Water-Quality Trend Analysis and Sampling Design for the Souris River, Saskatchewan, North Dakota, and Manitoba

By Aldo V. Vecchia

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

U.S. GEOLOGICAL SURVEY
CHARLES G. GROAT, Director

CONTENTS

FIGURES, Continued

Water-Quality Trend Analysis and Sampling Design for the Souris River, Saskatchewan, North Dakota, and Manitoba

By Aldo V. Vecchia

Abstract

The Souris River Basin is a 24,600-square-mile basin located in southeast Saskatchewan, north-central North Dakota, and southwest Manitoba. The Souris River Bilateral Water Quality Monitoring Group, formed in 1989 by the governments of Canada and the United States, is responsible for documenting trends in water quality in the Souris River and making recommendations for monitoring future water-quality conditions. This report presents results of a study conducted for the Bilateral Water Quality Monitoring Group by the U.S. Geological Survey, in cooperation with the North Dakota Department of Health, to analyze historic trends in water quality in the Souris River and to determine efficient sampling designs for monitoring future trends. U.S. Geological Survey and Environment Canada water-quality data collected during 1977-96 from four sites near the boundary crossings between Canada and the United States were included in the trend analysis.

A parametric time-series model was developed for detecting trends in historic constituent concentration data. The model can be applied to constituents that have at least 90 percent of observations above detection limits of the analyses, which, for the Souris River, includes most major ions and nutrients and many trace elements. The model can detect complex nonmonotonic trends in concentration in the presence of complex interannual and seasonal variability in daily discharge. A key feature of the model is its ability to handle highly irregular sampling intervals. For example, the intervals between concentration measurements may be as short as 10 days to as long as several months, and the number of samples in any given year can range from zero to 36.

Results from the trend analysis for the Souris River indicated numerous trends in constituent concentration. The most significant trends at the two sites located near the upstream boundary crossing between Saskatchewan and North Dakota consisted of increases in concentrations of most major ions, dissolved boron, and dissolved arsenic during 1987-91 and decreases in concentrations of the same constituents during 1992-96. Significant trends at the two sites located near the downstream boundary crossing between North Dakota and Manitoba included increases in dissolved sodium, dissolved chloride, and total phosphorus during 1977-86, decreases in dissolved oxygen and dissolved boron and increases in total phosphorus and dissolved iron during 1987-91, and a decrease in total phosphorus during 1992-96.

The time-series model also was used to determine the sensitivity of various sampling designs for monitoring future water-quality trends in the Souris River. It was determined that at least two samples per year are required in each of three seasons--March through June, July through October, and November through February--to obtain reasonable sensitivity for detecting trends in each season. In addition, substantial improvements occurred in sensitivity for detecting trends by adding a third sample for major ions and trace elements in March through June, adding a third sample for nutrients in July through October, and adding a third sample for nutrients, trace elements, and dissolved oxygen in November through February.

INTRODUCTION

The Souris River Basin is a 24,600-square-mile basin located in southeast Saskatchewan, north-central North Dakota, and southwest Manitoba (fig. 1). The Souris River originates in Saskatchewan, flows southeast across the international boundary into North Dakota, and then veers north and recrosses the international boundary into Manitoba before eventually flowing into the Assiniboine River near Brandon, Manitoba.
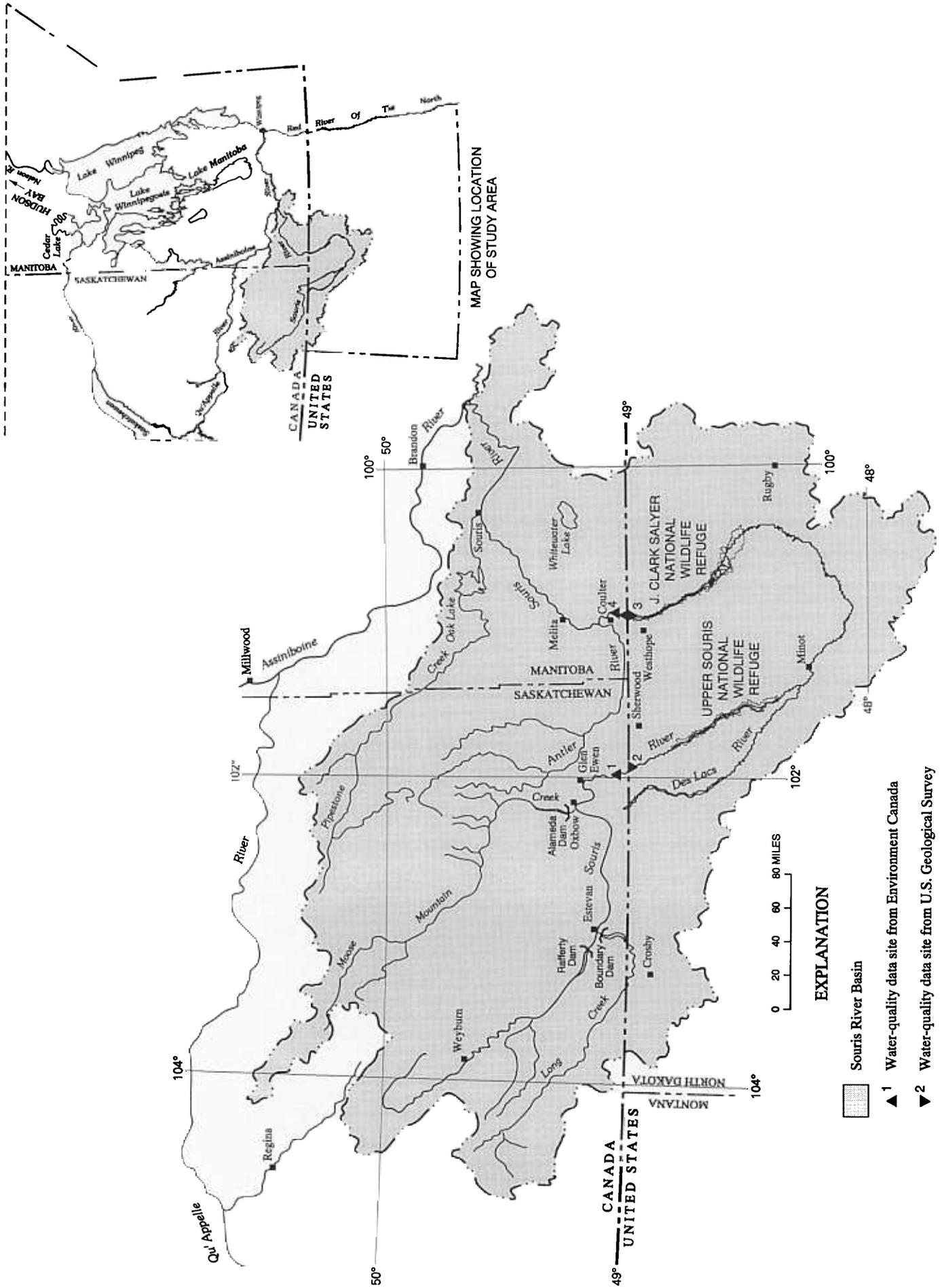
Figure 1. Location of Souris River Basin, Saskatchewan, North Dakota, and Manitoba, and locations of sites used in study.

Maintaining good water quality in the Souris River is important for human health and ecological concerns. Several communities in Saskatchewan, North Dakota, and Manitoba rely on the Souris River for all or part of their water supply. For example, Minot, N. Dak., with a population of about 35,000 (U.S. Department of Commerce, Bureau of the Census, 1981), uses ground water pumped from aquifers that are recharged by the Souris River for its water supply. The Souris River also is an important source of water for waterfowl, fish, and other aquatic organisms. For example, the Upper Souris and J. Clark Salyer National Wildlife Refuges in North Dakota rely on the Souris River to maintain prime habitat for migratory birds and fish. Because Souris River water eventually makes its way to Lake Winnipeg and Hudson Bay, maintaining good water quality in the Souris River also is important for maintaining the aquatic health of those prime fisheries.

The Souris River Bilateral Water Quality Monitoring Group was formed in 1989 by the governments of Canada and the United States. The Bilateral Group is responsible for documenting water-quality trends in the Souris River and making recommendations regarding sampling requirements for monitoring those trends. Therefore, under the request and guidance of the Bilateral Group, the U.S. Geological Survey (USGS), in cooperation with the North Dakota Department of Health, conducted a study to analyze historic data for water-quality trends in the Souris River and to determine efficient sampling designs for monitoring future trends. This report presents results of the study.

Water-quality data from the USGS and Environment Canada (EC) were analyzed for trends using a statistical time-series model described in this report. The time-series model filters out interannual and seasonal variability in concentration that results from trends and natural seasonal fluctuations in discharge. Thus, the model decreases the chance of identifying natural variability as a trend and increases the chance of detecting a true trend. The model also filters out serial persistence, or autocorrelation, between neighboring concentration measurements that can bias estimated trends and associated significance levels. For example, during low-flow conditions, discharge may be regulated by upstream reservoir releases. Constituent concentrations in a reservoir may be highly persistent in that measurements obtained several months apart may be highly correlated. Therefore, several water-quality samples collected during low-flow conditions in a given year may not contain much more information than a single sample collected during the same time period.

The time-series model also was used to determine efficient sampling designs for monitoring future trends to maintain or reduce sampling costs incurred during past monitoring efforts. Alternate sampling designs were determined, and information gained or lost by increasing or decreasing the number of samples collected was quantified.

The parametric time-series approach used in this study differs from more commonly used nonparametric approaches, such as seasonal Kendall's tau (Hirsch and Slack, 1984). Parametric tests are more powerful and flexible than nonparametric tests for detecting nonmonotonic trends, managing highly irregular sampling frequencies, and determining efficient sampling designs, but nonparametric tests are easier to apply and require fewer assumptions than parametric tests.

DATA USED FOR WATER-QUALITY TREND ANALYSIS

The focus of this study was on two international boundary crossings near Sherwood and Westhope, N. Dak. Data from four sites near the crossings were used for the trend analysis (table 1; fig. 1). The first two sites will be referred to as the Sherwood sites and the last two sites will be referred to as the Westhope sites.

Data from the USGS and EC sites were combined to obtain a reasonable period of record and sample size for calibrating the model. However, because USGS and EC data collected before 1987 were difficult to combine because of differences in sampling and laboratory techniques, only USGS data were used for 1977-86. Also, few data were available for sites 2 and 3 during 1987-91, and water-quality sampling was discontinued at sites 1 and 4 during 1992. Therefore, data used in the study were collected at sites 2 and 3 before 1987, sites 1 and 4 during 1987-91, and sites 2 and 3 after 1991.

Some data requirements of the model limited the set of potential constituents included in the analysis. For example, although sampling frequency and timing can vary from year to year, the model requires at least 60 concentration measurements during at least 15 years (the years may be nonconsecutive). Therefore, 36 measurements per year for 5 years (180 measurements) are not sufficient, but 4 measurements per year for 15 years (60 measurements) are sufficient

Table 1. Sites used for trend analysis

[EC, Environment Canada; USGS, U.S. Geological Survey]

Site number	Site	Agency	Location	Period of record used in study
	Souris River south of Glen Ewen, Saskatchewan	EC	4.1 miles upstream of first boundary crossing	[1]1987-91
	Souris River near Sherwood, N. Dak.	USGS	0.8 mile downstream of first boundary crossing	1977-86, 1992-96
	Souris River near Westhope, N. Dak.	USGS and EC	0.2 mile upstream of second boundary crossing	1977-86, 1992-96
	Souris River near Coulter, Manitoba	EC	5.1 miles downstream of second boundary crossing	[1]1987-91

[1]Water-quality data collected before 1987 are available from Environment Canada. However, because U.S. Geological Survey and Environment Canada data collected before 1987 were difficult to combine because of differences in sampling and laboratory techniques, only U.S. Geological Survey data were used for 1977-86.

Also, sampling frequency can be as often as three times per month (about a 10-day sampling interval) or as seldom as an average of four times per year (the sampling frequency for some years may be less than four). The most restrictive requirement is that fewer than 10 percent of the measurements are less than the detection limit. This requirement excluded most pesticides and many trace elements from the analysis. Methods currently are being developed for extending the parametric time-series approach used in this study to allow censoring rates greater than 10 percent, but, in this study, highly censored constituents were not analyzed.

After the available data were screened to determine which constituents were amenable to the parametric time-series approach, the Souris River Bilateral Water Quality Monitoring Group chose 13 constituents for the analysis. Those constituents are given in tables 2 and 3 along with selected summary statistics. Priority was given to constituents that are critical to the aquatic health of the river. The data for the constituents given in tables 2 and 3 were processed for input into the model and, therefore, may not include all data available for the given time periods and sites. For example, if several measurements were recorded for the same day, only the first measurement was used. Also, measurements obtained less than 10 days apart were deleted from the data used because the model requires sampling intervals of at least 10 days.

Trace-element data collected before 1992 were primarily dissolved concentrations, and trace-element data collected during 1992-96 were primarily total concentrations. Because dissolved and total concentrations of arsenic and boron are comparable, concentrations for each of those constituents were combined for the analysis. However, dissolved and total concentrations of iron and manganese are not comparable and, therefore, only dissolved concentrations were included in the analysis. Dissolved iron and manganese concentrations for the Sherwood sites also were not analyzed for trends because the data do not meet the model sampling requirement of at least 60 measurements during 15 years (table 2).

TIME-SERIES MODEL FRAMEWORK

The theory behind the time-series model is given in detail in appendix A. A simplified overview of the model is given in this section, and trend analysis results for all constituents are given in a later section.

The model was used to filter out as much natural variability in concentration as possible before analyzing for trends. Much of this natural variability usually is induced by variability in discharge, which is extremely large for streams (including the Souris River) in this prairie region. To illustrate the model, recorded daily discharges for the Souris River near Sherwood, N. Dak., for 1977-96 are shown in figure 2, and recorded dissolved sulfate concentrations for the Sherwood sites are shown in figure 3. The approximate 10-day sampling interval used in figure 2 was simulated by including three daily discharges per month. A 10-day sampling interval was used because that is the smallest interval that is computationally manageable. Use of a larger interval, such as monthly, would simplify the trend analysis but would not account for short-term variability in discharge.

Table 2. Summary of water-quality data for selected constituents, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

[USGS, U.S. Geological Survey; EC, Environment Canada; med, median; min, minimum; max, maximum; n, number of samples; mg/L, milligrams per liter; µg/L, micrograms per liter]

Constituent and USGS (EC) code	Unit	Souris River south of Glen Ewen, Saskatchewan (1987-91), med (min, max) n	Souris River near Sherwood, North Dakota (1977-86), med (min, max) n	Souris River near Sherwood, North Dakota (1992-96), med (min, max) n
Dissolved oxygen 00300 (08101W)	mg/L	8.5 (1; 14.2) 32	7.1 (0.2; 47) 133	8.6 (2.3; 16.2) 42
Dissolved calcium 00915 (20103L)	mg/L	53 (14; 183) 32	63 (26; 140) 95	49 (16; 110) 37
Dissolved sodium 00930 (11103L)	mg/L	173 (47; 401) 32	120 (27; 350) 95	160 (14; 440) 37
Dissolved sulfate 00945 (16306L)	mg/L	237 (75; 598) 32	210 (82; 550) 96	280 (61; 1,000) 37
Dissolved chloride 00940 (17206L)	mg/L	84 (17; 199) 32	49 (7; 190) 96	47 (4; 220) 37
Dissolved solids 70300 (00201L)	mg/L	839 (212; 1,859) 32	728 (236; 1,749) 127	822 (203; 2,540) 38
Total ammonia 00610 (07506P)	mg/L	0.035 (0.016; 0.77) 32	0.10 (0.01; 1.8) 97	0.045 (0.02; 1.3) 8
Total nitrogen 00600 (07603L)	mg/L	0.93 (0.39; 3.02) 32	.54 (0.66; 3.53) 86	.7 (1.2; 7.55) 37
Total phosphorus 00665 (15423L)	mg/L	0.25 (0.10; 0.93) 32	0.13 (0.02; 0.68) 106	0.25 (0.10; 1.9) 37
Dissolved arsenic[1] 01000 (33108D)	µg/L	3.8 (0.8; 15) 32	3 (1; 8) 26	3 (1; 18) 39
Dissolved boron[1] 01020 (05105D)	µg/L	490 (185; 4,638) 31	250 (70; 940) 96	325 (40; 2,399) 38
Dissolved iron 01046 (26104G)	µg/L	43 (11; 1,080) 32	50 (4; 3,200) 27	No data
Dissolved manganese 01056 (25104G)	µg/L	28 (3; 3,271) 32	125 (3; 4,000) 26	No data

[1]Concentrations for 1992-96 are total rather than dissolved.

The solid line in figure 2 represents smooth interannual and intra-annual (seasonal) fluctuations in discharge. The line was obtained using a smoothing algorithm that separates low-frequency (annual and seasonal) variability from high-frequency variability (appendix A). The smoothed data will be called the low-frequency component (LFC) of discharge, and the deviations of the recorded values from the LFC will be called the high-frequency component (HFC) of discharge. Both components are important in explaining variability in concentration.

The solid line in figure 3 represents the fitted LFC of concentration obtained from a nonlinear regression of concentration on the LFC of discharge (appendix A). A strong dilution effect is shown by the tendency for the LFC of concentration to mirror the LFC of discharge. Because the LFC consists of natural seasonal and discharge-related variability in concentration that may hide true trends in concentration, it is advantageous to remove the LFC before analyzing for trends. Therefore, flow-adjusted concentrations are obtained by subtracting the LFC of concentration (fig. 3) from the recorded data and then adding the mean of the recorded data. Flow-adjusted concentrations can be interpreted loosely as the concentrations that would have been observed if flow conditions had been uniform throughout the entire

Table 3. Summary of water-quality data for selected constituents, Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

[USGS, U.S. Geological Survey; EC, Environment Canada; med, median; min, minimum; max, maximum; n, number of samples; mg/L, milligrams per liter; μg/L, micrograms per liter]

Constituent and USGS (EC) code	Unit	Souris River near Westhope, North Dakota (1977-86), med (min, max) n	Souris River near Westhope, North Dakota (1992-96), med (min, max) n	Souris River near Coulter, Manitoba (1987-91), med (min, max) n
Dissolved oxygen 00300 (08101W)	mg/L	9.1 (0.1; 22) 124	7.8 (0.1; 17) 50	8.1 (0.2; 18) 50
Dissolved calcium 00915 (20103L)	mg/L	62 (29; 190) 66	53 (27; 160) 44	56 (17; 149) 47
Dissolved sodium 00930 (11103L)	mg/L	115 (28; 330) 66	145 (39; 330) 44	152 (13; 339) 47
Dissolved sulfate 00945 (16306L)	mg/L	220 (60; 760) 66	223 (76; 462) 44	200 (30; 538) 47
Dissolved chloride 00940 (17206L)	mg/L	30 (9; 100) 66	36 (11; 100) 44	47 (4; 122) 47
Dissolved solids 70300 (00201L)	mg/L	730 (230; 2,000) 99	745 (273; 1,919) 45	732 (129; 1,588) 51
Total ammonia 00610 (07506P)	mg/L	0.2 (0.005; 3.6) 77	0.12 (0.005; 1.34) 38	0.1 (0.032; 4.74) 48
Total nitrogen 00600 (07603L)	mg/L	2.2 (0.95; 36.1) 85	2.0 (0.58; 5.9) 44	2.2 (0.85; 8.2) 46
Total phosphorus 00665 (15423L)	mg/L	0.24 (0.02; 2) 99	0.34 (0.06; 3.2) 44	0.38 (0.13; 1.1) 46
Dissolved arsenic[1] 01000 (33108D)	μg/L	5 (1; 11) 26	3.6 (1.5; 8.6) 32	3.7 (1.2; 9.9) 51
Dissolved boron[1] 01020 (05105D)	μg/L	240 (80; 430) 21	163 (44; 305) 22	263 (160; 1,270) 51
Dissolved iron 01046 (26104G)	μg/L	30 (10; 310) 27	44 (15; 250) 12	35 (8; 190) 46
Dissolved manganese 01056 (25104G)	μg/L	65 (3; 2,700) 26	68 (7; 2,300) 12	80 (2; 2,581) 45

[1]Concentrations for 1992-96 are total rather than dissolved.

sampling period. Although the LFC of concentration accounts for much of the variability in concentration, the flow-adjusted concentration also contains considerable variability as shown in figure 4 for dissolved sulfate concentrations

Also shown in figure 4 is the fitted trend in flow-adjusted dissolved sulfate concentrations obtained using methods described in appendix A. The fitted trend consists of a combination of several simple trends--monotonic (log-linear) trends during 1977-86, 1987-91, and 1992-96; a step trend in 1987; and an equal but opposite step trend in 1992. The time intervals for the monotonic trends were selected on the basis of empirical analysis of the data, and the time interval of the step trend was selected on the basis of changes in sampling locations between USGS and EC sites. The magnitudes and associated significance levels of the fitted trends are given in a later section. The ability of the model to handle missing data allowed the use of a 10-day sampling interval although most of the 10-day intervals had missing concentration data. A larger interval, such as monthly, could have been used for concentration, but the 10-day interval accounted for the timing of water-quality samples within the month and allowed the full set of daily discharges shown in figure 2 to be used in the model to filter out discharge-related variability in concentration.
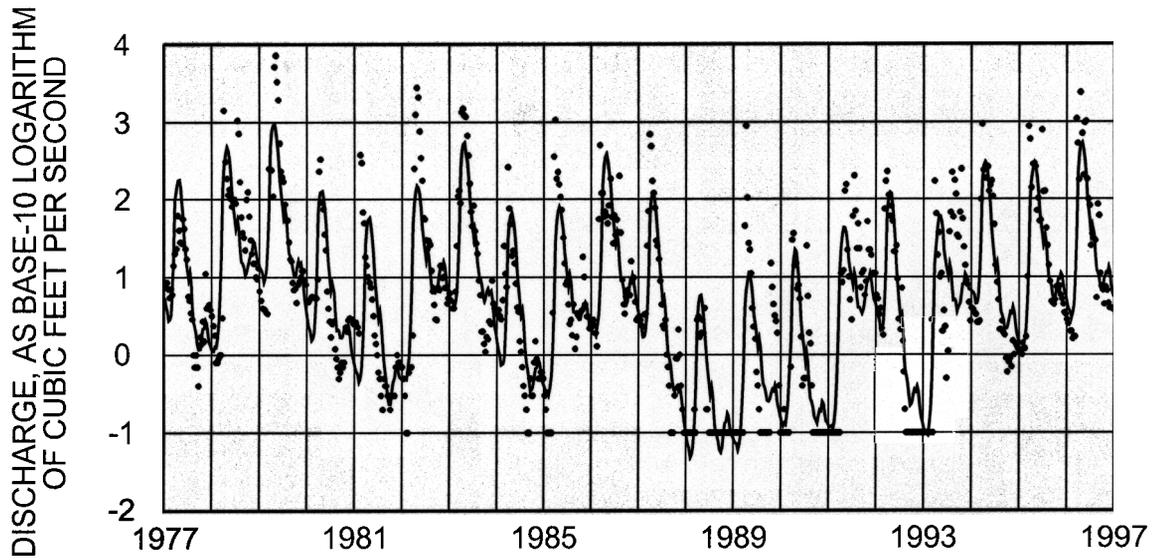
**Figure 2.** Recorded daily discharges (points, 36 per year) and fitted low-frequency component (line), Souris River near Sherwood, North Dakota, 1977-96.
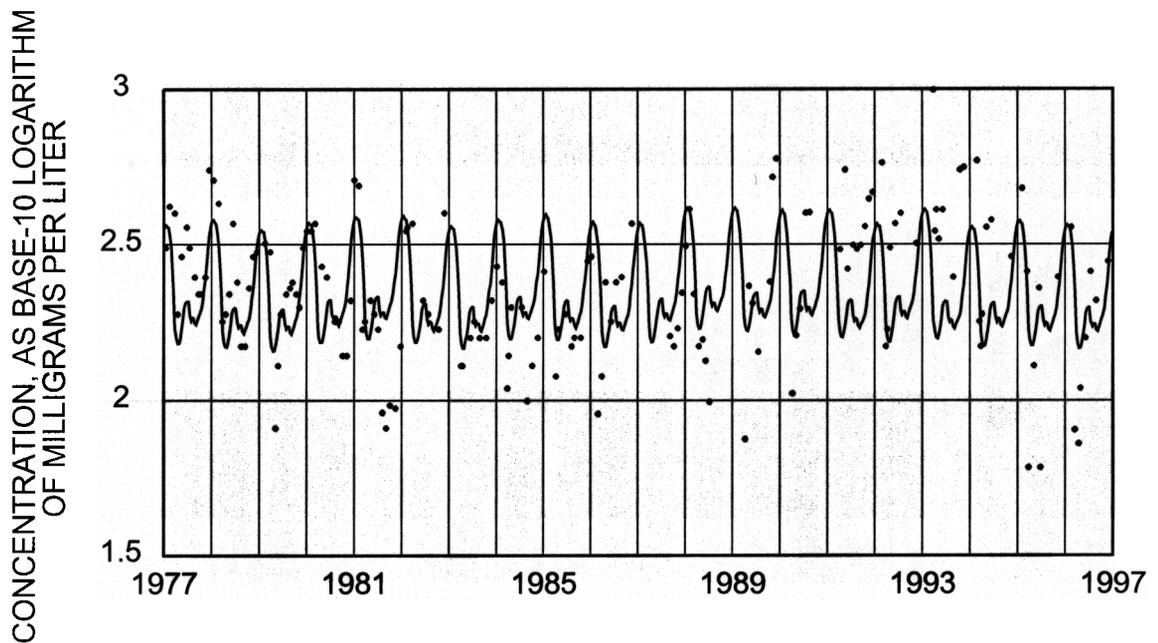

Figure 3. Recorded dissolved sulfate concentrations (points) and fitted low-frequency component showing discharge-related variability (line), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96.
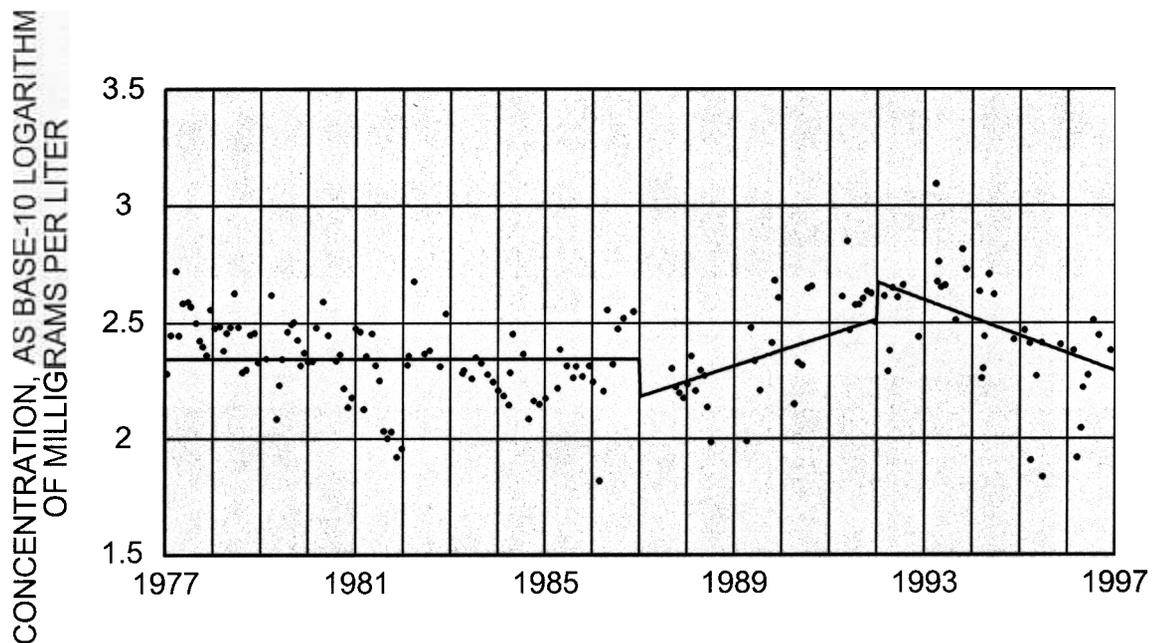
Figure 4. Flow-adjusted dissolved sulfate concentrations (points) and fitted trend (line), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96.

The model also was used to filter out the complicated serial persistence that exists in the HFC of concentration on a seasonal time scale. This was accomplished by calibrating a periodic autoregressive moving-average (PARMA) model to the HFCs of discharge and concentration using procedures given in appendix A. The HFC of concentration is the time series of deviations of flow-adjusted concentrations from a fitted trend line.

