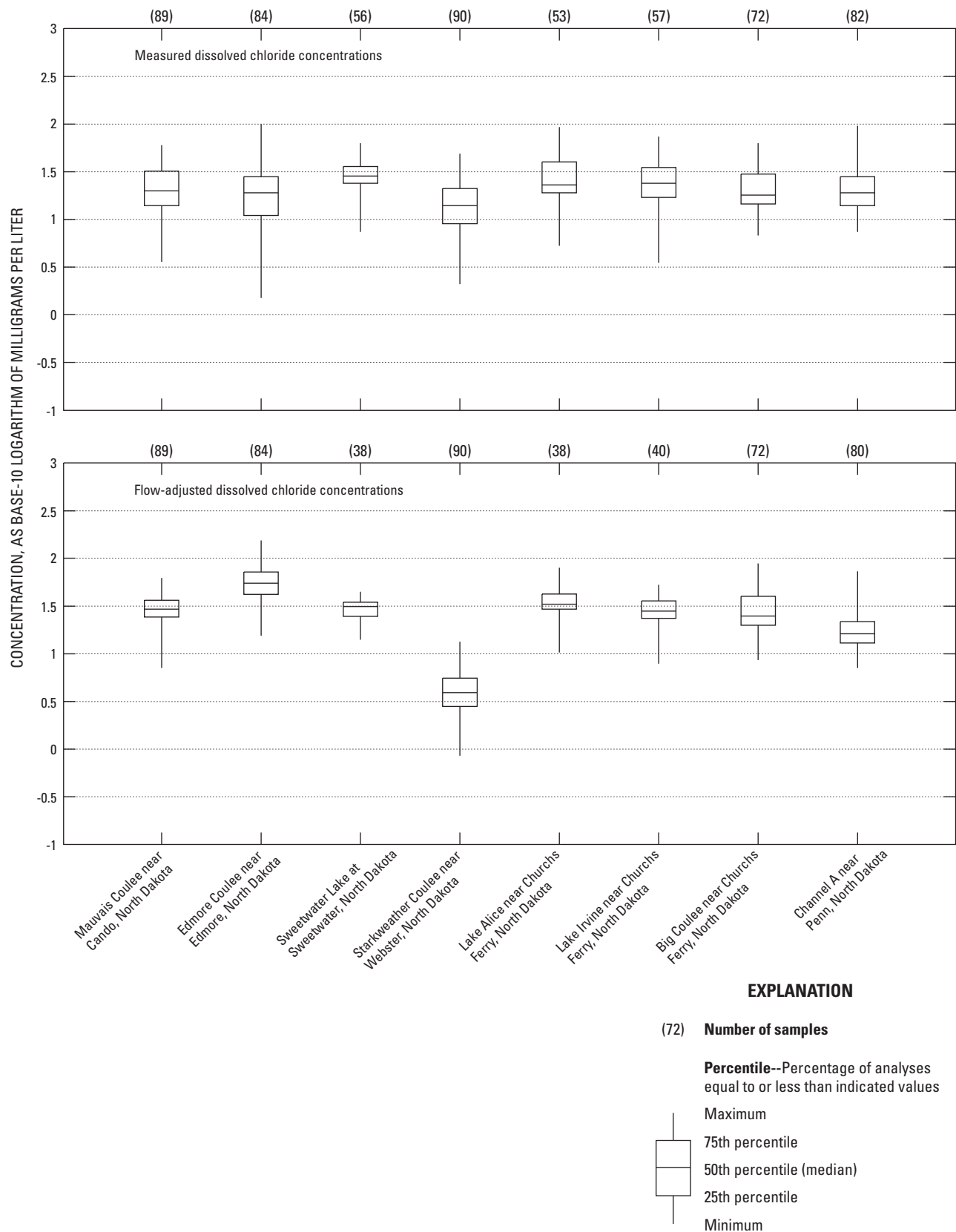


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.—Continued

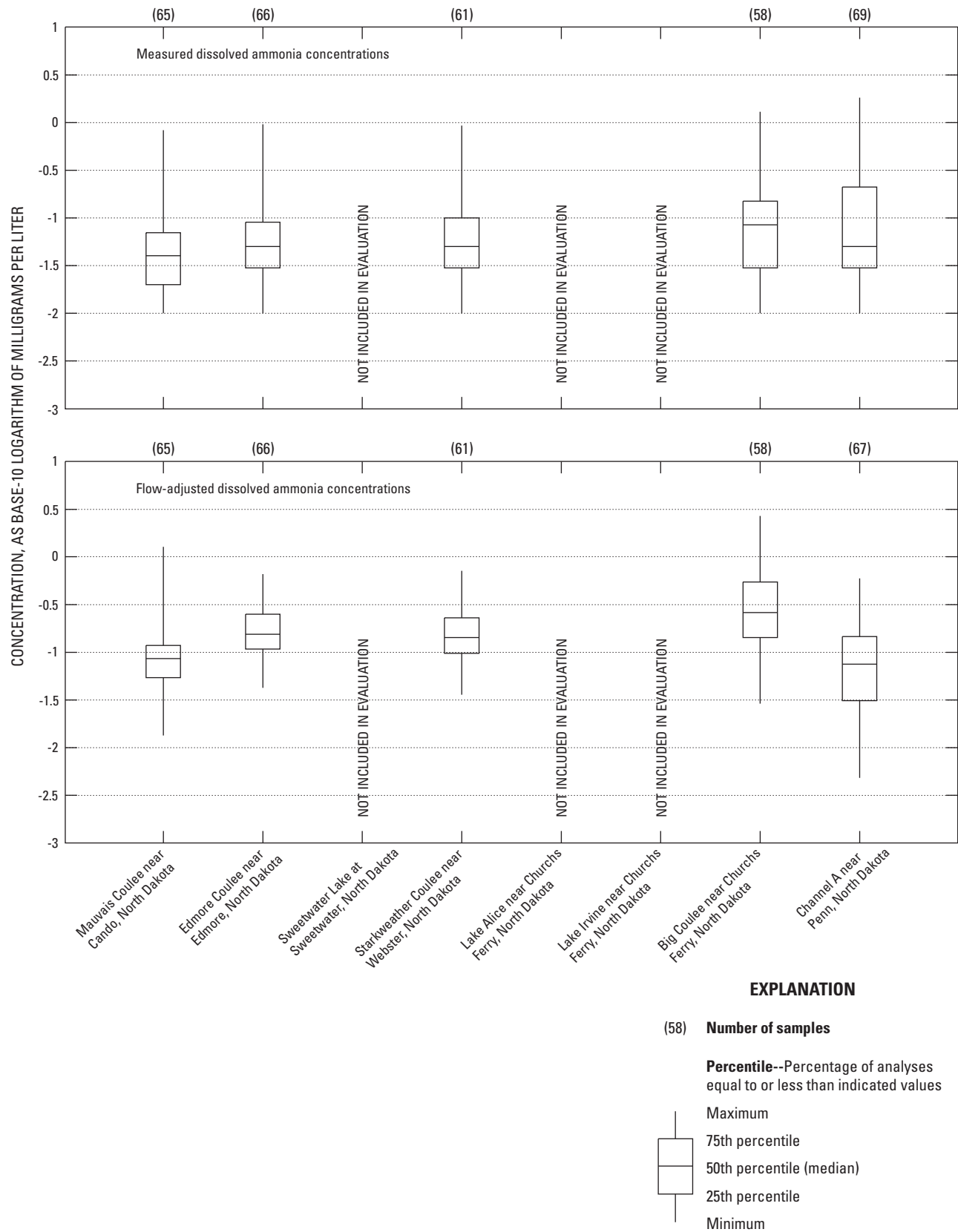


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.—Continued

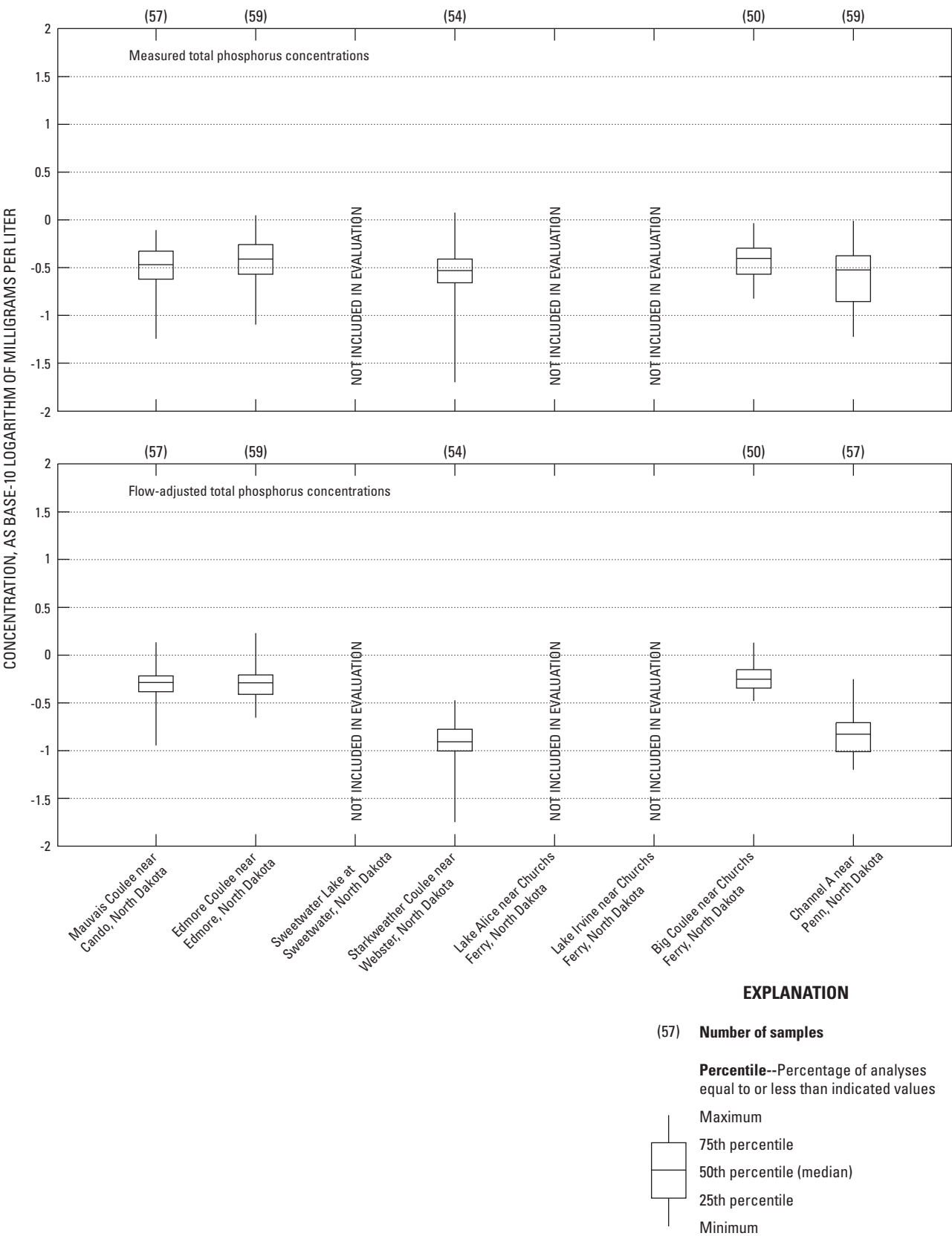


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.—Continued

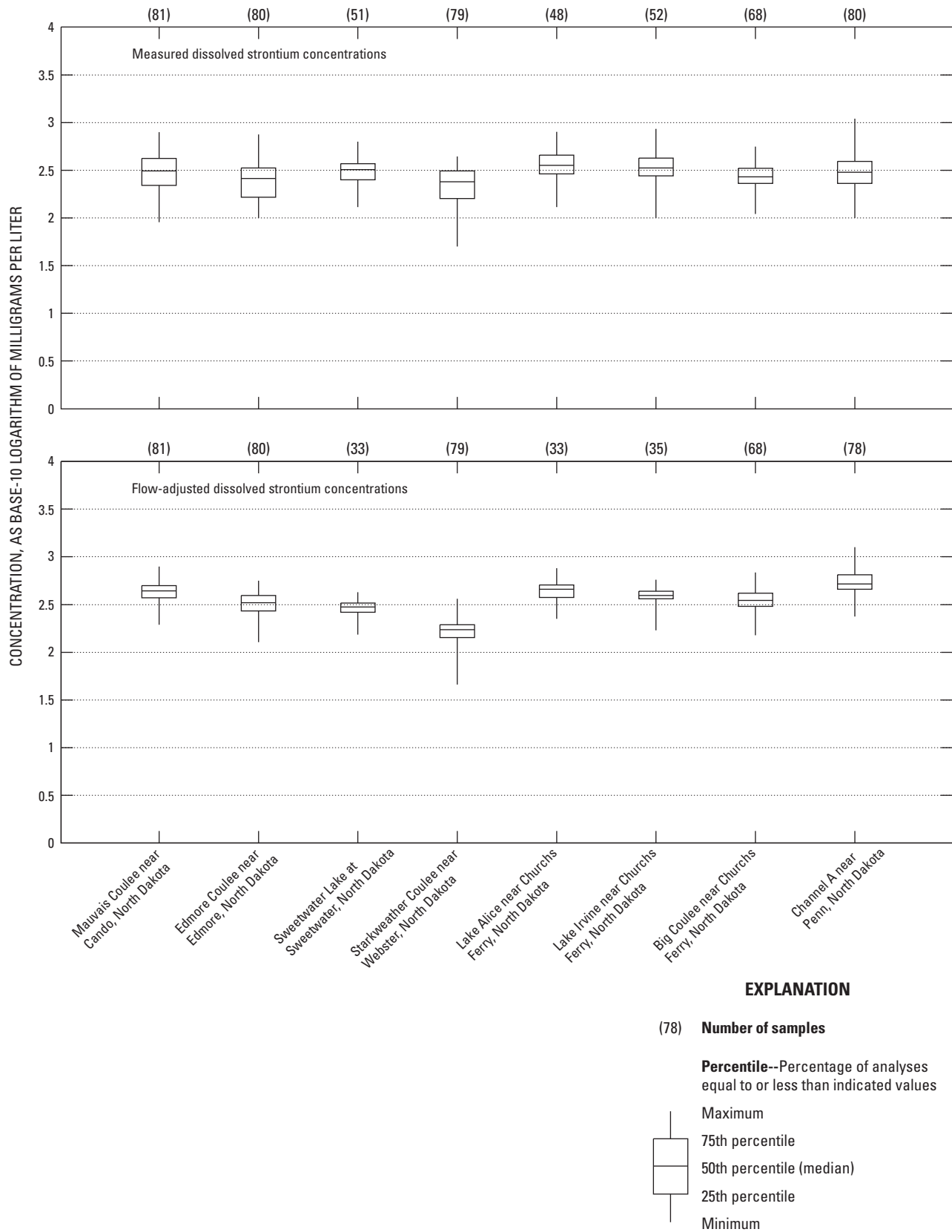


Figure 15. Measured and flow-adjusted concentrations for 1980 through September 2003 for stations used in water-quality trend analysis.—Continued

Water-Quality Trend Analysis

The constituents used for the water-quality trend analysis were evaluated for two specific trends, a step trend and a linear trend. The step trend was used to examine the effect of Channel A on water quality in the basin, and the linear trend was used to detect gradual changes with time from January 1980 through September 2003. Operation of Channel A began in 1979 to reduce spring-runoff-induced flooding around Dry Lake and around other lakes adjacent to Dry Lake by lowering the lake levels more rapidly than what would have occurred through Dry Lake's natural outlet (U.S. Army Corps of Engineers, 2003). Channel A connects Dry Lake to Sixmile Bay of Devils Lake, thus creating a shorter path to Devils Lake than through the upstream chain of lakes, Big Coulee, and Pelican Lake (fig. 2).

The constituents used for the water-quality trend analysis were evaluated by adding two trend variables to the flow-adjustment model—one for the step trend and one for the linear trend. The step trend variable represented whether the sample was collected before or after operation of Channel A began and was used only for those stations for which data are available before 1980. The step trend variable was included both for stations (such as Big Coulee near Churchs Ferry) that potentially were affected by Channel A and for stations (such as Edmore Coulee near Edmore) that were not affected by Channel A to evaluate whether trends unrelated to Channel A were confounded with the effect of Channel A.

Results of the trend analysis are given in table 4. For discussion purposes, a trend will be called significant if the p-value (attained significance level) for the trend is less than 0.05.

Dissolved Major Ions

The fitted step trends for flow-adjusted dissolved calcium concentrations (fig. 16, table 4) were small and insignificant. Therefore, the concentrations probably were not affected by Channel A or any other factor. The fitted linear trends for the upstream stream stations (sites 1, 2, and 4) also were small and insignificant. The upward trends that occurred for one lake station (site 5) and for the Big Coulee near Churchs Ferry and Channel A near Penn stations (sites 7 and 8, respectively) were small and insignificant, but the upward trends that occurred during 1980–2003 for the remaining lake stations (sites 3 and 6) were significant. Because no data are available for the Big Coulee station (site 7) for 1999–2003, some trends for that station may not have been detected. During 1999–2003, the concentrations for the Lake Irvine near Churchs Ferry station (site 6) were higher than the concentrations for the remaining stations. The differences among the fitted linear trends for the upstream stream stations and for the remaining stations indicate chemical or hydrologic processes occurring in the

lakes probably had a substantial effect on the concentrations. During the typical streamflow conditions that occurred during the 1970s and early 1980s, the concentrations for the Sweetwater Lake at Sweetwater and Lake Irvine stations (sites 3 and 6, respectively) generally were smaller than the concentrations for the Mauvais Coulee near Cando and Edmore Coulee near Edmore stations (sites 1 and 2, respectively). However, during the wet conditions from the mid-1990s to 2003, the concentrations for the lake stations were similar to the concentrations for the upstream stream stations.

The fitted step trends for flow-adjusted dissolved sulfate concentrations for the Lake Alice near Churchs Ferry, Lake Irvine, and Big Coulee stations (sites 5, 6, and 7, respectively) (fig. 17, table 4) were positive and similar in magnitude. Of the three upward trends, one, that for the Lake Irvine station (site 6), was significant. The upward trends for the three stations, along with the insignificant trends for the Mauvais Coulee, Edmore Coulee, and Sweetwater Lake stations (sites 1, 2, and 3, respectively), may indicate the concentrations for the Lake Alice, Lake Irvine, and Big Coulee stations (sites 5, 6, and 7, respectively) were affected by Channel A. The concentrations for the Edmore Coulee and Starkweather Coulee near Webster stations (sites 2 and 4, respectively) were considerably lower than the concentrations for the Mauvais Coulee station (site 1) so the lack of flow from the east to the west after 1979 probably caused the upward trends for sulfate. Downward linear trends occurred during 1980–2003 for most stations; however, the trends were not statistically significant. Because abundant sources of sulfur exist in soils and wetlands in the Devils Lake Basin, the decreasing concentrations may be related to changes in land use that may have reduced the amount of naturally occurring sulfate that reached the streams and lakes in the basin.

The fitted step trends for flow-adjusted dissolved chloride concentrations (fig. 18, table 4) were positive but insignificant. The upward trends for the Lake Alice and Big Coulee stations (sites 5 and 7, respectively) may indicate the concentrations for those stations were affected by Channel A. However, an upward trend also occurred for the Sweetwater Lake station (site 3) and concentrations for that station were not affected by Channel A. Therefore, although the results are inconclusive, the upward trends in 1979 may have been caused, in part, by factors other than Channel A. As for calcium, the fitted linear trends for the upstream stream stations (sites 1, 2, and 4) were small and insignificant. However, three of the downward trends that occurred during 1980–2003 for the remaining stations (those for sites 3, 7, and 8) were significant. The contrasting trends for calcium and chloride for the lake and downstream stream stations may indicate the chemical or hydrologic processes that caused increasing calcium concentrations also may have caused decreasing chloride concentrations. However, without a more detailed chemical, water-mass-balance model, the specific cause of the contrasting trends is difficult to determine.

Table 4. Stations and constituents used to test significance of Channel A step trend and linear trend from 1980 through September 2003.

[--, not tested for trend; <, less than; a p-value of less than 0.05 is significant]

Site number (figure 2)	U.S. Geological Survey station name	Coefficient for Channel A step trend	p-value for Channel A step trend	Coefficient for linear trend from 1980 through September 2003	p-value for linear trend from 1980 through September 2003
Calcium, dissolved					
1	Mauvais Coulee near Cando, North Dakota	0.020	0.622	0.001	0.751
2	Edmore Coulee near Edmore, North Dakota	-.069	.170	0	.924
3	Sweetwater Lake at Sweetwater, North Dakota	.018	.666	.011	.003
4	Starkweather Coulee near Webster, North Dakota	--	--	-.003	.360
5	Lake Alice near Churchs Ferry, North Dakota	.050	.240	.004	.279
6	Lake Irvine near Churchs Ferry, North Dakota	-.044	.292	.014	<.001
7	Big Coulee near Churchs Ferry, North Dakota	.012	.754	.005	.117
8	Channel A near Penn, North Dakota	--	--	.003	.608
Sulfate, dissolved					
1	Mauvais Coulee near Cando, North Dakota	-0.072	0.188	0.004	0.279
2	Edmore Coulee near Edmore, North Dakota	-.034	.670	-.008	.071
3	Sweetwater Lake at Sweetwater, North Dakota	.076	.102	-.005	.182
4	Starkweather Coulee near Webster, North Dakota	--	--	.001	.852
5	Lake Alice near Churchs Ferry, North Dakota	.113	.075	-.002	.719
6	Lake Irvine near Churchs Ferry, North Dakota	.090	.040	-.007	.062
7	Big Coulee near Churchs Ferry, North Dakota	.104	.053	-.005	.291
8	Channel A near Penn, North Dakota	--	--	-.002	.773
Chloride, dissolved					
1	Mauvais Coulee near Cando, North Dakota	0.062	0.377	-0.002	0.697
2	Edmore Coulee near Edmore, North Dakota	.046	.603	-.005	.286
3	Sweetwater Lake at Sweetwater, North Dakota	.059	.090	-.010	.001
4	Starkweather Coulee near Webster, North Dakota	--	--	.004	.478
5	Lake Alice near Churchs Ferry, North Dakota	.125	.066	-.012	.064
6	Lake Irvine near Churchs Ferry, North Dakota	.062	.249	-.007	.167
7	Big Coulee near Churchs Ferry, North Dakota	.105	.067	-.015	.001
8	Channel A near Penn, North Dakota	--	--	-.018	.002

Table 4. Stations and constituents used to test significance of Channel A step trend and linear trend from 1980 through September 2003.—Continued

[--, not tested for trend; <, less than; a p-value of less than 0.05 is significant]

Site number (figure 2)	U.S. Geological Survey station name	Coefficient for Channel A step trend	p-value for Channel A step trend	Coefficient for linear trend from 1980 through September 2003	p-value for linear trend from 1980 through September 2003
Nitrite plus nitrate as nitrogen, dissolved					
1	Mauvais Coulee near Cando, North Dakota	--	--	0.056	0.001
2	Edmore Coulee near Edmore, North Dakota	--	--	.016	.465
3	Sweetwater Lake at Sweetwater, North Dakota	0.078	0.734	.003	.867
4	Starkweather Coulee near Webster, North Dakota	--	--	-.026	.263
5	Lake Alice near Churchs Ferry, North Dakota	.422	.146	-.035	.092
6	Lake Irvine near Churchs Ferry, North Dakota	.110	.672	-.007	.757
7	Big Coulee near Churchs Ferry, North Dakota	-.115	.662	-.030	.100
8	Channel A near Penn, North Dakota	--	--	.082	.009
Ammonia, dissolved					
1	Mauvais Coulee near Cando, North Dakota	--	--	0.015	0.362
2	Edmore Coulee near Edmore, North Dakota	--	--	-.013	.278
4	Starkweather Coulee near Webster, North Dakota	--	--	-.039	.005
7	Big Coulee near Churchs Ferry, North Dakota	--	--	.026	.251
8	Channel A near Penn, North Dakota	--	--	.024	.466
Phosphorus, total					
1	Mauvais Coulee near Cando, North Dakota	--	--	0.053	0.030
2	Edmore Coulee near Edmore, North Dakota	--	--	.019	.203
4	Starkweather Coulee near Webster, North Dakota	--	--	.036	.159
7	Big Coulee near Churchs Ferry, North Dakota	--	--	-.013	.617
8	Channel A near Penn, North Dakota	--	--	.075	.003
Strontium, dissolved					
1	Mauvais Coulee near Cando, North Dakota	--	--	0.005	0.128
2	Edmore Coulee near Edmore, North Dakota	--	--	.005	.128
3	Sweetwater Lake at Sweetwater, North Dakota	-0.029	0.482	.005	.202
4	Starkweather Coulee near Webster, North Dakota	--	--	.007	.098
5	Lake Alice near Churchs Ferry, North Dakota	.058	.432	.006	.313

Table 4. Stations and constituents used to test significance of Channel A step trend and linear trend from 1980 through September 2003.—Continued

[--, not tested for trend; <, less than; a p-value of less than 0.05 is significant]

Site number (figure 2)	U.S. Geological Survey station name	Coefficient for Channel A step trend	p-value for Channel A step trend	Coefficient for linear trend from 1980 through September 2003	p-value for linear trend from 1980 through September 2003
Strontium, dissolved—Continued					
6	Lake Irvine near Churchs Ferry, North Dakota	-0.019	0.773	0.010	0.036
7	Big Coulee near Churchs Ferry, North Dakota	.073	.214	.006	.134
8	Channel A near Penn, North Dakota	--	--	-.005	.310

Nutrients

The fitted step trends for flow-adjusted dissolved nitrite plus nitrate concentrations (fig. 19, table 4) were insignificant. Therefore, either Channel A had no effect on the concentrations or the effect was too small to be detected. The fitted linear trends for the Mauvais Coulee near Cando and Channel A near Penn stations (sites 1 and 8, respectively) were significant. Upward trends occurred for those stations, but, because no data are available before 1987, the trends are for 1987-2003 rather than for 1980-2003. Most of the stations for which data are available before the mid-1980s had downward trends. Therefore, concentrations during recent years appear to be lower than those for the 1970s and early 1980s but higher than those for the late 1980s and early 1990s. Continued monitoring will be needed to determine if the recent trend toward higher concentrations continues in the future.

The trend results for flow-adjusted dissolved ammonia concentrations are shown in figure 20 and in table 4. Because no dissolved ammonia data are available before 1980, step trends were not included for that constituent. The fitted linear trends for the Mauvais Coulee, Big Coulee near Churchs Ferry, and Channel A stations (sites 1, 7, and 8, respectively) for 1980-2003 were small and insignificant. Upward trends occurred for those stations during 1980-2003, but downward trends occurred for the Edmore Coulee near Edmore and Starkweather Coulee near Webster stations (sites 2 and 4, respectively). The trend for the Starkweather Coulee station (site 4) was significant. From the mid- to late 1980s, the concentrations for the Channel A station (site 8) were substantially lower than the concentrations for the Edmore Coulee and Starkweather Coulee stations (sites 2 and 4, respectively). However, from the mid-1990s through 2003, the concentra-

tions for the Channel A station (site 8) were similar to the concentrations for the Edmore Coulee and Starkweather Coulee stations (sites 2 and 4, respectively). Therefore, chemical processes occurring in the lakes may have caused decreasing concentrations in Channel A during the late 1980s but not during the latter years of record. The Edmore Coulee and Starkweather Coulee stations (sites 2 and 4, respectively) are located upstream from the Channel A station (site 8).

The trend results for flow-adjusted total phosphorus concentrations are shown in figure 21 and in table 4. Because no total phosphorus data are available before 1989, step trends were not included for that constituent. The fitted linear trends for total phosphorus concentrations (fig. 21) for the Mauvais Coulee and Channel A stations (sites 1 and 8, respectively) were significant. Upward trends occurred from the late 1980s to 2003 for most stations, but a small and insignificant downward trend occurred for the Big Coulee station (site 7). However, few data are available for that station before 1993 and no data are available after 1998. Therefore, an upward trend may have occurred for that station, but the trend could not be detected. Continued monitoring will be needed to determine if the recent trend toward higher concentrations continues in the future.

Trace Element

The fitted step trends for flow-adjusted dissolved strontium concentrations (fig. 22, table 4) were small and insignificant. Therefore, the concentrations probably were not affected by Channel A. The fitted linear trends were insignificant except for the Lake Irvine near Churchs Ferry station (site 6). The upward trend for that station (site 6) was significant.

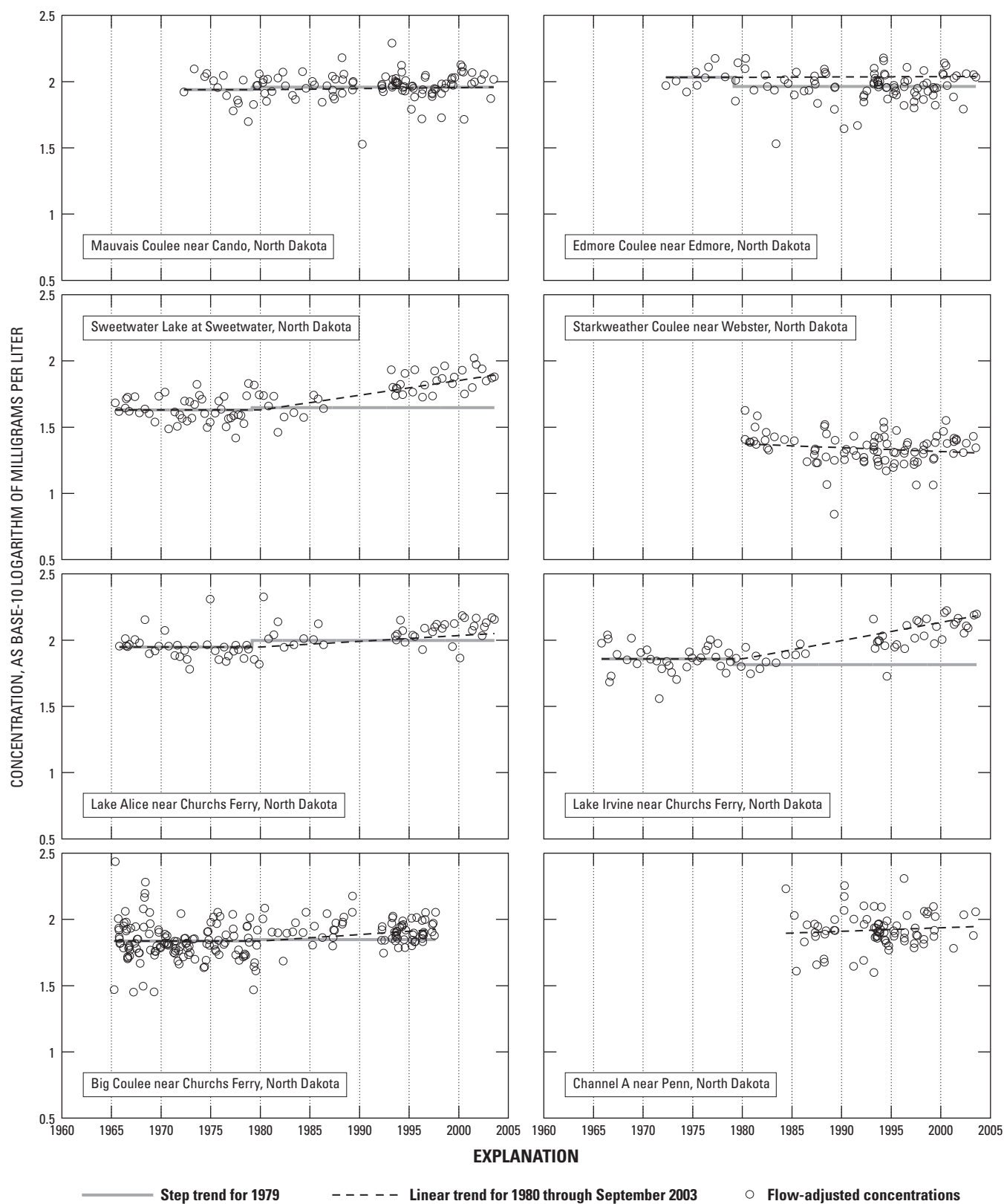


Figure 16. Fitted trends for flow-adjusted dissolved calcium concentrations.

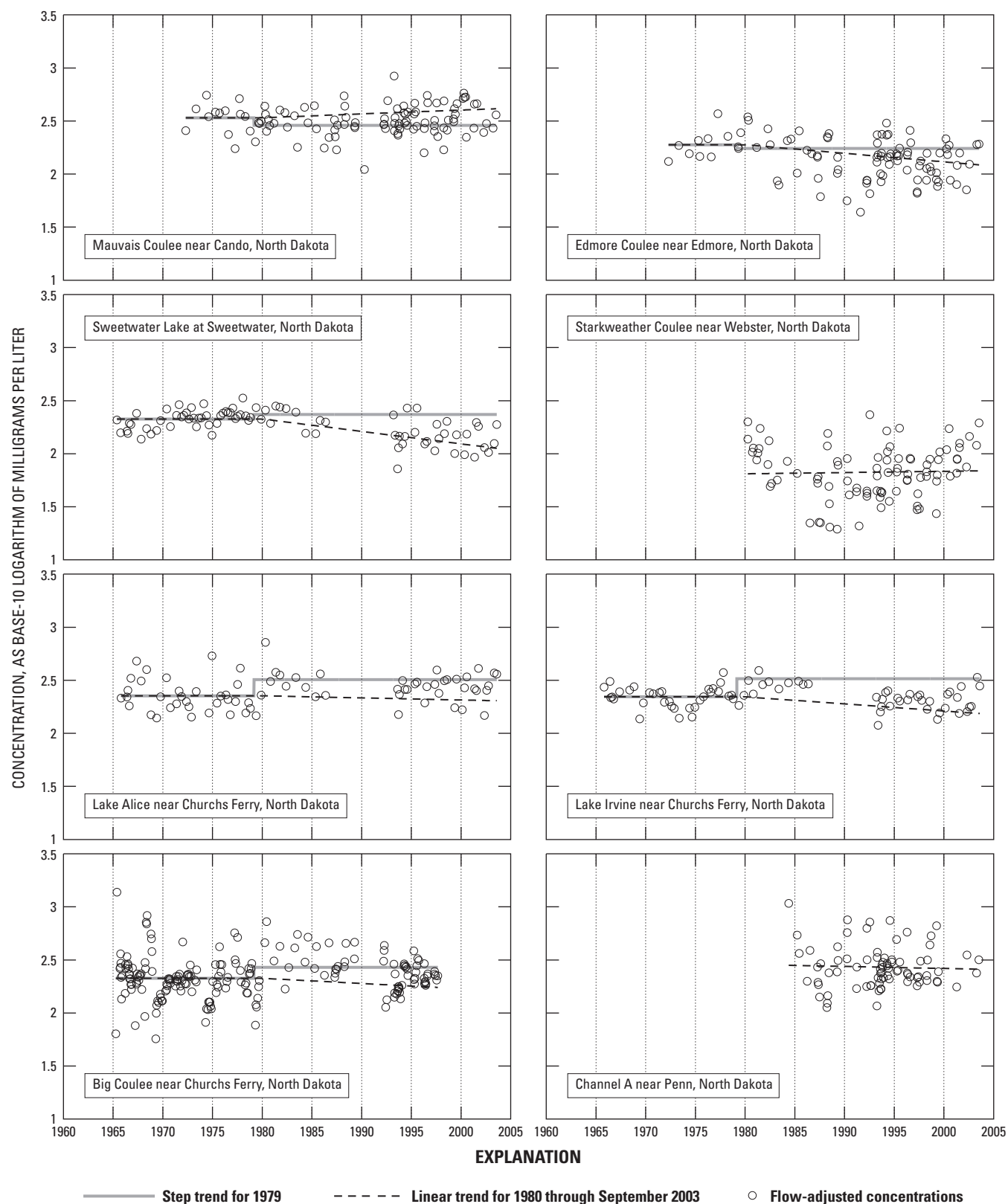


Figure 17. Fitted trends for flow-adjusted dissolved sulfate concentrations.