The PARMA model can be used to detect complex seasonal structure in the HFCs of concentration and discharge. For example, the fitted standard deviation for the HFC of dissolved sulfate concentration, the cross-correlation at a 0-day lag with the HFC of discharge, and the autocorrelation at a 30-day lag are shown in figure 5. The first plot indicates the variability in the HFC of concentration is largest during spring runoff, reaches a secondary peak in late fall, and is smallest in summer. The second plot indicates the correlation between the HFC of concentration and the HFC of discharge is negative in February through April, makes a transition from negative to positive in May, remains positive through the summer and fall, and becomes negative again in January. Thus, higher-than-normal discharge in February through April (for example, from an earlier-than-normal spring breakup) tends to result in lower-than-normal dissolved sulfate concentrations. Also, higher-than-normal discharge in the summer and fall (for example, from a large rainfall-runoff event) tends to produce higher-than-normal dissolved sulfate concentrations. The third plot indicates the seasonal structure in the autocorrelation between neighboring values of the HFC of concentration. For example, the HFC of concentration in February through June is only moderately correlated (autocorrelation less than 0.5) with the HFC of concentration 30 days earlier, but the HFC of concentration in August through November is highly correlated (a correlation coefficient greater than 0.7) with the HFC of concentration 30 days earlier. Although rigorous physical explanations for these autocorrelations are difficult to determine, some qualitative observations may explain some of the structure in the autocorrelations. For example, the low autocorrelation in June occurs during the transition between the spring runoff recession and the summer low-flow period. Spring runoff is dominated by snowmelt, and the summer low-flow period is dominated by base flow and reservoir releases. Therefore, concentrations in late June generally are not highly correlated with concentrations 30 days earlier because snowmelt tends to have concentrations that are not correlated with concentrations during low-flow conditions. However, as the summer low-flow period continues, concentrations in August to November become highly correlated because base flow and reservoir releases tend to have concentrations that vary
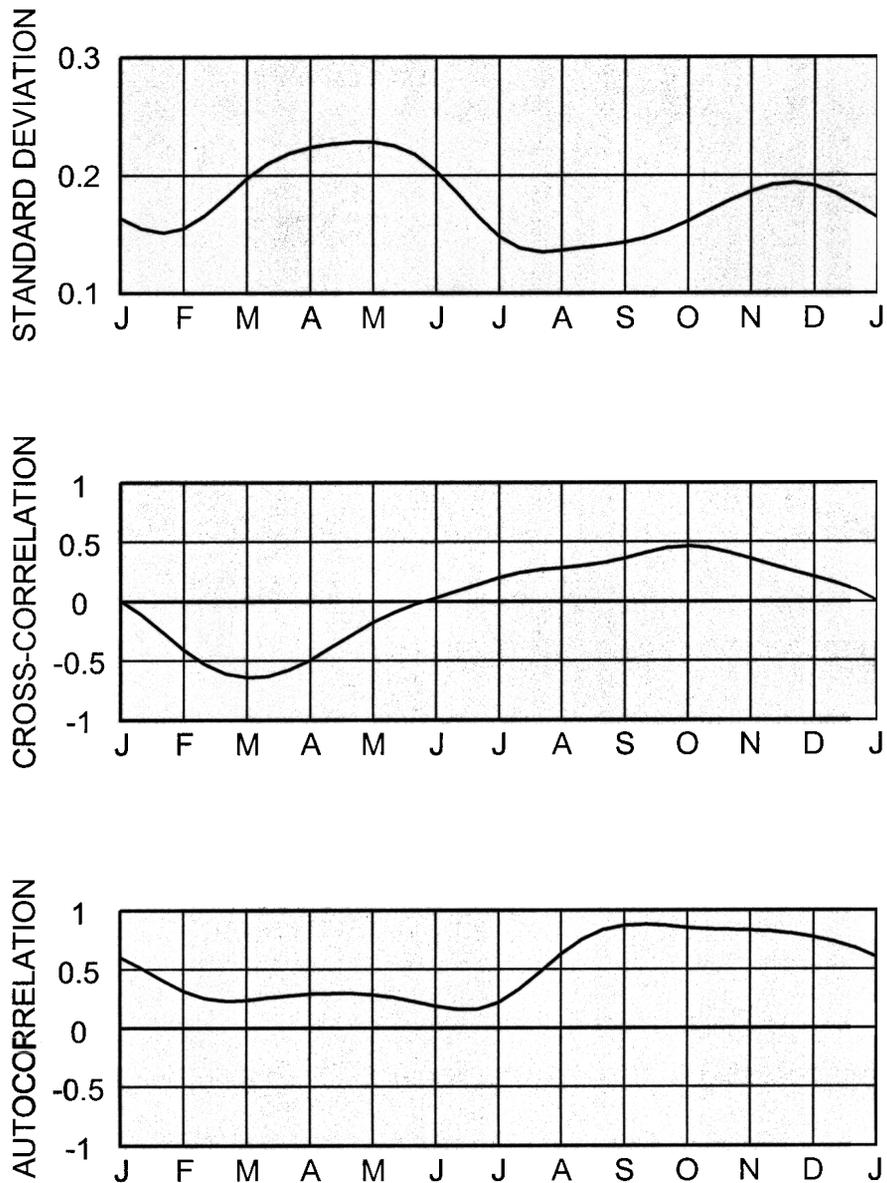
Figure 5. Computed seasonal standard deviation for the high-frequency component of dissolved sulfate concentration, cross-correlation at a 0-day lag with the high-frequency component of discharge, and autocorrelation at a 30-day lag, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota.

slowly with time. The high autocorrelation continues through the fall and winter and decreases at the start of spring breakup. Concentrations during February through June are only moderately correlated, probably because of year-to-year differences in the timing of spring runoff. For example, in years that have an early spring runoff, concentrations in March and April might be similar, but in years that have a late spring runoff, concentrations in March and April might not be similar.

Statistical properties such as those shown in figure 5 can bias estimated trends and significance levels if not properly accounted for in the trend analysis. However, if the PARMA model residuals, defined in appendix A, are approximately

uncorrelated with constant mean and variance, the model can be used to correct the bias in estimated trends. The PARMA model residuals for dissolved sulfate concentrations for the Sherwood sites are shown in figure 6. The nonrandomness in the HFC of dissolved sulfate concentration, as shown in figure 4 and quantified in figure 5, is not evident in the PARMA model residuals shown in figure 6. The PARMA model residuals shown in figure 6 were determined to be white noise (uncorrelated random variables with zero mean and constant variance). Furthermore, the residuals were determined to be approximately normally distributed with the exception of occasional outliers, none of which affected the calibrated model and associated trend analysis.
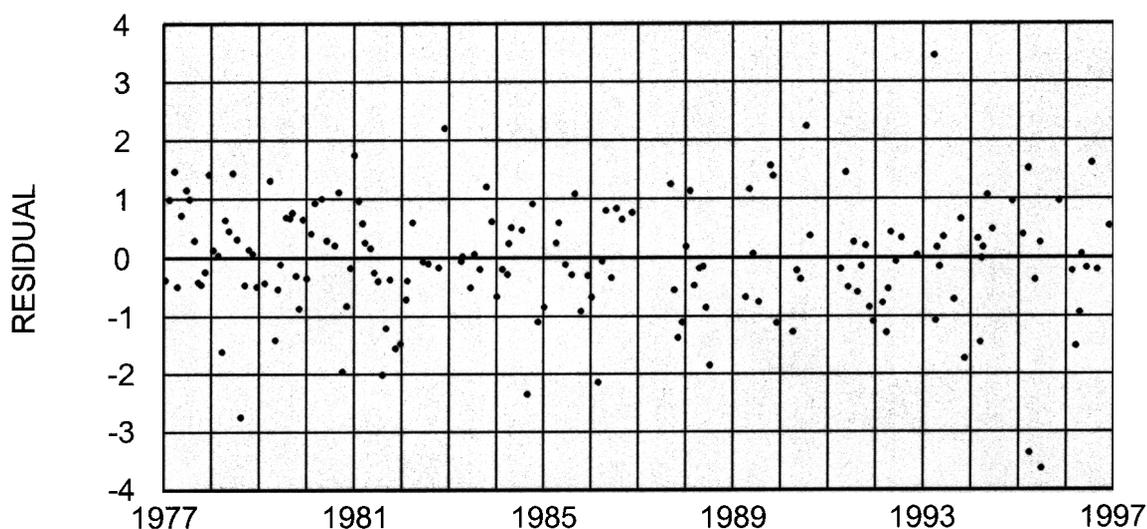

Figure 6. Residuals from periodic autoregressive moving-average (PARMA) model fitted to high-frequency component of dissolved sulfate concentrations, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96.

TREND ANALYSIS RESULTS

A flow-adjusted concentration data set, such as the dissolved sulfate concentration data set shown in figure 4, is said to be trend free if the expected value, or mean, of the data is constant for the entire period of record. Although trends in other data aspects, such as the variance, may be of interest, in this report, only trends in the mean were considered.

Because of autocorrelation in the data, the mean of a time series may be difficult to estimate even if the data are trend free. For example, a hypothetical 5-year continuous record of flow-adjusted concentrations is shown in figure 7. The data were generated from a time-series model using a zero mean and a stationary autocorrelation function (shown in the second plot in figure 7). The autocorrelation function was assumed to be constant throughout each year in contrast to the seasonally varying autocorrelation shown in figure 5. The sampling designs for the hypothetical concentrations also are shown in figure 7--one design (the o's) shows 20 measurements in a single year and the other (the +'s) shows 20 measurements in 5 years. Because of the autocorrelation, 20 measurements in a single year are not sufficient for estimating the mean. However, 20 measurements in 5 years are sufficient because those measurements included interannual variability in the data. Generally, the record length should be at least five times as long as the correlation length scale (the lag beyond which the autocorrelation function is essentially equal to zero) to distinguish a true trend from the effects of autocorrelation. For example, years 2 through 5 of the data shown in figure 7 appear to indicate an increasing trend, but the increase is a result of serial persistence and not a true trend.
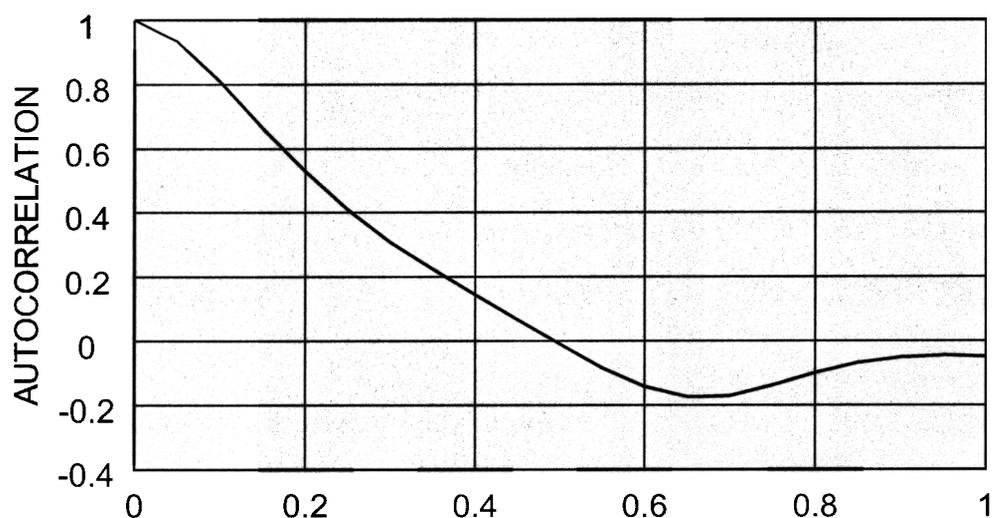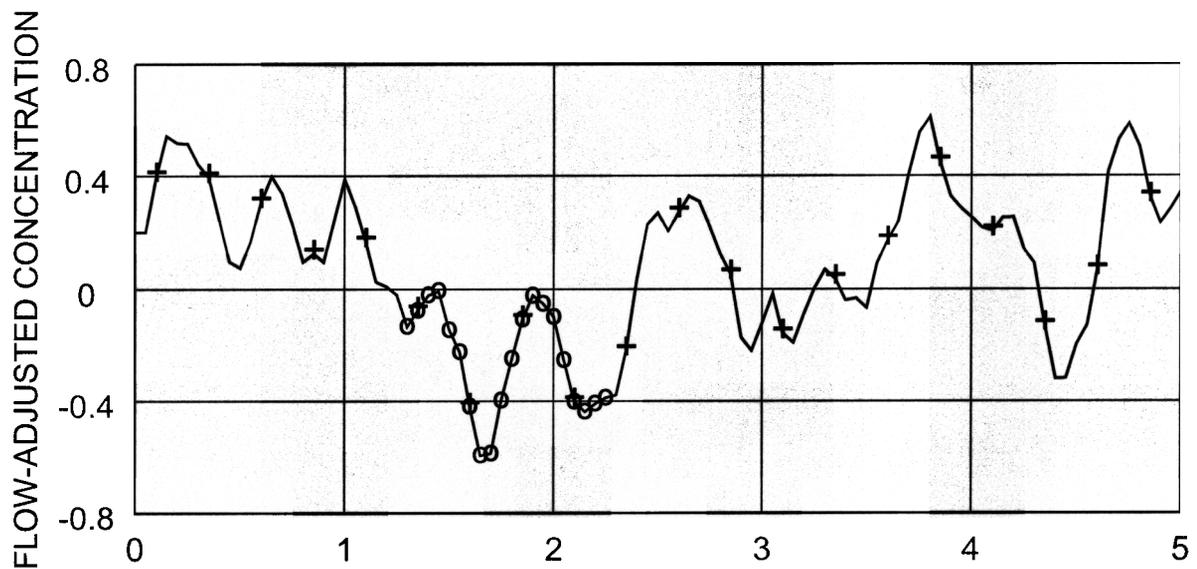
Figure 7. Sampling designs for hypothetical flow-adjusted concentrations (O, 20 measurements in 1 year; +, 20 measurements in 5 years) and autocorrelation function used to generate data.

Preliminary analysis of the data from the Sherwood and Westhope sites indicated a significant autocorrelation out to a lag of several months for many constituents but a negligible autocorrelation beyond a lag of 1 year for all constituents. Therefore, a minimum record length of 5 years is required for detecting trends. This record length is different from the minimum record length of 15 years used to calibrate the PARMA time-series model. The longer record length is required to calibrate the PARMA model, which, in turn, is used to determine the correlation structure of the data.

Trends in the mean can be classified as simple monotonic trends, compound monotonic trends, or step trends. A simple monotonic trend is a consistent increase or decrease in the mean with time, a compound monotonic trend is a series of two or more simple monotonic trends of different direction (for example, see figure 4), and a step trend is an abrupt change, or discontinuity, in the mean at a particular time. Most trends during long time periods are not simple monotonic trends because consistent increases or decreases usually are not persistent. For example, few monotonic trends in the data

for the Sherwood and Westhope sites persisted for the entire 20-year period of record. However, as indicated later, numerous monotonic trends persisted for 5 years or longer, particularly during 1977-86, 1987-91, and 1992-96. Step trends also occurred in 1987 and in 1992, reflecting the changes in sampling location (table 1). Therefore, for this study, the same trend combinations--monotonic trends for 1977-86, 1987-91, and 1992-96; a step trend in 1987; and an equal but opposite step trend in 1992--were assumed for each constituent and each site. The monotonic trends were assumed to consist of linear trends in log-transformed data or, equivalently, exponential trends in untransformed data. This assumption was determined to be a reasonable approximation.

Using the theory of maximum likelihood estimation for PARMA models, as detailed in appendix A, statistical significance levels were determined for each trend variable. For this study, only those monotonic trends that are significant at the 5-percent level (probability or p-value less than 0.05) were reported. However, all step trends were reported regardless of the significance level. A one-tailed test of significance was used throughout the analysis. For example, an estimated increasing trend of 10 percent per year with a p-value of 0.015 indicates a probability of 0.015, or 1.5 percent, that a positive trend of 10 percent or larger could have been erroneously identified even if no actual trend exists. Thus, the smaller the p-value, the larger the confidence level that an actual or real trend has been detected.

Significant Trends for Sherwood Sites

Trend analysis results for the Sherwood sites indicate a lack of trends for most constituents during 1977-86 (figs. 8 through 20). Except for dissolved calcium, total ammonia, and total nitrogen, flow-adjusted concentrations were stable
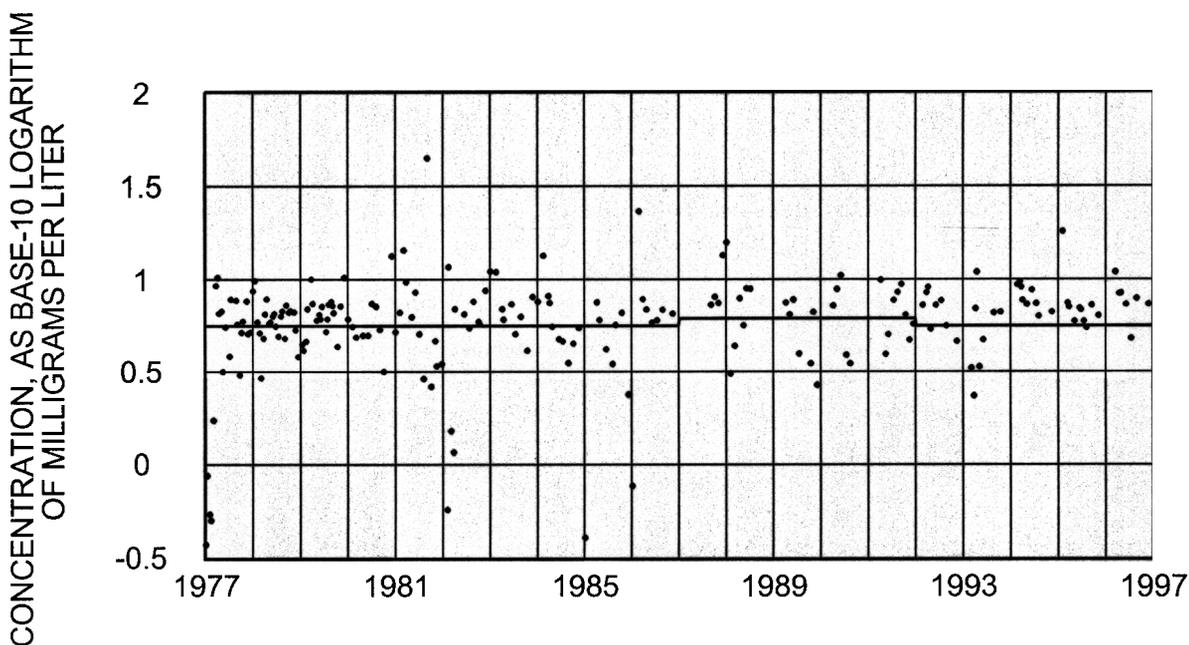

Figure 8. Trends (lines) in flow-adjusted dissolved oxygen concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]

during that period. A significant decrease occurred in dissolved calcium and significant increases occurred in total ammonia and total nitrogen, but, in terms of the remaining constituents, water quality at the Sherwood sites changed little from 1977 to 1986.
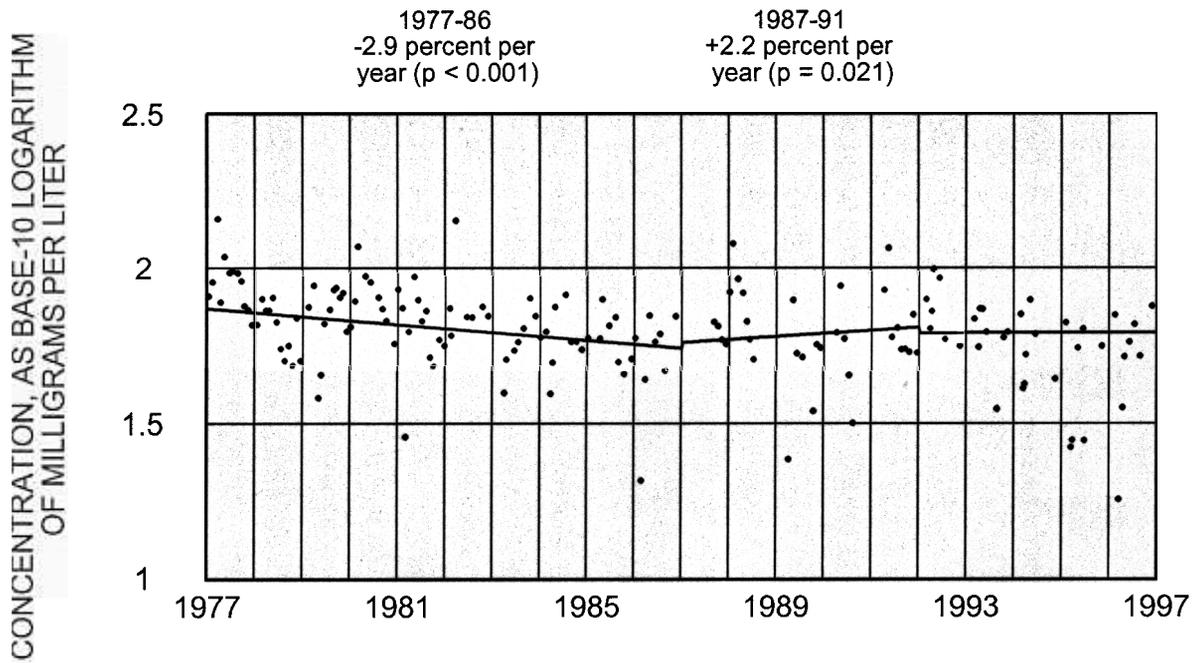
Figure 9. Trends (lines) in flow-adjusted dissolved calcium concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
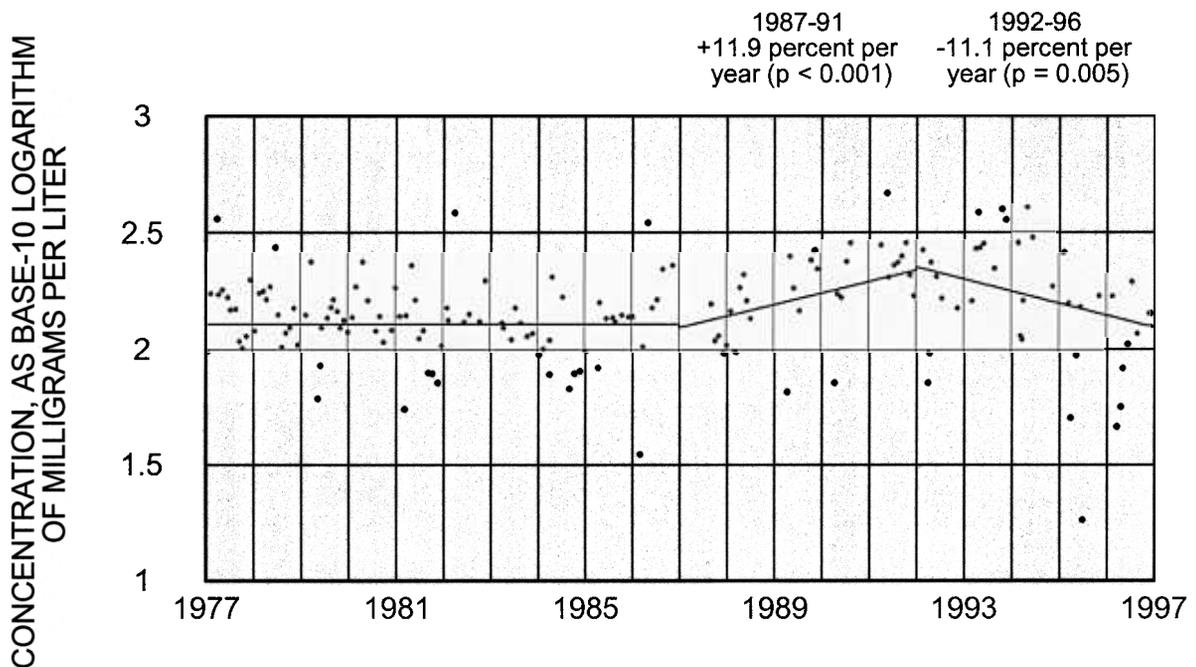
Figure 10. Trends (lines) in flow-adjusted dissolved sodium concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]

1987-91
+16.5 percent per
year (p < 0.001)

1992-96
-16.1 percent per
year (p < 0.001)
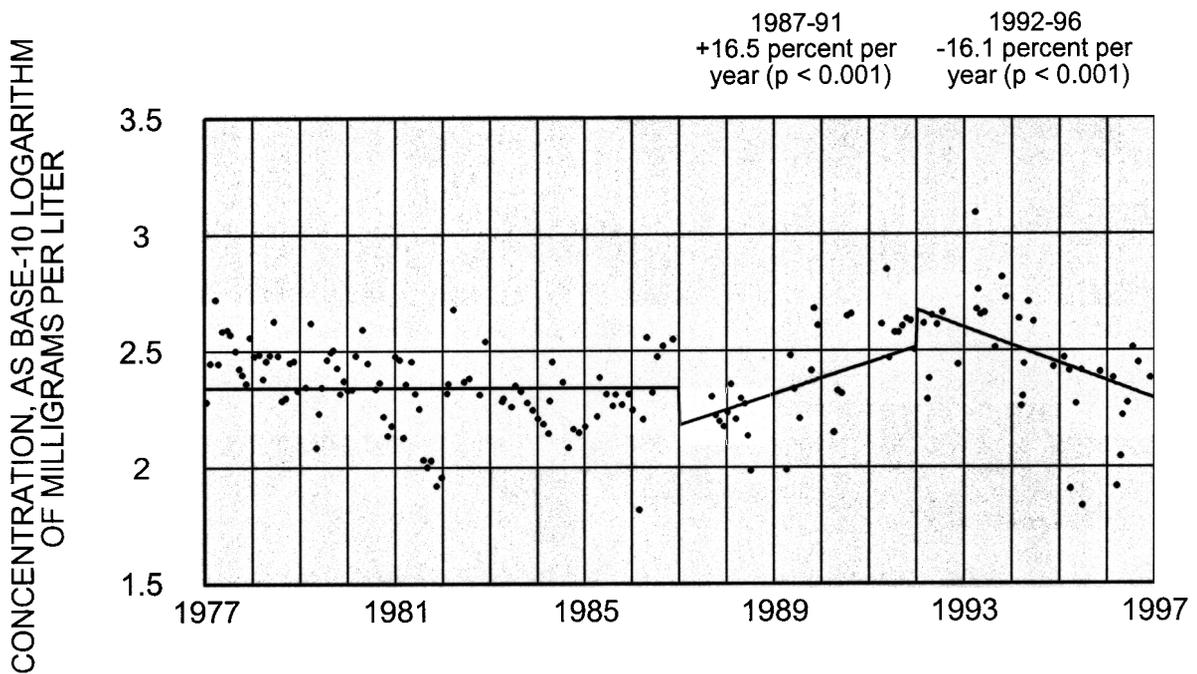
Figure 11. Trends (lines) in flow-adjusted dissolved sulfate concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]

1987-91
+17.1 percent per
year (p < 0.001)