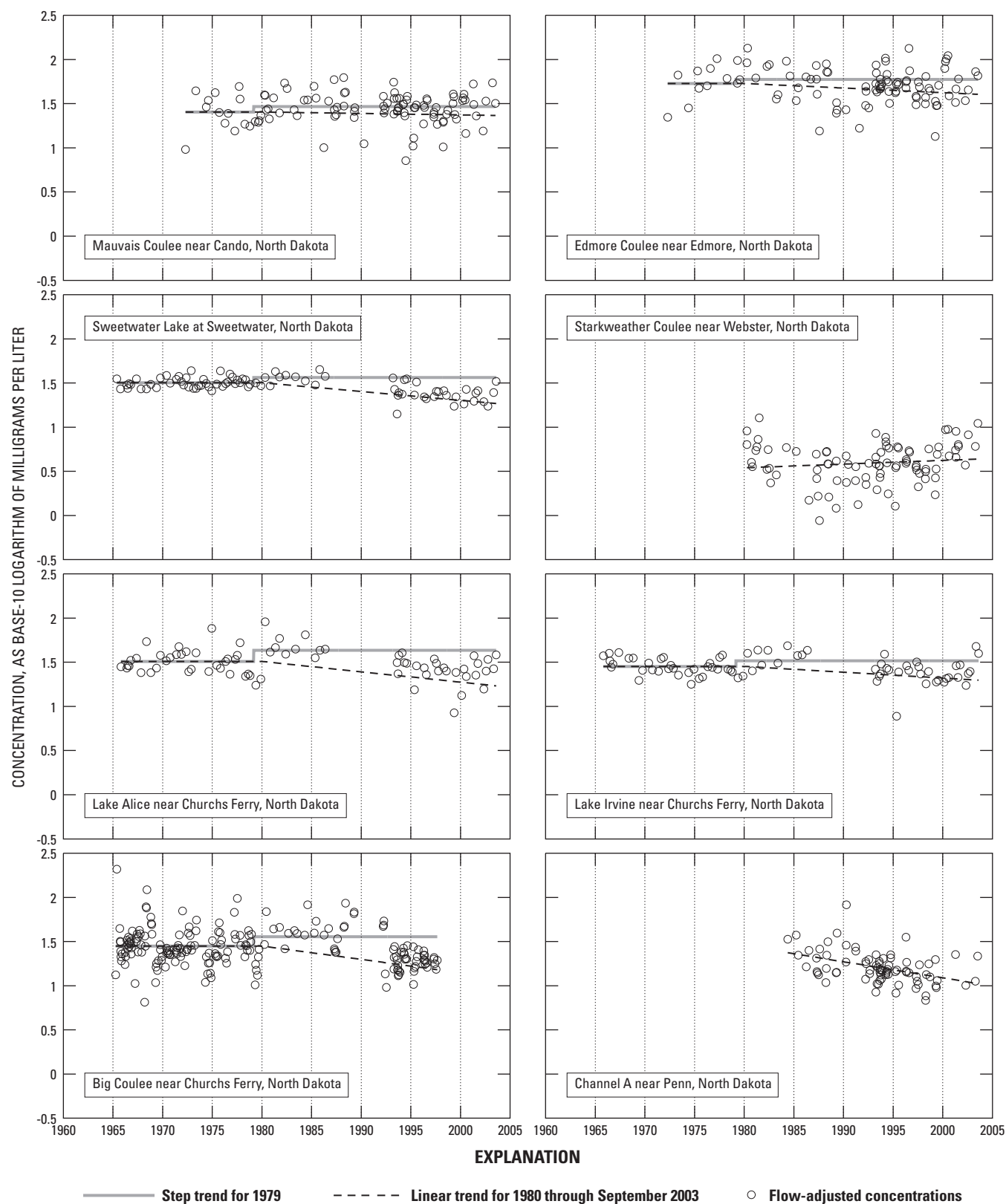


Figure 18. Fitted trends for flow-adjusted dissolved chloride concentrations.

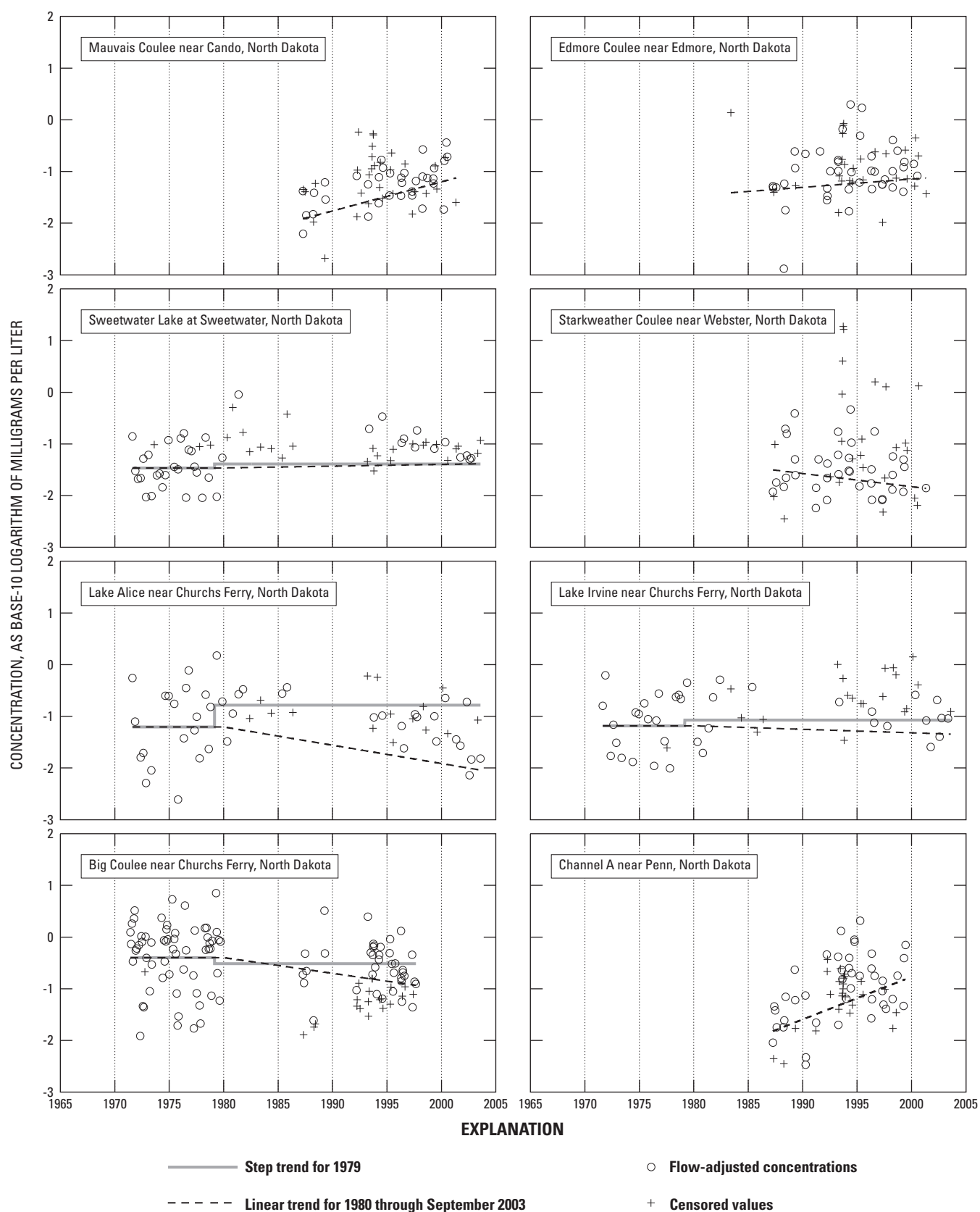


Figure 19. Fitted trends for flow-adjusted dissolved nitrite plus nitrate concentrations.

Sampling Design

The need for continued monitoring of water-quality trends in the upper Devils Lake Basin must be weighed against the cost of obtaining the monitoring data. Therefore, various sampling designs were evaluated with respect to two related concepts—sensitivity and efficiency. Sensitivity measures the ability of a design to detect a trend—the smaller the trend that can be detected, the more sensitive the design. An efficient design maximizes the sensitivity to detect a trend for a given cost, which for this report was measured in terms of the number of samples collected per year.

To evaluate sampling designs, the number of samples collected (on average) each year needs to be considered along with the sampling dates and whether the number of samples or sampling dates should be fixed (the same year after year) or variable. Because trends can occur anywhere, at any time, and the causes of the trends generally are not known in advance, the most efficient designs for monitoring trends in concentrations of multiple constituents generally are fixed designs in which the sampling dates and sampling locations are approximately the same year after year (Vecchia, 2005). Therefore, only fixed designs were considered for this report. The designs were evaluated with respect to their sensitivity to detect trends during two seasons—March through May, which generally includes the spring breakup and snowmelt season, and June through November, which generally includes the summer and fall low-flow season. Chemical source and transport mechanisms tend to differ for the two seasons, and, thus, trends that occur in one season may not necessarily occur in the other. For example, nutrients that accumulate during the winter, such as wastes from confined livestock operations, tend to reach streams in early spring. However, nutrients that are a result of fertilizer application or soil erosion tend to reach streams in late spring or early summer when streamflow consists mostly of rainfall runoff. Other constituents, such as dissolved chloride or dissolved nitrite plus nitrate, may originate from ground-water sources and, thus, may require monitoring during summer and fall low-flow conditions.

For this report, a trend in concentration was defined as a change in the mean of the log-transformed concentrations. Using the null hypothesis that the mean of log-transformed concentrations is constant, the log-transformed concentrations were expressed as follows:

$$\log C(t) = \mu + E[\log C(t) - \mu \mid \log Q(u), u \leq t] + \varepsilon(t)$$

where

\log denotes the base-10 logarithm;
 $C(t)$ is the concentration, in milligrams or micrograms per liter, for time t , in decimal years;

μ is the mean (expected value) of the log-transformed concentrations, as the base-10 logarithm of milligrams or micrograms per liter;
 $E[\]$ is the conditional expectation of $\log C(t) - \mu$ given log-transformed streamflow for all times μ up to and including time t (dimensionless);
 Q is streamflow, in cubic feet per second, for time u ;
 u is time, in decimal years;
 and
 $\varepsilon(t)$ is the residual, which, by definition, is uncorrelated with current and past values of log-transformed streamflow (dimensionless).

Flow-adjusted concentrations were expressed as follows:

$$FAC(t) = \log C(t) - E[\log C(t) - \mu \mid \log Q(u), u \leq t] = \mu + \varepsilon(t) \quad (2)$$

where

$FAC(t)$ is the flow-adjusted concentration, as the base-10 logarithm of milligrams or micrograms per liter for time t , in decimal years.

Because the mean of the log-transformed concentrations is identical to the mean of the flow-adjusted concentrations, the trends in log-transformed concentrations are equivalent to the trends in flow-adjusted concentrations. Trends are easier to detect using flow-adjusted concentrations rather than log-transformed concentrations because flow-adjusted concentrations generally have less variability than log-transformed concentrations.

For this report, the following statistic was used to determine potential sampling designs:

$$X_{n,d} = \text{Ave}\{FAC(u_j), j = 1, 2, \dots, nd\} \quad (3)$$

where

X is a statistic used to detect trends in concentrations;
 n is the number of samples per year;
 d is the duration, in years, of the time interval used to compute the statistic (the beginning and ending times of the interval are not important);
 Ave is the average of the values in braces;
 and
 u_1, u_2, \dots, u_{nd} are the specific times of sample collection.

An increase in d increases the sensitivity of a sampling design but not the efficiency of the design. Therefore, $d = 3$ was selected as a benchmark to compare the efficiency of various designs. Substituting equation 2 into equation 3 and setting $d = 3$ yielded the following equation:

$$X_{n,3} = \mu + E_{n,3} \quad (4)$$

where

$E_{n,3} = \text{Ave} \{ \varepsilon(u_j), j = 1, 2, \dots, 3n \}$ is the estimation error.

The sensitivity of a design for detecting a trend depends on the variance of the estimation error, which, in turn, depends on the variance and serial correlation of the errors in the flow-adjusted concentrations. Therefore, variogram analyses were

used to explore the variance and serial correlation structure of the errors in flow-adjusted concentrations (see appendix). The errors generally were uncorrelated for time lags of at least 2 weeks for season 1 and 6 weeks for season 2. Furthermore, the error variance was constant within each of two seasons—March through May and June through November. However, for some stations and some constituents, substantial differences occurred in the error variance between the two seasons. Therefore, provided that water-quality samples are collected at least 2 weeks apart for season 1 and 6 weeks apart for season 2,

$$\text{Var}\{E_{n,3}\} = \frac{n_1 v_1 + n_2 v_2}{3(n_1 + n_2)^2} \quad (5)$$

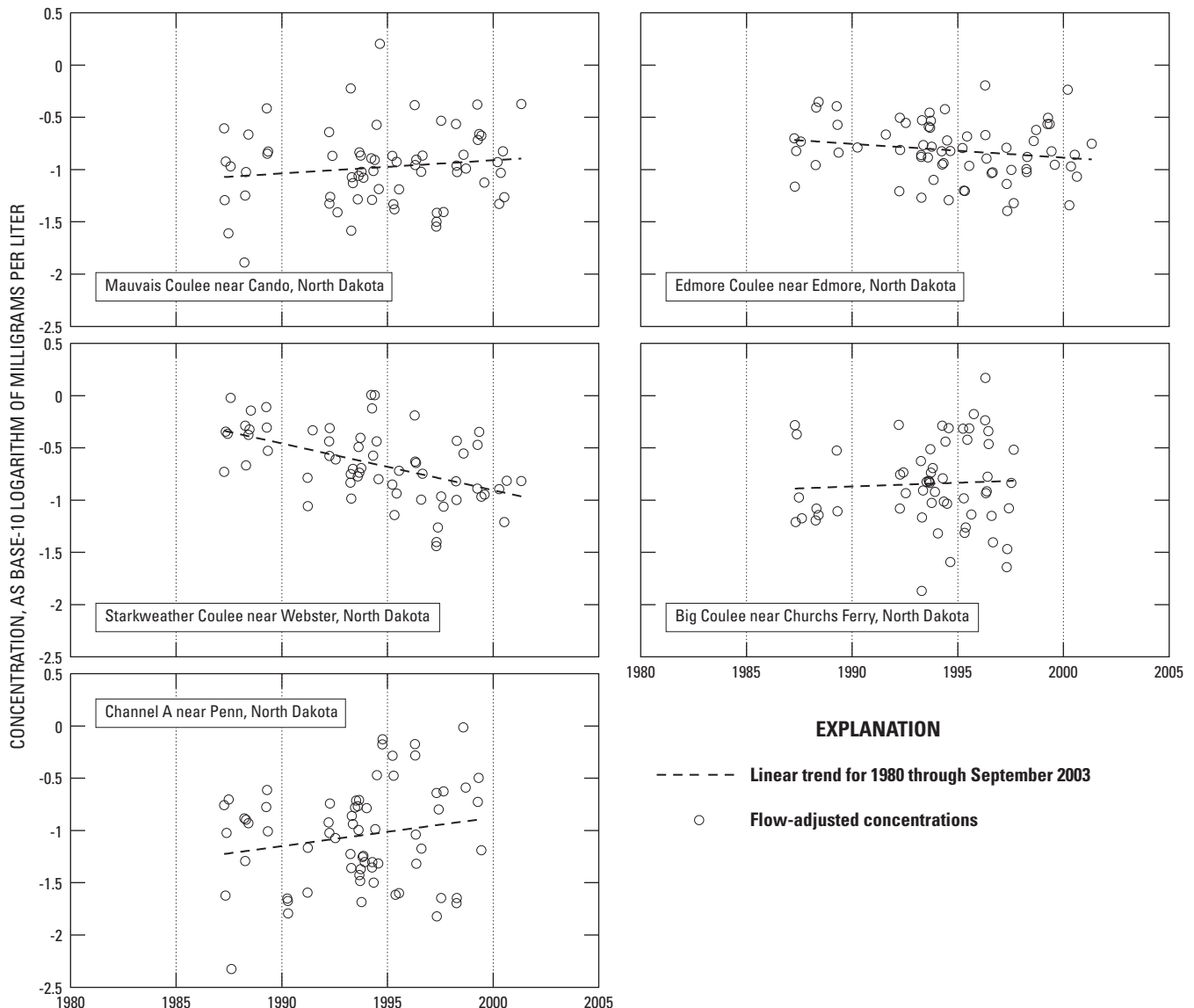


Figure 20. Fitted trends for flow-adjusted dissolved ammonia concentrations.

where

$\text{Var}\{ \}$ is the variance of the term in braces,
 n_1 is the number of samples per year for season 1
(March through May),
 v_1 is the variance of $\varepsilon(t)$ for t in season 1,
 n_2 is the number of samples per year for season 2
(June through November),

and

v_2 is the variance of $\varepsilon(t)$ for t in season 2.

For the sample design analysis, design efficiencies were evaluated separately for each season to monitor trends during both the snowmelt-runoff season (season 1) and the summer and fall low-flow season (season 2). Design efficiencies were not evaluated for December through February because too few data are available for those months. Evaluating equation 5 separately for each season and omitting the subscripts yielded the following equation:

$$\text{Var}(E) = \frac{v}{3n} \quad (6)$$

where

E is the estimation error for either season 1 or
season 2,
 v is the variance of the flow-adjusted
concentrations for either season 1 or
season 2,

and

n is the number of samples per year for either
season 1 or season 2.

For equation 6, samples were assumed to be collected at least 2 weeks apart for season 1 and 6 weeks apart for season 2. At most, 7 equally spaced samples per year for season 1 and 6 equally spaced samples per year for season 2 can be collected to satisfy this assumption. If more samples are collected, equation 6 should not be used to evaluate design efficiencies.

For future water-quality monitoring, a goal was set to have the estimated median concentration for any 3-year interval be within 25 percent of the actual median without increasing sampling costs. Using equation 4 and the log-transformed concentrations, the estimated median concentration was $10^{\mu + E}$ and the actual median concentration was 10^{μ} . Therefore, the probability that the estimated median was within 25 percent of the actual median was given by the following equation:

$$\begin{aligned} & \text{Prob} \left\{ -25 < 100 \left[\frac{10^{\mu + E} - 10^{\mu}}{10^{\mu}} \right] < 25 \right\} \\ & = \text{Prob} \{ -25 < 100(10^E - 1) < 25 \} \end{aligned} \quad (7)$$

where

$\text{Prob}\{ \}$ is the probability of the event in braces.

Assuming E is approximately normally distributed and substituting equation 6 into equation 7 yielded the following equation:

$$\begin{aligned} \text{Prob} \{ -25 < PE < 25 \} = & \text{Prob} \left[Z < \log(1.25) \left(\frac{3n}{v} \right)^{\frac{1}{2}} \right] \\ & - \text{Prob} \left[Z < \log(0.75) \left(\frac{3n}{v} \right)^{\frac{1}{2}} \right] \end{aligned} \quad (8)$$

where

PE is the percent error,

and

Z is a standard normal random variable.

For this report, the smallest integer value for n was determined such that the probability (eq. 8) was at least 0.8 or 0.6—that is, for which at least an 80 or 60 percent chance existed that the estimated median was within the error tolerance. A separate determination for n was made for season 1 (that is, assuming all samples were collected from March through May) and for season 2 (that is, assuming all samples were collected from June through November) so the same error tolerance was achieved for both seasons.

Because the sampling specifications used to develop the sampling designs for this report are lenient, the sampling designs may not be sensitive enough for some potential applications. For example, a study to evaluate the mass balance of certain chemical constituents may require a larger number of samples (thus increasing the cost) than indicated in this report. Similarly, a study to evaluate the changes in and causes of the complex relation between antecedent streamflow and water quality would require a larger number of samples than indicated in this report. Therefore, given the constraint that sampling costs remain at or below those in the 1992–2003 sampling program, an 80- or 60-percent chance that the estimated median was within 25 percent of the actual median was a realistic goal.

The number of samples required per year for each constituent is given in table 5. Dissolved nitrite plus nitrate was not included because of the large percentage of censored values for that constituent. Rather, dissolved ammonia was used as a surrogate for dissolved nitrite plus nitrate in the design analysis. Design results for the Lake Irvine near Churchs Ferry station (site 6) also were not included because those results were similar to the results for the Lake Alice near Churchs Ferry station (site 5).

The smallest numbers of samples required per year were for dissolved calcium and dissolved strontium (table 5). To obtain at least an 80-percent chance of being within the error tolerance for those constituents, one sample per year per season generally was sufficient. The numbers of samples required per year for dissolved sulfate and dissolved chloride were similar. To obtain at least an 80-percent chance of being within the error tolerance, two or three samples per year per season generally were sufficient. However, for the Starkweather Coulee station (site 4), three samples were required per year for season 2 to obtain at least a 60-percent chance of being within the error tolerance for dissolved sulfate. The largest numbers of samples required per year were for dissolved ammonia. To obtain at least an 80-percent chance of being within the error tolerance for that constituent, 4 to 11 samples were required

per year for season 1, and 3 to 14 samples were required per year for season 2. If the chance of being within the error tolerance is lowered to 60 percent, then two to five samples were required per year for season 1, and two to six samples were required per year for season 2. To obtain at least an 80-percent chance of being within the error tolerance for total phosphorus, two samples were required per year for season 1 and one sample generally was required per year for season 2. Exceptions were for the Starkweather Coulee near Webster station (site 4), for which four samples were required per year for season 2, and the Channel A near Penn station (site 8), for which three samples were required per year for season 2. For those stations, two samples were required per year for season 2 to obtain at least a 60-percent chance of being within the error tolerance.

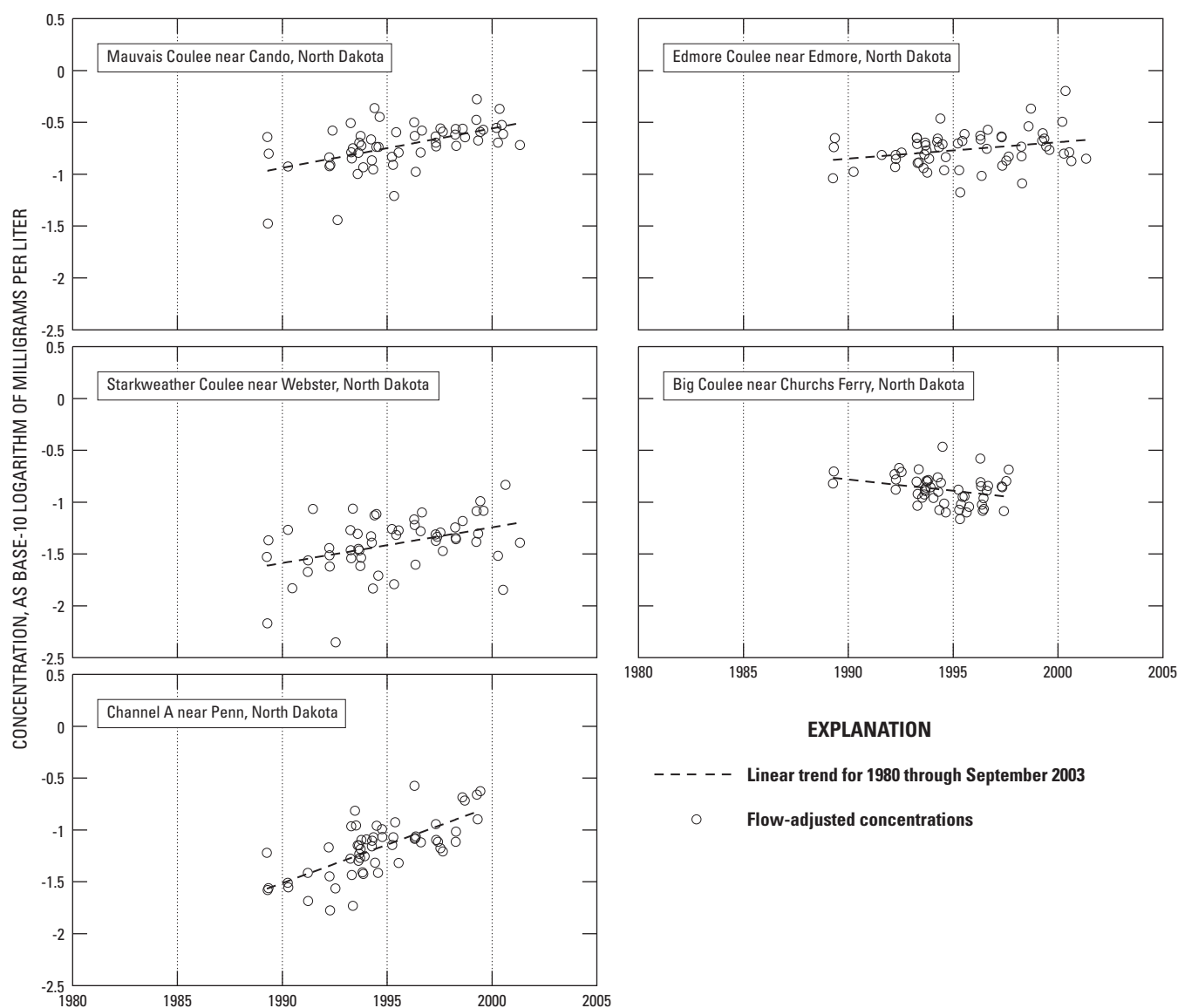


Figure 21. Fitted trends for flow-adjusted total phosphorus concentrations.

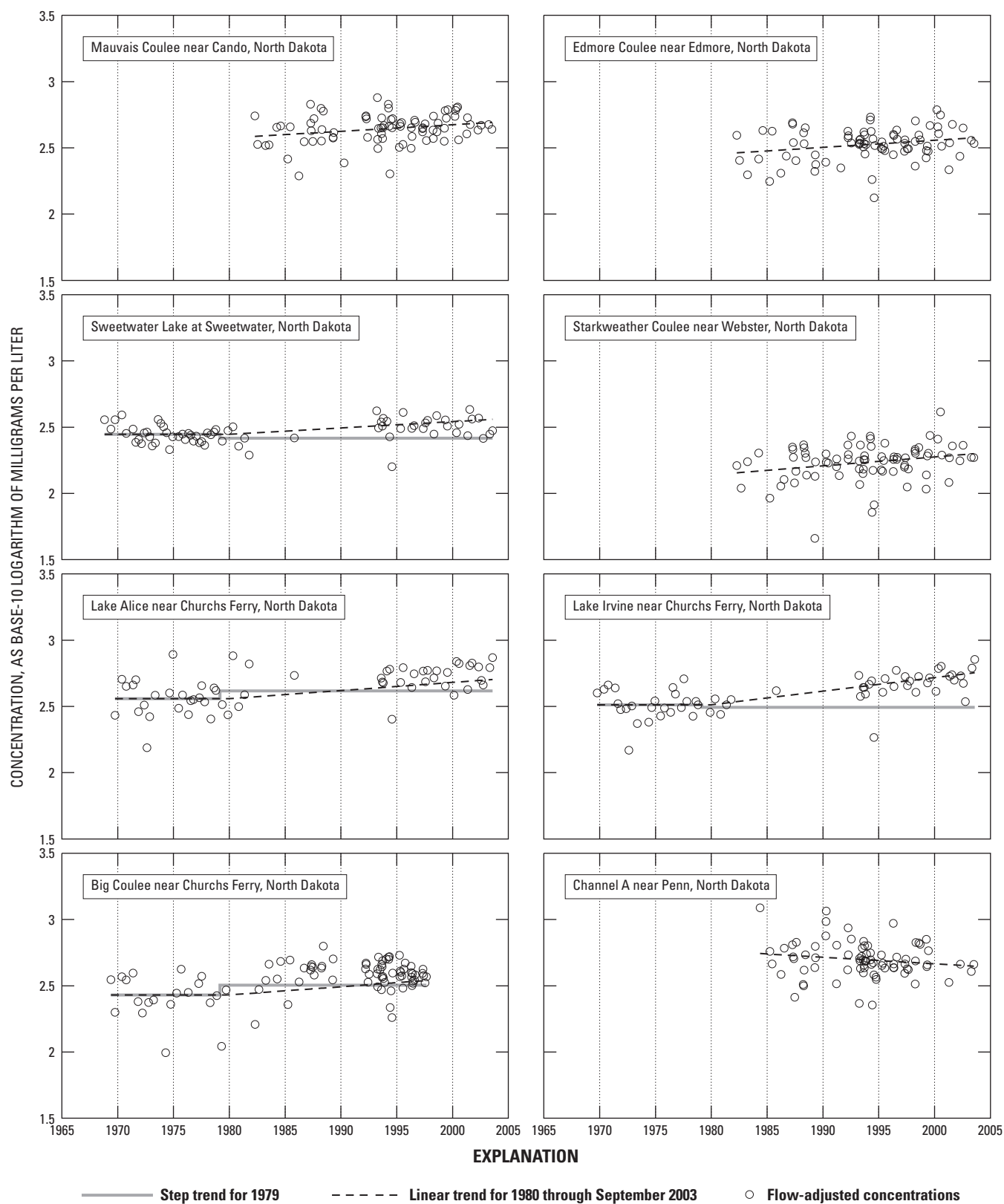


Figure 22. Fitted trends for flow-adjusted dissolved strontium concentrations.

The collection of water-quality samples can be costly. Therefore, an efficient plan is needed for the collection of samples that are representative of various seasonal and hydrologic conditions in the basin. Results of this study indicate a good overall design for the stations used in the water-quality trend analysis might consist of six samples per year per station, with three samples collected during season 1 and three samples collected during season 2. Except for dissolved ammonia for the Big Coulee near Churchs Ferry and Channel A stations (sites 7 and 8, respectively), this six-sample design generally resulted in at least a 60-percent chance, and in most cases at least an 80-percent chance, of being within the error tolerance for all constituents. For dissolved ammonia for the Big Coulee station (site 7), four samples were required per year for season 1 to obtain a 60-percent chance of being within the error tolerance. For dissolved ammonia for the Channel A station (site 8), five samples were required per year for season 1 and six samples were required per year for season 2 to obtain a 60-percent chance of being within the error tolerance.

Although constituents were evaluated separately for each season, in any given year a smooth transition occurs from season 1 to season 2 when, in late May or early June, streamflows change from primarily snowmelt to primarily rainfall runoff. Therefore, a sample collected in late May or early June generally could be considered “shared” between the two seasons, thus reducing the number of samples required per year from six to five. To minimize the potential effects of serial correlation, the remaining samples should be collected as far apart as possible within each season. Therefore, a five-sample design with samples collected in about early April, May, and June; late July or early August; and late September or early October would be a reasonable design for all stations and constituents. Except for dissolved ammonia and probably dissolved nitrite plus nitrate, this five-sample design generally would result in at least an 80-percent chance of being within the error tolerance for most stations. For dissolved ammonia and dissolved nitrite plus nitrate, the five-sample design generally would result in at least a 60-percent chance of being within the error tolerance for most stations.

Assuming a five-sample design, as described previously, sampling possibly could be discontinued at some of the existing stream or lake stations or initiated at additional stations. In the USGS lake-sampling program for 1992-2003, four major-ion, nutrient, and trace-element samples were collected per year for each of five lake stations (Sweetwater Lake, Morrison Lake, Dry Lake, Lake Alice, and Lake Irvine) (Ryberg and others, 2005). In addition, four major-ion, nutrient, and trace-element samples were collected per year for each of three stream stations (Mauvais Coulee, Edmore Coulee, and Starkweather Coulee), and two major-ion and trace-element samples were collected per year for two stream stations (Edmore Coulee Tributary and Mauvais Coulee Tributary) (Ryberg and others, 2005). For two of the lake stations (Morrison Lake and Dry Lake), no data are available before 1992. Therefore, those stations were not included in the trend analysis but were included in the comparison of data to determine

at which stations sampling possibly could be discontinued or initiated. Concentration data for the five stream stations and the five lake stations were compared using the graphical user interface described by Ryberg and others (2005). Results of the comparisons indicate sampling could be discontinued at two lake stations (Sweetwater Lake and Lake Irvine) because of similarities with other lake stations (Morrison Lake and Lake Alice, respectively). In addition, sampling could be discontinued at two stream stations (Edmore Coulee Tributary and Mauvais Coulee Tributary) because of redundancy of data with data for other stream and/or lake stations. For example, concentration data for the Edmore Coulee station and the Morrison Lake station can be used to determine if water-quality trends occur for Edmore Coulee Tributary (fig. 2). In 1997, the long-term stream gage for the Big Coulee station (site 7) was inundated with water from Devils Lake, and, in 1979, the long-term stream gage for the Channel A station (site 8) was inundated. Therefore, water-quality sampling has been discontinued at those stations. However, concentration data for Big Coulee were closely correlated with concentration data for Lake Alice, and concentration data for Channel A were closely correlated with concentration data for Dry Lake.

The previous discussion indicates that, for continued monitoring of water-quality trends in the upper Devils Lake Basin, an efficient sampling design consists of five major-ion, nutrient, and trace-element samples per year (collected in about early April, May and June; late July or early August; and late September or early October) at three existing stream stations (Mauvais Coulee, Edmore Coulee, and Starkweather Coulee) and at three existing lake stations (Morrison Lake, Dry Lake, and Lake Alice). This sampling design requires the collection of 15 stream samples and 15 lake samples per year rather than 16 stream samples and 20 lake samples per year as in the 1992-2003 program. Thus, the design would result in a program that is less costly and more efficient than the 1992-2003 program but that still would provide the data needed to monitor water-quality trends in the Devils Lake Basin.

Summary

This report presents the results of a study conducted by the U.S. Geological Survey, in cooperation with the North Dakota State Water Commission, the Devils Lake Basin Joint Water Resource Board, and the Red River Joint Water Resource District, to analyze historical water-quality trends in three dissolved major ions (calcium, sulfate, and chloride), three nutrients (dissolved nitrite plus nitrate as nitrogen, dissolved ammonia, and total phosphorus), and one dissolved trace element (strontium) for eight stations in the Devils Lake Basin in North Dakota and to develop an efficient sampling design to monitor the future trends. Streamflows for the Mauvais Coulee near Cando, North Dakota, and Edmore Coulee near Edmore, North Dakota, stations were selected as

being representative of streamflows in the Devils Lake Basin because of the long period of record and geographic location of each station.

A multiple-regression model was used to detect and remove streamflow-related variability in constituent concentrations. To separate the natural variability in concentration as a result of variability in streamflow from the variability in concentration as a result of other factors, the base-10 logarithm of daily streamflow for each station was divided into four components—a 5-year streamflow anomaly, an annual streamflow anomaly, a seasonal streamflow anomaly, and a daily streamflow anomaly. The constituent concentrations then were adjusted for streamflow-related variability by removing the 5-year, annual, seasonal, and daily variability.

Much of the variability in the fitted dissolved nitrite plus nitrate concentrations probably can be attributed to streamflow-related variability. However, because of the large percentage of censored values for this constituent, a survival regression technique was used to fit the model rather than the multiple-regression model used for the remaining constituents.

Constituents used for the water-quality trend analysis were evaluated for a step trend to examine the effect of Channel A on water quality in the basin and a linear trend to detect gradual changes with time from January 1980 through September 2003. The fitted upward linear trends for dissolved calcium concentrations during 1980-2003 for two stations were significant. The fitted step trends for dissolved sulfate concentrations for three stations were positive and similar in magnitude. Of the three upward trends, one was significant. The fitted step trends for dissolved chloride concentrations

were positive but insignificant. The fitted linear trends for the upstream stations were small and insignificant, but three of the downward trends that occurred during 1980-2003 for the remaining stations were significant. The fitted upward linear trends for dissolved nitrite plus nitrate as nitrogen concentrations during 1987-2003 for two stations were significant. However, concentrations during recent years appear to be lower than those for the 1970s and early 1980s but higher than those for the late 1980s and early 1990s. The fitted downward linear trend for dissolved ammonia concentrations for one station was significant. The fitted linear trends for total phosphorus concentrations for two stations were significant. Upward trends occurred from the late 1980s to 2003 for most stations, but a small and insignificant downward trend occurred for one station. The fitted step trends for dissolved strontium concentrations were small and insignificant. The fitted linear trends were insignificant except for one station. Continued monitoring will be needed to determine if the recent trend toward higher dissolved nitrite plus nitrate as nitrogen and total phosphorus concentrations continues in the future.

For continued monitoring of water-quality trends in the upper Devils Lake Basin, an efficient sampling design consists of five major-ion, nutrient, and trace-element samples per year at three existing stream stations and at three existing lake stations. This sampling design requires the collection of 15 stream samples and 15 lake samples per year rather than 16 stream samples and 20 lake samples per year as in the 1992-2003 program. Thus, the design would result in a program that is less costly and more efficient than the 1992-2003 program but that still would provide the data needed to monitor water-quality trends in the Devil Lake Basin.

Table 5. Number of samples required to obtain specified error tolerance for estimated median concentrations.