1992-96
-9.8 percent per
year (p = 0.015)
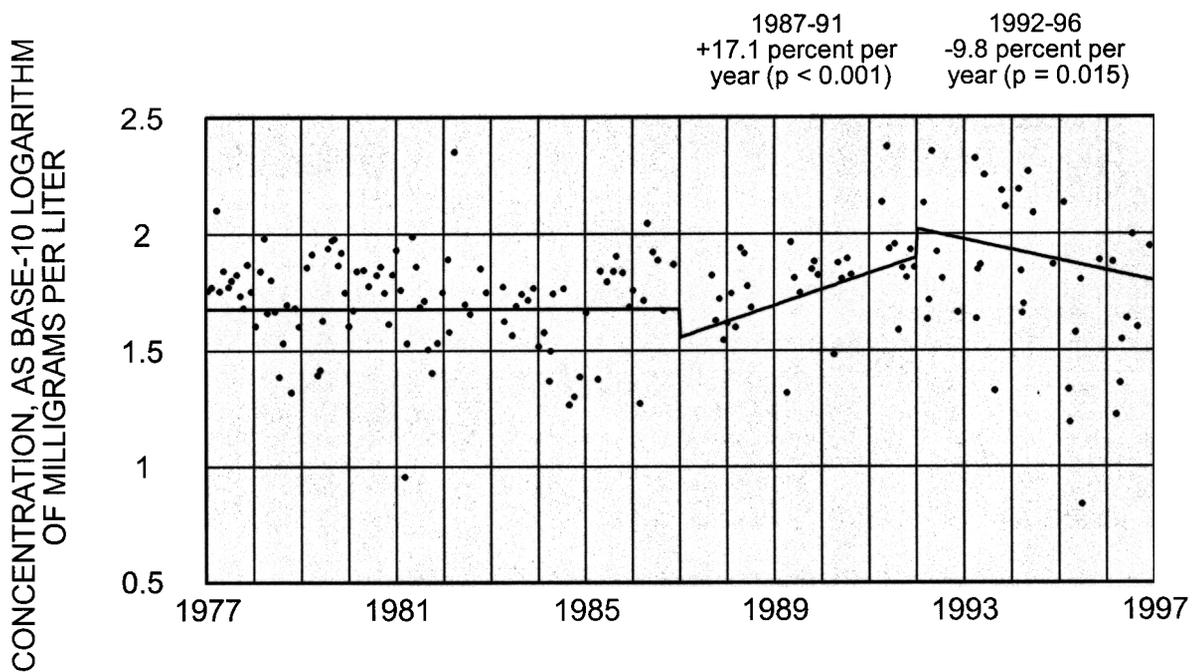
Figure 12. Trends (lines) in flow-adjusted dissolved chloride concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
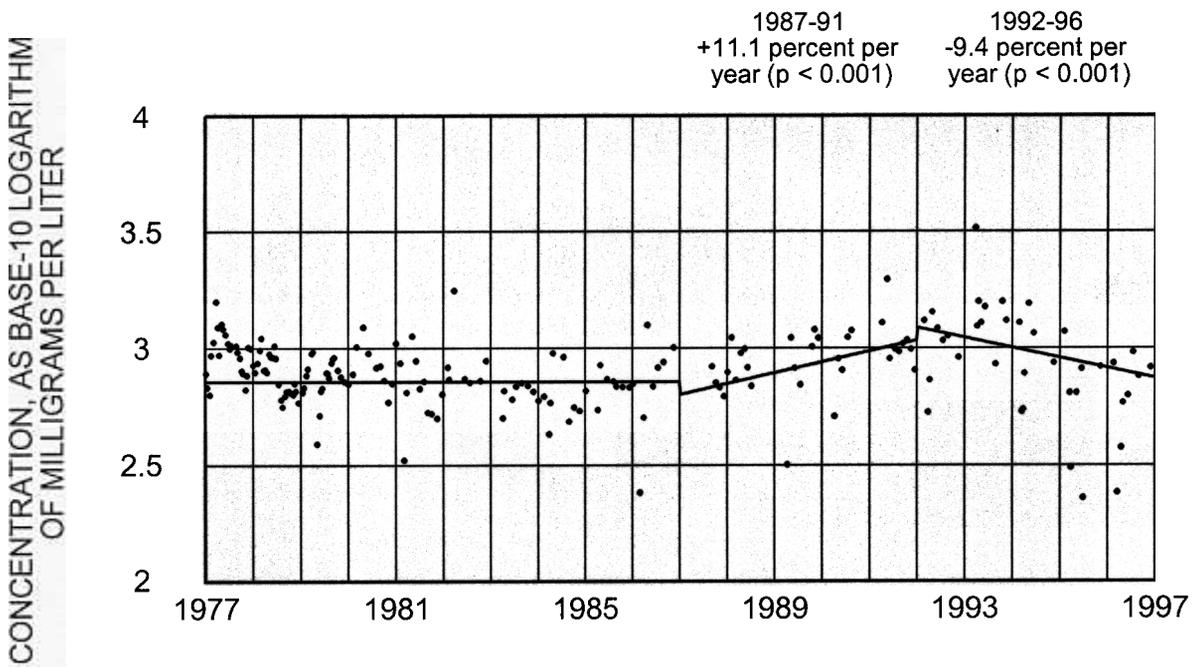
Figure 13. Trends (lines) in flow-adjusted dissolved solids concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
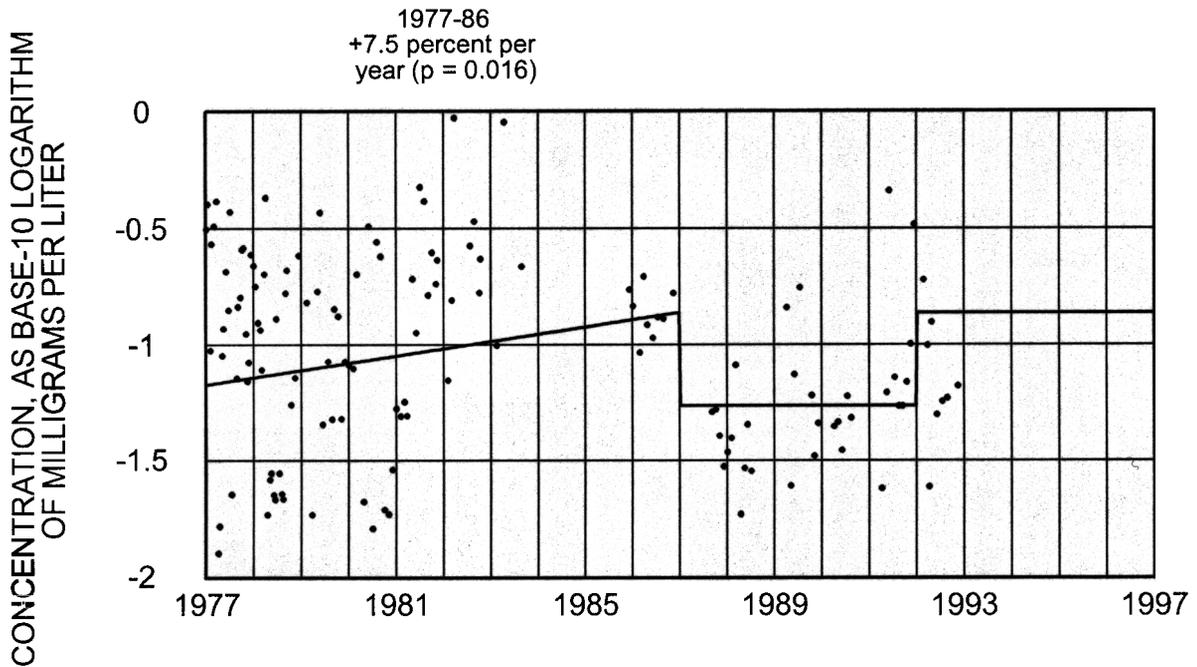
Figure 14. Trends (lines) in flow-adjusted total ammonia concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
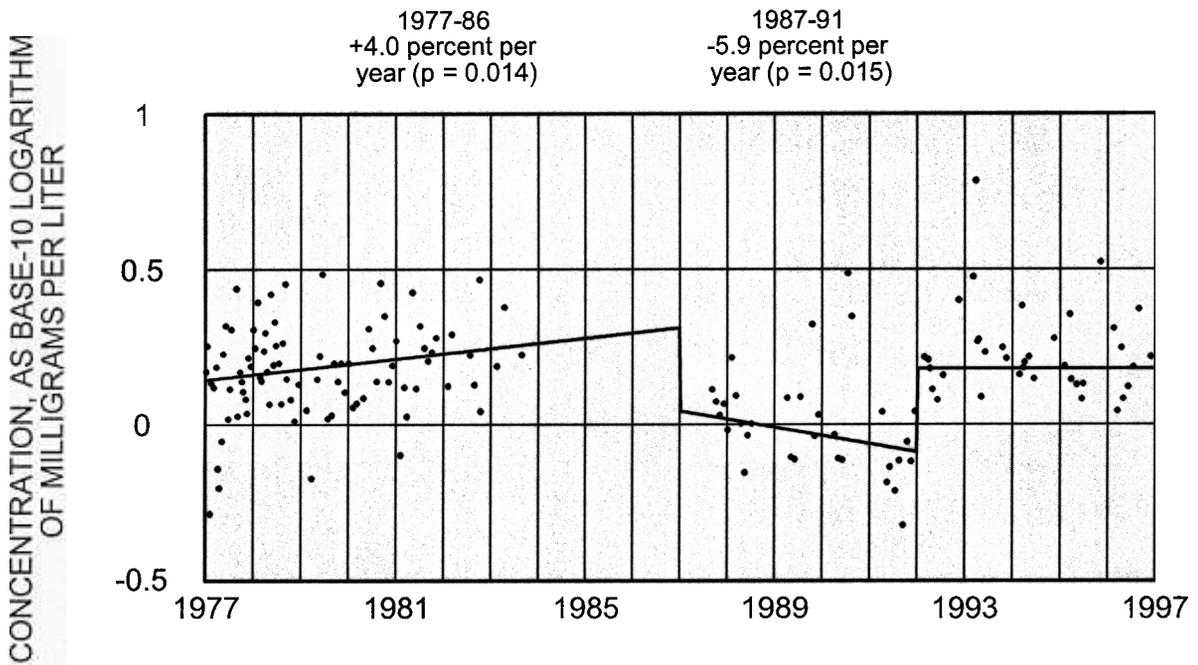
Figure 15. Trends (lines) in flow-adjusted total nitrogen concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
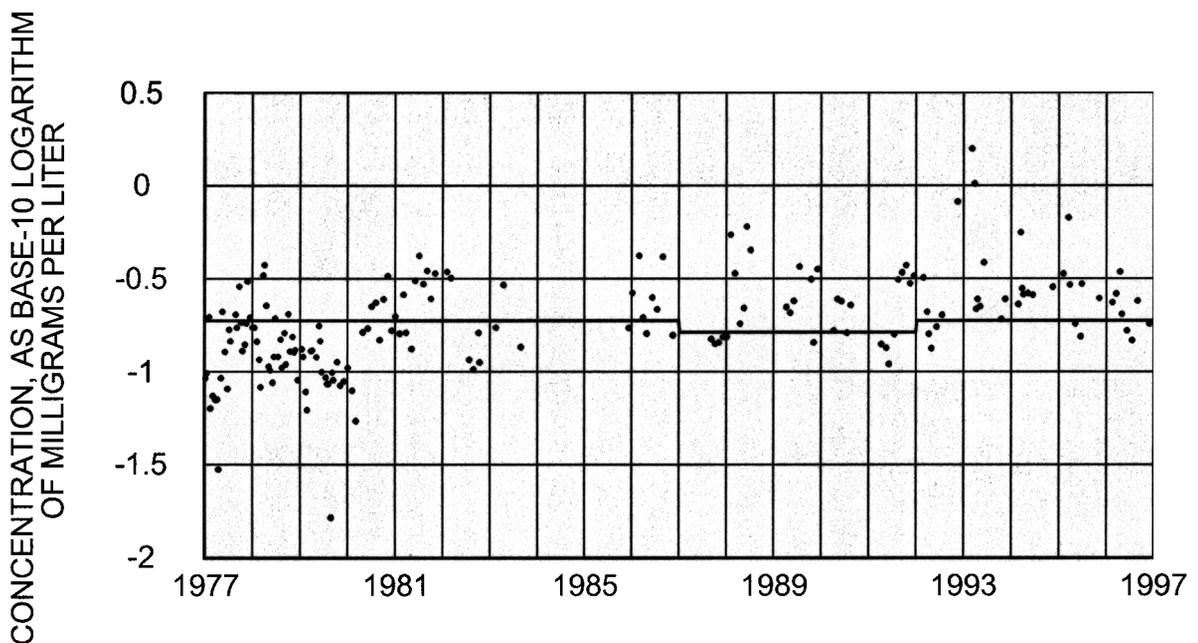

Figure 16. Trends (lines) in flow-adjusted total phosphorus concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
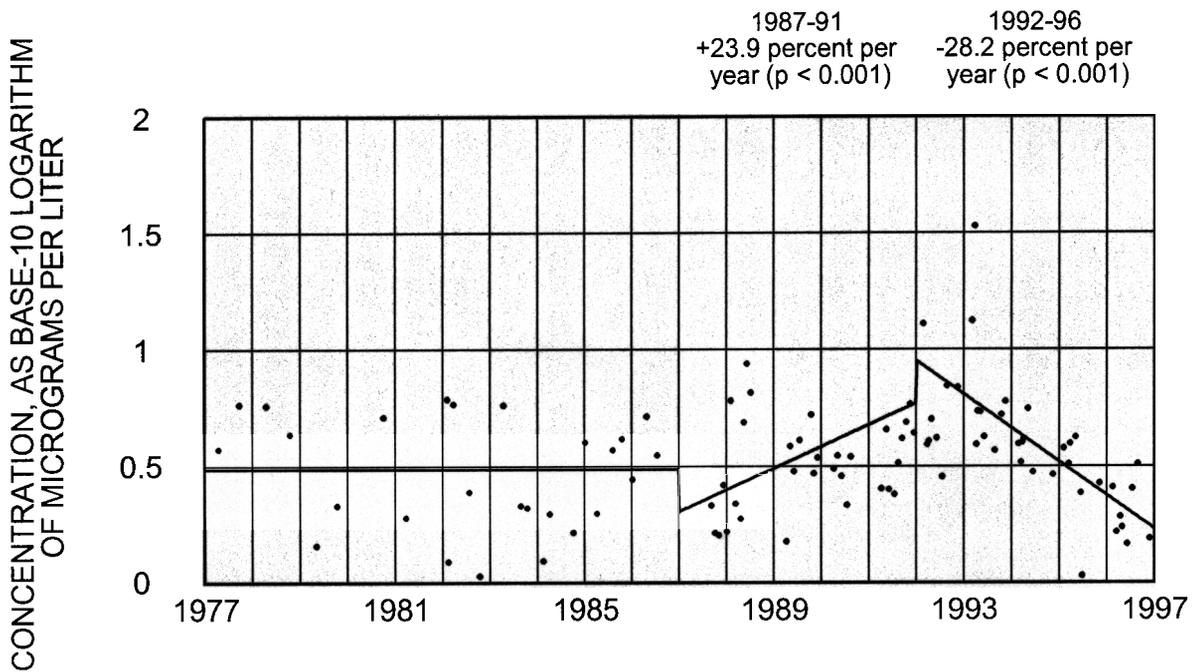
Figure 17. Trends (lines) in flow-adjusted dissolved and total arsenic concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
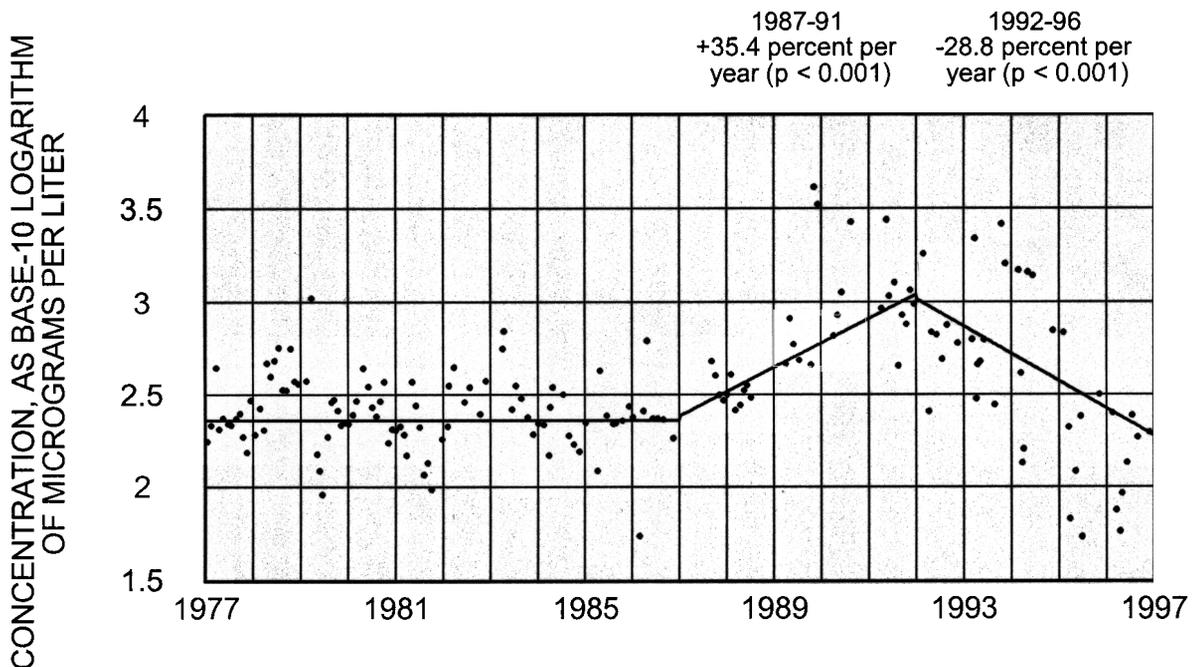

Figure 18. Trends (lines) in flow-adjusted dissolved and total boron concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
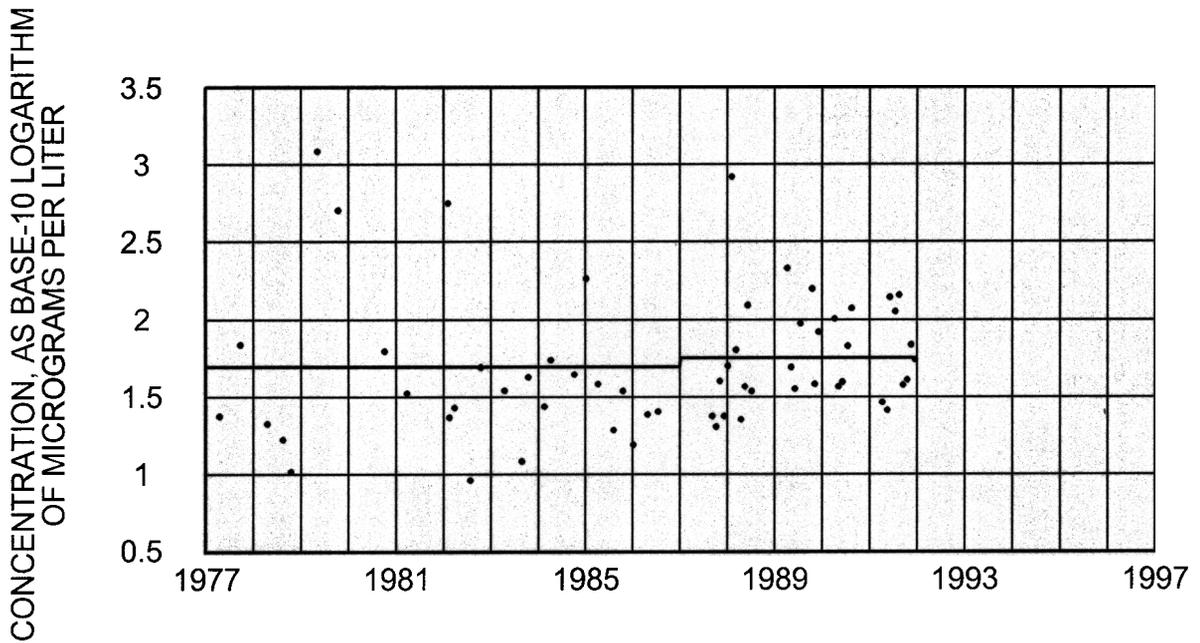
Figure 19. Trends (lines) in flow-adjusted dissolved iron concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]
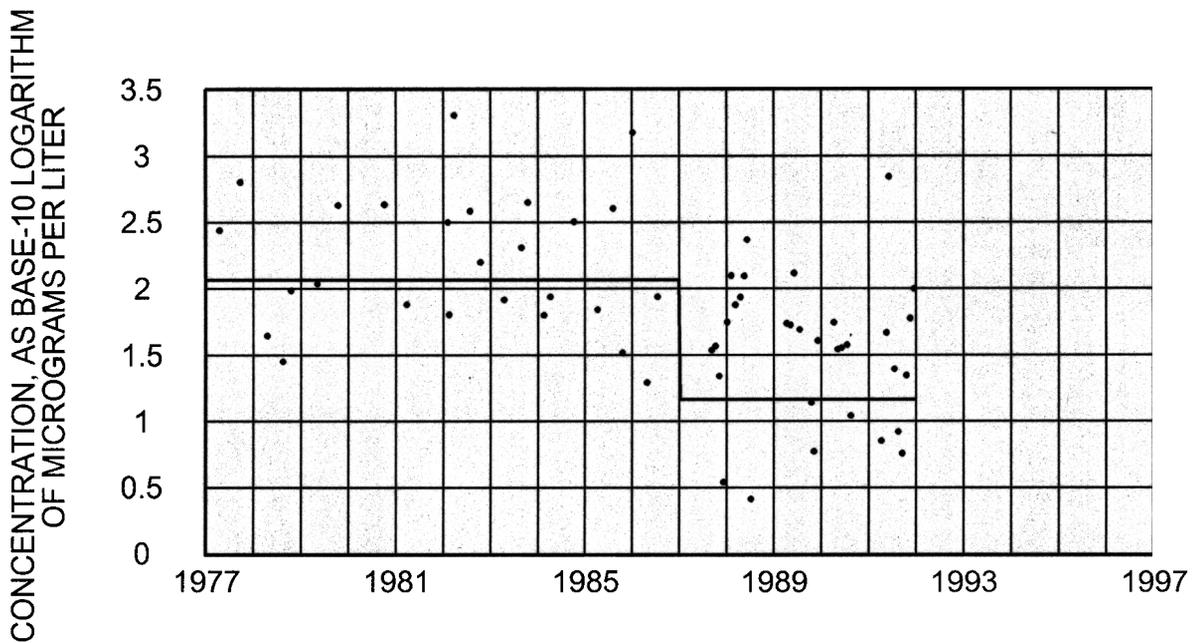
Figure 20. Trends (lines) in flow-adjusted dissolved manganese concentrations (points), Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96. [Significant trends, if any, are indicated at top of graph.]

Trend analysis results also indicate a strong and highly significant increase in flow-adjusted concentrations for most major ions, arsenic, and boron at the Sherwood sites beginning in about 1987. This increase was followed by an equally strong and highly significant decrease in concentrations for the same constituents beginning in about 1992. Determining the reasons for these changes is beyond the scope of this study.

The differences in concentrations between site 1 and site 2 are indicated by the step trends for 1987-91. Although many of these step trends are not statistically significant, flow-adjusted concentrations tend to be less at site 1 than at site 2. This could be a result of true differences in concentration, small differences in discharge between the two sites, or a combination of the two.

Significant Trends for Westhope Sites

Trend analysis results for the Westhope sites indicate many significant trends during 1977-86 (figs. 21 through 33). Significant increases occurred in dissolved oxygen, dissolved sodium, dissolved chloride, total phosphorus, and dissolved
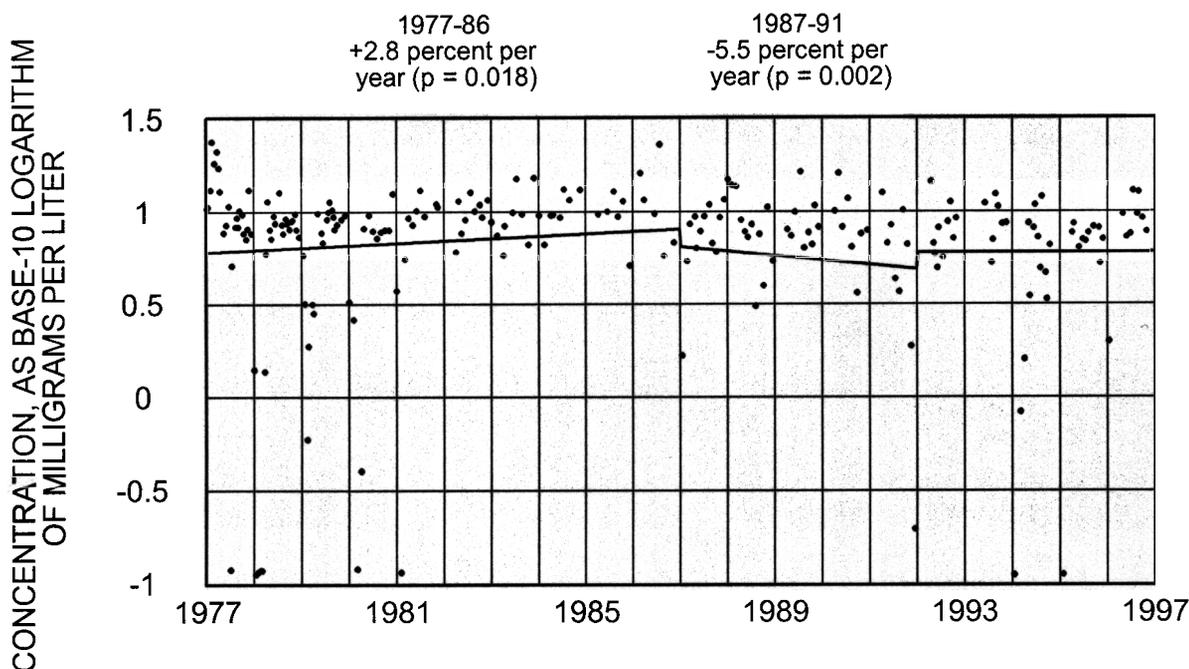
Figure 21. Trends (lines) in flow-adjusted dissolved oxygen concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]

boron, and a significant decrease occurred in dissolved arsenic during that period. These trends are in contrast to the trends for the Sherwood sites during 1977-86.

Significant trends for the Westhope sites during 1987-91 and 1992-96 were not as numerous or as strong as trends for the Sherwood sites during those periods. During 1987-91, significant decreases occurred in dissolved oxygen, dissolved sulfate, and dissolved boron, and significant increases occurred in total phosphorus, dissolved iron, and dissolved manganese. During 1992-96, a significant decrease occurred in total phosphorus, and significant increases occurred in dissolved sulfate and dissolved solids.
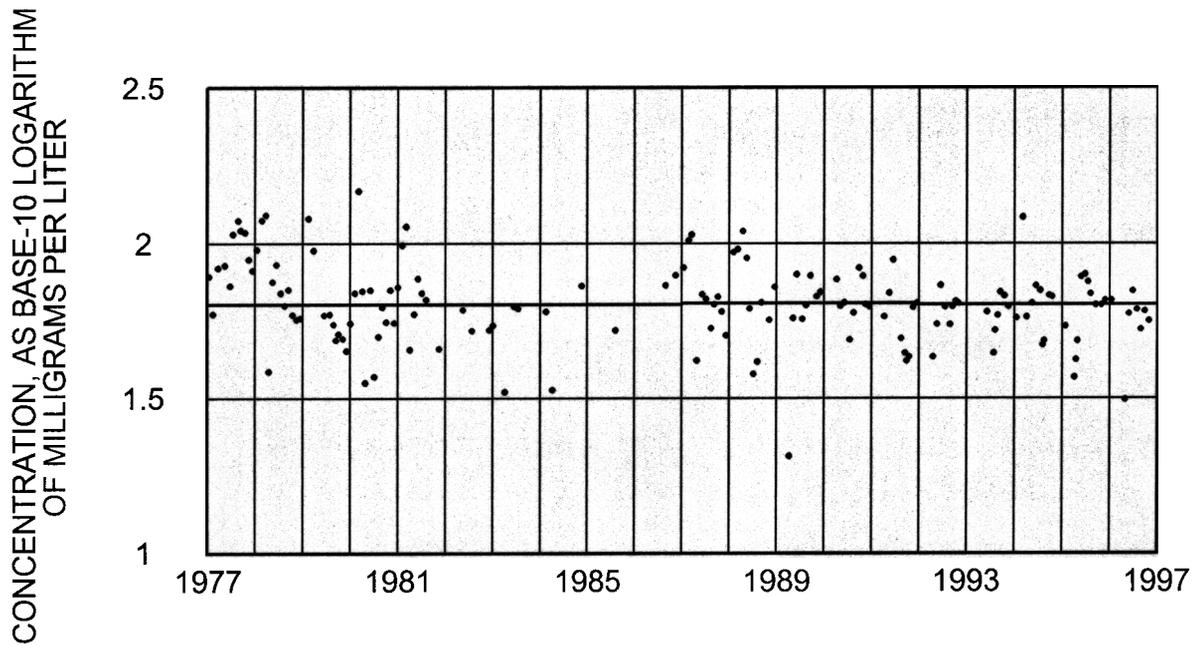
Figure 22. Trends (lines) in flow-adjusted dissolved calcium concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
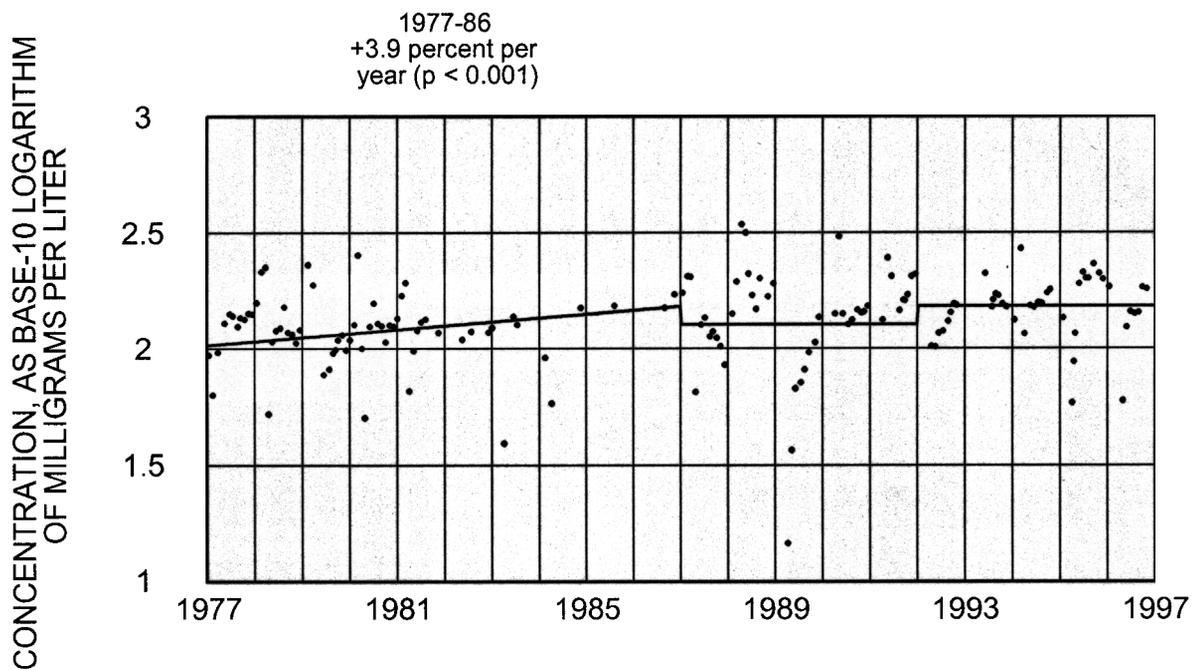