Probability that estimated median for a 3-year interval is within 25 percent of actual median	Season ¹	Site number	U.S. Geological Survey station name	Number of samples required per year for						
				Dissolved calcium	Dissolved sulfate	Dissolved chloride	Dissolved ammonia	Total phosphorus	Dissolved strontium	
0.8	1	1	Mauvais Coulee near Cando, North Dakota	1	1	2	5	2	1	
	1	2	Edmore Coulee near Edmore, North Dakota	1	2	2	4	2	1	
	1	3	Sweetwater Lake at Sweetwater, North Dakota	1	1	1	--	--	1	
	1	4	Starkweather Coulee near Webster, North Dakota	1	2	2	6	2	1	
	1	5	Lake Alice near Churchs Ferry, North Dakota	1	1	2	--	--	1	
	1	7	Big Coulee near Churchs Ferry, North Dakota	1	1	2	92	2	1	
	1	8	Channel A near Penn, North Dakota	1	2	2	11 ²	2	1	
	2	1	Mauvais Coulee near Cando, North Dakota	1	1	1	3	1	1	
	2	2	Edmore Coulee near Edmore, North Dakota	1	2	1	3	1	1	
	2	3	Sweetwater Lake at Sweetwater, North Dakota	1	1	1	--	--	1	
	2	4	Starkweather Coulee near Webster, North Dakota	1	6	3	3	4	1	
	2	5	Lake Alice near Churchs Ferry, North Dakota	1	1	1	--	--	1	
	2	7	Big Coulee near Churchs Ferry, North Dakota	1	2	3	6	1	1	
	2	8	Channel A near Penn, North Dakota	1	2	1	14 ²	3	1	
	1	1	Mauvais Coulee near Cando, North Dakota	1	1	1	3	1	1	
	1	2	Edmore Coulee near Edmore, North Dakota	1	1	1	2	1	1	
0.6	1	3	Sweetwater Lake at Sweetwater, North Dakota	1	1	1	--	--	1	
	1	4	Starkweather Coulee near Webster, North Dakota	1	1	1	3	1	1	
	1	5	Lake Alice near Churchs Ferry, North Dakota	1	1	1	--	--	1	
	1	7	Big Coulee near Churchs Ferry, North Dakota	1	1	1	4	1	1	
	1	8	Channel A near Penn, North Dakota	1	1	1	5	1	1	
	2	1	Mauvais Coulee near Cando, North Dakota	1	1	1	3	1	1	
	2	2	Edmore Coulee near Edmore, North Dakota	1	2	1	3	1	1	
	2	3	Sweetwater Lake at Sweetwater, North Dakota	1	1	1	--	--	1	

Table 5. Number of samples required to obtain specified error tolerance for estimated median concentrations.—Continued

Probability that estimated median for a 3-year interval is within 25 percent of actual median	Season ¹	Site number	U.S. Geological Survey station name	Number of samples required per year for						
				Dissolved calcium	Dissolved sulfate	Dissolved chloride	Dissolved ammonia	Total phosphorus	Dissolved strontium	
	2	1	Mauvais Coulee near Cando, North Dakota	1	1	1	2	1	1	
	2	2	Edmore Coulee near Edmore, North Dakota	1	1	1	2	1	1	
	2	3	Sweetwater Lake at Sweetwater, North Dakota	1	1	1	--	--	1	
	2	4	Starkweather Coulee near Webster, North Dakota	1	3	1	2	2	1	
	2	5	Lake Alice near Churchs Ferry, North Dakota	1	1	1	--	--	1	
	2	7	Big Coulee near Churchs Ferry, North Dakota	1	1	2	3	1	1	
	2	8	Channel A near Penn, North Dakota	1	1	1	6	2	1	

¹Season 1 is March through May, and season 2 is June through November.²Minimum spacing between samples is less than 2 weeks for season 1 and less than 6 weeks for season 2.

References Cited

- Bluemle, J.P., 1991, Radiocarbon dating of beaches and outlets of Devils Lake: North Dakota Geological Survey Miscellaneous Series 75, 10 p.
- Cressie, N., 1991, Statistics for spatial data: Wiley, New York, 900 p.
- Cressie, N., and Hawkins, D.M., 1980, Robust estimation of the variogram: Journal of the International Association for Mathematical Geology, v. 12, p. 115-125.
- Helsel, D.R., 2005, Nondetects and data analysis: Hoboken, New Jersey, John Wiley & Sons, Inc., 250 p.
- Helsel, D.R., and Hirsch, R.M., 1995 (3rd printing), Studies in environmental science 49—Statistical methods in water resources: Amsterdam, Elsevier Science B.V., 529 p.
- Insightful Corporation, 2001, Overview of survival analysis, in S-PLUS 6 for Windows guide to statistics, volume 2: Seattle, Washington, Insightful Corporation, p. 297-309.
- Murphy, E.C., Fritz, A.K., and Fleming, R.F., 1997, The Jerusalem and Tolna outlets in the Devils Lake Basin, North Dakota: North Dakota Geological Survey Report of Investigation No. 100, 36 p.
- Neter, J., Kutner, M.H., Nachtsheim, C.J., and Wasserman, W., 1996, Applied linear statistical models, Fourth Edition: Boston, WCB McGraw-Hill, 1408 p.
- North Dakota State Water Commission, n.d., Outlet quick facts: accessed October 17, 2005, at http://www.swc.state.nd.us/projects/newdevilslake/pdfs/Outlet_Quick_Facts.pdf
- Ryberg, K.R., Vecchia, A.V., and Damschen, W.C., 2005, Graphical user interface for accessing water-quality data for the Devils Lake Basin, North Dakota, 2005: U.S. Geological Survey Open-File Report 2005-1419, online at <http://pubs.water.usgs.gov/ofr20051419>
- Sether, B.A., Vecchia, A.V., and Berkas, W.R., 1999, Spatial and temporal variability of dissolved sulfate in Devils Lake, North Dakota, 1998: U.S. Geological Survey Fact Sheet FS-096-99, 4 p., online at <http://nd.water.usgs.gov/pubs/fs/fs09699/index.html>
- U.S. Army Corps of Engineers, 2003, Final - Devils Lake, North Dakota Integrated Planning Report and Environmental Impact Statement, Volume 1, Chapter 3, online at <http://www.health.state.nd.us/PublicComment/DevilsLakeOutletDischargePermit/DevilsLakeEIS/>
- U.S. Geological Survey, various dates, Water analysis: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chaps. A1-A6, variously paginated, accessed February 7, 2006, at <http://pubs.usgs.gov/twri/>
- Vecchia, A.V., 2000, Water-quality trend analysis and sampling design for the Souris River, Saskatchewan, North Dakota, and Manitoba: U.S. Geological Survey Water-Resources Investigations Report 00-4019, 77 p., online at <http://nd.water.usgs.gov/pubs/wri/wri004019/index.html>
- Vecchia, A.V., 2003, Water-quality trend analysis and sampling design for streams in North Dakota, 1971-2000: U.S. Geological Survey Water-Resources Investigations Report 03-4094, 73 p., online at <http://nd.water.usgs.gov/pubs/wri/wri034094/index.html>
- Vecchia, A.V., 2005, Water-quality trend analysis and sampling design for streams in the Red River of the North Basin, Minnesota, North Dakota, and South Dakota, 1970-2001: U.S. Geological Survey Scientific Investigations Report 2005-5224, 54 p., online at <http://pubs.water.usgs.gov/sir2005-5224/>
- Wiche, G.J., Hoetzer, S.M., and Rankl, J.G., 1986, Hydrology of the Devils Lake Basin, northeastern North Dakota: North Dakota State Water Commission Water-Resources Investigations Report 3, 86 p., online at <http://nd.water.usgs.gov/pubs/wri/wri3/index.html>
- Wiche, G.J., Vecchia, A.V., Osborne, Leon, Wood, C.M., and Fay, J.T., 2000, Climatology, hydrology, and simulation of an emergency outlet, Devils Lake Basin, North Dakota: U.S. Geological Survey Water-Resources Investigations Report 00-4174, 16 p., online at <http://nd.water.usgs.gov/pubs/wri/wri004174/index.html>
- Wilde, F.D., Radtke, D.B., Gibs, J., and Iwatsubo, R.T., eds., various dates, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A1-A9, variously paginated, online at <http://water.usgs.gov/owq/FieldManual/>

Appendix

Parametric Flow-Adjustment Model

To model streamflow-related variability, concentrations for each constituent were expressed using the following equation:

$$Y(t) = \alpha_0 + \alpha_1 L_x(t) + \alpha_2 A_x(t) + \alpha_3 S_x(t) + \alpha_4 D_x(t) + \alpha_5 \cos(2\pi t) + \alpha_6 \sin(2\pi t) + \alpha_7 \cos(4\pi t) + \alpha_8 \sin(4\pi t) + \alpha_9 D_x(t) \cos(2\pi t) + \alpha_{10} D_x(t) \sin(2\pi t) + \alpha_{11} D_x(t) \cos(4\pi t) + \alpha_{12} D_x(t) \sin(4\pi t) + U(t) \quad (A1)$$

where

$Y(t)$ is the base-10 logarithm of concentration, in milligrams or micrograms per liter, for time t , in decimal years;

$\alpha_0, \alpha_1, \dots$, and α_{12} are regression parameters to be estimated;

$L_x(t)$ is the 5-year streamflow anomaly (dimensionless) for time t ;

$A_x(t)$ is the annual streamflow anomaly (dimensionless) for time t ;

$S_x(t)$ is the seasonal streamflow anomaly (dimensionless) for time t ;

$D_x(t)$ is the daily streamflow anomaly (dimensionless) for time t ;

$\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$ are periodic functions for describing seasonal variability not explained by variability in streamflow;

and

$U(t)$ is the residual for time t .

Terms 5 through 12 in equation A1 were selected through an exploratory analysis of the residuals computed from the model that had only the first four terms. Terms 5 through 8 model seasonal variability in concentration that seemingly is unrelated to seasonal variability in streamflow. For example, fertilizer application may cause an increase in dissolved ammonia concentrations during some months, irrespective of streamflow conditions. Terms 9 through 12 model seasonal variability in the slope of the line that relates $Y(t)$ to $D_x(t)$. For example, a high daily streamflow anomaly that occurs during the early spring (when soils are frozen) may cause a decrease in total phosphorus concentrations whereas a high daily streamflow anomaly that occurs during the summer may cause an increase in total phosphorus concentrations. The various terms in equation A1 are necessary to model the complex relations between concentration, streamflow, and time of year for most constituents. In some cases, additional terms, such as those that model interactions between $L_x(t)$ and $D_x(t)$ or between time and the various cosine and sine terms, could improve the model. However, such cases tend to occur only for certain isolated stations or constituents.

For constituents for which less than 20 percent of the values were censored, an exhaustive (all-subsets) regression procedure was performed to determine the best potential models for each constituent at each station. The procedure examined subsets of the 12 potential explanatory variables from size 1 to size 11 and the entire set of 12 explanatory variables. The procedure returned as many as 100 of the best potential models for each possible number of explanatory variables (1 through 12). The models were ranked by standard error, adjusted coefficient of multiple determination (R_a^2), Mallows's C_p , and the predicted error sum of squares ($PRESS$). R_a^2 allows for the comparison of models that have differing numbers of explanatory variables by penalizing models that have additional coefficients (Helsel and Hirsch, 1995). The C_p criterion is a measure of the total mean squared error and an indicator of model bias (Neter and others, 1996). The $PRESS$ criterion is a measure of how well the fitted values from a potential model predict the measured values (Neter and others, 1996).

For lake stations, changes in inflows to the lake can take several days to cause substantial changes in lake volumes because of the large volume of water stored in the lakes. Therefore, to determine if time delays could improve the model, potential lake-station models were examined for all constituents at all lake stations with time lags of 7 days ($t - 7$ days) and 14 days ($t - 14$ days).

All potential models were examined to determine one stream-station model for each constituent and one lake-station model for each constituent. Selecting a common model for each station grouping (streams versus lakes) was beneficial for comparing trends among the various stations. The number of potential models was reduced from the initial list created by the all-subsets regression procedure by examining the models that had R_a^2 values that were greater than a limiting value (large R_a^2 values were desirable) and models that had C_p and $PRESS$ values that were less than limiting values (small C_p and $PRESS$ values were desirable). The limits for the criterion varied depending upon how closely the constituents were estimated by the model. The number of potential models also was limited by requiring that the cosine (cos) and sine (sin) terms were included as pairs. For example, if a potential model included $\cos(2\pi t)$, the model also was required to include $\sin(2\pi t)$. Including pairs of cos/sin terms may result in models where one member of the pair is significant while the other is not. However, using only one member of the cos/sin pair forces an arbitrary phase shift rather than a phase shift determined by the data (Helsel and Hirsch, 1995). Finally, the number of potential models was reduced by requiring the individual terms in the interaction terms to be included in the model if the interaction terms were used. For example, including $D_x(t)\cos(2\pi t)$ meant that the individual terms $D_x(t)$ and $\cos(2\pi t)$ also had to be in the model and that the corresponding sine terms, $D_x(t)\sin(2\pi t)$ and $\sin(2\pi t)$, had to be in the model.

The fitted model coefficients for each constituent for stream stations and for each constituent for lake stations are

given in table A1. The lake-station models are for explanatory variables lagged by 7 days because that lag was determined to result in the best overall lake-station models. Because lakes are the primary sources of streamflow for the Big Coulee near Churchs Ferry and Channel A near Penn stations (sites 7 and 8, respectively), the explanatory variables also were lagged by 7 days for those stations. R_a^2 , which adjusts for the number of predictor variables, and the p-value for the F -test for a regression relation between the response variable and the set of predictor variables also are given in table A1.

At an α -level, or significance level, of 0.05, all of the models given in table A1 are statistically significant. That is, the p-value, based on an F -test for the regression relation, is less than 0.05 in all of the models, indicating the probability of obtaining the computed F -test statistic if the models were not statistically significant is less than 0.05 in all cases. A p-value of less than 0.05 indicates that, for each model, a significant regression relation exists between the base-10 logarithm of concentration and the model predictor variables. However, a p-value of less than 0.05 does not ensure that useful predictions can be made using the model. The model removes as much streamflow-related variability in measured concentrations as possible but does not predict actual concentrations.

At an α -level of 0.01, the model for total phosphorus for the Starkweather Coulee near Webster station (site 4) is the only model that is not significant—the p-value for that model is 0.0240.

R_a^2 , which indicates the proportionate reduction in the variability of the base-10 logarithm of concentration obtained using the model predictor variables and which is adjusted for the number of predictor variables to allow for the comparison of models, can be used to compare the effectiveness of the models in accounting for the variability in the measured concentrations. For the stream stations, especially the Mauvais Coulee near Cando and Edmore Coulee near Edmore stations (sites 1 and 2, respectively), R_a^2 values generally were higher for the major ions than for the nutrients and trace element. For the lake stations, R_a^2 values were higher for the major ions than for the trace element. Because too few nutrient data are available for the lake stations, multiple-regression models could not be determined for nutrients.

The highest R_a^2 value for the stream stations was 76.3 for dissolved sulfate for the Mauvais Coulee station (site 1). The highest R_a^2 value for the lake stations was 70.0 for dissolved sulfate for the Lake Irvine near Churchs Ferry station (site 6). The lowest R_a^2 value for the stream stations was 18.3 for dissolved strontium for the Big Coulee near Churchs Ferry station (site 7). The lowest R_a^2 value for the lake stations was 33.6 for dissolved strontium for the Lake Alice near Churchs Ferry station (site 5).

Survival Regression for Dissolved Nitrite Plus Nitrate as Nitrogen

Dissolved nitrite plus nitrate concentrations (fig. A1) were highly censored for the stations used for the water-quality trend analysis (table 3). The censored values (more than 20 percent of the concentrations) indicate nitrite plus nitrate was present in the water-quality samples and the concentrations were between zero and a particular laboratory reporting level. The laboratory reporting levels were 0.05, 0.06, and 0.10. The multiple reporting levels are a result of different laboratories analyzing the samples with time and of changes in laboratory equipment.

Because parametric multiple regression does not work well for highly censored data, a parametric regression method called survival regression was used to estimate the parameters for the regression equation (eq. A1). Survival regression was developed for the analysis of time to death for patients in medical studies in which some patients lived beyond the cutoff point of the study, such as 10 years (Insightful Corporation, 2001). The times to death for patients living beyond 10 years, in a 10-year study, are right censored at 10 (time to death > 10) because the true times of death are unknown. Survival analysis methods have since been applied to other fields of study, such as water-quality monitoring, and the term survival analysis has expanded to apply to problems other than time to death or time to failure. In this study, censored data are left censored—that is, the data are less than a particular laboratory reporting level.

For nitrite plus nitrate, one model was selected for the stream stations and one model was selected for the lake stations. The models were chosen by examining potential models such as those used for other constituents, using analysis of variance (anova) to test the significance of terms in the potential models, and visually inspecting plots of the measured concentrations and the concentrations predicted by the potential models. As with the multiple-regression models, the cosine and sine terms were required to be included as pairs and the individual terms in the interaction terms were required to be included in the model if the interaction terms were used.

Parametric survival regression uses maximum likelihood estimation to compute parameter estimates based on measured concentrations and the observed proportions of data below one or more censoring levels (Helsel, 2005). Maximum likelihood estimation assumes that the measured concentrations are from a particular underlying distribution. For this study, the assumption was that the base-10 logarithm of concentration followed a normal distribution. This assumption was checked before performing survival regression by examining probability plots that check the similarity of the measured concentrations to the

specified distribution. In the probability plots, the assumed distribution is represented by a straight line and the concentrations are plotted individually. Departures from the straight line show how the data differ from the assumed distribution. Concentrations that closely follow the line are assumed to be from the specified distribution (Helsel, 2005). The normal distribution assumption was reasonable for the measured concentrations, as shown for the Mauvais Coulee near Cando station (site 1) (fig. A2), and the distribution used for the survival regression was the normal distribution.

Another important consideration in maximum likelihood estimation is the number of samples collected. Maximum likelihood estimation methods in survival analysis work poorly for a small number of samples. According to Helsel (2005), the method works best when $n > 50$. The number of samples for dissolved nitrite plus nitrate was greater than 50 for all of the stations used for this study (table 3).

After survival regression was performed, the residuals were checked using probability plots to determine whether the residuals followed the assumed normal distribution. A normal distribution was suitable for survival regression of nitrite plus nitrate for the Mauvais Coulee station (site 1) (fig. A3).