Figure 23. Trends (lines) in flow-adjusted dissolved sodium concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba 1977-96. [Significant trends, if any, are indicated at top of graph.]
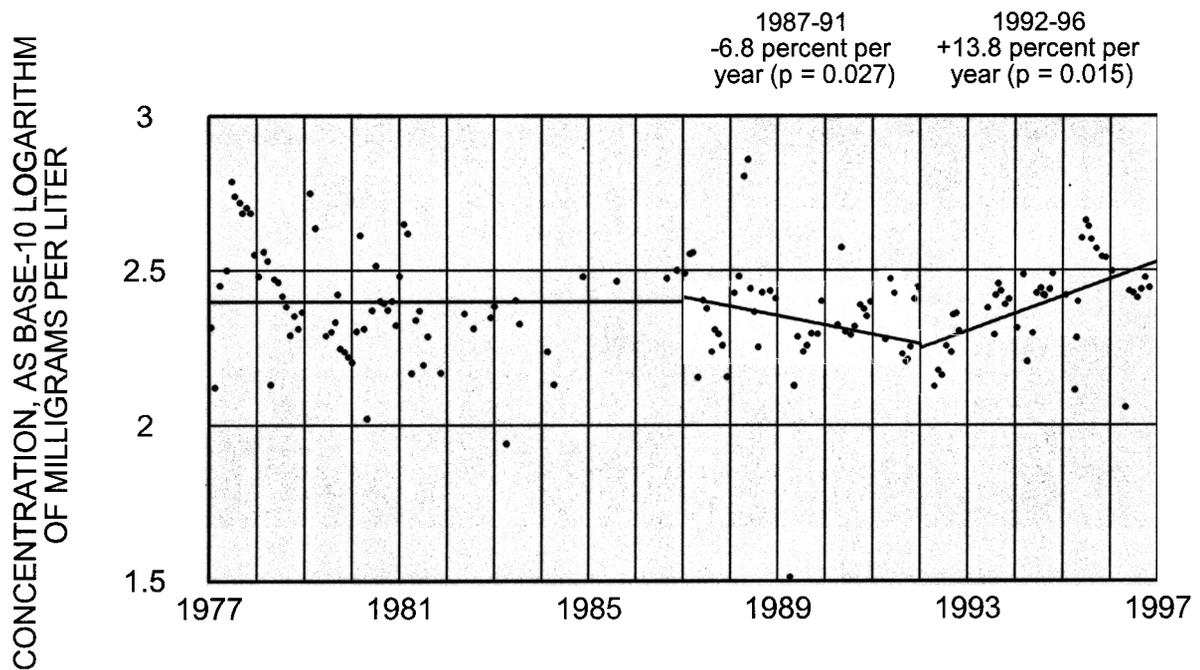
Figure 24. Trends (lines) in flow-adjusted dissolved sulfate concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
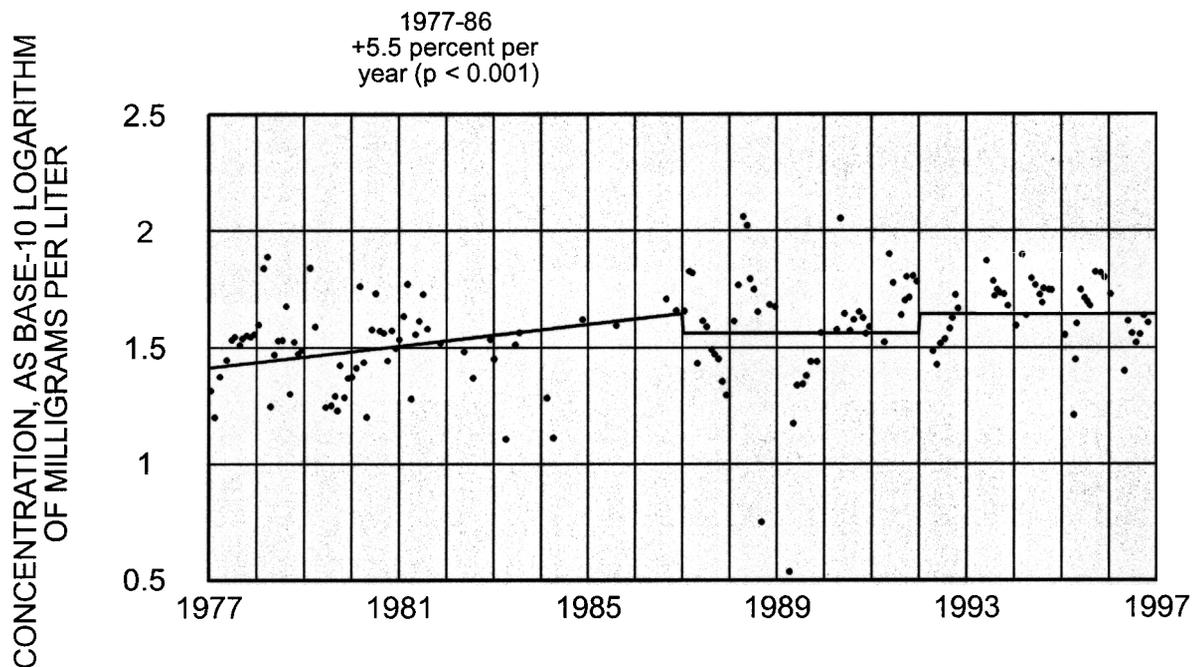
Figure 25. Trends (lines) in flow-adjusted dissolved chloride concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
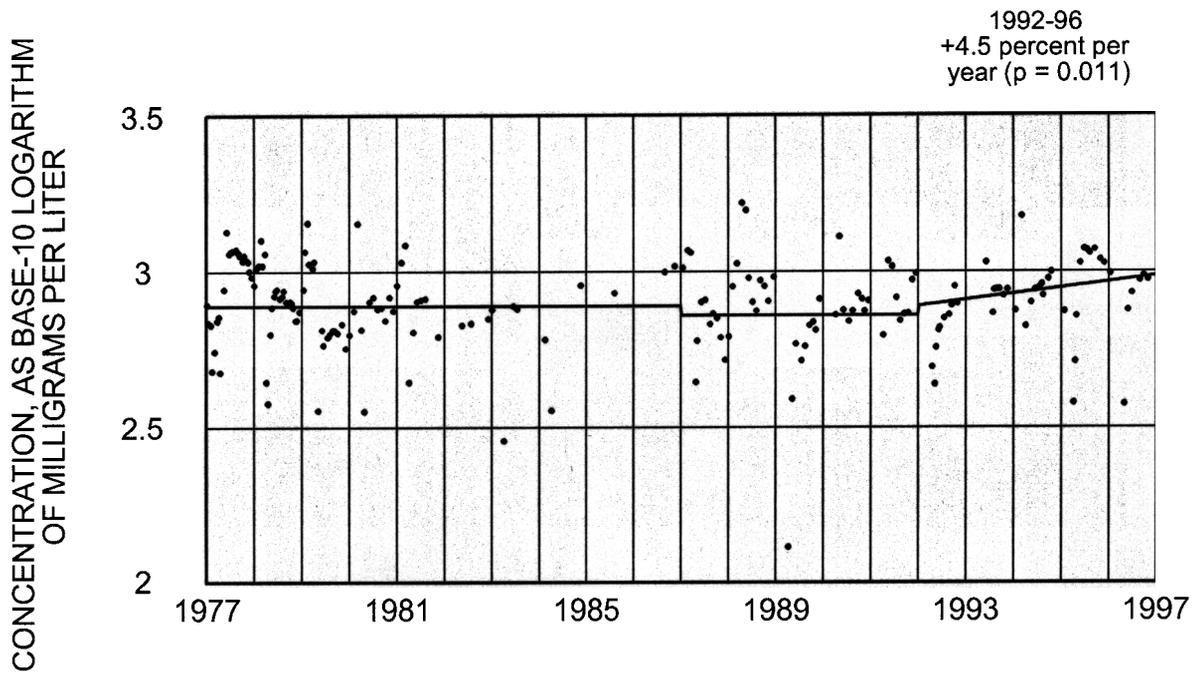
Figure 26. Trends (lines) in flow-adjusted dissolved solids concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
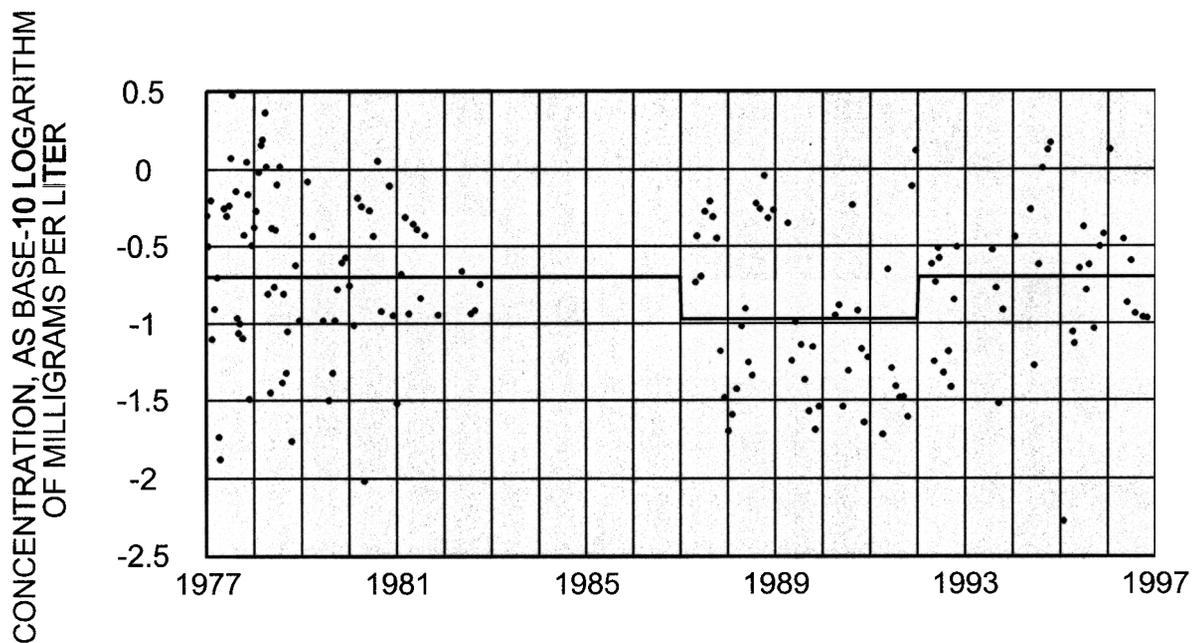
Figure 27. Trends (lines) in flow-adjusted total ammonia concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
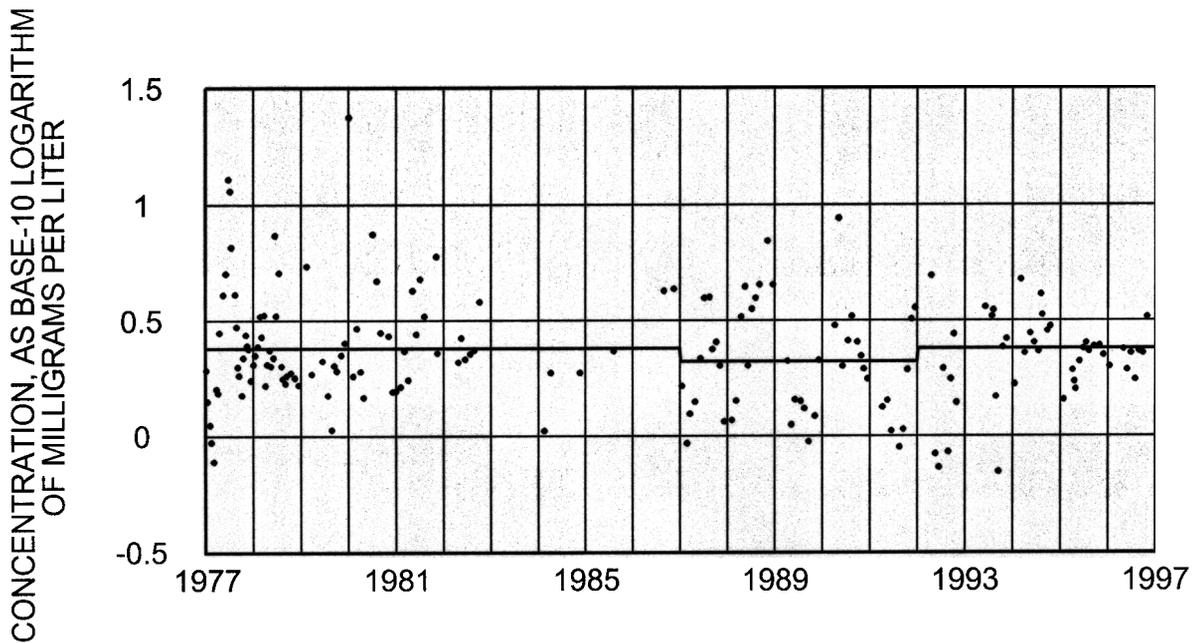
Figure 28. Trends (lines) in flow-adjusted total nitrogen concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
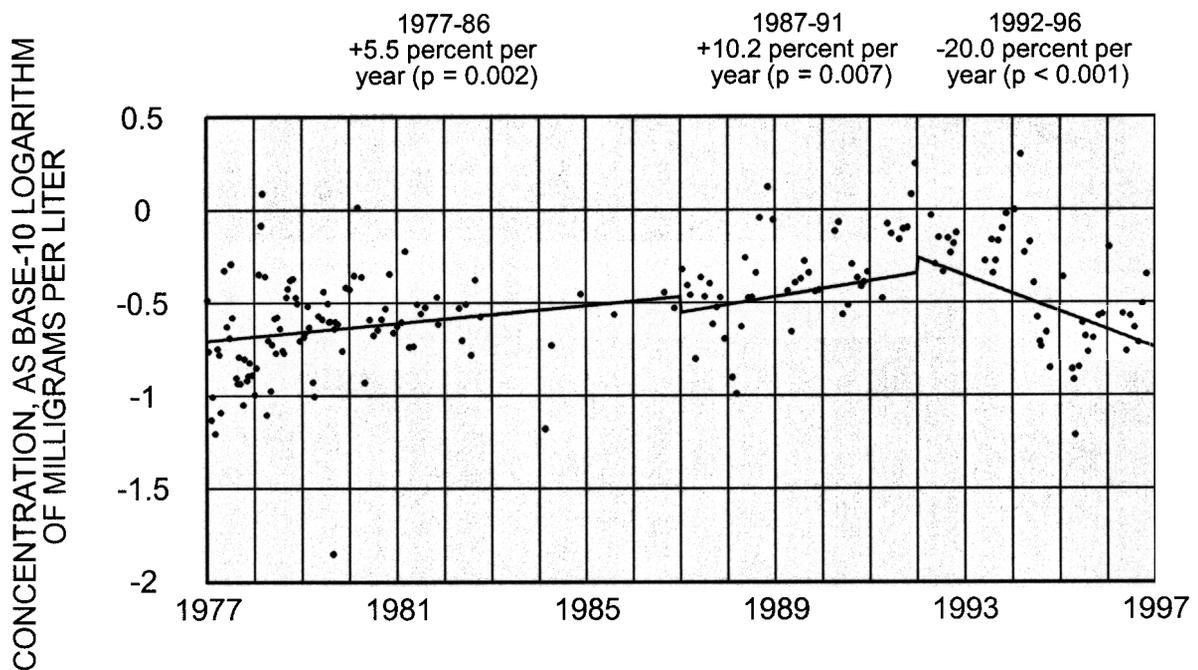

Figure 29. Trends (lines) in flow-adjusted total phosphorus concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
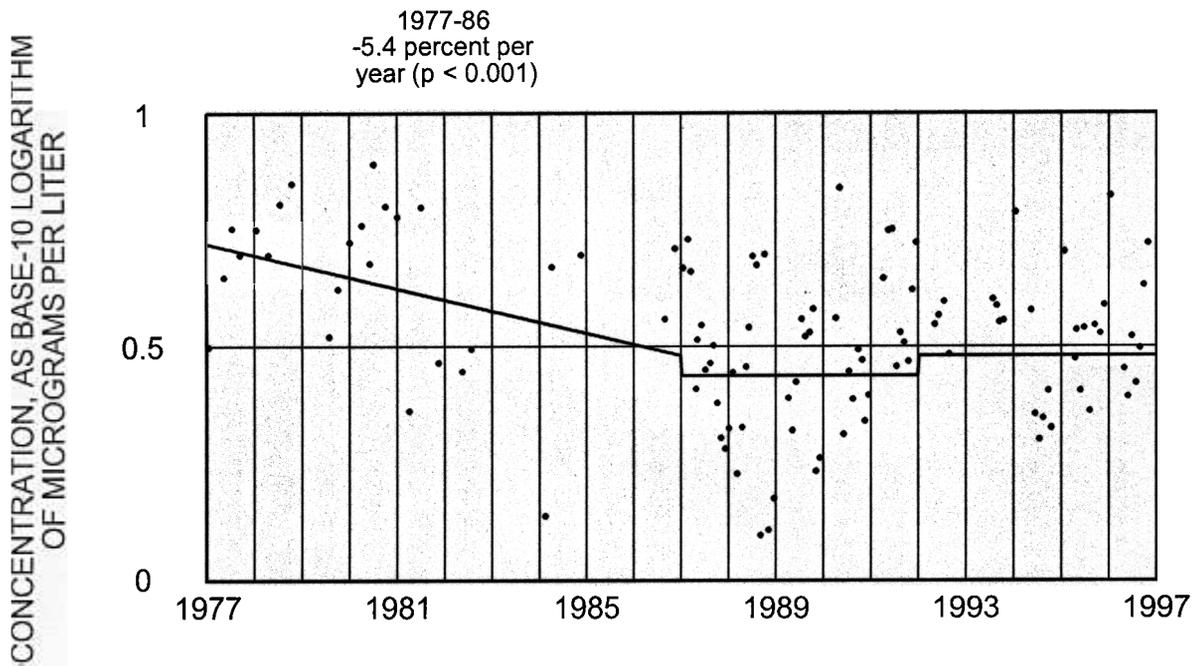
Figure 30. Trends (lines) in flow-adjusted dissolved and total arsenic concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
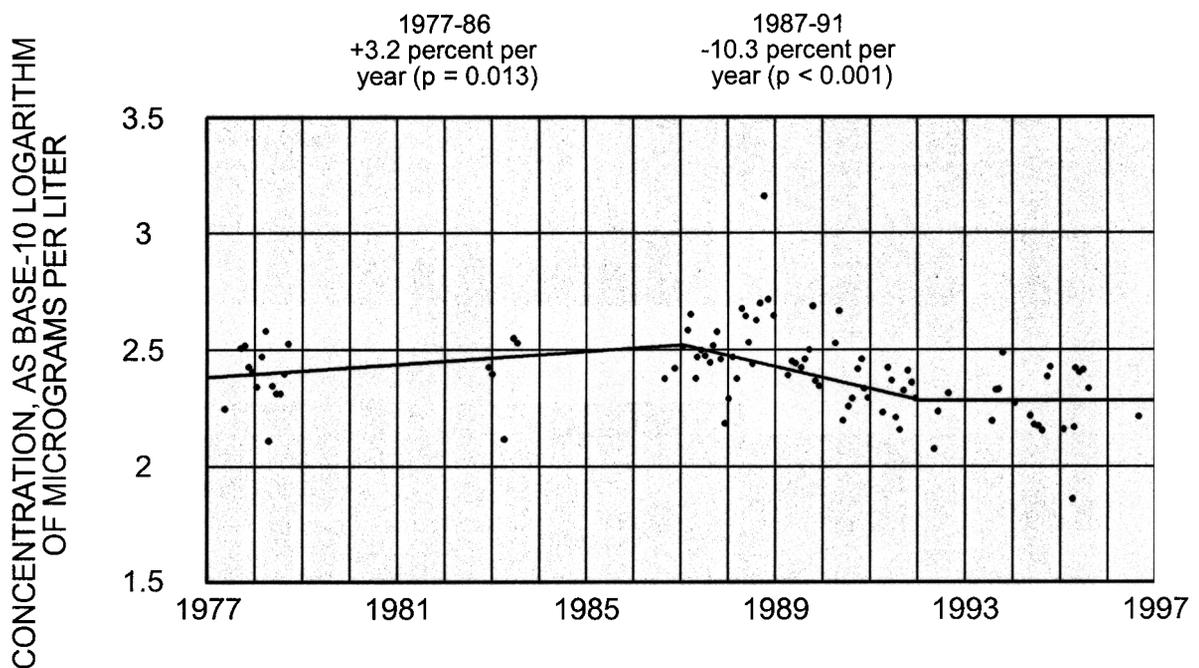

Figure 31. Trends (lines) in flow-adjusted dissolved and total boron concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
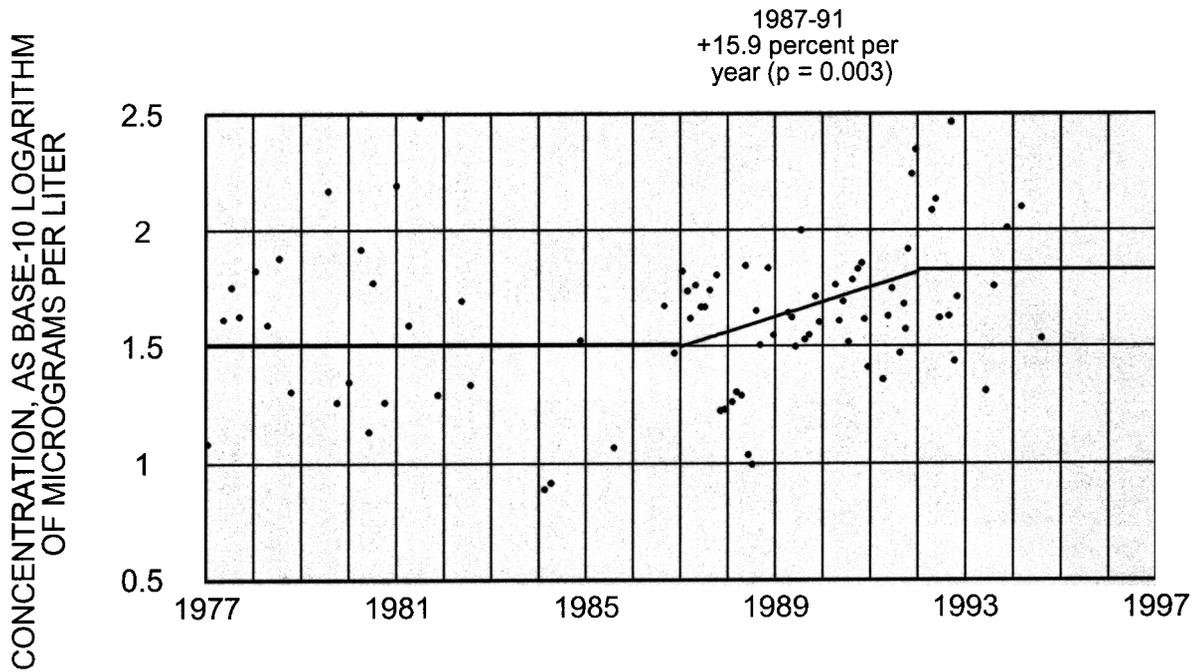
Figure 32. Trends (lines) in flow-adjusted dissolved iron concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]
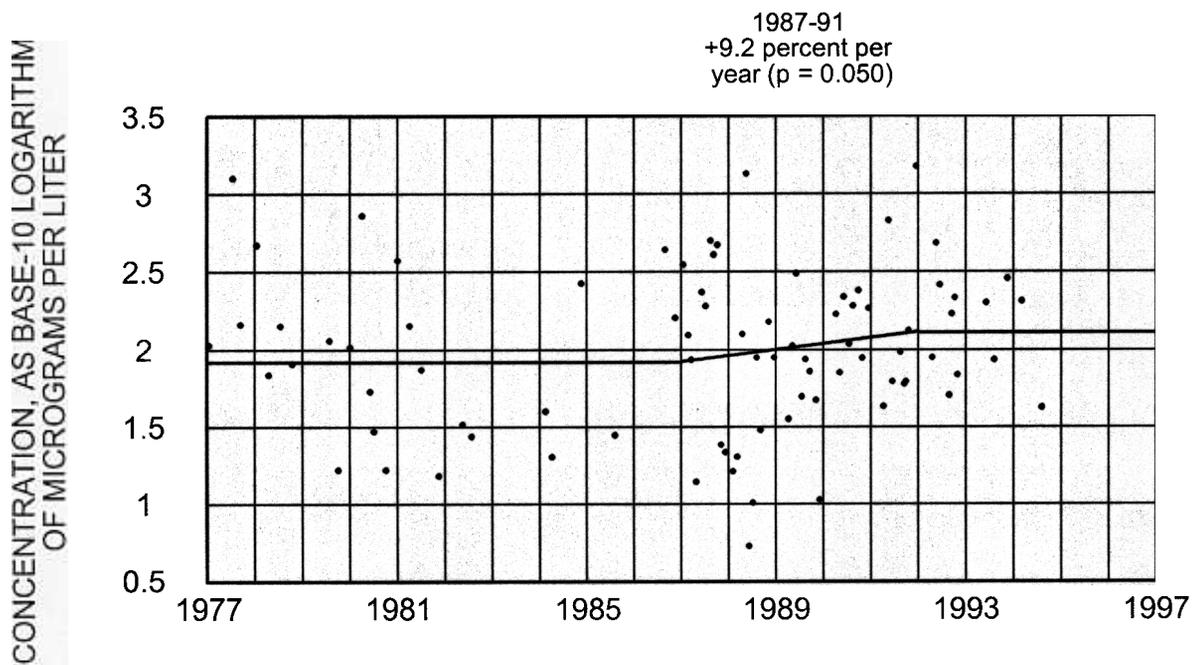

Figure 33. Trends (lines) in flow-adjusted dissolved manganese concentrations (points), Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba, 1977-96. [Significant trends, if any, are indicated at top of graph.]

The direction (increasing or decreasing) of the significant trends that occurred during 1987-91 and 1992-96 differed between the Sherwood sites and the Westhope sites. Significant trends often occurred at one site but not at the other. Also, conflicting trends occurred several times. For example, during 1987-91, decreases occurred in dissolved sulfate and dissolved boron at the Westhope sites, but increases occurred at the Sherwood sites. During 1992-96, significant increases occurred in dissolved sulfate and dissolved solids at the Westhope sites, and significant decreases occurred at the Sherwood sites. This indicates that factors that affect water quality at the Sherwood sites may be different from factors that affect water quality at the Westhope sites.

The step trends shown in figures 21 through 33 indicate flow-adjusted concentrations at site 4 tend to be slightly lower than flow-adjusted concentrations at site 3. As with the Sherwood sites, the step trends could be, in part, a result of small differences in discharge between the two sites.

SAMPLING DESIGN CONSIDERATIONS

The PARMA time-series model can be used not only to analyze historic data for water-quality trends but also to determine efficient sampling designs for monitoring future trends to maintain or reduce sampling costs incurred during past monitoring efforts. Implementation of the optimal sampling design is not logistically feasible because every constituent may have a different optimal sampling schedule. Therefore, alternate sampling designs that work well for most constituents, are logistically feasible to implement, and maintain an ability or power to detect trends that is similar to that which currently exists are given in this section, and the information gained or lost by increasing or decreasing the number of samples from current levels (about eight samples per year) is quantified.

The procedure for using the calibrated model to evaluate the efficiency of various sampling designs for detecting trends is given in appendix A. Information such as that shown in figure 5 and similar information on autocorrelations and cross-correlations at other lags have an effect on the efficiency of a given design. The results given in this section are based on the following assumptions:

1. The statistical properties of future flow-adjusted concentrations (after removing trends) are similar to the statistical properties of flow-adjusted concentrations for the model calibration period (1977-96).

2. The operation of streamflow gages near Sherwood and Westhope is continued so that future daily discharges are available for both sites.

3. The sampling design is efficient for detecting monotonic (log-linear) trends in flow-adjusted concentrations

4. The sampling design is reasonably efficient for detecting more complicated trends.

5. The sampling design is efficient for detecting seasonal trends (for example, trends that may occur only during certain parts of the year).

Assumption 1 was made to draw conclusions about the efficiency of a sampling design on the basis of the calibrated model, which was calibrated using data from 1977 to 1996. Assumption 2 was made because the efficiency of a design is affected by the availability of discharge data, and a full set of daily discharge data simplifies the analysis. Assumption 3, which also simplifies the analysis, was made because the minimum requirement for any design should be to detect monotonic trends. Assumption 4 was made to avoid weighting the design too much toward monotonic trends. If a design is set up to detect only monotonic trends, more samples are collected near the ends of the period than in the middle. With such a design, the true trend (an increase followed by a decrease) (fig. 11) in dissolved sulfate at the Sherwood sites during 1987-96 might not be recognized. A balanced design (one that maintains the same sampling frequency every year) gives reasonable power for detecting most types of trends. Assumption 5 was made to balance seasonal information. If a design is set up to detect annual trends with no regard to season, more samples are collected during months that have low variability in concentration than during months that have high variability. With such a design, trends in, for example, dissolved oxygen that occur during the winter might not be recognized because most samples were collected during the spring when concentrations are high and variability is low.

To determine a sampling design that is efficient for detecting seasonal trends, the year was divided into three 4-month seasons--March through June, July through October, and November through February. To have at least some sensitivity in each season, at least one sample per season per year was required. Then, to quantify the information gained by increasing the number of samples collected, the single-sample-per-season design was compared to designs that had two through five samples per season. The sampling schedules considered in the trend analysis are given in table 4. For example, if three samples are collected in season 2 (July through October), the assumption was made that one sample is collected during each of three periods--July 21-31, August 21-31, and September 21-30. The day on which the sample is collected is not critical, but, to maximize information, the sampling interval should be about 30 days. For example, sampling on July 25, August 25, and September 25 is preferable to sampling on July 31, August 25, and September 21. For four samples, the sampling interval also should be about 30 days, and, for five samples, the sampling interval should be about 20 days. For three, four, or five samples, the sampling interval within the season is more important than the starting and ending dates of sample collection. For example, sampling on July 15, August 15, and September 15 essentially is equivalent to sampling on July 25, August 25, and September 25. For one or two samples, the date of sample collection also becomes important. For example, if only one sample is collected in season 1, the degree of sensitivity will depend on whether the sample is collected at the beginning of spring runoff, near the peak discharge, or near the end of spring runoff.

The number of samples collected each season may differ. For example, three samples can be collected in season 1 (on March 25, April 25, and May 25); two samples can be collected in season 2 (on July 25 and September 25); and one sample can be collected in season 3 (on December 21).

The sampling designs given in this section were based on table 4 and should be used only as a general guide for evaluating the effect of increasing or decreasing the number of samples. The designs are not meant to be used as rigid schedules that must be followed each and every year. For example, for three samples in season 1, the first sample may be collected near the beginning of spring runoff, the second sample may be collected near the peak discharge, and the third sample may be collected about 30 days after the peak discharge. In years that have a rapidly rising hydrograph, the first two samples may be collected less than 30 days apart, but the data still can be used for trend analysis with little loss of sensitivity. Similarly, for two samples in season 3, both samples could be collected during winter ice-covered conditions although, in some years, the samples may need to be collected in January and February rather than in November and January.

A quantitative measure of the efficiency of a given sampling design is needed to distinguish efficient designs from inefficient designs. In this study, the efficiency of a given design was quantified using the following definition:

> The **characteristic trend** of a design is the size of monotonic trend, in percent per year, that has a 50-percent chance of being detected with 10 years of data, using a trend test with a 5-percent significance level.

Using this definition, the characteristic trend has a 50-percent chance of being detected (for example, of being flagged as significant at the 5-percent level). The smaller the characteristic trend, the more sensitive the design. The characteristic trend is inversely proportional to the length of the sampling period; for this study, a 10-year sampling period was chosen as a benchmark.

Efficient Sampling Designs for Sherwood Sites

To develop a sampling design that balances information between major constituent groups, the list of 13 constituents was broken into four groups--dissolved oxygen; major ions (dissolved calcium, sodium, sulfate, and chloride and dissolved solids); nutrients (total ammonia, nitrogen, and phosphorus); and trace elements (dissolved arsenic, boron, iron, and manganese). A composite trend was computed for each group by taking the geometric mean of the characteristic trends for individual constituents in the group (fig. 34).

The first logistical consideration for any sampling design is the number of site visits required per year. General observations regarding site visits can be made from figure 34:

Table 4. Sampling schedules considered for trend analysis

[n, number of samples per season; =, equals; O, no sample in given period; X, one sample in given period]

	Season 1						Season 2						Season 3				
Period	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	Period	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	Period	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
March 1-10	O	O	O	O	O	July 1-10	O	O	O	O	O	November 1-10	O	O	O	O	O
March 11-20	O	O	O	X	X	July 11-20	O	O	O	X	X	November 11-20	O	O	O	X	X
March 21-31	O	X	X	O	O	July 21-31	O	X	X	O	O	November 21-30	O	X	X	O	O
April 1-10	O	O	O	O	X	August 1-10	O	O	O	O	X	December 1-10	O	O	O	O	X
April 11-20	O	O	O	X	O	August 11-20	O	O	O	X	O	December 11-20	O	O	O	X	O
April 21-30	X	O	X	O	X	August 21-31	X	O	X	O	X	December 21-31	X	O	X	O	X
May 1-10	O	O	O	O	O	September 1-10	O	O	O	O	O	January 1-10	O	O	O	O	O
May 11-20	O	O	O	X	X	September 11-20	O	O	O	X	X	January 11-20	O	O	O	X	X
May 21-31	O	X	X	O	O	September 21-30	O	X	X	O	O	January 21-31	O	X	X	O	O
June 1-10	O	O	O	O	O	October 1-10	O	O	O	O	O	February 1-10	O	O	O	O	O
June 11-20	O	O	O	X	X	October 11-20	O	O	O	X	X	February 11-20	O	O	O	X	X
June 21-30	O	O	O	O	O	October 21-31	O	O	O	O	O	February 21-28	O	O	O	O	O
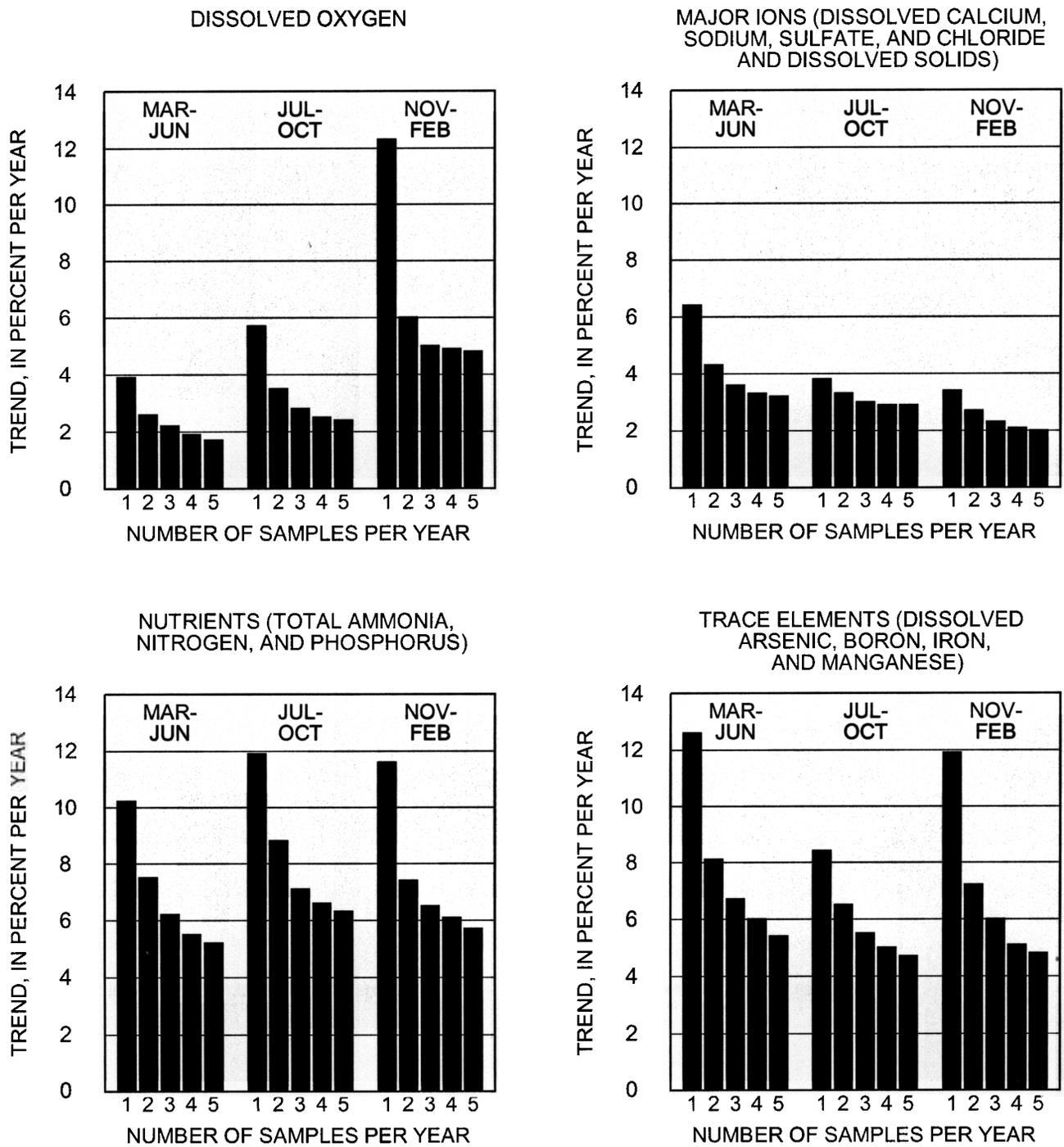
Figure 34. Characteristic trends in dissolved oxygen, major ions, nutrients, and trace elements for different sampling designs, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota.

. At least one sample is required each season to maintain the ability to distinguish trends separately for each season. The incremental information gained from the first sample essentially is infinite because with no samples a trend cannot be identified, no matter how large.

2. The incremental information gained from the second sample each season for two or more of the constituent groups is high, as indicated by a large decrease in the characteristic trend. To take advantage of the high incremental information, at least two site visits should be made each season.

3. The incremental information gained from the third sample each season is considerably less than the incremental information gained from the second sample. However, some incremental information is gained from the third sample each season. Although the benefits of the third sample are not as dramatic as the benefits of the second sample, three site visits in any of the seasons can be justified, depending on prioritization of trend issues.

4. The benefits of a fourth or fifth sample in any of the seasons are small from a trend-monitoring perspective However, monitoring goals not included in this study may require more samples.

5. For monitoring trends, the most efficient sampling designs are either two or three site visits per season per year. Less than two samples per season results in a substantial loss of information although a single sample per season still is better than no samples. Similarly, the information gained from a fourth or fifth sample indicates little reason for more than three samples per season.

Assuming two or three site visits per year, the information shown in figure 34 was used to determine several sampling designs that are efficient for monitoring trends and in which sampling efforts are maintained near current levels. The designs are given in table 5. Generally, without logistical considerations (for example, difficulty in winter sampling) or prioritization of trend issues (for example, trace elements are a higher priority than major ions), the most efficient designs balance information between seasons and between constituent groups. Therefore, the designs given in table 5 were selected such that the characteristic trends were as uniform as possible given a fixed number of samples. The assumption was made that dissolved oxygen is measured each time a site is visited.

Sampling design 1 in table 5 is the minimal design needed to maintain a reasonable power to detect trends in all three seasons and all four constituent groups. However, sampling design 1 is not very sensitive for detecting trends in either nutrients or trace elements. For example, the characteristic trends for nutrients in season 2 (8.8) and trace elements in season 1 (8.1) are high. In sampling design 2, a nutrient sample added in season 2 and a trace-element sample added in season 1 resulted in substantial gains in efficiency (from 8.8 to 7.1 for nutrients and from 8.1 to 6.7 for trace elements), but a major-ion sample added in season 1 resulted in only a slight gain in efficiency (from 4.3 to 3.6). The additional major-ion and trace-element samples in season 1 should be collected near the peak discharge, and the additional nutrient sample in season 2 should be collected in midsummer.

In sampling design 3 (table 5), a third nutrient sample added in season 1 reduced the characteristic trend from 7.5 (sampling design 2) to 6.2; and a third trace-element sample added in season 3 reduced the characteristic trend from 7.2 (sampling design 2) to 6.0. The dissolved oxygen sample added in season 3 also reduced the characteristic trend from 6.0 (sampling design 2) to 5.0. In sampling design 4 (table 5), the sensitivity of sampling design 3 was increased by adding a midwinter nutrient sample and a midsummer trace-element sample. As discussed earlier, increasing the number of samples beyond that in sampling design 4 probably is not warranted for trend analysis.

Efficient Sampling Designs for Westhope Sites

The characteristic trends for the Westhope sites were composited over the same constituent groups as for the Sherwood sites and are shown in figure 35. The general observations regarding site visits for the Sherwood sites also apply to site visits for the Westhope sites--the most efficient sampling designs from a trend-monitoring perspective are either two or three site visits per season per year. The characteristic trends for the Sherwood and Westhope sites are similar (figs. 34 and 35), indicating the four sites have similar optimal sampling designs that have similar sensitivity for detecting trends.

Table 5. Efficient sampling designs, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

Sampling design	Season	Number of site visits	Number of samples per year[1]			
			Dissolved oxygen	Major ions	Nutrients	Trace elements
1	1 (March through June)	2	2 (2.6)	2 (4.3)	2 (7.5)	2 (8.1)
	2 (July through October)	2	2 (3.5)	2 (3.3)	2 (8.8)	2 (6.5)
	3 (November through February)	2	2 (6.0)	2 (2.7)	2 (7.4)	2 (7.2)
		6	6	6	6	6
2	1 (March through June)	3	3 (2.2)	3 (3.6)	2 (7.5)	3 (6.7)
	2 (July through October)	3	3 (2.8)	2 (3.3)	3 (7.1)	2 (6.5)
	3 (November through February)	2	2 (6.0)	2 (2.7)	2 (7.4)	2 (7.2)
	Total	8	8	7	7	7
3	1 (March through June)	3	3 (2.2)	3 (3.6)	3 (6.2)	3 (6.7)
	2 (July through October)	3	3 (2.8)	2 (3.3)	3 (7.1)	2 (6.5)
	3 (November through February)	3	3 (5.0)	2 (2.7)	2 (7.4)	3 (6.0)
	Total	9	9	7	8	8
4	1 (March through June)	3	3 (2.2)	3 (3.6)	3 (6.2)	3 (6.7)
	2 (July through October)	3	3 (2.8)	2 (3.3)	3 (7.1)	3 (5.5)
	3 (November through February)	3	3 (5.0)	2 (2.7)	3 (6.5)	3 (6.0)
		9	9	7	9	9

[1] Values in parentheses indicate the size of monotonic trend, in percent per year, that has a 50-percent chance of being detected with 10 years of data and a trend test with a 5-percent significance level.

Several efficient sampling designs for the Westhope sites are given in table 6. From sampling design 1 to sampling design 2, a third sample is added in midspring (near the peak discharge) for dissolved oxygen, major ions, and trace elements and in midsummer for dissolved oxygen and nutrients. The additional samples resulted in substantial gains in efficiency for nutrients and trace elements and slight gains in efficiency for dissolved oxygen and major ions.

From sampling design 3 to sampling design 4 (table 6), a third sample is added in midwinter for dissolved oxygen, nutrients, and trace elements. The additional samples resulted in substantial gains in efficiency for all three constituent groups. From sampling design 3 to sampling design 4, a third sample is added in midspring for nutrients and midsummer for trace elements. These additional samples also resulted in gains in efficiency for both constituent groups.

A comparison of the sampling designs for the Sherwood and Westhope sites (tables 5 and 6) indicates the designs and sensitivities are similar. However, as indicated by the characteristic trends, detecting trends in major ions, nutrients, and trace elements for the Sherwood sites is more difficult than for the Westhope sites.

DISSOLVED OXYGEN

MAJOR IONS (DISSOLVED CALCIUM,
SODIUM, SULFATE, AND CHLORIDE
AND DISSOLVED SOLIDS)

NUTRIENTS (TOTAL AMMONIA,
NITROGEN, AND PHOSPHORUS)

TRACE ELEMENTS (DISSOLVED
ARSENIC, BORON, IRON,
AND MANGANESE)
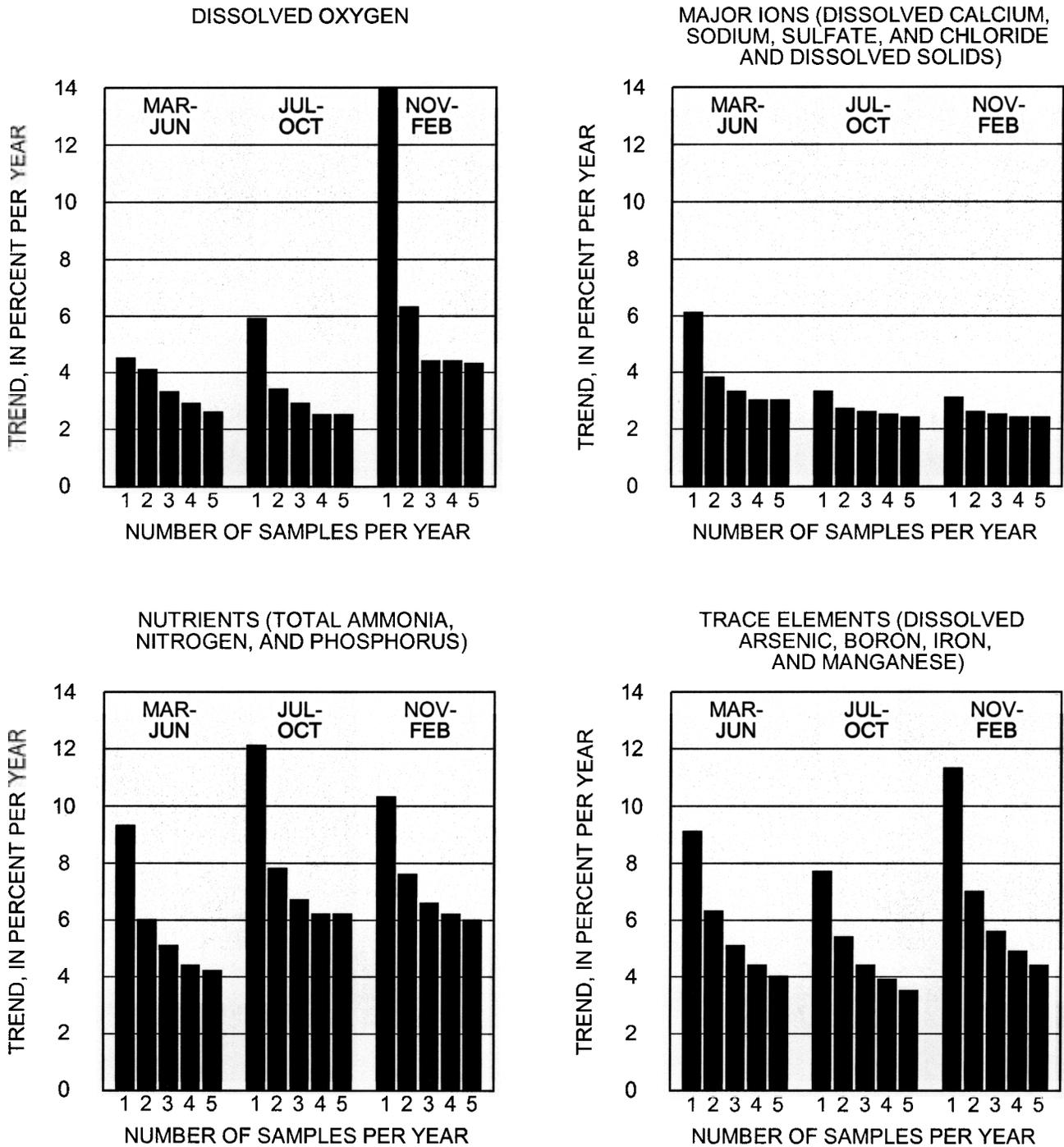
Figure 35. Characteristic trends in dissolved oxygen, major ions, nutrients, and trace elements for different sampling designs, Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba.

Table 6. Efficient sampling designs, Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

Sampling design	Season	Number of site visits	Number of samples per year[1] Dissolved oxygen	Major ions	Nutrients	Trace elements
	1 (March through June)	2	2 (4.1)	2 (3.8)	2 (6.0)	2 (6.3)
	2 (July through October)	2	2 (3.4)	2 (2.7)	2 (7.8)	2 (5.4)
	3 (November through February)	2	2 (6.3)	2 (2.6)	2 (7.6)	2 (7.0)
	Total	6	6	6	6	6
2	1 (March through June)	3	3 (3.3)	3 (3.3)	2 (6.0)	3 (5.1)
	2 (July through October)	3	3 (2.9)	2 (2.7)	3 (6.7)	2 (5.4)
	3 (November through February)	2	2 (6.3)	2 (2.6)	2 (7.6)	2 (7.0)
	Total	8	8	7	7	7
	1 (March through June)	3	3 (3.3)	3 (3.3)	2 (6.0)	3 (5.1)
	2 (July through October)	3	3 (2.9)	2 (2.7)	3 (6.7)	2 (5.4)
	3 (November through February)	3	3 (4.4)	2 (2.6)	3 (6.6)	3 (5.6)
	Total	9	9	7	8	8
4	1 (March through June)	3	3 (3.3)	3 (3.3)	3 (5.1)	3 (5.1)
	2 (July through October)	3	3 (2.9)	2 (2.7)	3 (6.7)	3 (4.4)
	3 (November through February)	3	3 (4.4)	2 (2.6)	3 (6.6)	3 (5.6)
	Total	9	9	7	9	9

[1]Values in parentheses indicate the size of monotonic trend, in percent per year, that has a 50-percent chance of being detected with 10 years of data and a trend test with a 5-percent significance level.

SUMMARY

The Souris River Basin is a 24,600-square-mile basin located in southeast Saskatchewan, north-central North Dakota, and southwest Manitoba. The Souris River Bilateral Water Quality Monitoring Group, formed in 1989 by the governments of Canada and the United States, is responsible for documenting trends in water quality in the Souris River and making recommendations for monitoring future water-quality conditions. This report presents results of a study conducted for the Bilateral Water Quality Monitoring Group by the U.S. Geological Survey, in cooperation with the North Dakota Department of Health, to analyze historic trends in water quality in the Souris River and to determine efficient sampling designs for monitoring future trends.

U.S. Geological Survey and Environment Canada water-quality data collected during 1977-96 from four sites on the Souris River were analyzed for trends. The two Sherwood sites (the Souris River south of Glen Ewen, Saskatchewan, and the Souris River near Sherwood, N. Dak.) are near the upstream boundary crossing between Saskatchewan and North Dakota, and the two Westhope sites (the Souris River near Westhope, N. Dak., and the Souris River near Coulter, Manitoba) are near the downstream boundary crossing between North Dakota and Manitoba. Constituents analyzed for

trends included dissolved oxygen; dissolved calcium, sodium, sulfate and chloride; dissolved solids; total ammonia, nitrogen, and phosphorus; and dissolved arsenic, boron, iron, and manganese. Daily discharges for the Souris River near Sherwood, N. Dak., and the Souris River near Westhope, N. Dak., also were used in the trend analysis.

A parametric time-series model was developed for detecting trends in historic constituent concentration data. The model can be applied to constituents that have at least 90 percent of observations above detection limits of the analyses, which, for the Souris River, includes most major ions and nutrients and many trace elements. The model can detect complex nonmonotonic trends in concentration in the presence of complex interannual and seasonal variability in daily discharge. A key feature of the model is its ability to handle highly irregular sampling intervals. For example, the intervals between concentration measurements may be as short as 10 days to as long as several months, and the number of samples in any given year can range from zero to 36.

The parametric time-series model for constituent concentrations separates variability in log-transformed concentrations into three components--a low-frequency component, a trend component, and a high-frequency component. The low-frequency component consists of a nonlinear regression model that relates concentration to daily discharge and is used to isolate and remove as much of the natural seasonal and discharge-related variability in concentration as possible before analyzing for a trend. Following subtraction of the low-frequency component, the resulting concentration data are called flow-adjusted concentrations. The trend component consists of a combination of several simple trends--monotonic (log-linear) trends during 1977-86, 1987-91, and 1992-96; a step trend during 1987; and an equal but opposite step trend in 1992. The time intervals for the monotonic trends were selected on the basis of empirical analysis of the data, and the time interval of the step trend was selected on the basis of changes in sampling locations between U.S. Geological Survey and Environment Canada sites. The high-frequency component consists of noisy, short-term variability in concentration and has a complex seasonal autocorrelation structure. A periodic autoregressive moving-average (PARMA) time-series model was used to model the seasonal autocorrelation structure in the high-frequency component.

Trend analysis results indicated significant trends in flow-adjusted concentrations for many of the constituents. For example, at the two Sherwood sites, highly significant (p-value less than 0.01) increases occurred in dissolved sodium, sulfate, and chloride; dissolved solids; and dissolved arsenic and boron during 1987-91, and highly significant decreases occurred in the same constituents during 1992-96 (except for dissolved chloride, for which the p-value was slightly higher than 0.01 for 1992-96). Significant (p-value less than 0.05) increases also occurred in total ammonia and nitrogen during 1977-86, and a significant decrease occurred in total nitrogen during 1987-91.

Trend analysis results for the two Westhope sites contrasted with results for the Sherwood sites. For example, significant increases occurred in dissolved sodium, dissolved chloride, and total phosphorus during 1977-86. During 1987-91, highly significant decreases occurred in dissolved oxygen and dissolved boron, and highly significant increases occurred in total phosphorus and dissolved iron. During 1992-96, a significant decrease occurred in total phosphorus, and significant increases occurred in dissolved sulfate and dissolved solids.

The time-series model also was used to determine efficient sampling designs for monitoring future water-quality trends. To determine a design that was efficient for detecting seasonal trends, the year was divided into three seasons--March through June, July through October, and November through February. Sampling designs and sensitivities were similar for the Sherwood and Westhope sites. It was determined that at least two samples per year are required in each season to maintain reasonable sensitivity for detecting trends in most constituents. In addition, improvements occurred in sensitivity for detecting trends by adding a third sample for some seasons and some constituents, depending on prioritization of trend issues. For example, significant benefits occurred by adding a third sample for major ions and trace elements in March through June, adding a third sample for nutrients in July through October, and adding a third sample for nutrients, trace elements, and dissolved oxygen in November through February. However, increasing to more than three samples in any one season resulted in small gains in sensitivity and probably is not necessary for trend-monitoring purposes.

Details of the statistical methods used to conduct the water-quality trend analysis are given in an appendix. The computer program, users' manual, and discharge and water-quality data sets used can be accessed on the home page of the U.S. Geological Survey in North Dakota (http://nd.water.usgs.gov).

REFERENCES

Brockwell, P.J., and Davis, R.A., 1987, Time series--Theory and methods: New York, Springer-Verlag, 519 p.

Dennis, J.E., and Schnabel, R.B., 1983, Numerical methods for unconstrained optimization and nonlinear equations: Englewood Cliffs, New Jersey, Prentice-Hall, 319 p.

Hirsch, R.M., and Slack, J.R., 1984, A nonparametric trend test for seasonal data with serial dependence: Water Resources Research, v. 20, p. 727-732.

Jimenez, C., McLeod, A.C., and Hipel, K.W., 1989, Kalman filter estimation for periodic autoregressive moving average models: Stochastic Hydrology and Hydraulics, v. 3, p. 227-240.

Salas, J.D., Tabios, G.Q., and Bartolini, P., 1985, Approaches to multivariate modeling of water resources time series: Water Resources Bulletin, v. 21, no. 4, p. 683-708.

Tasker, G.D., 1997, Bootstrap position analysis for forecasting low-flow frequency: Journal of Water Resources Planning and Management, v. 123, no. 6, p. 359-367.

U.S. Department of Commerce, Bureau of the Census, 1981, Number of inhabitants, North Dakota: PC80-1-A36, 37 p.

Vecchia, A.V., 1985, Periodic autoregressive moving average (PARMA) modeling with applications to water resources: Water Resources Bulletin, v. 21, p. 721-730.

APPENDIX A. STATISTICAL METHODS USED TO CONDUCT WATER-QUALITY TREND ANALYSIS AND SAMPLING DESIGN

The purpose of this appendix is to describe in detail the statistical methods used to conduct the water-quality trend analysis for this study. The computer program for calibrating the time-series model, users' manual, and discharge and water-quality data sets used to conduct the analysis can be accessed on the home page of the U.S. Geological Survey in North Dakota (http://nd.water.usgs.gov). The computer program can be executed on a PC using Windows 95 or a newer operating system. Plotting programs written in S-PLUS (MathSoft) also are included so output from the computer program can be used to produce plots similar to those given in this report.

Time-Series Model Framework

The general form of the time-series model used to analyze trends and natural variability in discharge and concentration data can be expressed as

$$X(t) = C_X + A_X(t) + S_X(t) + U(t) \tag{A1}$$

X is the base-10 logarithm of discharge;

t is the time index, in years;

C_X is a constant;

$A_X(t)$ is a time series that represents annual variability in X;

$S_X(t)$ is a time series that represents seasonal variability in X; and

$U(t)$ is a time series of deviations of $X(t)$ from $C_X + A_X(t) + S_X(t)$

and

$$Y(t) = C_Y + A_Y(t) + S_Y(t) + Tr(t) + W(t) \tag{A2}$$

Y is the base-10 logarithm of concentration,

C_Y is a constant,

$A_Y(t)$ is a time series that represents annual variability in Y,

$S_Y(t)$ is a time series that represents seasonal variability in Y,

$Tr(t)$ is a trend in Y, and

$W(t)$ is a time series of deviations of $Y(t)$ from $C_Y + A_Y(t) + S_Y(t) + Tr(t)$.

The model was applied to log-transformed discharge and concentration data rather than to untransformed data because log-transformed data usually are more normally distributed than untransformed data. Because discharge and concentration data usually vary over several orders of magnitude, log-transformed data also are more numerically stable than untransformed data for estimating parameters of the time-series model. To apply the log-transform, zero values were replaced by the smallest nonzero value of the corresponding time series. The time index, t, was expressed in years, and decimal values corresponded to seasons within the year. For this study, the

model calibration period was 1977-96, and the seasons corresponded approximately to 10-day time intervals, or 36 time intervals per year. Therefore, for this study,

$$t = j + \frac{k}{36}$$

    j is the year and is equal to 1977, 1978, ..., 1996; and

    k is the season and is equal to 1, 2, ..., 36.

Separating the data into components of annual variability, seasonal variability, and noise often is referred to as classical decomposition (Brockwell and Davis, 1987). Because no unique method exists to separate the data, the best form of decomposition needed to be determined empirically. The following forms worked well for the discharge data analyzed in this study:

$$C_X = \text{mean}\{X(t), 1977 < t \leq 1997\}$$

$$A_X(t) = \text{mean}\{X(s) - C_X, (t-0.5) < s \leq (t+0.5)\},$$

$$S_X\left(j + \frac{k}{36}\right) = \text{mean}\left\{X(s) - C_X - A_X(s), s = 1977 + \frac{k}{36}, ..., 1996 + \frac{k}{36}\right\}$$

mean{.} is the arithmetic mean of the terms in the braces.

The starting and ending values for $A_X(t)$ can be obtained using discharge data from before 1977 and after 1996 or by setting

$$A_X(t) = A_X(1977.5) \text{ for } t < 1977.5$$

$$A_X(t) = A_X(1996.5) \text{ for } t > 1996.5$$

The constant, C_X, is the overall mean of the data. The annual time series, $A_X(t)$, was obtained by applying a 1-year moving average to the mean-centered time series, and the seasonal time series, $S_X(t)$, was obtained by subtracting the annual time series from the mean-centered time series and computing the means of the resulting values for each season. $S_X\left(j + \frac{k}{36}\right)$ takes on a different value for each season but does not depend on j.

The individual components of equation A1 are shown in figure A1 for the Souris River near Sherwood, N. Dak. (site 2). The recorded data and the constant, C_X, are shown in the upper-left graph; the mean-centered data, $X(t) - C_X$, and the annual component, $A_X(t)$, are shown in the upper-right graph; the deviations from the annual component, $X(t) - C_X - A_X(t)$, and the seasonal component, $S_X(t)$, are shown in the lower-left graph; and the deviations from the seasonal component, $U(t) = X(t) - C_X - A_X(t) - S_X(t)$, are shown in the lower-right graph. The first three terms on the right-hand side of the equation are called the low-frequency component
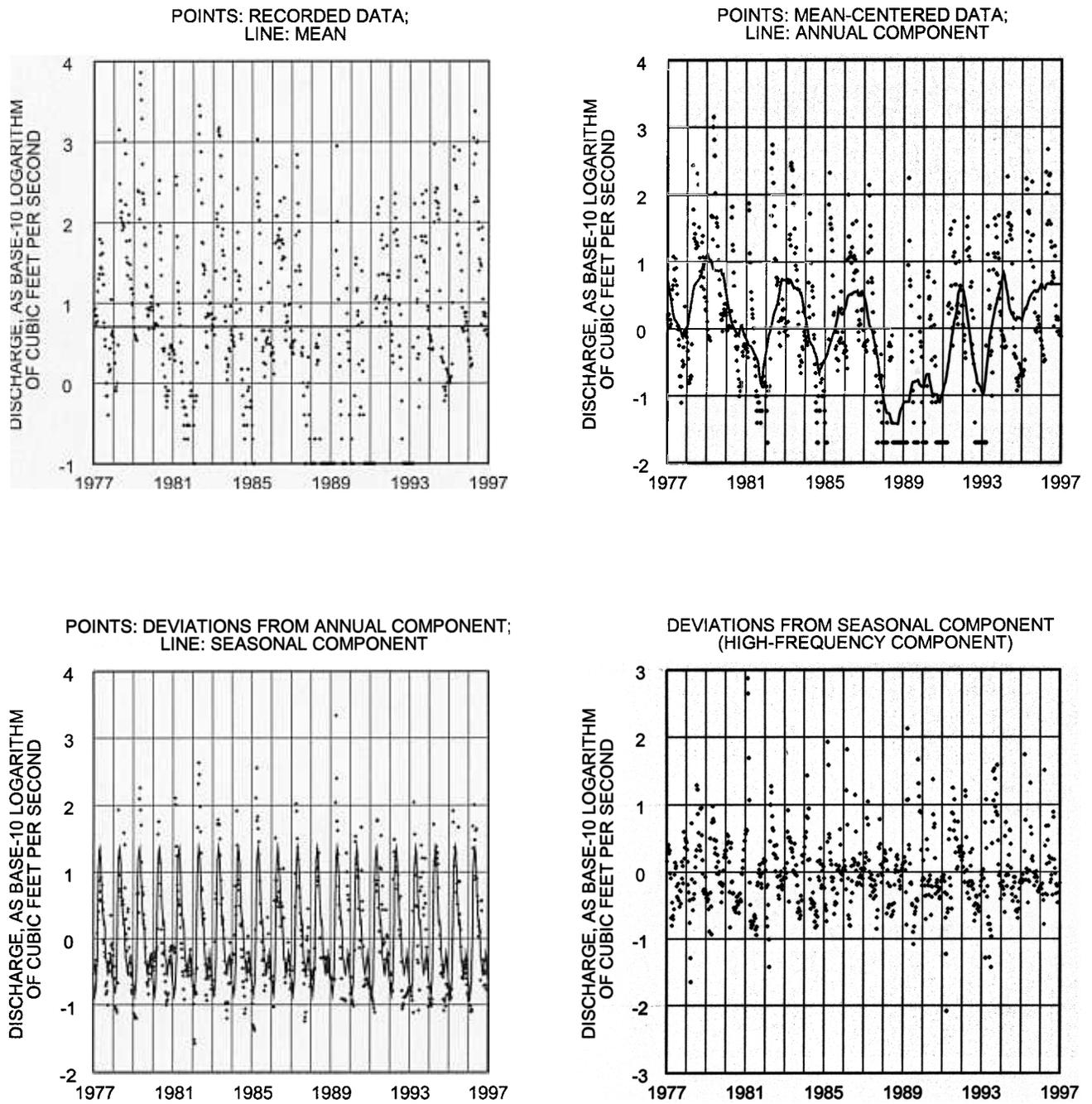
Figure A1. Mean, annual component, seasonal component, and high-frequency component of discharge, Souris River near Sherwood, North Dakota, 1977-96.

(LFC) of discharge because much of the noise, or high-frequency variation, in the data has been smoothed out. The LFC of discharge, which can be expressed as

$$LFC_X(t) = C_X + A_X(t) + S_X(t),$$

is shown in figure 2 along with the recorded data for the Souris River near Sherwood, N. Dak. (site 2). The last term on the right-hand side of the equation is called the high-frequency component (HFC) of discharge and can be expressed as

$$HFC_X(t) = U(t) = X(t) - LFC_X(t).$$

The individual components of equation A2 are shown in figure A2 for dissolved sulfate for the Souris River south of Glen Ewen, Saskatchewan (site 1), and the Souris River near Sherwood, N. Dak. (site 2). The first three terms on the right-hand side of the equation are called the LFC of concentration and can be expressed as

$$LFC_Y(t) = C_Y + A_Y(t) + S_Y(t).$$

The LFC of dissolved sulfate concentration for the Sherwood sites is shown in figure 3. The last term on the right-hand side of the equation is called the HFC of concentration and can be expressed as

$$HFC_Y(t) = W(t) = Y(t) - LFC_Y(t) - Tr(t). \tag{A3}$$

Although the trend term, $Tr(t)$, can be considered part of the LFC of concentration, the term was kept separate so that trend tests could be developed as described later. Estimation of the trend term also is described later.

The LFC of concentration was computed using a regression procedure because most of the concentration data are missing and using simple smoothing procedures was difficult. The constant, C_Y, was estimated by fitting an intercept as follows:

$$C_Y = \beta_0$$

where

β_0 is a regression parameter.

Next, a reasonable approximation for the annual component of concentration was obtained by assuming the component was linearly related to the annual component of discharge as follows:

$$A_Y(t) = \beta_1 A_X(t)$$

where

β_1 is a regression parameter.