The fitted survival regression model coefficients for nitrite plus nitrate for each stream station and for each lake station are given in table A2. As with the multiple-regression analysis, the explanatory variables for the Big Coulee near Churchs Ferry and Channel A near Penn stations (sites 7 and 8, respectively) and for the lake stations were lagged by 7 days. Blank cells in the table indicate explanatory variables that were not significant and, thus, were not used in the model. The model coefficients obtained from survival regression describe the distribution with the maximum likelihood of

having produced the measured concentrations and the proportions of censored values below each censoring level.

The p-value for the chi-squared test of the overall significance of the regression model also is given in table A2. The test determines whether the entire model is a statistically significant improvement over the null model, the model in which only the intercept term is given and all other coefficients are equal to zero.

At an α -level of 0.01, the model for nitrite plus nitrate was significant for all stream stations. That is, the p-value, based on a chi-squared test for regression relation, is less than 0.01 for each station, indicating the probability of obtaining the computed chi-squared test statistic if the model was not statistically significant is less than 0.01. A p-value of less than 0.01 indicates a significant regression relation exists between the base-10 logarithm of concentration and the model predictor variables for stream stations. However, a p-value of less than 0.01 does not ensure that useful predictions can be made using the model.

At an α -level of 0.01, the model was significant for the Lake Alice near Churchs Ferry and Lake Irvine near Churchs Ferry stations (sites 5 and 6, respectively). The p-values for those models indicate a significant regression relation exists between the base-10 logarithm of concentration and the model predictor variables for those stations. The Sweetwater Lake at Sweetwater station (site 3) had a high p-value of 0.720, indicating the explanatory variables in the model do not provide a statistically significant improvement for predicting the base-10 logarithm of nitrite plus nitrate concentration over the null model.

Partial significance tests were performed for each of the coefficients in the survival regression. The p-values for those tests also are given in table A2.

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).

$[L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; R_a^2 , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Site number (figure 2)	Regression term													P-value
	Model coefficient													
	Intercept	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$	
	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	
Calcium, dissolved														
1	1.960 ($<.001$)	0.002 (.917)	0.008 (.734)	-0.099 ($<.001$)	-0.060 (.223)	-0.027 (.673)	-0.051 (.069)	0.057 (.111)	-0.068 (.021)	0.011 (.879)	-0.022 (.474)	0.027 (.437)	0.016 (.611)	67.2 $<.001$
2	1.972 ($<.001$)	.099 (.001)	0 (.999)	-.033 (.187)	-.004 (.946)	.219 (.120)	-.105 (.132)	.092 (.099)	-.158 (.003)	.118 (.129)	-.091 (.073)	.007 (.841)	-.057 (.174)	61.9 $<.001$
3	1.700 ($<.001$)	.184 ($<.001$)	.066 (.027)	-.018 (.539)	-.023 (.339)	.092 (.004)	.098 ($<.001$)	-.006 (.838)	.082 ($<.001$)	-.002 (.963)	.045 (.056)	--	--	56.6 $<.001$
4	1.342 ($<.001$)	.087 (.008)	.042 (.128)	.028 (.243)	.166 (.002)	-.689 ($<.001$)	-.133 (.001)	-.340 ($<.001$)	-.151 ($<.001$)	.309 (.002)	-.018 (.564)	.135 (.014)	.006 (.862)	48.6 $<.001$
5	1.997 ($<.001$)	.011 (.694)	-.061 (.046)	-.080 (.004)	.005 (.858)	.162 ($<.001$)	.028 (.361)	.056 (.095)	.054 (.048)	.031 (.468)	-.028 (.246)	--	--	53.7 $<.001$
6	1.925 ($<.001$)	.022 (.430)	-.074 (.020)	-.047 (.096)	-.012 (.721)	.116 (.014)	.057 (.067)	-.018 (.584)	.049 (.067)	.036 (.451)	-.063 (.016)	--	--	43.9 $<.001$
7	1.857 ($<.001$)	-.034 (.323)	-.054 (.023)	.054 (.002)	.017 (.323)	.163 ($<.001$)	-.035 (.080)	.085 ($<.001$)	.004 (.816)	-.004 (.871)	.003 (.862)	-.012 (.538)	.010 (.551)	31.4 $<.001$
8	1.910 ($<.001$)	-.109 (.021)	-.110 (.009)	-.021 (.528)	.020 (.774)	.141 (.248)	.051 (.503)	.091 (.109)	.034 (.581)	.033 (.713)	.058 (.230)	.042 (.342)	.050 (.191)	23.5 .002

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).—Continued

[$L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; R^2_a , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Site number (figure 2)	Regression term													p-value
	Model coefficient													
	Intercept	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$	
	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	
Sulfate, dissolved														
1	2.494 ($<.001$)	0.028 (.345)	-0.126 ($<.001$)	-0.241 ($<.001$)	-0.124 ($<.001$)	-0.022 (.723)	-0.078 (.009)	0.173 ($<.001$)	-0.072 (.010)	--	--	--	--	76.3 <.001
2	2.179 ($<.001$)	.273 ($<.001$)	.044 (.323)	-.143 ($<.001$)	-.144 ($<.001$)	-.073 (.532)	.143 (.016)	.150 (.030)	-.085 (.071)	--	--	--	--	67.9 <.001
3	2.330 ($<.001$)	.171 ($<.001$)	-.133 ($<.001$)	-.085 (.006)	-.005 (.856)	.036 (.277)	.047 (.019)	--	--	0.062 (.135)	-0.013 (.583)	--	--	55.8 <.001
4	1.818 ($<.001$)	.277 ($<.001$)	.099 (.064)	-.051 (.298)	0 (.992)	-.246 (.222)	-.103 (.136)	.004 (.966)	-.173 (.010)	--	--	--	--	39.3 <.001
5	2.394 ($<.001$)	.020 (.611)	-.189 ($<.001$)	-.195 ($<.001$)	.078 (.063)	.009 (.858)	-.044 (.233)	--	--	.160 (.010)	-.045 (.173)	--	--	56.1 <.001
6	2.351 ($<.001$)	.052 (.059)	-.240 ($<.001$)	-.088 (.001)	.070 (.033)	.047 (.201)	-.038 (.141)	--	--	.180 ($<.001$)	-.054 (.024)	--	--	7.0 <.001
7	2.343 ($<.001$)	.053 (.285)	-.257 ($<.001$)	.040 (.103)	.011 (.518)	.167 ($<.001$)	-.052 (.048)	.043 (.091)	.001 (.968)	--	--	--	--	34.0 <.001
8	2.436 ($<.001$)	-.046 (.433)	-.284 ($<.001$)	0 (.996)	-.074 (.006)	.077 (.337)	.225 (.001)	.074 (.152)	.111 (.030)	--	--	--	--	54.2 <.001

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).—Continued

$[L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; R^2_a , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Site number (figure 2)	Regression term													p-value
	Model coefficient													
	Chloride, dissolved													
Intercept	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$	R^2_a	
B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12		
1	1.447 ($<.001$)	-0.003 (.944)	--	-0.226 ($<.001$)	-0.011 (.890)	-0.043 (.697)	-0.123 (.011)	0.076 (.216)	-0.080 (.113)	0.114 (.344)	-0.061 (.250)	0.055 (.352)	0.020 (.697)	61.4 $<.001$
2	1.741 ($<.001$)	.196 ($<.001$)	--	-.112 (.010)	.071 (.460)	.616 (.013)	-.216 (.062)	.277 (.006)	-.280 (.002)	.391 (.004)	-.165 (.064)	.042 (.465)	-.225 (.003)	55.9 $<.001$
3	1.470 ($<.001$)	.048 (.030)	-0.129 ($<.001$)	-.054 (.026)	-.019 (.325)	.065 (.012)	.037 (.036)	.024 (.264)	.024 (.128)	.029 (.344)	-.014 (.454)	--	--	55.2 $<.001$
4	.586 (.003)	.227 ($<.001$)	--	-.014 (.765)	.214 (.038)	-.890 (.006)	-.169 (.021)	-.434 (.006)	-.158 (.031)	.442 (.018)	.012 (.833)	.184 (.077)	-.023 (.721)	32.3 $<.001$
5	1.500 ($<.001$)	-.127 (.004)	-.173 ($<.001$)	-.200 ($<.001$)	-.001 (.976)	-.014 (.837)	-.059 (.223)	.006 (.900)	-.001 (.978)	.092 (.171)	-.075 (.050)	--	--	51.8 $<.001$
6	1.442 ($<.001$)	-.055 (.100)	-.209 ($<.001$)	-.115 (.001)	.004 (.919)	-.014 (.798)	-.060 (.096)	-.061 (.118)	-.019 (.541)	.106 (.064)	-.102 (.001)	--	--	59.2 $<.001$
7	1.408 ($<.001$)	.019 (.722)	--	.030 (.254)	.026 (.331)	.179 ($<.001$)	-.098 (.002)	.038 (.199)	.040 (.173)	.040 (.320)	-.009 (.764)	-.015 (.615)	.035 (.198)	24.7 $<.001$
8	1.226 ($<.001$)	.096 (.071)	--	-.015 (.678)	-.060 (.470)	-.013 (.930)	.251 (.003)	.066 (.339)	.134 (.064)	.044 (.690)	.126 (.041)	.037 (.500)	.089 (.065)	27.2 $<.001$

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).—Continued

$[L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; R_a^2 , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Regression term													
Site number (figure 2)	Model coefficient												
	Intercept	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$
p-value													
R_a^2													
Ammonia, dissolved													
B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	
1	-1.068 (.001)	--	-0.089 (.249)	0.111 (.415)	0.461 (.093)	0.146 (.380)	0.300 (.098)	0.194 (.178)	0.234 (.235)	-.002 (.984)	--	--	20.5 0.008
2	-.806 (.001)	--	-.295 (.001)	-.026 (.786)	.628 (.007)	.254 (.094)	.299 (.026)	.340 (.010)	-.001 (.996)	-.003 (.972)	--	--	50.7 <.001
4	-.817 (.014)	--	.048 (.540)	.133 (.398)	.873 (.091)	.201 (.304)	.479 (.049)	.072 (.680)	.281 (.099)	.012 (.930)	--	--	39.4 <.001
7	-.713 (.001)	--	-.245 (.009)	-.007 (.950)	.368 (.089)	-.014 (.917)	.086 (.561)	.288 (.032)	.141 (.316)	-.088 (.387)	--	--	28.3 .002
8	-1.169 (.001)	--	-.226 (.058)	-.217 (.176)	-.053 (.852)	.036 (.855)	-.054 (.704)	.319 (.103)	0 (.998)	.198 (.138)	--	--	23.9 .003

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).—Continued

$[L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$; periodic functions for describing seasonal variability; R_a^2 , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Site number (figure 2)	Regression term												p-value	
	Model coefficient													
	Phosphorus, total													
	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	
Intercept	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$	R_a^2	
1	-0.702 (.006)	0.032 (.800)	-0.093 (.583)	--	0.375 (.046)	0.487 (.165)	-0.357 (.145)	0.451 (.055)	0.029 (.886)	0.537 (.064)	-0.147 (.278)	--	--	22.8 0.010
2	-0.676 (.005)	.091 (.385)	-.293 (.013)	--	.088 (.503)	.534 (.085)	-.114 (.615)	.386 (.033)	.325 (.090)	.181 (.349)	.122 (.232)	--	--	34.8 <.001
4	-2.156 (.002)	.202 (.279)	-.154 (.400)	--	.333 (.331)	-1.068 (.296)	.115 (.797)	-.931 (.053)	.232 (.565)	.491 (.222)	-.232 (.440)	--	--	19.6 .024
7	-.671 (<.001)	-.400 (.006)	-.134 (.350)	--	.278 (.211)	.396 (.130)	-.242 (.494)	.228 (.228)	.242 (.019)	.298 (.265)	-.037 (.902)	--	--	37.8 .001
8	-.835 (<.001)	.228 (.004)	.102 (.213)	--	.064 (.472)	-.179 (.222)	-.083 (.523)	-.095 (.216)	.108 (.359)	-.013 (.920)	.038 (.582)	--	--	37.8 <.001

Table A1. Fitted model coefficients for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).—Continued

[$L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; R^2_a , adjusted coefficient of multiple determination; <, less than; numbers in parentheses indicate p-values for individual model coefficients; --, indicates regression terms not used for a particular site]

Site number (figure 2)	Regression term												R^2_a	p-value
	$L_x(t)$	$A_x(t)$	$S_x(t)$	$D_x(t)$	$\cos(2\pi t)$	$\sin(2\pi t)$	$\cos(4\pi t)$	$\sin(4\pi t)$	$D_x(t)$ $\cos(2\pi t)$	$D_x(t)$ $\sin(2\pi t)$	$D_x(t)$ $\cos(4\pi t)$	$D_x(t)$ $\sin(4\pi t)$		
	Model coefficient													
	Strontium, dissolved													
	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11		
1	2.631 (.930)	-0.041 (0.195)	-0.092 (.001)	-0.053 (.169)	-0.012 (.882)	-0.064 (.166)	0.031 (.538)	-0.050 (.204)	0.018 (.737)	-0.066 (.031)	--	--	61.1 <.001	
2	2.505 (.001)	-.017 (.605)	.008 (.788)	.027 (.503)	.042 (.657)	-.183 (.005)	.004 (.943)	-.218 (.001)	.146 (.009)	-.073 (.027)	--	--	54.9 <.001	
3	2.465 (.001)	-.024 (.304)	--	-.015 (.271)	.073 (.001)	.065 (.001)	-.014 (.497)	.065 (.001)	--	--	-0.019 (.478)	0.024 (.142)	47.1 <.001	
4	2.214 (.001)	.029 (.387)	.040 (.188)	.086 (.203)	-.289 (.167)	-.194 (.031)	-.183 (.053)	-.182 (.018)	.153 (.037)	-.085 (.158)	--	--	36.5 <.001	
5	2.617 (.001)	-.084 (.075)	--	.033 (.168)	.222 (.001)	.046 (.255)	.066 (.135)	.064 (.088)	--	--	-.043 (.342)	.051 (.088)	33.6 <.001	
6	2.570 (.001)	-.108 (.007)	--	-.007 (.746)	.128 (.001)	.061 (.052)	-.037 (.302)	.069 (.022)	--	--	-.036 (.336)	.076 (.005)	37.3 <.001	
7	2.496 (.001)	-.080 (.026)	.054 (.028)	.014 (.549)	.123 (.017)	-.064 (.040)	.083 (.022)	.025 (.406)	.039 (.256)	.036 (.142)	--	--	18.3 <.001	
8	2.724 (.001)	-.172 (.001)	-.019 (.548)	.037 (.427)	.190 (.020)	-.033 (.633)	.058 (.144)	-.034 (.538)	.141 (.016)	-.029 (.430)	--	--	43.9 <.001	

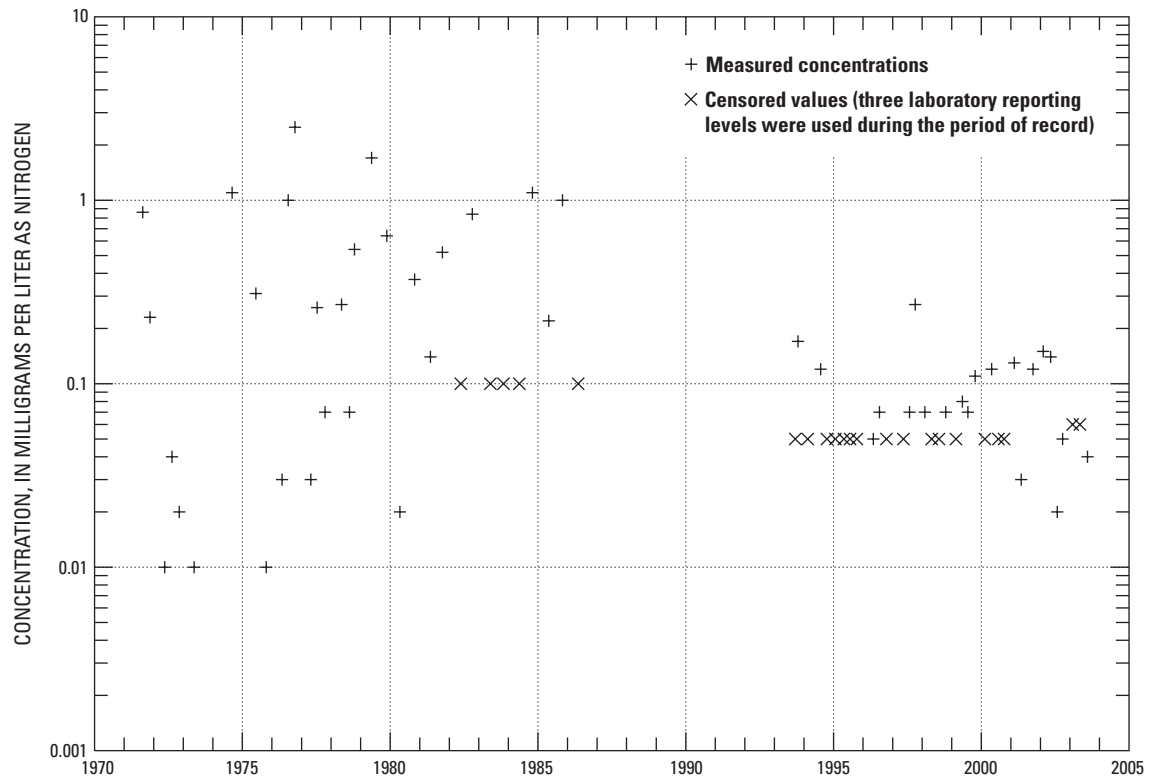


Figure A1. Dissolved nitrite plus nitrate concentrations for the Lake Alice near Church's Ferry, North Dakota, station.

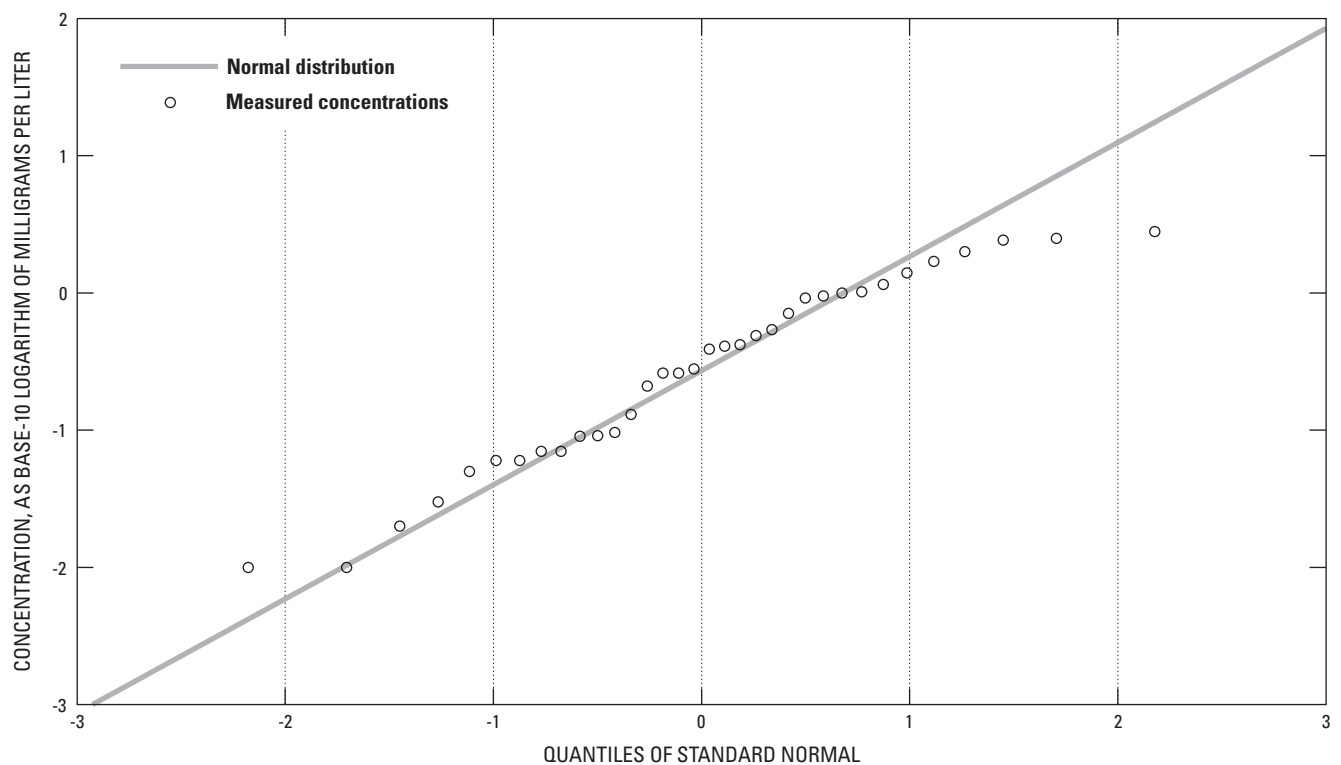


Figure A2. Dissolved nitrite plus nitrate concentrations for the Mauvais Coulee near Cando, North Dakota, station.

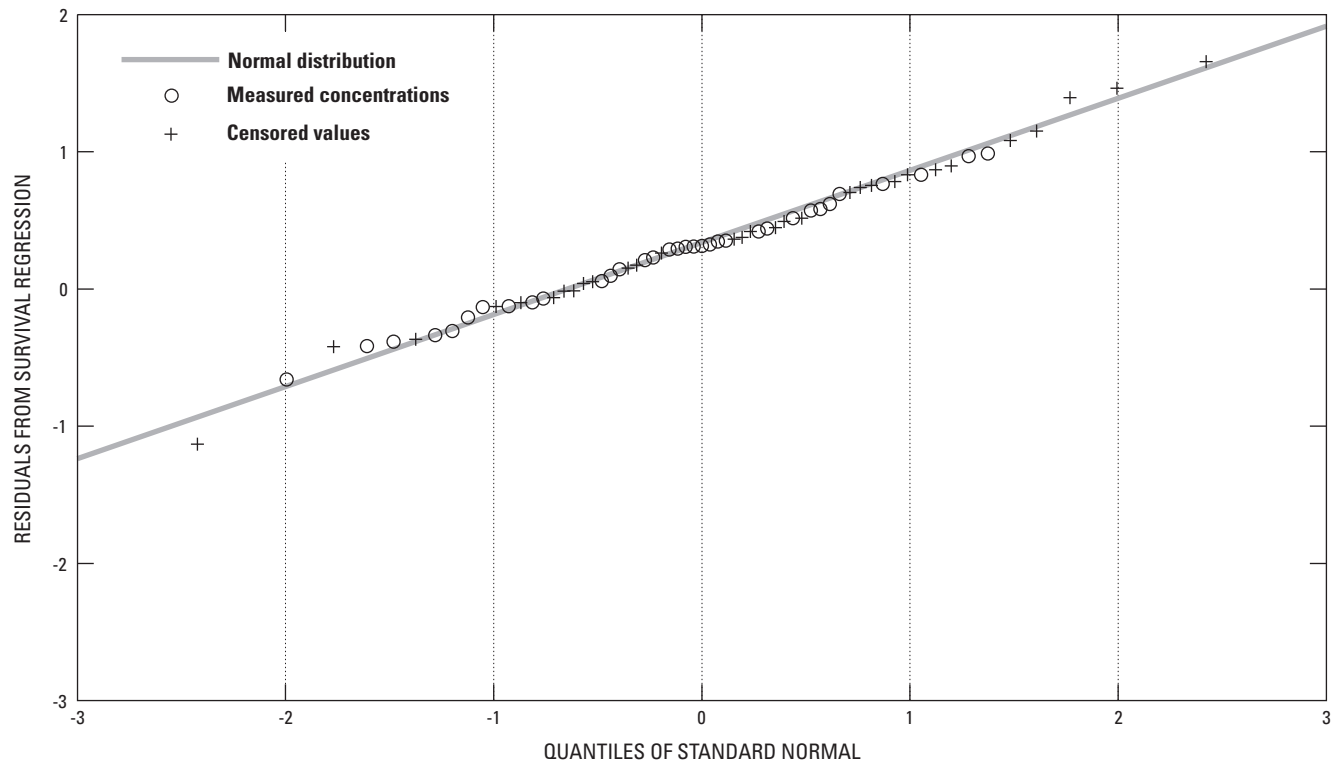


Figure A3. Dissolved nitrite plus nitrate concentrations and censored values for the Mauvais Coulee near Cando, North Dakota, station.

Table A2. Fitted survival regression model coefficients for dissolved nitrite plus nitrate for stream stations (sites 1, 2, 4, 7, and 8) and lake stations (sites 3, 5, and 6).

$[L_x(t)$, 5-year streamflow anomaly for time t ; $A_x(t)$, annual streamflow anomaly for time t ; $S_x(t)$, seasonal streamflow anomaly for time t ; $D_x(t)$, daily streamflow anomaly for time t ; $\cos(2\pi t)$, $\sin(2\pi t)$, $\cos(4\pi t)$, and $\sin(4\pi t)$, periodic functions for describing seasonal variability; --, indicates regression terms not used for a particular site; <, less than; numbers in parentheses indicate p-values for individual model coefficients]

Site number (figure 2)	Regression term													p-value for chi-squared test of model significance
	Model coefficient													
	B0	B1	B2	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12	
Nitrite plus nitrate as nitrogen, dissolved														
1	-1.463 ($<.001$)	--	--	-0.052 (.630)	-0.173 (.524)	0.864 (.080)	0.686 (.099)	0.815 (.020)	0.885 (.004)	-0.518 (.160)	0.772 (.001)	--	--	$<.001$
2	-1.243 (.002)	--	--	-0.013 (.920)	-0.057 (.782)	.831 (.126)	1.064 (.008)	.686 (.032)	.933 (.004)	-.357 (.234)	.468 (.010)	--	--	$<.001$
3	-1.416 ($<.001$)	-0.086 (.489)	-0.072 (.650)	--	-.314 (.082)	-.055 (.591)	.087 (.437)	-.006 (.963)	-.072 (.415)	-.428 (.108)	-.111 (.340)	--	--	.720
4	-1.659 (.490)	--	--	-.067 (.570)	-.514 (.191)	.836 (.778)	2.513 (.220)	1.262 (.048)	1.406 (.242)	-.918 (.012)	.808 (.030)	--	--	$<.001$
5	-1.253 ($<.001$)	-.363 (.029)	-.222 (.297)	--	.011 (.956)	-.139 (.649)	-.599 (.023)	.210 (.288)	-.384 (.121)	.137 (.662)	.176 (.288)	--	--	.008
6	-1.183 ($<.001$)	-.583 ($<.001$)	-.667 (.001)	--	-.504 (.015)	.171 (.506)	-.368 (.087)	.404 (.027)	-.540 (.012)	-.460 (.150)	.302 (.069)	--	--	$<.001$
7	-.636 ($<.001$)	--	--	-.024 (.823)	-.005 (.968)	.469 (.004)	.098 (.448)	.328 (.008)	.112 (.389)	.174 (.318)	.313 (.005)	--	--	$<.001$
8	-1.506 ($<.001$)	--	--	-.591 (.001)	-.144 (.557)	-.362 (.384)	-.188 (.574)	-.120 (.587)	.597 (.062)	-.356 (.278)	.245 (.237)	--	--	$<.001$

Variogram Analysis of Flow-Adjusted Concentrations

Techniques used to estimate the variance and serial correlation of the flow-adjusted concentrations are described in this section. Standard time-series techniques such as correlograms are not readily applied to unequally spaced data. Therefore, variograms, which are easily applied to unequally spaced data, were used to analyze the flow-adjusted concentrations.

Given the flow-adjusted concentration residuals, $\varepsilon(t)$ (eq. 2), a determination needed to be made as to whether the variance of the residuals changed depending on the time of year the water-quality sample was collected. Therefore, graphs of the fourth root of the squared residuals in relation to the time of year were used to detect seasonality in the residual variance (see Cressie and Hawkins, 1980, for the rationale for using such plots). The residuals for the flow-adjusted sulfate concentrations for the Starkweather Coulee near Webster station (site 4) are shown in figure A4. An abrupt increase occurred in the central tendency of the transformed residuals from late May to early June, and the mean of the values for season 2 (June through November) was significantly (p-value of less than 0.001) higher than the mean for season 1 (March through May). The residuals for the flow-adjusted dissolved chloride concentrations for the Edmore Coulee near Edmore station (site 2) (fig. A5) indicated a significant (p-value of 0.039) decrease in variability from season 1 to season 2.

Results of the variability analysis for all stations and all constituents are given in table A3. The estimated seasonal standard deviations given in table A3 are robust estimates obtained by squaring the means of the fourth root of the squared residuals for each season and applying a bias correction factor (Cressie and Hawkins, 1980). The p-values were obtained using a two-sided, two-sample *t*-test for the difference between the means of the fourth root of the squared residuals for seasons 1 and 2. As indicated in table A3, only sulfate for the Starkweather Coulee station (site 4) and chloride for the Mauvais Coulee near Cando and Edmore Coulee stations (sites 1 and 2, respectively) indicated a significant difference between the standard deviations for seasons 1 and 2. However, in the sample design analysis, the seasonal standard deviations given in table A3 were used for each constituent-station combination regardless of the significance of the difference between the two seasons. The seasonal standard deviations were used in the design analysis for the following reasons—first, differences between season 1 and season 2 were expected because streamflow for season 1 consisted primarily of snowmelt or precipitation runoff from frozen soils but streamflow for season 2 consisted primarily of subsurface flow or ground-water discharge; and second, the variogram analysis indicated seasonal differences existed in the serial correlation structure for many constituent-station combinations.

To detect potential serial correlation in the flow-adjusted concentration residuals, the residuals were standardized by

dividing the residuals by the seasonal standard deviations given in table A3. The (semi-)variogram of the standardized residuals was defined as

$$\gamma(h) = \frac{1}{2}EV\{[\varepsilon^*(t+h) - \varepsilon^*(t)]^2\} \quad (A2)$$

where

γ is the variogram;
 h is the time lag, in weeks;
 EV denotes the expected value;
 ε^* is the standardized residual for the flow-adjusted concentration;

and

t is time, in decimal years.

The variogram and correlogram of the standardized residuals were related through the following equation (Cressie, 1991):

$$\rho(h) = 1 - \gamma(h) \quad (A3)$$

where

$\rho(h)$ is the correlogram [the correlation between $\varepsilon^*(t+h)$ and $\varepsilon^*(t)$].

A robust estimator of the variogram (Cressie, 1991, eq. 2.4.12) was computed for each constituent-station pair. To detect potential differences between the two seasons, two estimated variograms were computed. For the first estimated variogram, only residual pairs with at least one of the residuals from season 1 were used in the computations, and, for the second estimated variogram, only residual pairs with at least one residual from season 2 were used. The estimated seasonal variograms for dissolved sulfate for the Big Coulee near Churchs Ferry station (site 7) are shown in figures A6 and A7. The points in the graph were computed using pairs of residuals with time lags binned into 2-week intervals—for example, the first point was computed using residual pairs separated by less than 2 weeks, the second using residual pairs separated by 2 to 4 weeks, etc. As indicated in figure A6, for season 1, the variogram was approximately constant at 1.0, which indicates no discernible serial correlation existed among residuals separated by 2 weeks or more. In contrast, as indicated in figure A7, the variogram for season 2 was substantially less than 1.0 for small lags and increased gradually until reaching a sill of 1.0 at about a 12-week lag. Therefore, in season 2, serial correlation existed between neighboring residuals up to a lag of about 12 weeks. Most of the flow for Big Coulee is from Lakes Alice and Irvine (fig. 2) and the flow-adjusted dissolved sulfate concentrations for those lakes were expected to be serially correlated, especially during the summer months. Because too few concentration data are available for the lakes, the variogram for time lags of less than 12 weeks could not be estimated. Therefore, the variograms for the lake data could

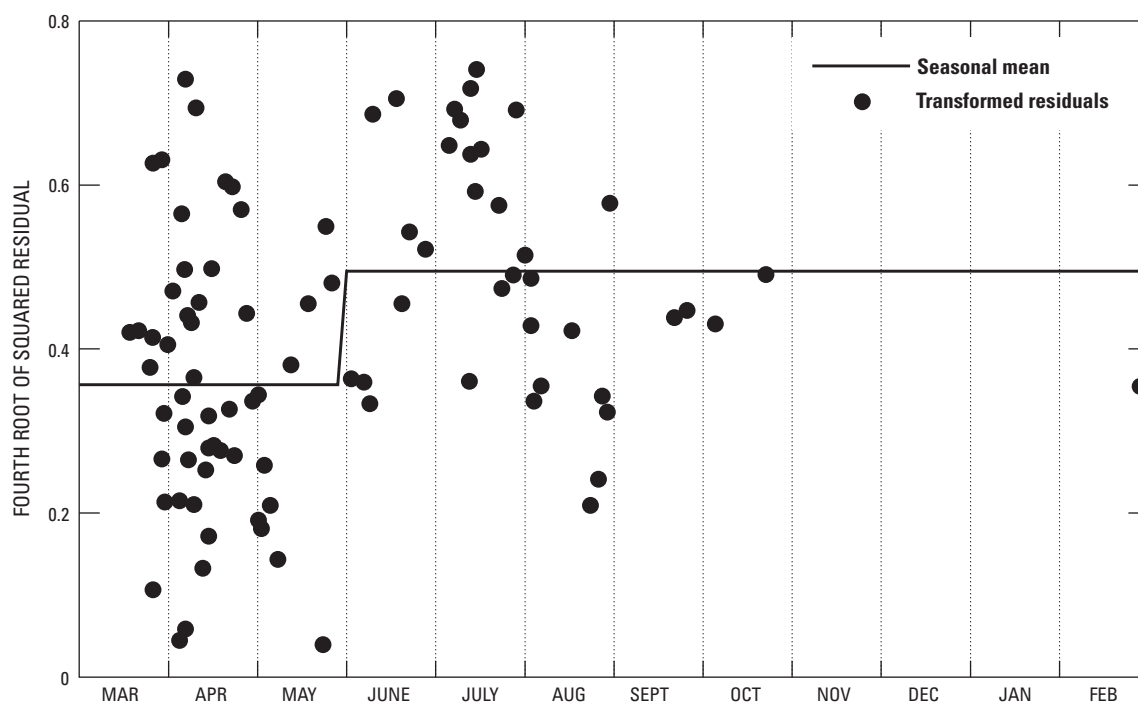


Figure A4. Fourth root of squared residuals for flow-adjusted dissolved sulfate concentrations for the Starkweather Coulee near Webster, North Dakota, station.

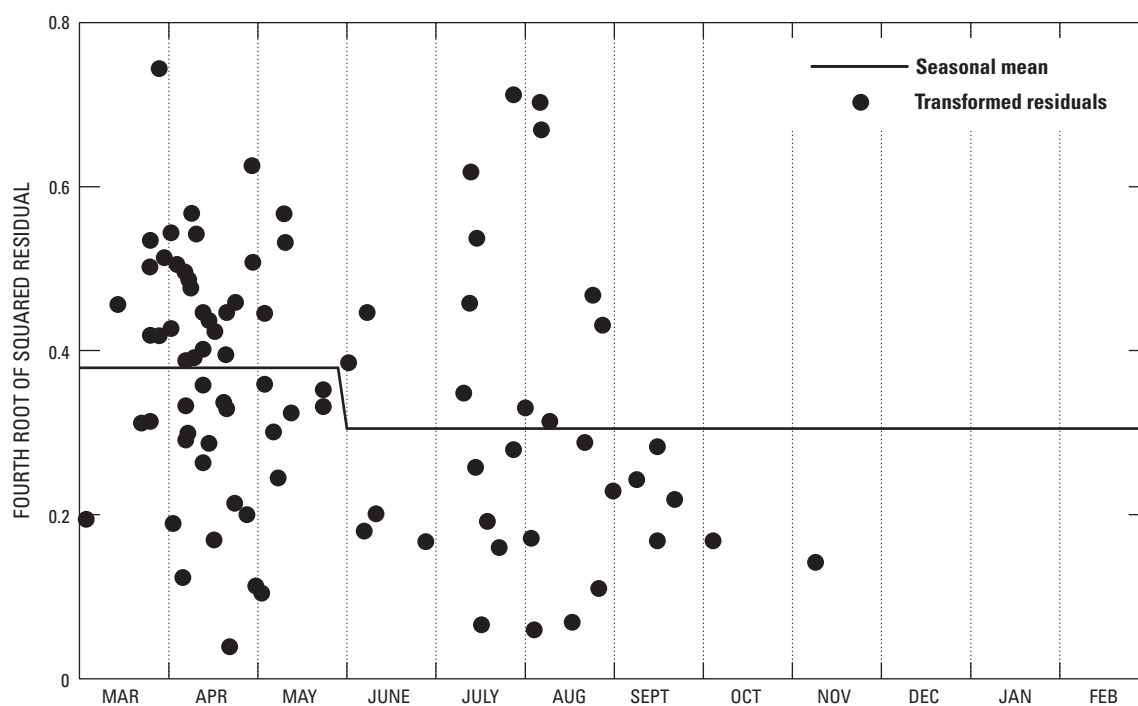


Figure A5. Fourth root of squared residuals for flow-adjusted dissolved chloride concentrations for the Edmore Coulee near Edmore, North Dakota, station.

not be compared easily to the variogram for the Big Coulee station (site 8).