Then, because a simple linear relation between the seasonal components of discharge and concentration was not adequate, a more complicated regression model was developed for the seasonal component:

$$S_Y(t) = \beta_2 [S_X(t)]^+ + \beta_3 [S_X(t)]^- + \beta_4 \cos(2\pi t) + \beta_5 \sin(2\pi t) + \beta_6 \cos(4\pi t) + \beta_7 \sin(4\pi t)$$

where

$\beta_2, ..., \beta_7$ are regression parameters;

$[.]^+$ denotes the value in brackets is replaced by zero if the value is negative;

$[.]^-$ denotes the value in brackets is replaced by zero if the value is positive;
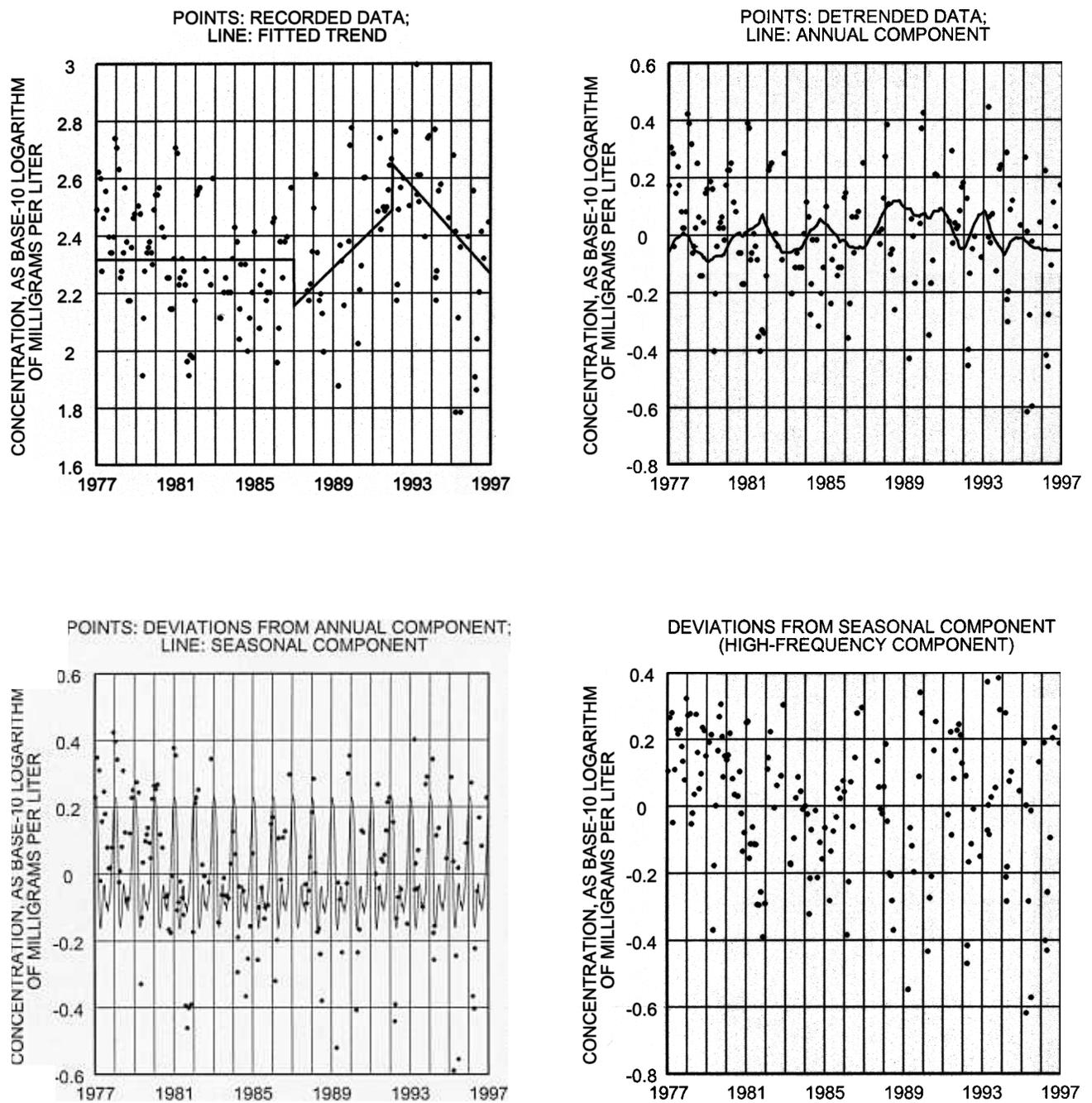
Figure A2. Fitted trend, annual component, seasonal component, and high-frequency component of dissolved sulfate, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96.

cos(a) is the cosine of angle a, where a is in radians; and

sin(a) is the sine of angle a, where a is in radians.

Equation A6 was determined through trial and error by fitting various functions of $S_X(t)$ and various trigonometric functions to the concentration data and then choosing the combination of terms that resulted in the best overall fit to the data. The model was used to explain as much of the natural seasonal variability in the concentration data as possible by using only discharge and known seasonal

functions of t as explanatory variables. The procedure used to estimate the regression parameters in equations A4 through A6 are described later.

Because the LFC of concentration minus the constant, $LFC_Y(t) - C_Y$, is obtained by regressing concentration data against various functions of discharge data, the deviation of $Y(t)$ from $LFC_Y(t) - C_Y$ is called the flow-adjusted concentration and is denoted by

$$Y_{FA}(t) = Y(t) - LFC_Y(t) + C_Y.$$

Rearranging equation A3 yields

$$Y_{FA}(t) = C_Y + Tr(t) + W(t),$$

indicating the trend in the flow-adjusted concentrations is identical to the trend in the recorded concentrations (see eq. A2). Thus, fitted trends in concentration usually are plotted with flow-adjusted data (as in figs. 8 through 20) rather than recorded data because the flow-adjustment process removes much of the variability in concentration and makes the trends easier to discern.

Flow-adjusted dissolved sulfate concentrations for the Sherwood sites are shown in figure 4 along with the fitted trend. The flow-adjusted concentration time series can be loosely interpreted as an estimate of the concentrations that would have been observed if flow conditions had been uniform throughout the entire sampling period. Although, as defined earlier, the trend consists only of $Tr(t)$, in figure 4, the trend line consists of $C_Y + Tr(t)$.

Periodic Autoregressive Moving-Average (PARMA) Time-Series Model

If the HFC of discharge [$U(t)$ in eq. A1] and the HFC of concentration [$W(t)$ in eq. A2] were uncorrelated and if both time series consisted of white noise (uncorrelated random variables with constant variance), the data could have been analyzed using a classical regression analysis. However, for most water-quality constituents, $U(t)$ and $W(t)$ are not uncorrelated, white-noise time series because significant cross-correlation and autocorrelation occur in the data. Also, the time series are not stationary because the cross-correlation, autocorrelation, and variance of the data may change depending on the time of year. For example, the variance of the HFC of concentration often is higher at the beginning of spring breakup than at other times of the year, and the autocorrelation between neighboring values of the HFC often is higher during summer low-flow conditions than during other times of the year. Therefore, a time-series model that can filter out complex seasonal structure in the HFCs is needed to detect trends in concentration. The periodic autoregressive moving-average (PARMA) time-series model described in this section was designed to accomplish that task and has been used extensively to analyze seasonal hydrologic time-series data (see, for example, Salas and others, 1985; Vecchia, 1985; Tasker, 1997). The PARMA model used for this study is a special version of the general multivariate PARMA model and was designed for use with discharge and concentration data. The model consists of two equations:

$$U(t) = \phi_1(t)U(t-\delta) + \theta_1(t)\sigma_1(t-\delta)\varepsilon_1(t-\delta) + \sigma_1(t)\varepsilon_1(t) \tag{A7}$$

and

$$W(t) = \phi_2(t)U(t) + \phi_3(t)W(t-\delta) + \theta_2(t)\sigma_1(t)\varepsilon_1(t) + \theta_3(t)\sigma_2(t-\delta)\varepsilon_2(t-\delta) + \sigma_2(t)\varepsilon_2(t) \tag{A8}$$

$\phi_1(t)$, $\phi_2(t)$, and $\phi_3(t)$ are periodic autoregressive coefficients;

δ is the sampling interval (one divided by the number of seasons);

$\theta_1(t)$, $\theta_2(t)$, and $\theta_3(t)$ are periodic moving-average coefficients;

$\sigma_1(t)$ and $\sigma_2(t)$ are periodic residual standard deviation coefficients;

$\epsilon_1(t)$ is the residual time series for discharge; and

$\epsilon_2(t)$ is the residual time series for concentration.

The autoregressive coefficients relate values of the time-series variables, $U(t)$ and $W(t)$, to current or past values of themselves; the moving-average coefficients relate values of the time-series variables to current or past values of the residuals; and the residual standard deviation coefficients are multiplied by the residuals to account for seasonal fluctuations in the variance of the residuals. The residual time series are assumed to be uncorrelated, and each residual time series is assumed to consist of white noise with zero mean and variance equal to one.

The model coefficients are assumed to be periodic functions of the time index with a period of 1 year. For example, for any given season, k, the autoregressive coefficient, $\phi_1(j + k\delta)$, that relates $U(j + k\delta)$ to $U(j + k\delta - \delta)$ depends on k but is the same for every year, j. Thus, given 36 seasons, each coefficient in the model can have one of 36 different values. To simplify the model, the coefficients are assumed to be represented by a trigonometric series with harmonics corresponding to periods of 1 year and one-half year. For example,

$$\phi_1(t) = a_1 + a_2\cos(2\pi t) + a_3\sin(2\pi t) + a_4\cos(4\pi t) + a_5\sin(4\pi t)$$

a_1, a_2, ..., a_5 are model parameters

is the trigonometric series representation for the first autoregressive coefficient. Thus, although the coefficient takes on a different value for each of the 36 seasons, only 5 parameters need to be estimated. Similar representations for the remaining autoregressive coefficients, the moving-average coefficients, and the residual standard deviation coefficients are assumed. Therefore, 40 parameters (5 for each of the 8 coefficients) need to be estimated to calibrate the model.

Terms that relate the HFC of discharge, $U(t)$, to the HFC of concentration, $W(t)$, are not included in equation A7 because discharge should not be causally related to concentration. However, this does not imply that discharge is uncorrelated with concentration because, in the second model equation (eq. A8), discharge and concentration are cross-correlated.

Algorithms used to estimate the PARMA model parameters are readily available for when no data are missing. However, algorithms used to estimate parameters when data are missing are less common. The procedure used in this study to estimate parameters is based on a periodic Kalman filtering algorithm that can easily manage missing data. The first step was to estimate the parameters in equation A7 using discharge data. The parameters were estimated by minimizing the following quantity:

$$L_1(\phi_1, \theta_1, \sigma_1) = \text{Sum}\left\{\frac{[U(t) - CE_U(t)]^2}{CV_U(t)} + \ln CV_U(t)\right\} \tag{A9}$$

$L_1(\phi_1, \theta_1, \sigma_1)$ is the quantity to be minimized with respect to ϕ_1, θ_1, and σ_1;

Sum{.} is the summation of the terms in braces for all observation times;

$CE_U(t)$ is the conditional expectation of $U(t)$ given $U(s)$ for all $s < t$; and

$CV_U(t)$ is the conditional variance of $U(t)$ given $U(s)$ for all $s < t$.

For the first observation time, the conditional expectation, $CE_U(t)$, was equal to zero and the conditional variance, $CV_U(t)$, was equal to the unconditional variance of $U(t)$. For each successive observation time, the conditional expectations and variances were computed using a modified periodic Kalman filtering algorithm for PARMA models (Jimenez and others, 1989). Equation A9 was minimized using a modified Gauss-Newton nonlinear minimization algorithm (Dennis and Schnabel, 1983). As mentioned previously, the minimization was with respect to the parameters in the Fourier series representation of the coefficients, so 15 parameters (5 for each of the model coefficients, ϕ_1, θ_1, and σ_1) were estimated.

If the discharge residuals in equation A7 are assumed to be independent, normally distributed random variables and the autoregressive and moving-average coefficients satisfy certain technical constraints, parameter estimates obtained by minimizing equation A9 are equivalent to maximum likelihood parameter estimates (Vecchia, 1985). However, even if the residuals are not normally distributed, the parameter estimates are asymptotically (as the number of years of data gets large) efficient (Brockwell and Davis, 1987).

To obtain parameter estimates for the entire PARMA model, which consists of equations A7 and A8, the discharge and concentration data were combined and the parameters of both equations were estimated jointly. However, because $U(t)$ and $W(t)$ are cross-correlated, this method resulted in different parameter estimates for discharge than those described earlier. Also, the parameter estimates for discharge were different for each constituent. Therefore, to simplify the analysis and increase the computational efficiency of the algorithm, parameter estimates for equation A7 were fixed at the values obtained using only discharge data. Then, the parameters of equation A8 were estimated by minimizing the following quantity:

$$L_2(\phi_2, \phi_3, \theta_2, \theta_3, \sigma_2) = \text{Sum}\left\{\frac{[U(t) - C\hat{E}_U(t)]^2}{C\hat{V}_U(t)} + \ln\ C\hat{V}_U(t)\right\} + \text{Sum}\left\{\frac{[W(t) - CE_W(t)]^2}{CV_W(t)} + \ln\ CV_W(t)\right\} \qquad \text{(A10)}$$

$L_2(\phi_2, \phi_3, \theta_2, \theta_3, \sigma_2)$ is the quantity to be minimized with respect to ϕ_2, ϕ_3, θ_2, θ_3, and σ_2;

Sum{.} is the summation of the terms in braces for all observation times with nonmissing data;

$C\hat{E}_U(t)$ is the conditional expectation of $U(t)$ given $U(s)$ for all $s < t$ and $W(s)$ for all $s < t$;

$C\hat{V}_U(t)$ is the conditional variance of $U(t)$ given $U(s)$ for all $s < t$ and $W(s)$ for all $s < t$;

$CE_W(t)$ is the conditional expectation of $W(t)$ given $U(s)$ for all $s \leq t$ and $W(s)$ for all $s < t$; and

$CV_W(t)$ is the conditional variance of $W(t)$ given $U(s)$ for all $s \leq t$ and $W(s)$ for all $s < t$.

The conditional expectations and variances were computed using a multivariate version of the periodic Kalman filtering algorithm described in Jimenez and others (1989). Equation A10 was minimized with respect to 25 parameters (5 for each of the coefficients, ϕ_2, ϕ_3, θ_2, θ_3, and σ_2). The conditional expectation and variance terms, $C\hat{E}_U(t)$ and $C\hat{V}_U(t)$, in equation A10 depend on past values of both U and W, but the corresponding terms in equation A9 depend only on past values of U. Thus, the first term on the

right-hand side of equation A10 depended on ϕ_2, ϕ_3, θ_2, θ_3, and σ_2 and improved the parameter estimates obtained by minimizing only the second term.

If the concentration residuals are assumed to be independent, normally distributed variables and if ϕ_2, ϕ_3, θ_2, and θ_3 satisfy certain technical constraints, the parameter estimates described earlier are equivalent (as the number of years of data increases) to maximum likelihood parameter estimates. The parameter estimates also are efficient for non-normal residual distributions. However, as with any procedure used to estimate parameters, the residuals needed to be examined to ensure the absence of extreme outliers or influential observations that may adversely affect the estimates.

Given the PARMA model (eqs. A7 and A8) and parameter estimates obtained as described earlier, the calibrated model can be used to calculate virtually any statistical property associated with the HFCs. For example, one important statistical property for trend-analysis purposes is the seasonal standard deviation of the HFC of concentration,

$$SD_W(k) = \{\text{Var}[W(j+k\delta)]\}^{\frac{1}{2}}$$

where

$SD_W(k)$ is the standard deviation of W for season k, and

$\text{Var}[.]$ is the variance of the quantity in brackets.

Another important statistical property is the seasonal cross-correlation between the HFCs of discharge and concentration,

$$\text{Corr}_{WU}(k, m) = \frac{\text{Cov}[W(j+k\delta),\, U(j+k\delta-m\delta)]}{SD_W(k)SD_U(k-m)}$$

where

$\text{Corr}_{WU}(k, m)$ is the cross-correlation between W and U for season k and lag m, and

$\text{Cov}[.\,,.]$ is the covariance of the quantities in brackets.

Yet another important statistical property is the autocorrelation of the HFC of concentration,

$$\text{Corr}_{WW}(k, m) = \frac{\text{Cov}[W(j+k\delta),\, W(j+k\delta-m\delta)]}{SD_W(k)SD_W(k-m)}$$

where

$\text{Corr}_{WW}(k, m)$ is the autocorrelation of W for season k and lag m.

The seasonal statistics, $SD_W(k)$, $\text{Corr}_{WU}(k, m)$ with $m = 0$, and $\text{Corr}_{WW}(k, m)$ with $m = 3$, for dissolved sulfate for the Sherwood sites are shown in figure 5. The first plot indicates the standard deviation of the HFC of concentration is highest during March through May and lowest during July through September. This implies that trends in concentration will be more difficult to detect during spring breakup than during summer low-flow conditions because more natural variability, or noise, occurs in the HFC of concentration during spring breakup. The second plot indicates a significant negative cross-correlation between the HFCs of discharge

and concentration during February through April. Thus, much of the increased variability in concentration during spring breakup is related to short-term variability in discharge. Higher-than-normal discharge during February through April usually results from an earlier-than-normal spring breakup, which produces lower-than-normal dissolved sulfate concentrations because fresh water from snowmelt does not penetrate the frozen soils. However, higher-than-normal discharge during the summer and fall produced higher-than-normal dissolved sulfate concentrations because increased runoff in the summer and fall may cause increased leaching of sulfate from the soils. The third plot indicates autocorrelation between neighboring values of the HFC of concentration (at a lag of about 30 days) is highest during August through December and lowest during February through June. Thus, concentrations generally vary more slowly during August through December low-flow conditions, when discharge is predominately from base flow and reservoir releases, than during the rest of the year, when flow conditions can change rapidly between low flow, ice-covered flow, and high flow.

Detection of Trends in Flow-Adjusted Concentrations

The procedure for detecting trends in flow-adjusted concentrations is described in this section. However, before the procedure is described, a summary of the model structure up to this point is given. The general model is expressed by equations A1 and A2. The constant, annual variability, and seasonal variability in discharge are computed directly from the discharge data as described earlier, and the constant, annual variability, and seasonal variability in concentration are given in terms of regression equations A4, A5, and A6. The HFCs are modeled using the PARMA model (eqs. A7 and A8). The only component still needing to be defined explicitly is the trend in concentration, $Tr(t)$, in equation A2. The trend is expressed in terms of a linear combination of known functions of time as

$$Tr(t) = \gamma_1 f_1(t) + \gamma_2 f_2(t) + \quad \gamma_p f_p(t)$$

where

$\gamma_1, \gamma_2, ..., \gamma_p$ are trend parameters;

$f_1(t), f_2(t), ..., f_p(t)$ are specified functions of t; and

p is the number of trend parameters.

An intercept term (eq. A4) already is included in the model and, therefore, is not included in equation A11. The general trend function (eq. A11) provides considerable flexibility in the types of trends that can be modeled. For example, to model a step trend in concentration resulting from the different sampling locations given in table 1, a single trend function that has a value of one for 1987-91 and zero for other periods is appropriate:

$$f_1(t) = 1 \text{ for } 1987 < t \le 1992; f_1(t) = 0 \text{ for } 1977 < t \le 1987 \text{ and } 1992 < t \le 1997.$$

To model a linear trend in log-transformed concentration during 1977-86, the following function is appropriate:

$$f_2(t) = t - 1977 \text{ for } 1977 < t \le 1987; f_2(t) = 10 \text{ for } 1987 < t \le 1997$$

Similarly, a linear trend during 1987-91 is expressed as

$$f_3(t) = 0 \text{ for } 1977 < t \le 1987; f_3(t) = t - 1987 \text{ for } 1987 < t \le 1992; f_3(t) = 5 \text{ for } 1992 < t \le 1997$$

and a linear trend during 1992-96 is expressed as

$$f_4(t) = 0 \text{ for } 1977 < t \le 1992; f_4(t) = t - 1992 \text{ for } 1992 < t \le 1997$$

The trend model used in this report is, thus, given by equation A11 (with $p = 4$) and equations A12 through A15 (the four trend functions). Although more complicated trends were not modeled in this report, the general trend function (eq. A11) can be used to construct more complicated trends. For example, if an increase in agricultural production is suspected of causing an increase in total nitrogen concentration, one of the trend functions could be an index of total agricultural production (for example, total acres of irrigated crops in the basin). Other exogenous variables, such as gallons of water released from hydroelectric plants, total coal production, etc., also could be included. In addition, trends in concentration that occur only during certain parts of the year could be modeled.

To detect trends in flow-adjusted concentrations, a determination needed to be made as to whether the trend parameters in equation A11 are significantly different from zero. This determination was made by expressing the time-series model (eqs. A1 and A2) in the form of a generalized, multiple linear regression model. The linear regression model was obtained by combining the discharge and concentration data (eqs. A1 and A2) into a single vector of observations. For a particular year, j, the terms on the left-hand sides of the equations were combined as follows:

$$\boldsymbol{D}_j^T = [X(j+\delta) \quad Y(j+\delta) \quad X(j+2\delta) \quad Y(j+2\delta) \qquad X(j+1) \quad Y(j+1)]$$

where

$\boldsymbol{D}_j$ is a data vector for year j, and

$[.]^T$ denotes the vector or matrix transpose.

Thus, for 36 seasons, $\boldsymbol{D}_j$ has 72 elements (all 36 values of Y are included at this point although many values are missing). Using the regression equations A4 through A6, the trend equation A11, and some matrix algebra, the following equation then was derived:

$$\boldsymbol{D}_j = \boldsymbol{A}_{j0} + \boldsymbol{A}_{j1}\beta + \boldsymbol{A}_{j2}\gamma + \boldsymbol{H}_j$$

where

$$\boldsymbol{A}_{j0}^T = [LFC_X(j+\delta) \quad 0 \quad LFC_X(j+2\delta) \quad 0 \qquad LFC_X(j+1) \quad 0],$$

$$\boldsymbol{A}_{j1}^T = \begin{bmatrix} 0 & 0 & 0 & \\ 0 & A_X(j+\delta) & 0 & A_X(j+2\delta) & 0 & A_X(j+1) \\ 0 & [S_X(j+\delta)]^+ & 0 & [S_X(j+2\delta)]^+ & 0 & [S_X(j+1)]^+ \\ 0 & [S_X(j+\delta)]^- & 0 & [S_X(j+2\delta)]^- & 0 & [S_X(j+1)]^- \\ 0 & \cos[2\pi(j+\delta)] & 0 & \cos[2\pi(j+2\delta)] & 0 & \cos[2\pi(j+1)] \\ 0 & \sin[2\pi(j+\delta)] & 0 & \sin[2\pi(j+2\delta)] & 0 & \sin[2\pi(j+1)] \\ 0 & \cos[4\pi(j+\delta)] & 0 & \cos[4\pi(j+2\delta)] & 0 & \cos[4\pi(j+1)] \\ 0 & \sin[4\pi(j+\delta)] & 0 & \sin[4\pi(j+2\delta)] & 0 & \sin[4\pi(j+1)] \end{bmatrix},$$

$$\beta^T = [\beta_0 \; \beta_1 \qquad \beta_7],$$

$$A_{j2}^T = \begin{bmatrix} 0 & f_1(j+\delta) & 0 & f_1(j+2\delta) & 0 & f_1(j+1) \\ 0 & f_2(j+\delta) & 0 & f_2(j+2\delta) & 0 & f_2(j+1) \\ 0 & f_3(j+\delta) & 0 & f_3(j+2\delta) & 0 & f_3(j+1) \\ 0 & f_4(j+\delta) & 0 & f_4(j+2\delta) & 0 & f_4(j+1) \end{bmatrix}$$

$$\gamma^T = [\gamma_1 \; \gamma_2 \; \gamma_3 \; \gamma_4], \text{ and}$$

$$H_j^T = [U(j+\delta) \; W(j+\delta) \; U(j+2\delta) \; W(j+2\delta) \qquad U(j+1) \; W(j+1)]$$

Next, using equation A16, a single equation was formed for all years:

$$D = A_0 + A_1\beta + A_2\gamma + H \tag{A17}$$

where

$D = \text{Stack}\{D_j\};$

$A_0 = \text{Stack}\{A_{j0}\};$

$A_1 = \text{Stack}\{A_{j1}\};$

$A_2 = \text{Stack}\{A_{j2}\};$

$H = \text{Stack}\{H_j\};$ and

Stack{.} denotes that the matrices in braces are stacked, by rows, for $j = 1977, \; ., 1996$.

This equation (eq. A17) is in the form of a generalized, multiple linear regression model with correlated errors. The errors in the equation refer to the vector of HFCs, H, that are correlated as a result of the PARMA model (eqs. A7 and A8).

To detect trends in flow-adjusted concentrations, the parameter vectors, β and γ, were estimated jointly with the PARMA model parameters. This was accomplished using the periodic Kalman filter described earlier. The periodic Kalman filter is a linear filter and, thus, could be applied to equation A17 to obtain the following filtered version of the equation:

$$D^* = A_0^* + A_1^*\beta + A_2^*\gamma + H^* \tag{A18}$$

where

* denotes that the Kalman filter has been applied to the given vector or matrix.

If a matrix has several columns, the Kalman filter is applied separately to each column. Also, the Kalman filter automatically adjusts for missing data in the original data vector, D, by reducing the row dimension of equation A18. The elements of the Kalman-filtered vector, H^*, are uncorrelated with zero mean and variance equal to one. Therefore, equation A18 is in the form of an ordinary, multiple linear regression model.

The PARMA model parameters then were estimated jointly with the regression parameters. Because the estimation equation for discharge (eq. A9) has no regression parameters, the equation was not modified. However, equation A10 was modified to include the regression parameters. The values of the HFC, $W(t)$, in equation A10 depend on the unknown parameter vectors, β and γ, as indicated in equation A17. Therefore, the HFC was estimated by replacing β and γ in equation A17 with estimated values and solving for $\hat{H}$ as follows:

$$\hat{H} = D - A_0 - A_1\hat{\beta} - A_2\hat{\gamma} \tag{A19}$$

$\hat{H}$ is the estimated value of H,

$\hat{\beta}$ is the estimated value of β, and

$\hat{\gamma}$ is the estimated value of γ.

The estimated values of β and γ were obtained using the ordinary least-squares regression model (eq. A18),

$$[\hat{\beta}^T \ \hat{\gamma}^T]^T = \text{inv}\{M^T M\} M^T (D^* - A_0^*)$$

inv denotes the matrix inversion, and

$M = [A_1^* \ A_2^*]$.

The values of $W(t)$ in equation A10 (which corresponds to particular elements of H as indicated in equation A17) were replaced with the corresponding elements, $\hat{W}(t)$, of $\hat{H}$ in equation A19. The PARMA model parameter estimates for equation A8 then were obtained by minimizing the modified version of equation A10.

Estimates of the PARMA model residuals, H^* in equation A18, were computed by applying the periodic Kalman filter to H in equation A19. The PARMA model residuals for dissolved sulfate for the two Sherwood sites are shown in figure 6. The effectiveness of the Kalman filter in removing autocorrelation in the HFC is shown by a comparison of figure 6 and the lower right graph in figure A2. The deviations from the seasonal component in figure A2 are the nonmissing values of the HFC, $\hat{W}(t)$, contained in the vector, $\hat{H}$, in equation A19. As shown in figure A2, a high degree of autocorrelation exists in the HFC. For example, large groups of positive values occur in 1977, 1978, and 1991, and large groups of negative values occur in 1981, 1983 through 1985, 1992, and 1995. However, for the PARMA model residuals (fig. 6), groups of positive or negative values are not apparent, indicating that the Kalman filter is effective in removing autocorrelation in the HFC.

Although the parameter estimates are fairly robust with respect to the distribution of the PARMA model residuals (because the record length is large), the residuals were evaluated to ensure that they were not highly skewed or heavy tailed and that they did not contain any extreme outliers or influential observations. For example, the residuals shown in figure 6 have an approximate standard normal distribution except for a possible positive outlier in 1993 and two possible negative outliers in 1995. After removing the outliers and recalibrating the model to the reduced data set, virtually no change occurred in the calibrated model. Therefore, an assumption was made that the outliers do not affect the model or any conclusions regarding significance levels of fitted trends. The

model residuals also were evaluated in relation to season (fig. A3). No obvious bias or heteroskedasticity existed in the residuals, and the nonmissing residuals were scattered fairly uniformly among the seasons. Residual plots for the 13 constituents chosen for analysis for both the Sherwood and Westhope sites are given in appendix B. In all cases, the residuals appear to be approximately normally distributed except for occasional outliers, none of which unduly influenced the calibrated model.
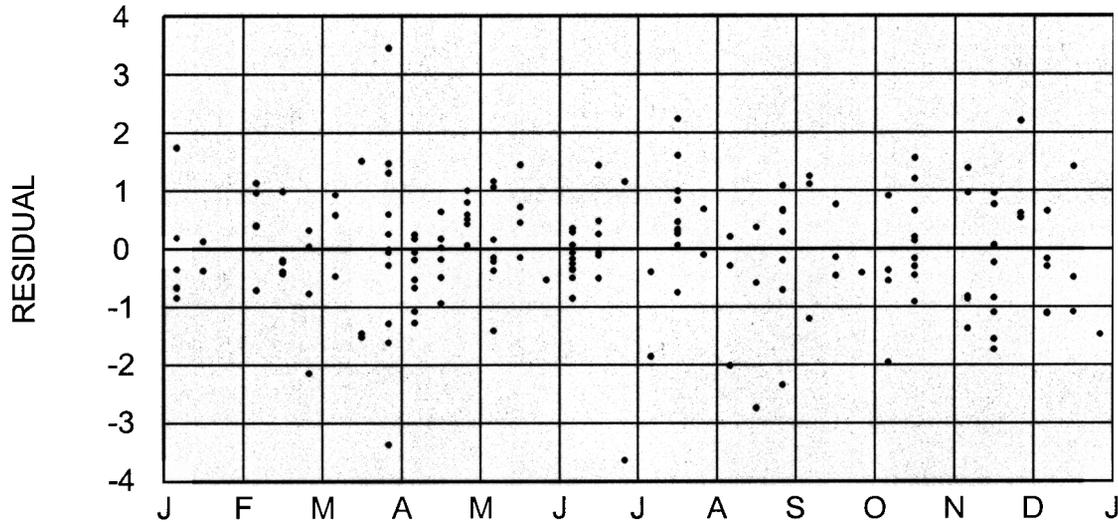
Figure A3. Periodic autoregressive moving-average (PARMA) model residuals versus season for dissolved sulfate, Souris River south of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota, 1977-96.