The estimated variograms for total phosphorus concentrations for the Channel A near Penn station (site 8) (figs. A8 and A9) were similar to the variograms for dissolved sulfate. No discernible serial correlation existed for season 1, but some serial correlation may have occurred for season 2 for time lags of less than about 8 weeks. Estimated variograms for dissolved calcium, sulfate, and chloride for most stations were similar to the variograms shown in figures A6 and A7 as were the estimated variograms for total phosphorus for the Big Coulee station (site 7). For total phosphorus for the remaining stations and for dissolved ammonia and strontium for all stations, no discernible serial correlation existed in either season.

Plots of the estimated variograms were examined as described previously to select a correlation range (an approximate time lag beyond which the flow-adjusted concentrations could be assumed to be uncorrelated). The correlation ranges, which were used in the design analysis, are given in table A3.

Model Robustness

The fitted trends and p-values given in table 4 were based on the assumptions that the residuals from the regression model, with both the flow-adjustment variables and the trend variables included, were uncorrelated and had constant variance. However, as indicated previously, in a few cases, serial correlation existed among neighboring residuals and/or a significant difference existed between the residual variances for the two seasons (table A3).

The potential difference between the residual variances for the two seasons was expected to have a negligible effect on the fitted trends and p-values. In all but a few cases, the difference between the residual variances for the two seasons was small and insignificant (table A3). In cases when a significant difference did occur, time-series plots of the residuals did not indicate a systematic pattern of change in the residual variances from year to year (only from season to season) because the apportionment of samples among seasons did not change appreciably with time.

Unlike the seasonal residual variances, serial correlation in the residuals was expected to have an effect on the fitted trends and p-values in some cases. Ignoring serial correlation can inflate the error variance and thus reduce the significance of a fitted trend or inflate the effective number of samples and thus artificially increase the significance of a fitted trend. The effects of serial correlation can be removed by estimating

the correlation matrix of the residuals along with the regression parameters and generalized least squares used to fit the regression model or by “thinning” the data so the residuals are spaced far enough apart to be uncorrelated. Using the correlation ranges given in table A3, original data needed to be thinned only for the Big Coulee near Churchs Ferry station (site 7). Therefore, that station likely would be most affected by serial correlation.

The trend results for dissolved chloride and dissolved nitrite plus nitrate for the Big Coulee station (site 7), both with and without data thinning, are given in table A4. For dissolved chloride, the data set contained 198 measurements before thinning and 103 measurements after thinning. The estimated coefficient for the Channel A step trend increased from 0.105 without thinning to 0.160 with thinning, and the p-value decreased from 0.067 to 0.025. The estimated coefficient for the linear trend essentially was unchanged, but the p-value of the trend increased from 0.001 to 0.014. Thus, the step trend became more significant and the linear trend less significant as a result of data thinning. For dissolved chloride, the large group of low concentrations that occurred from 1992-97 had a substantial effect on the fitted trends (see figure 18). Removing some of the low concentrations probably resulted in the increased significance of the step trend and the decreased significance of the linear trend. However, the thinning did not have a drastic effect on the fitted trends.

For dissolved nitrite plus nitrate, the data set contained 118 measurements before thinning and 63 measurements after thinning. For both the step trend and the linear trend, the sign of the trends was unchanged but the magnitude of the trends decreased and the p-values increased as a result of the thinning. As for dissolved chloride, the differences probably resulted from thinning of the low concentrations that occurred at the end of the period of record (see figure 19). However, because few high concentrations occurred during 1980-90 to counter the effects of the low concentrations, both the step trend and the linear trend became less significant as a result of the data thinning rather than more significant as for dissolved chloride.

As shown in the previous examples, long sampling gaps or highly variable sampling frequencies from year to year probably have more of an effect on trend results than potential serial correlation. Also, as mentioned previously, fixed sampling designs in which sampling frequencies remain constant from year to year generally are more efficient for analyzing long-term trends than variable sampling designs in which the sampling frequencies change. Sampling at too high a frequency during some years may introduce redundancy and thus decrease the efficiency of a design without drastically changing the direction or significance of the fitted trends.

Table A3. Estimated seasonal standard deviations and correlation ranges for flow-adjusted concentrations.

U.S. Geological Survey station name	Season 1 (March through May)		Season 2 (June through November)		p-value for difference in seasonal standard deviations
	Standard deviation	Correlation range	Standard deviation	Correlation range	
Calcium, dissolved					
Mauvais Coulee near Cando, North Dakota	0.094	2	0.062	4	0.056
Edmore Coulee near Edmore, North Dakota	.108	2	.076	4	.117
Sweetwater Lake at Sweetwater, North Dakota	.119	2	.075	4	.106
Starkweather Coulee near Webster, North Dakota	.104	2	.107	2	.899
Lake Alice near Churchs Ferry, North Dakota	.093	2	.069	6	.342
Big Coulee near Churchs Ferry, North Dakota	.108	2	.091	6	.395
Channel A near Penn, North Dakota	.139	2	.090	6	.059
Sulfate, dissolved					
Mauvais Coulee near Cando, North Dakota	0.135	2	0.116	2	0.458
Edmore Coulee near Edmore, North Dakota	.188	2	.183	4	.893
Sweetwater Lake at Sweetwater, North Dakota	.126	2	.147	4	.522
Starkweather Coulee near Webster, North Dakota	.185	2	.357	8	< .001
Lake Alice near Churchs Ferry, North Dakota	.121	2	.106	12	.686
Big Coulee near Churchs Ferry, North Dakota	.148	2	.193	12	.180
Channel A near Penn, North Dakota	.207	2	.155	4	.158
Chloride, dissolved					
Mauvais Coulee near Cando, North Dakota	0.177	2	0.118	6	0.047
Edmore Coulee near Edmore, North Dakota	.210	2	.136	4	.039
Sweetwater Lake at Sweetwater, North Dakota	.098	2	.116	4	.448
Starkweather Coulee near Webster, North Dakota	.201	2	.226	6	.554
Lake Alice near Churchs Ferry, North Dakota	.181	2	.100	6	.097
Big Coulee near Churchs Ferry, North Dakota	.171	2	.233	6	.148
Channel A near Penn, North Dakota	.181	2	.124	4	.083
Ammonia, dissolved					
Mauvais Coulee near Cando, North Dakota	0.327	2	0.246	2	0.327
Edmore Coulee near Edmore, North Dakota	.276	2	.232	2	.462
Starkweather Coulee near Webster, North Dakota	.348	2	.243	2	.119
Big Coulee near Churchs Ferry, North Dakota	.437	2	.361	2	.345
Channel A near Penn, North Dakota	.486	2	.534	2	.551

Table A3. Estimated seasonal standard deviations and correlation ranges for flow-adjusted concentrations.—Continued

U.S. Geological Survey station name	Season 1 (March through May)		Season 2 (June through November)		p-value for difference in seasonal standard deviations
	Standard deviation	Correlation range	Standard deviation	Correlation range	
Phosphorus, total					
Mauvais Coulee near Cando, North Dakota	0.162	2	0.105	2	0.178
Edmore Coulee near Edmore, North Dakota	.151	2	.122	2	.417
Starkweather Coulee near Webster, North Dakota	.171	2	.276	2	.096
Big Coulee near Churchs Ferry, North Dakota	.148	2	.118	8	.323
Channel A near Penn, North Dakota	.188	2	.230	8	.384
Strontium, dissolved					
Mauvais Coulee near Cando, North Dakota	0.113	2	0.079	2	0.098
Edmore Coulee near Edmore, North Dakota	.125	2	.097	2	.225
Sweetwater Lake at Sweetwater, North Dakota	.065	2	.086	2	.350
Starkweather Coulee near Webster, North Dakota	.113	2	.119	2	.809
Lake Alice near Churchs Ferry, North Dakota	.103	2	.101	2	.940
Big Coulee near Churchs Ferry, North Dakota	.114	2	.101	2	.601
Channel A near Penn, North Dakota	.120	2	.102	2	.406

Table A4. Results of test of significance for Channel A step trend and linear trend from 1980 through September 2003 for the Big Coulee near Churchs Ferry, North Dakota, station.

[Numbers in parentheses are for data that were thinned to remove serial correlation.]

Site number (figure 2)	U.S. Geological Survey station name	Coefficient for Channel A step trend	p-value for Channel A step trend	Coefficient for linear trend from 1980 through September 2003	p-value for linear trend from 1980 through September 2003
Chloride, dissolved					
7	Big Coulee near Churchs Ferry, North Dakota	0.105 (.160)	0.067 (.025)	-0.015 (-.016)	0.001 (.014)
Nitrite plus nitrate as nitrogen, dissolved					
7	Big Coulee near Churchs Ferry, North Dakota	-0.115 (-.018)	0.662 (.951)	-0.030 (-.025)	0.100 (.235)

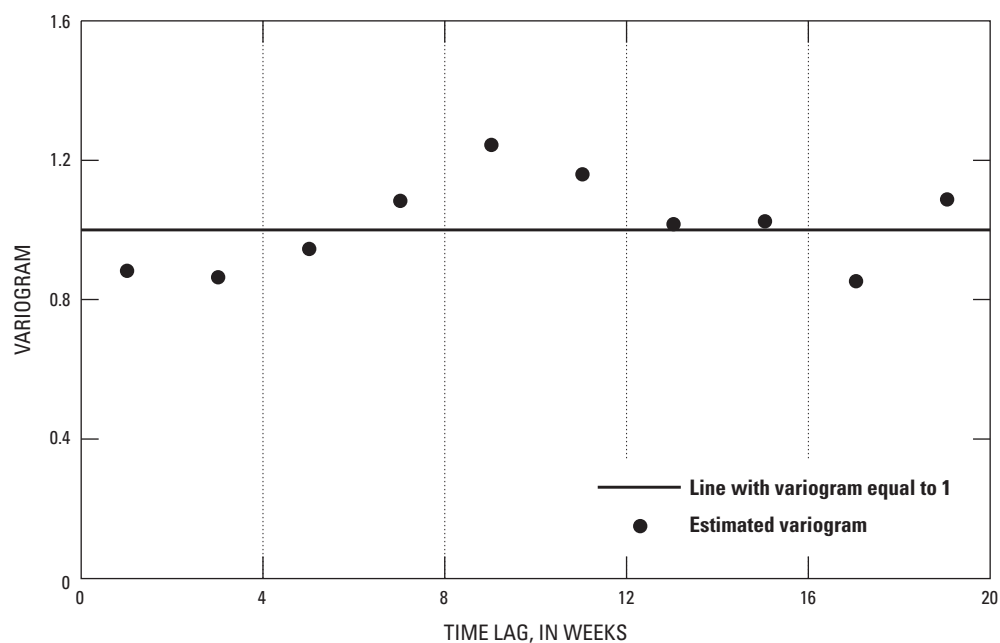


Figure A6. Estimated variogram for flow-adjusted dissolved sulfate concentrations for the Big Coulee near Churchs Ferry, North Dakota, station for season 1 (March through May).

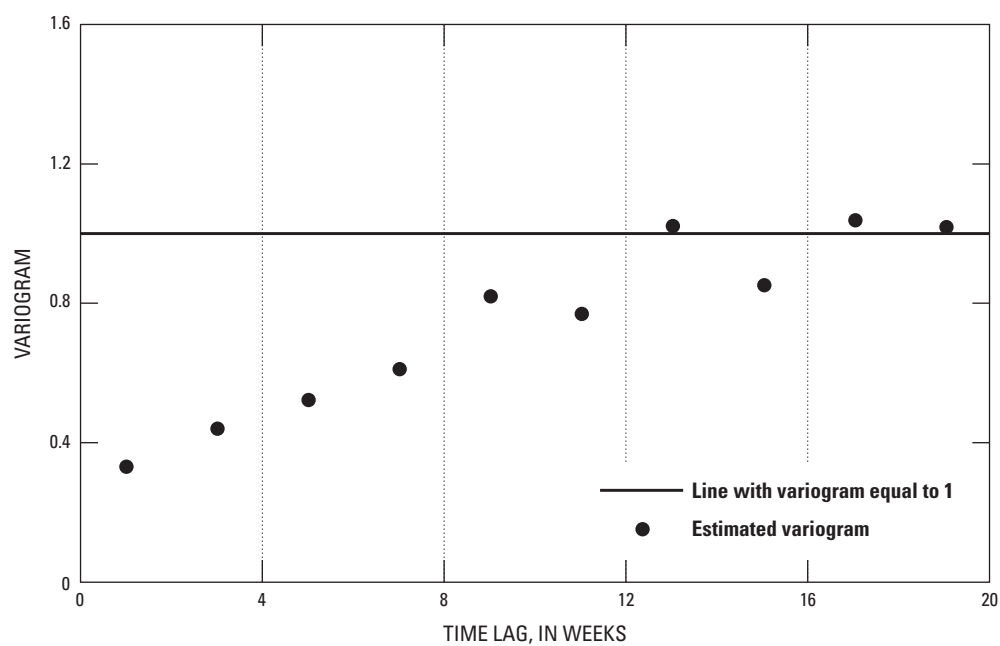


Figure A7. Estimated variogram for flow-adjusted dissolved sulfate concentrations for the Big Coulee near Churchs Ferry, North Dakota, station for season 2 (June through November).

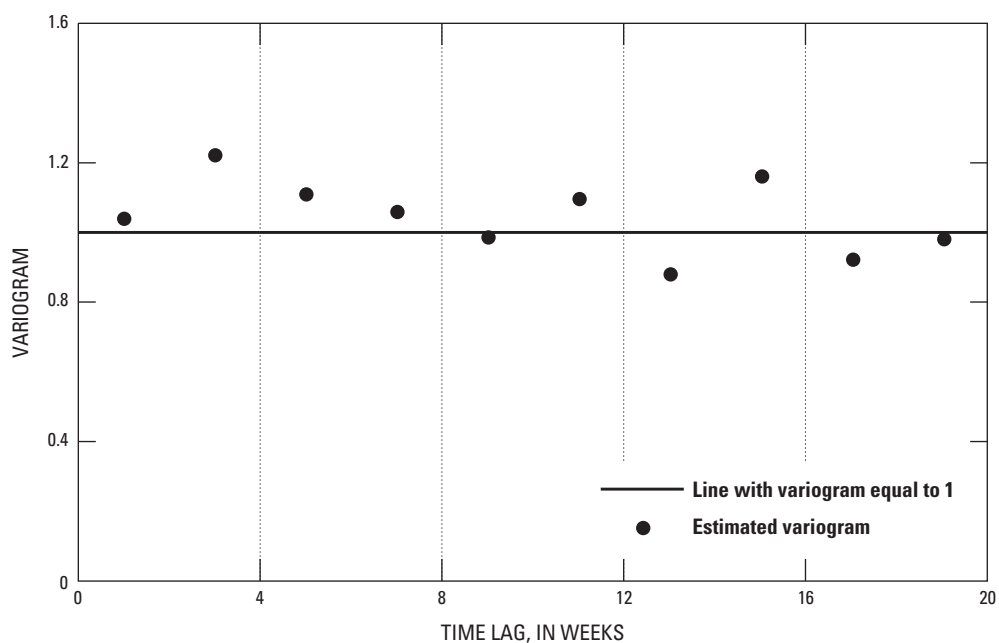


Figure A8. Estimated variogram for flow-adjusted total phosphorus concentrations for the Channel A near Penn, North Dakota, station for season 1 (March through May).

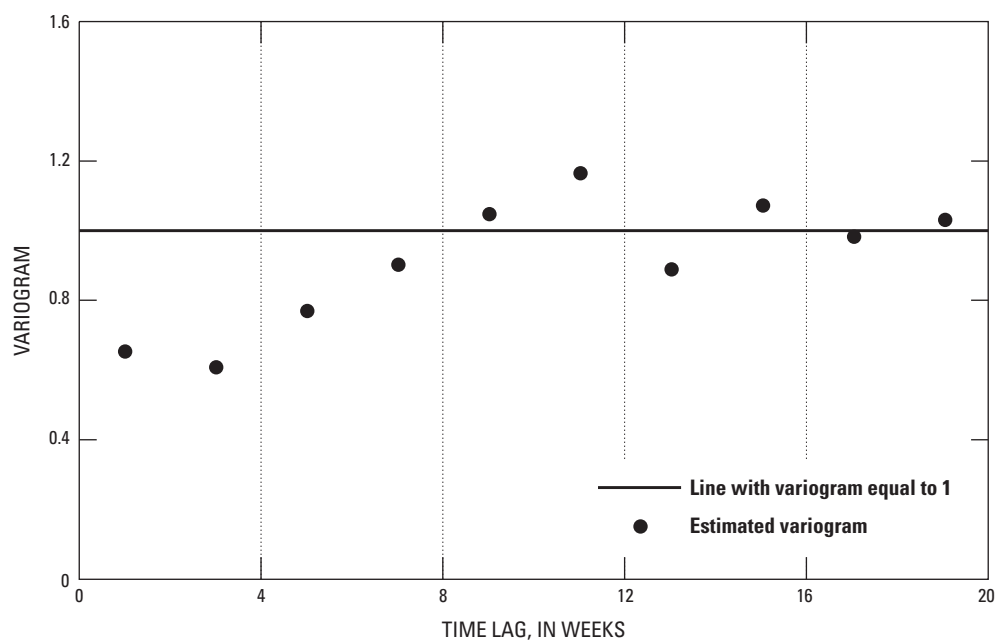


Figure A9. Estimated variogram for flow-adjusted total phosphorus concentrations for the Channel A near Penn, North Dakota, station for season 2 (June through November).

For more information concerning the
research in this report, contact:

Director, U.S. Geological Survey
North Dakota Water Science Center
821 East Interstate Avenue
Bismarck, North Dakota 58503
(701) 250-7400

<http://nd.water.usgs.gov/>