Hypothesis tests regarding trends in concentration can be constructed using the estimated regression and PARMA model parameters. Using standard results from asymptotic theory (Brockwell and Davis, 1987, ch. 6), the estimated trend parameters have an approximate multivariate normal distribution as shown by the following equation:

$$\hat{\gamma} \Rightarrow MVN(\mu, \Sigma)$$

where

$\Rightarrow$ denotes convergence in distribution (as the record length increases);

$MVN(\mu, \Sigma)$ denotes a multivariate normal distribution with mean vector μ and covariance matrix Σ;

$\mu = \gamma$;

$\Sigma = inv\left\{ A_2 *^T [I - A_1 * \; inv(A_1 *^T A_1 *) A_1 *^T] A_2 * \right\}$; and

I is an identity matrix.

The covariance matrix in equation A20 was computed using PARMA model parameter estimates. As indicated previously, the PARMA model parameter estimates are equivalent (as the record length increases) to maximum likelihood parameter estimates. Thus the PARMA model parameter estimates converge in probability to the true PARMA model parameters as the record length increases and equation A20 holds whether the covariance matrix is computed using the estimates or the true parameters. Generally, the

covariance matrix in equation A20 provides an adequate approximation to the true covariance matrix provided 15 or more years of data are used to estimate the PARMA model parameters.

The estimated covariance matrix in equation A20 is robust to the actual distribution of the PARMA model residuals. However, if the PARMA model residuals are not normally distributed, each column of $A_2{}^*$ needs to have a sufficient number of nonzero elements for equation A20 to be valid. The number of nonzero elements in the kth column of $A_2{}^*$ is the number of nonmissing concentration observations that depend on γ_k. For example, to estimate the trend parameter, γ_4, that corresponds to a linear trend during 1992-96 (eq. A15), a sufficient number of nonmissing values needs to be available for 1992-96. The sufficient number depends on the actual distribution of the PARMA model residuals. If the residuals are normally distributed, only one nonmissing value is required, but, if the residuals are not normally distributed, more than one nonmissing value is required. Generally, if the residuals are not highly skewed or heavy tailed, 10 nonmissing values should be adequate to estimate γ_4. The data need to be sufficient to estimate each trend parameter. For example, to estimate the size of a step trend, at least five concentration measurements should be available both before and after the trend. The smaller the number of data points, the greater the amount of care needed to ensure that the PARMA model residuals are normally distributed.

To determine if a particular trend is significant, the estimated trend parameter is divided by its standard error as follows

$$z_k = \frac{\hat{\gamma}_k}{SE(\hat{\gamma}_k)}$$

where

z_k is the normal score for the kth trend parameter, and

SE (.) denotes the standard error of the estimate.

The standard error is the square root of the corresponding diagonal element of the covariance matrix in equation A20. The one-sided p-value corresponding to the kth parameter is defined as

$$p_k = \Phi[\text{abs}(z_k)]$$

where

Φ [.] is the standard normal quantile function, and

abs(.) denotes absolute value.

For this report, linear trend parameters that correspond to equations A13 through A15 were kept in the model only if the trends were significant at the 5-percent level (p-values less than 0.05). Thus, a backward elimination procedure was used to remove the parameters, one at a time, until all remaining trends were significant at that level. For example, after the model was calibrated for dissolved sulfate for the Sherwood sites, the linear trend for 1977-86 was not significant. Thus, the trend parameter was removed from the model, and the model was recalibrated. The linear trends for 1987-91 and 1992-96 were highly significant (p-values less than 0.001; fig. 11) and, thus, were kept in the model. Because the data are log-transformed, the fitted trends are interpreted as the percent change per year in the median concentration. For example, the estimated increasing trend of 16.5 percent per year for 1987-91 (fig. 11) translates to an

increase of $100[(1.165)^5 - 1]$, or 115 percent in the median concentration from the beginning to the end of the 5-year period. To draw an analogy with economics, the trends are "compounded annually."

Efficiency of Sampling Designs for Detecting Trends

The covariance matrix of the multivariate normal distribution in equation A20 depends only on the calibrated PARMA model for the HFC and on the trend functions in equation A11. Thus, the PARMA model parameters contain all the information that is needed to evaluate the efficiency of various sampling designs for detecting trends. For example, given the calibrated PARMA model for 1977-96, a sampling design for detecting a linear trend in log-transformed concentration for the future 10-year period 1997-2006 can be developed using equation A11 with $p = 1$:

$$f_1(t) = t - 1997 \text{ for } 1997 < t \le 2007$$

Furthermore, to detect a trend that occurs only during the spring high-flow conditions of March through June (season 1 in table 4), the Kalman filter can be used to calculate the standard error of the estimated trend. Let the potential data set,

$$D_j^T = [X(j + \delta) \quad Y(j + \delta) \quad X(j + 2\delta) \quad Y(j + 2\delta) \qquad X(j + 1) \quad Y(j + 1)]$$

consist of the full set of 36 discharge and 36 concentration measurements for each future year, $j = 1997, ..., 2006$. To detect a trend in season 1 using two samples per year (see table 4), the nonmissing data for a single year are specified by a vector,

$$V_j^T = [V1 \quad V2 \quad V3],$$

where

$V1 = [\text{x o x o x x x o x o x o x o x o x x x o x o x o}]$,
$V2 = [\text{x o x o x o x o x o x o x o x o x o x o x o x o}]$,
$V3 = [\text{x o x o x o x o x o x o x o x o x o x o x o x o}]$,
x denotes a nonmissing value, and
o denotes a missing value.

$V1$ specifies that season 1 has two nonmissing concentration values (in the third and ninth 10-day periods), and $V2$ and $V3$ specify that seasons 3 and 4 have no concentration data. Thus, the data set for the sampling design consists of 360 discharge measurements (36 per year) and 20 concentration measurements (2 per year). The periodic Kalman filter with the appropriate missing value codes then is used in equation A20 to calculate the standard error of the estimated trend.

The efficiency of the sampling design is related inversely to the standard error. For example, suppose that the null hypothesis is that no trend exists ($\gamma_1 = 0$) and the alternative hypothesis is that an increasing trend exists ($\gamma_1 > 0$). Also, suppose that a 5-percent significance level is used. The null hypothesis is rejected if

$$\frac{\hat{\gamma}}{\hat{\gamma}} \quad \Phi \qquad .64$$

efficient design   the   of the   the probability of rejecting the null hypothesis,   that the alternativ
hypothes   In equation form, the   er of the

$$\text{Power}(\quad \text{Prob}\left[\frac{\hat{\gamma}_1}{SE(\hat{\gamma}} \quad .64 \right.$$

which   be shown with   algebraic manipulation   be equivalent

$$\text{Power}(\quad \Phi\left(.64: \quad \frac{\gamma_1}{SE(\hat{\gamma}_1)} \right)$$

Therefore, for any value of   greater than   the power of the   maximized whenever the standard   of the estimate
minimized.   quantify the efficiency of the sampling design using   single number   note that the power of the design   equal
that is,   i0-50   hance holds for rejecting the null hypothesis   provided that

$$.64 \quad SE($$

The characteristic trend used   evaluate the sampling designs   this   therefore, defined

$$C \qquad 100($$

where   the characteristic trend of the des gn,   percent

APPENDIX B. WATER-QUALITY DATA AND TIME-SERIES MODEL OUTPUT

The water-quality data and selected time-series model output used in this report are shown in the figures in this appendix. The figures can be reproduced by accessing the home page of the U.S. Geological Survey in North Dakota (http://nd.water.usgs.gov) and following instructions in the users' manual. The graphs in each of the figures are interpreted as follows:

Upper-left graph

Recorded data, $Y(t)$, and fitted trend, $C_Y + Tr(t)$ (see eq. A2);

Upper-right graph

Detrended data, $Y(t) - Tr(t)$, and fitted low-frequency component, $C_Y + A_Y(t) + S_Y(t)$;

Lower-left graph

Flow-adjusted data, $Y(t) - A_Y(t) - S_Y(t)$, and fitted trend, $C_Y + Tr(t)$. (These graphs correspond to figures 8 through 33, except that nonsignificant trends have not been removed.)

Lower-right graph

PARMA model residuals, which are the estimates of H^* in equation A18.

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED OXYGEN, IN MILLIGRAMS PER LITER
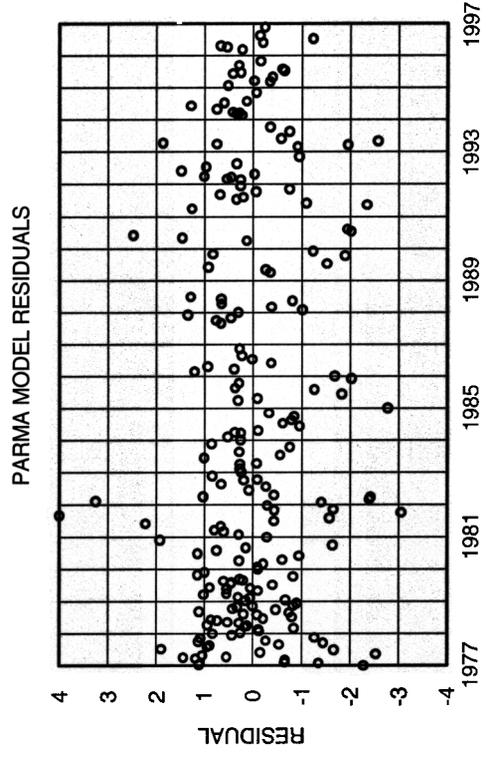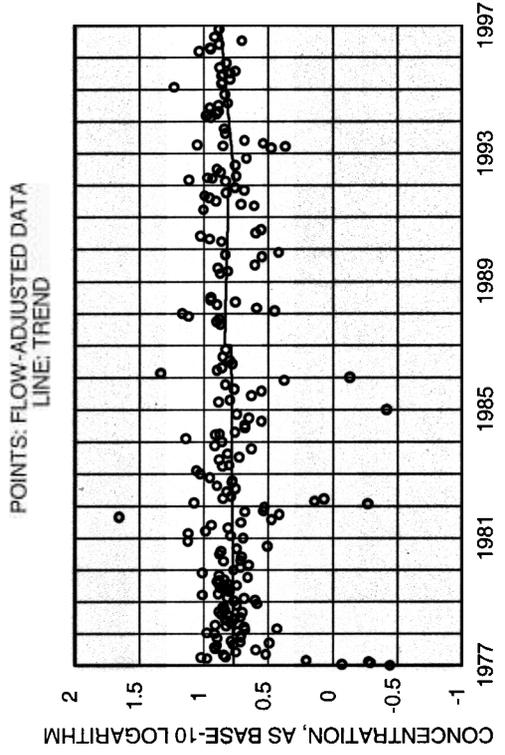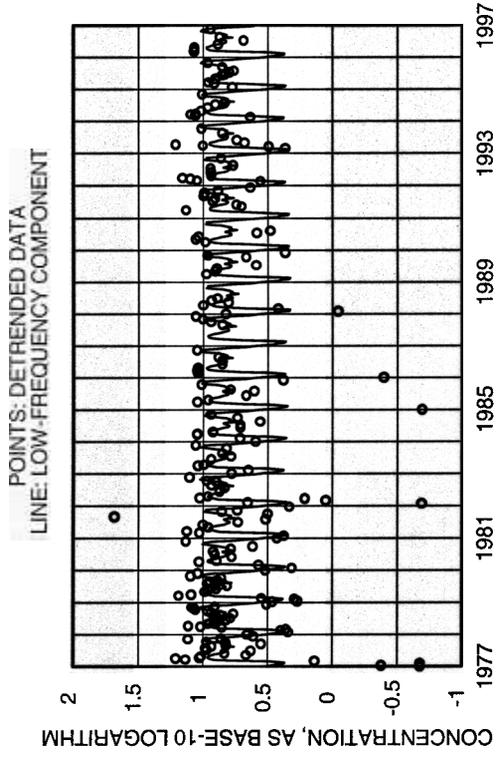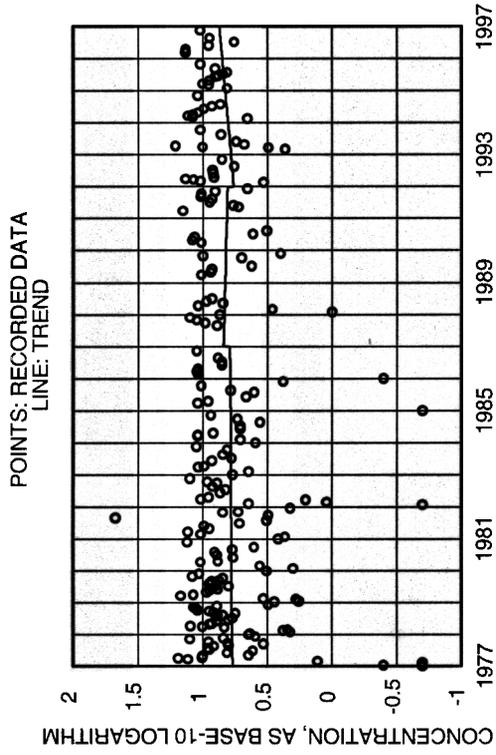
POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED CALCIUM, IN MILLIGRAMS PER LITER

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED SODIUM, IN MILLIGRAMS PER LITER
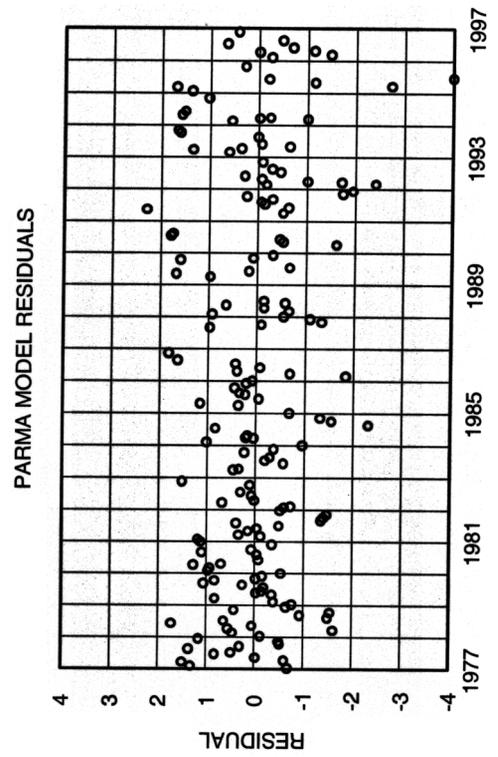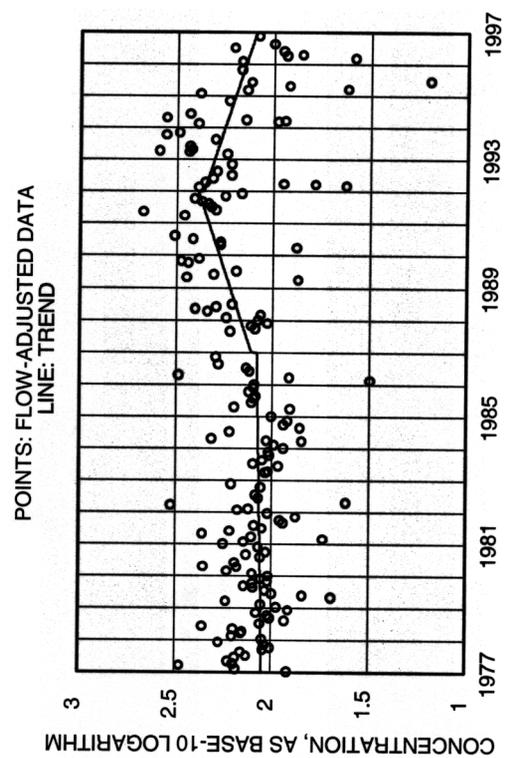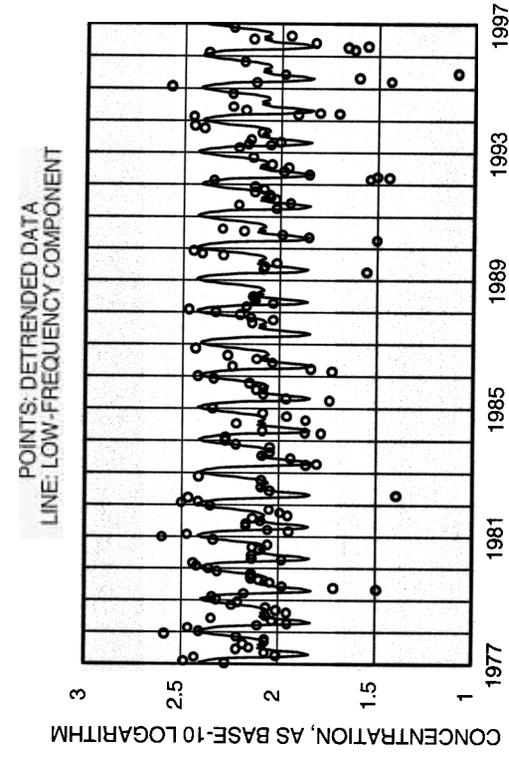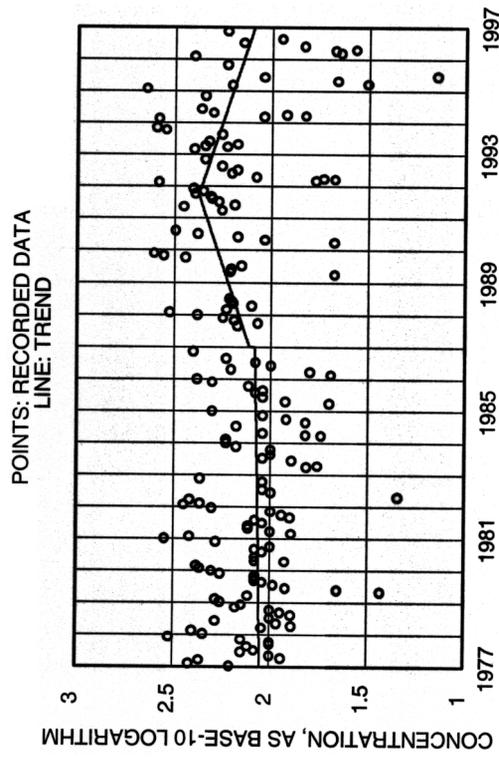
POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED SULFATE, IN MILLIGRAMS PER LITER
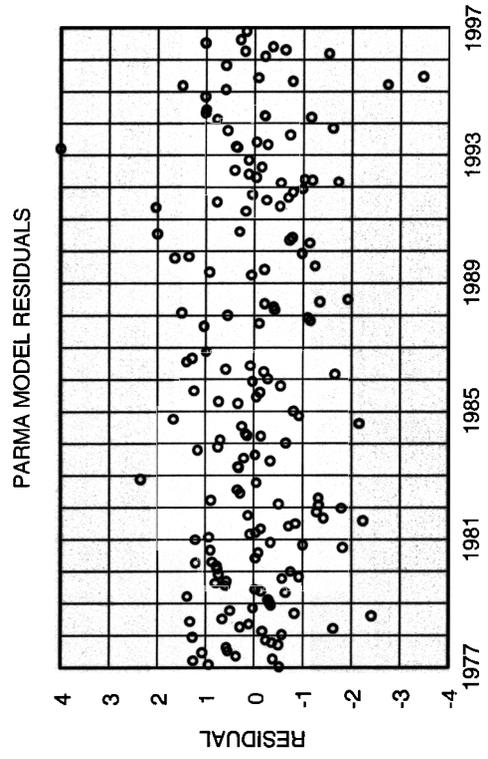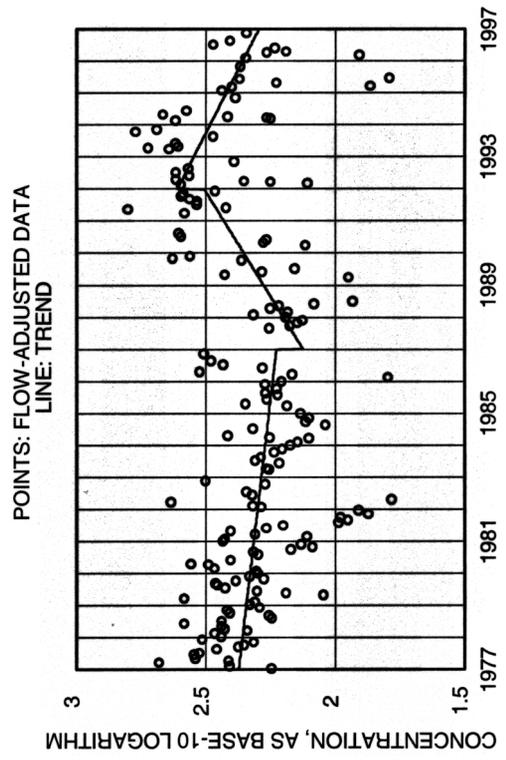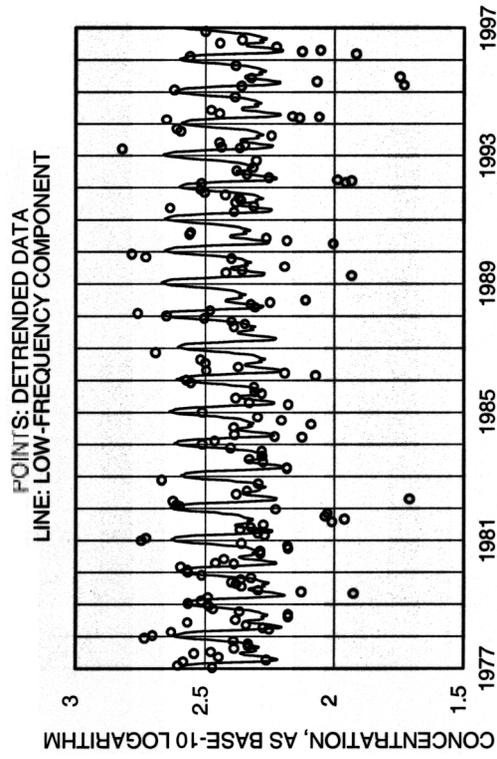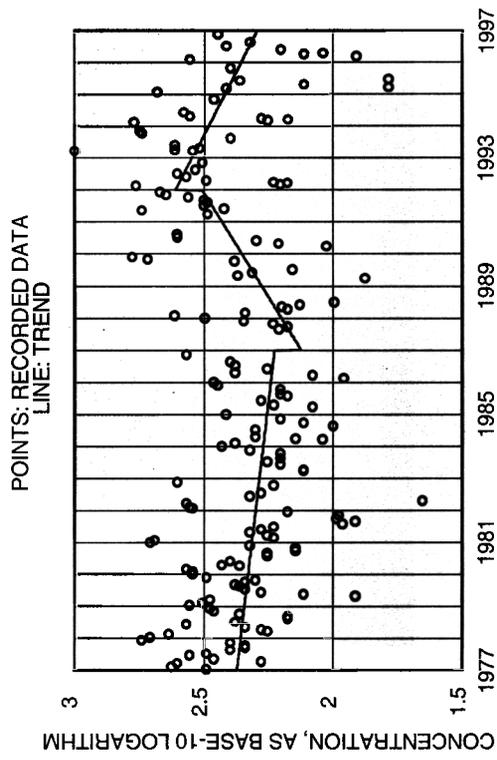
Souris River South | Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED CHLORIDE, IN MILLIGRAMS PER LITER
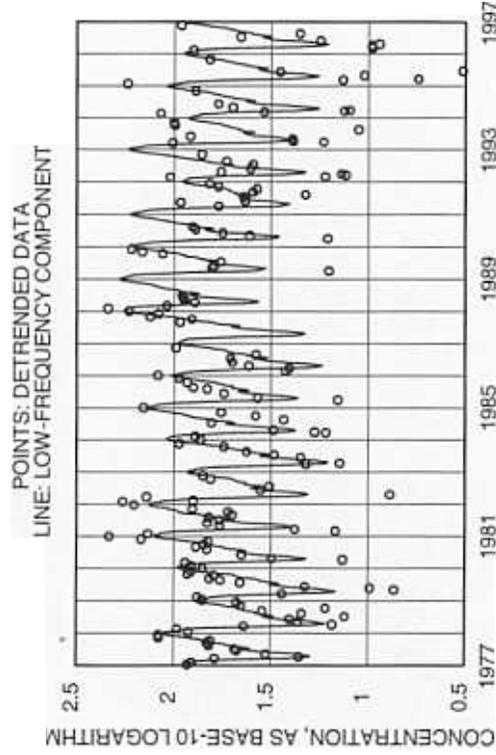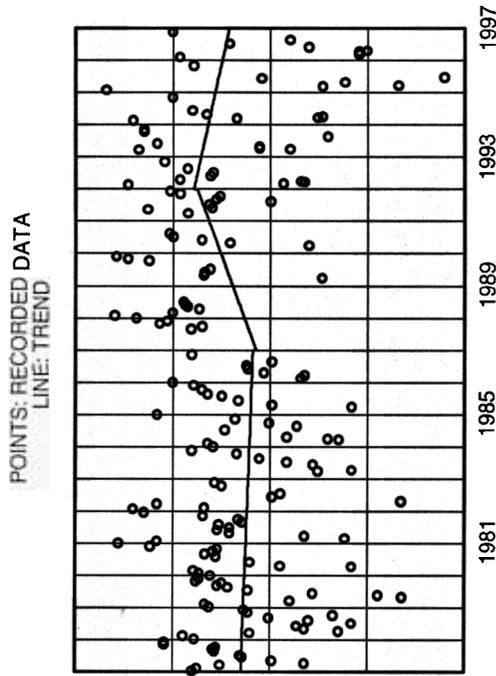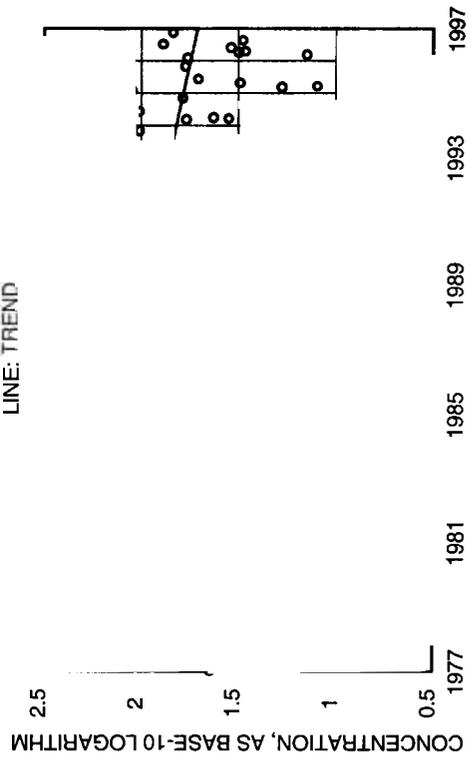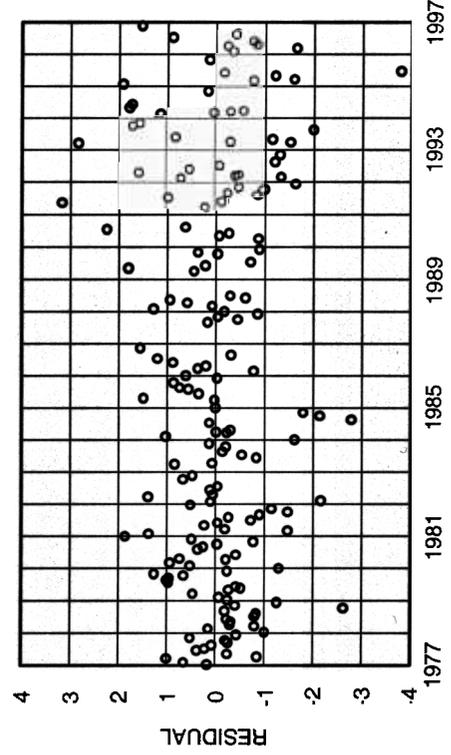
POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL

FLOW-ADJUSTED DATA
LINE: TREND

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED SOLIDS, IN MILLIGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS
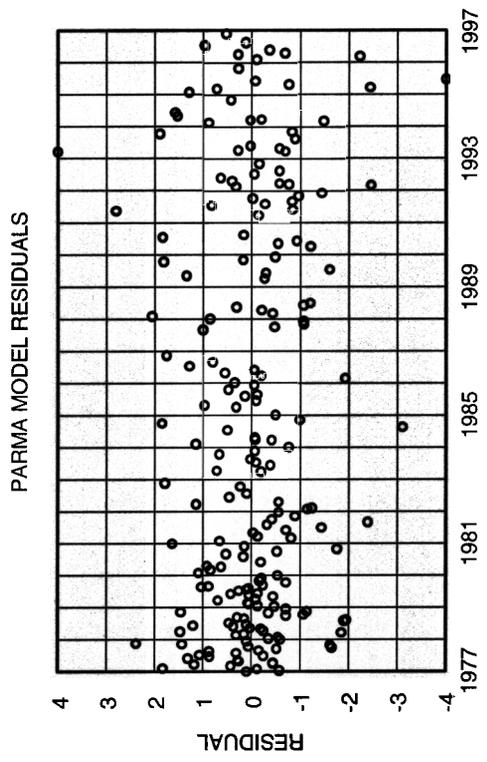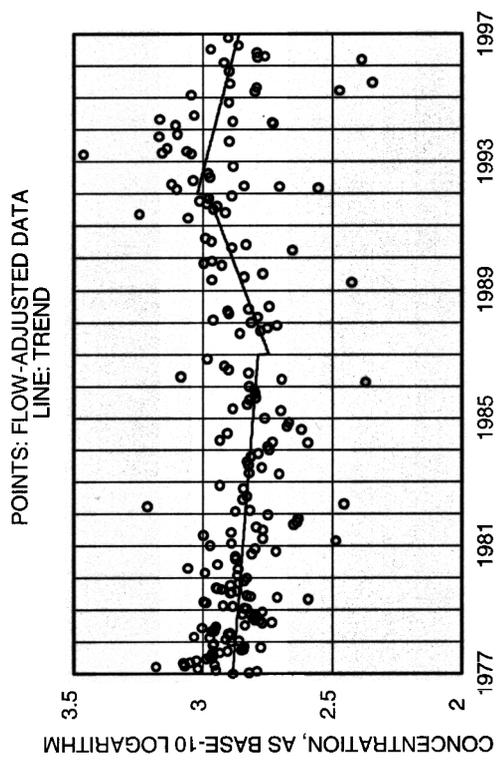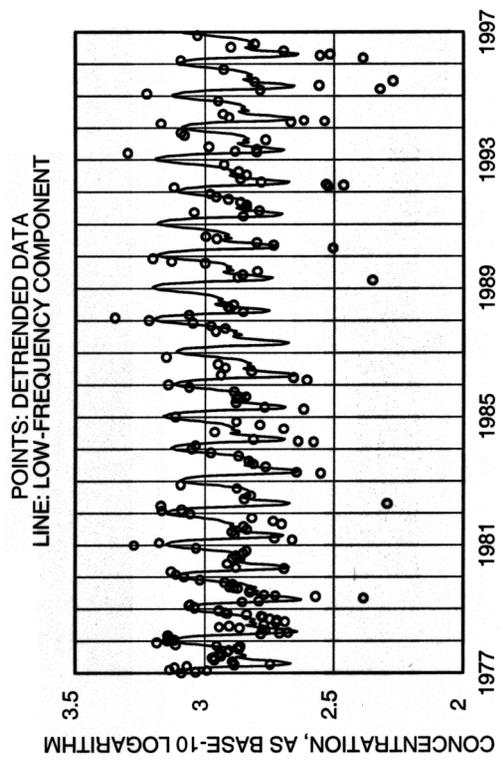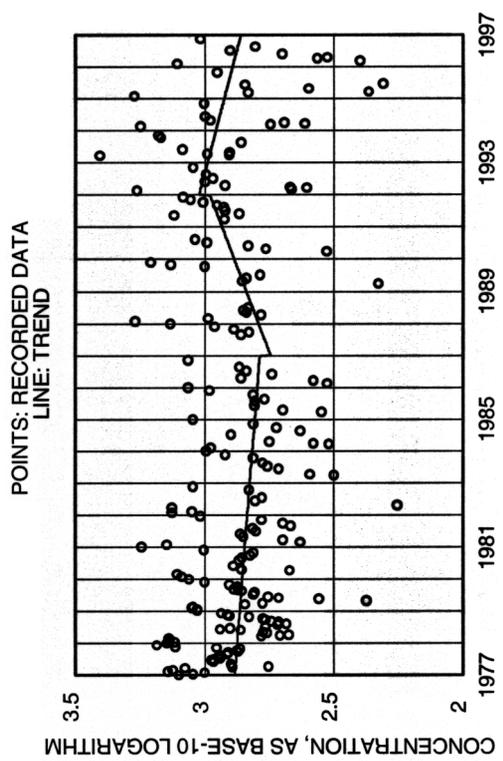
Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

TOTAL AMMONIA, IN MILLIGRAMS PER LITER
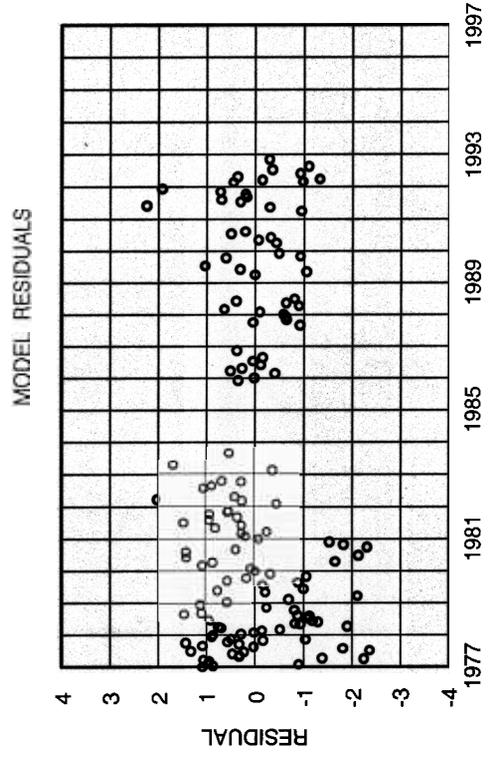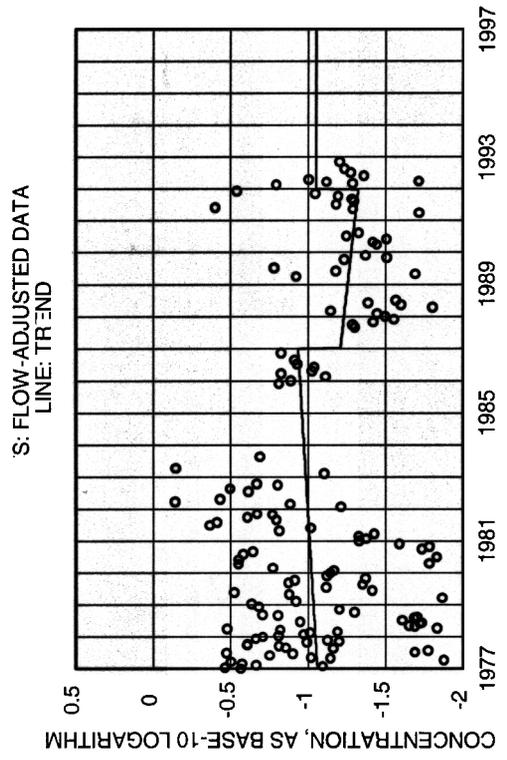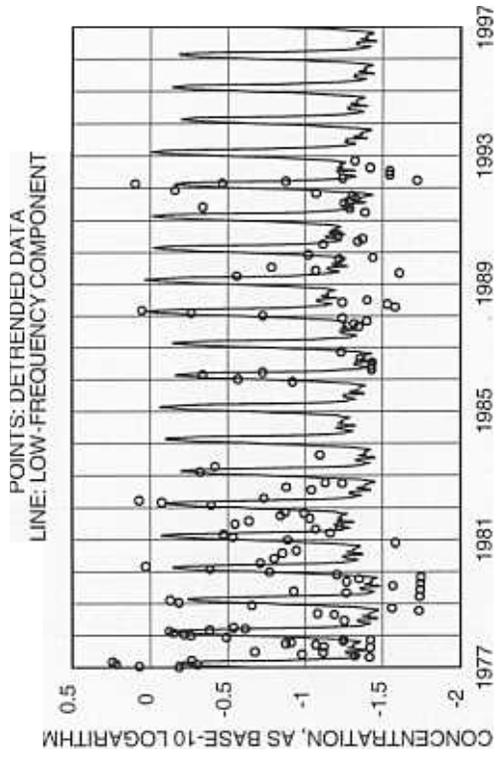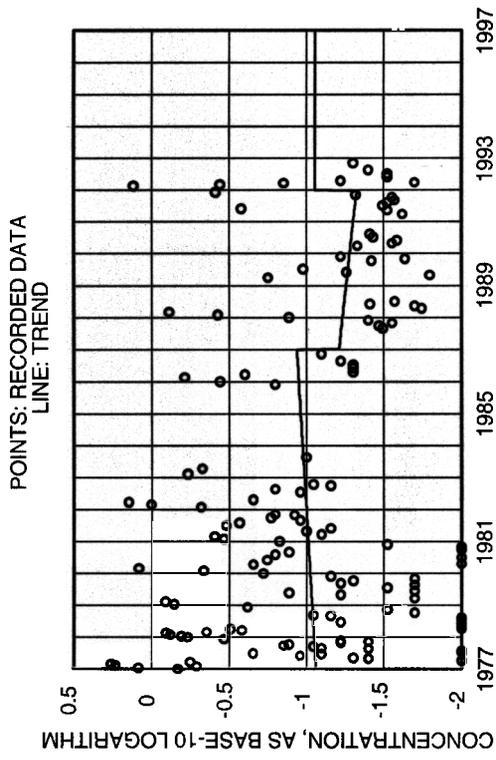
POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

CONCENTRATION, AS BASE-10 LOGARITHM

MODEL RESIDUALS

RESIDUAL

POINTS: RECORDED DATA
LINE: TREND

CONCENTRATION, AS BASE-10 LOGARITHM

S: FLOW-ADJUSTED DATA
LINE: TREND

CONCENTRATION, AS BASE-10 LOGARITHM

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

TOTAL NITROGEN, IN MILLIGRAMS PER LITER
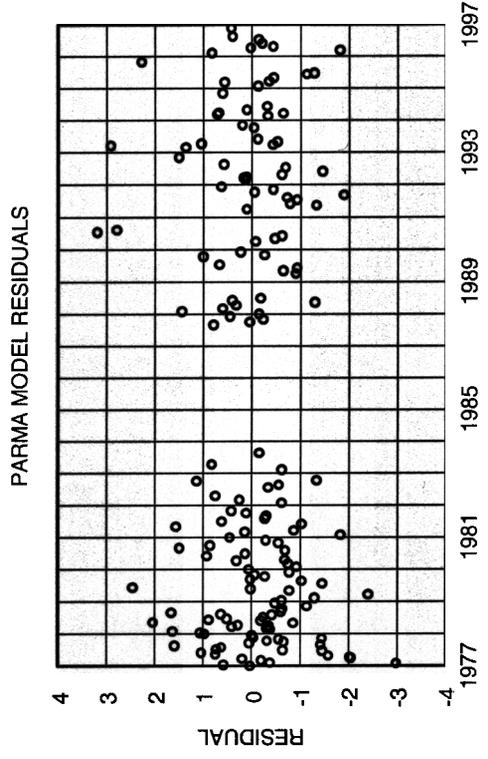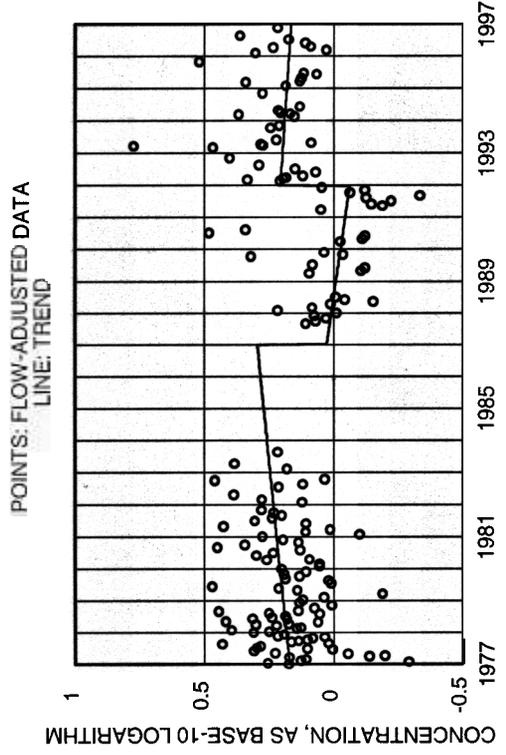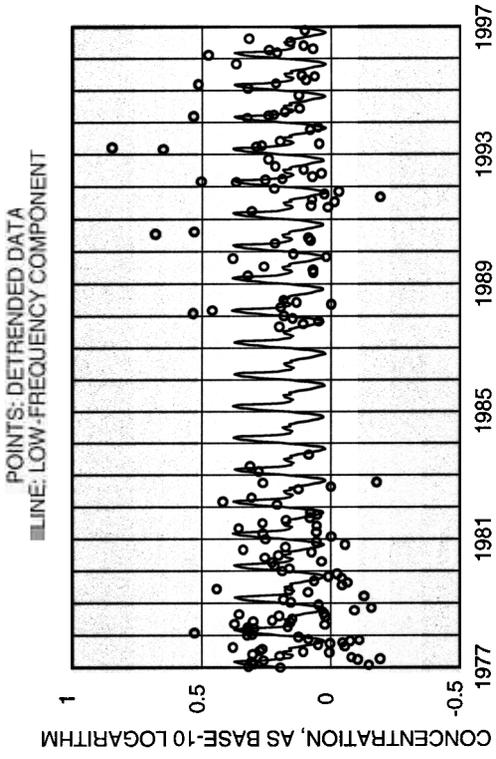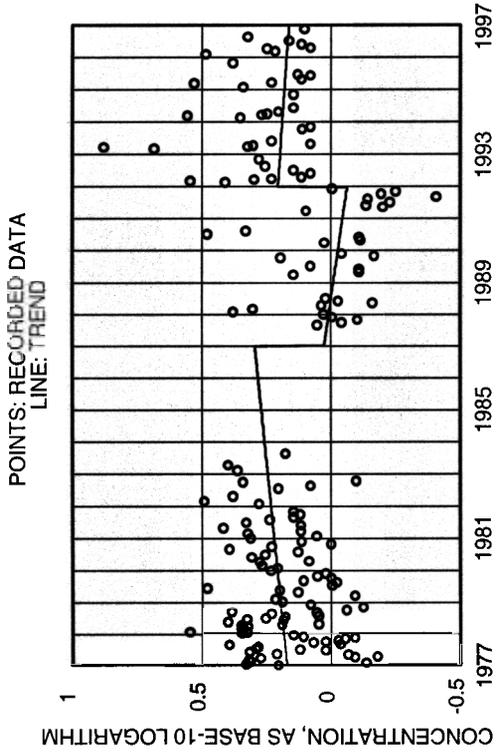
POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

TOTAL PHOSPHORUS, IN MILLIGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL RESIDUALS

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

62
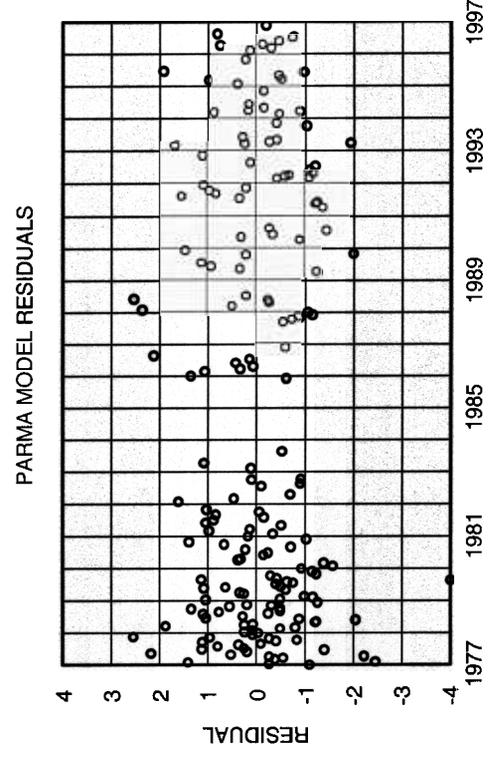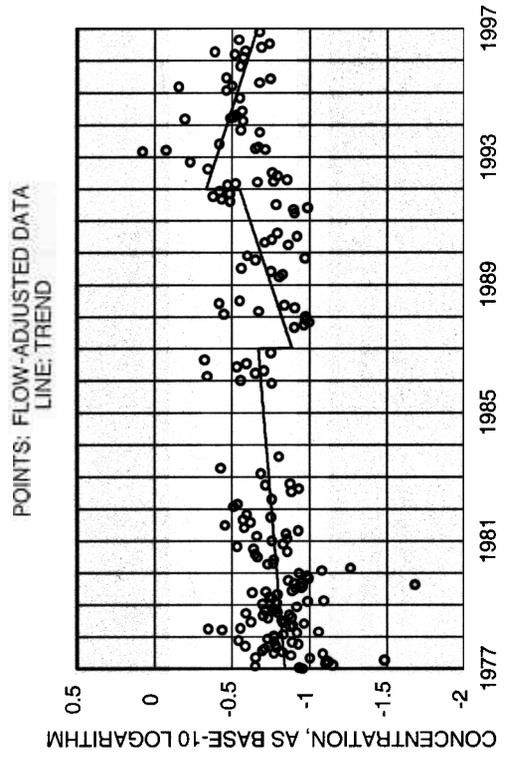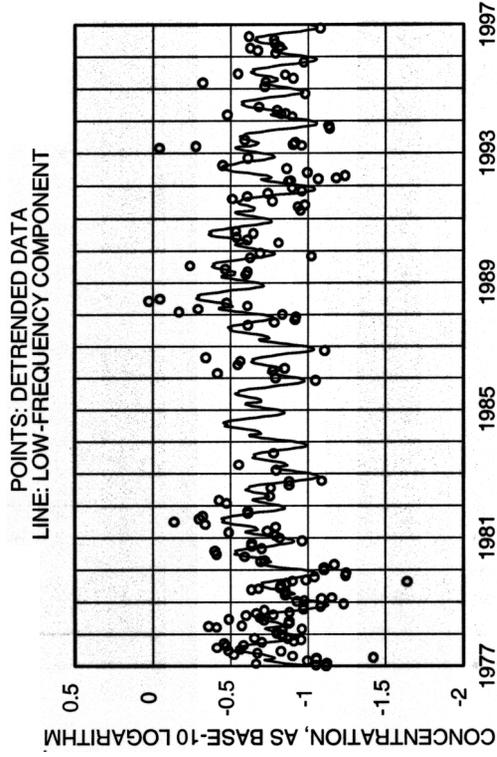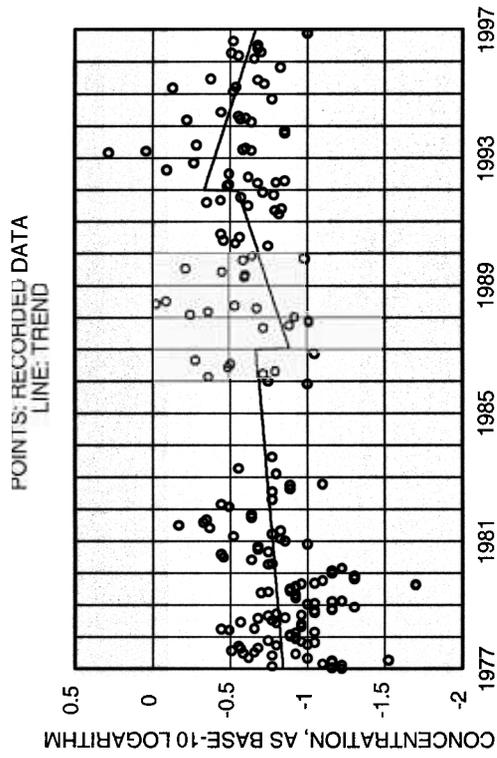
Souris River ° of Glen Ewen, Saskatchewan, and Souris River near S d, ⌐akota

MICROGRAMS PER LITER

POINTS: RECORDED
LINE: TREND

'OTAL AF

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

PARMA MODEL RESIDUALS

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

Souris River South of Glen Ewen, Saskatchewan, and Souris River near Sherwood, North Dakota

DISSOLVED/TOTAL BORON, IN MICROGRAMS PER LITER

POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
TREND

PARMA MODEL RESIDL

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED OXYGEN, IN MILLIGRAMS PER LITER

POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED CALCIUM, IN MILLIGRAMS PER LITER

POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED SODIUM, IN MILLIGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL RESIDUALS

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED SULFATE, IN MILLIGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

PARMA MODEL RESIDUALS

POINTS: RECORDED DATA
LINE: TREND

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED CHLORIDE, IN MILLIGRAMS PER LITER

# Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED SOLIDS, IN MILLIGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL RESIDUALS

POINTS: FLOW-ADJUSTED DATA
LINE: TREND
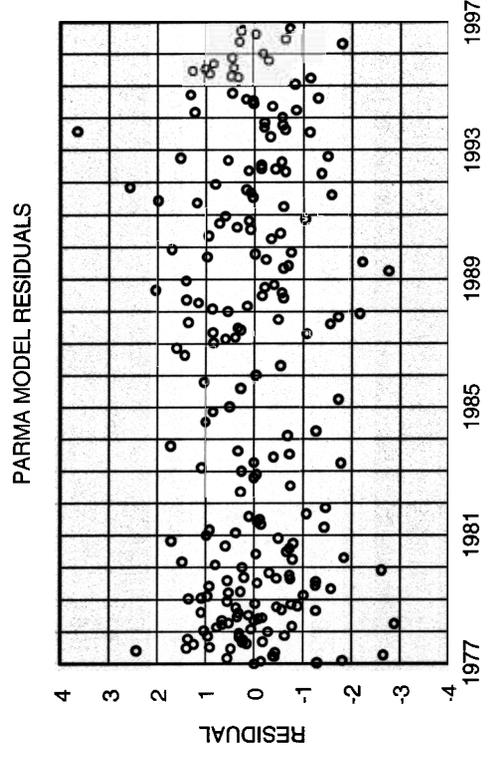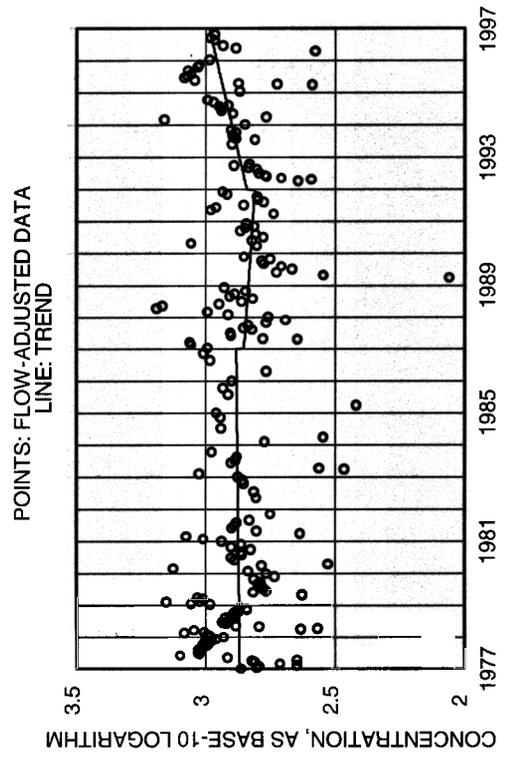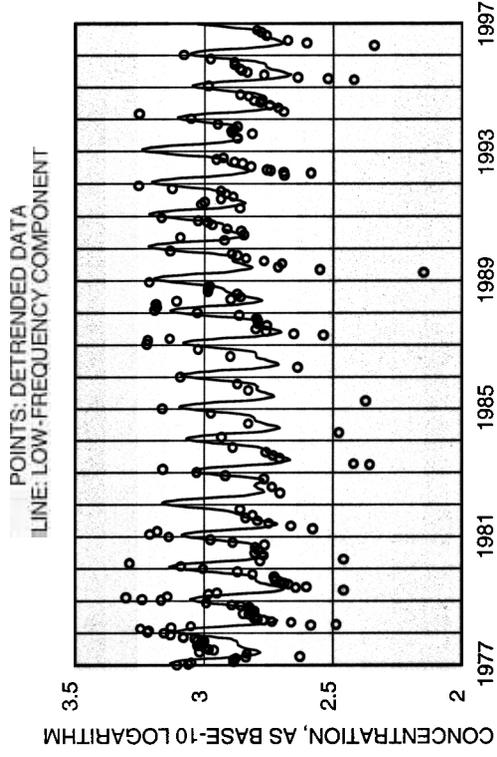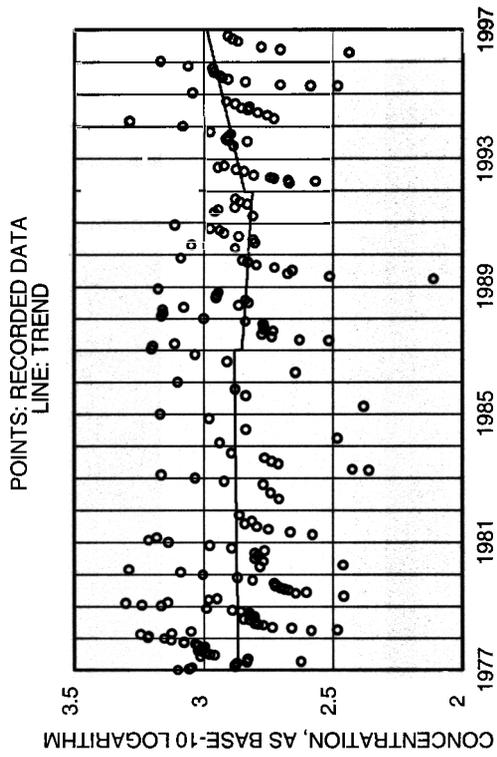
Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

TOTAL AMMONIA, IN MILLIGRAMS PER LITER
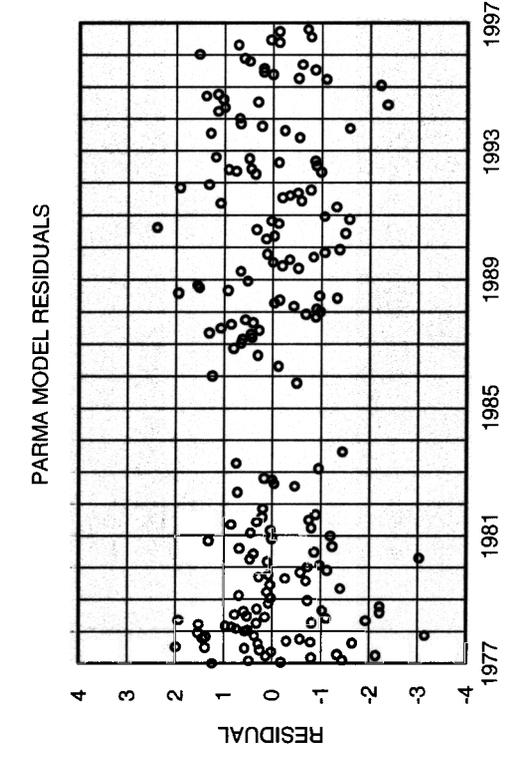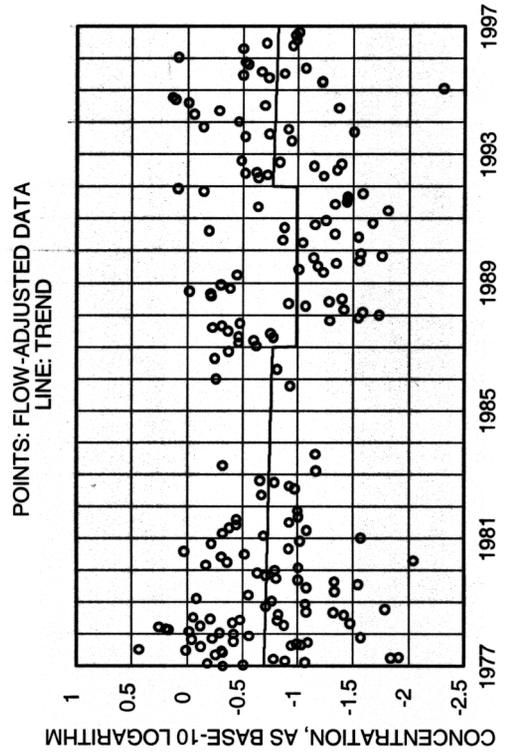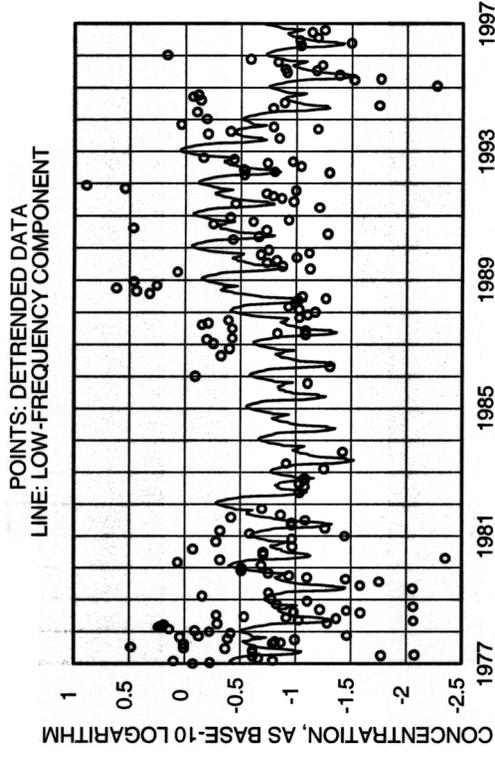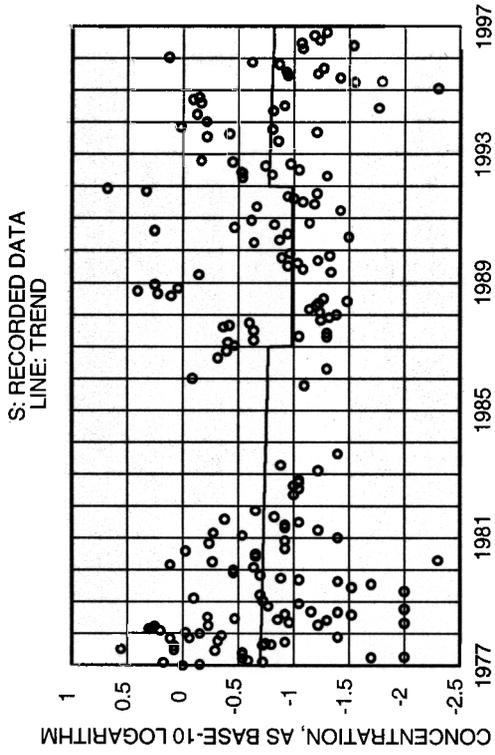
Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

TOTAL NITROGEN, IN MILLIGRAMS PER LITER
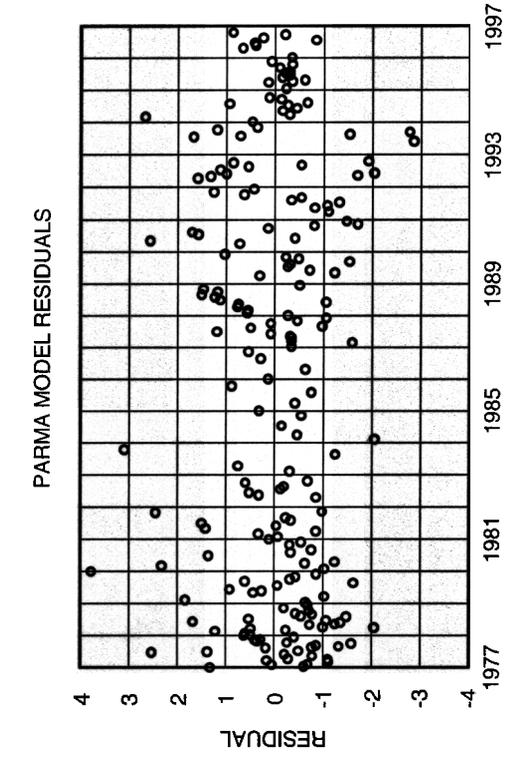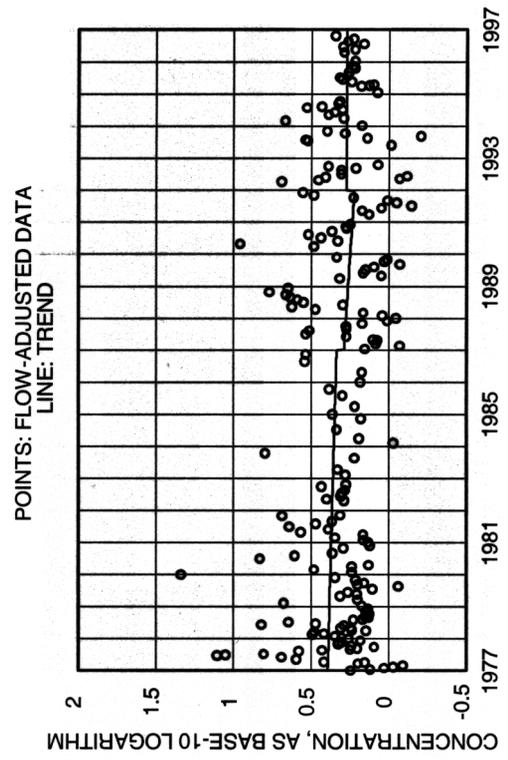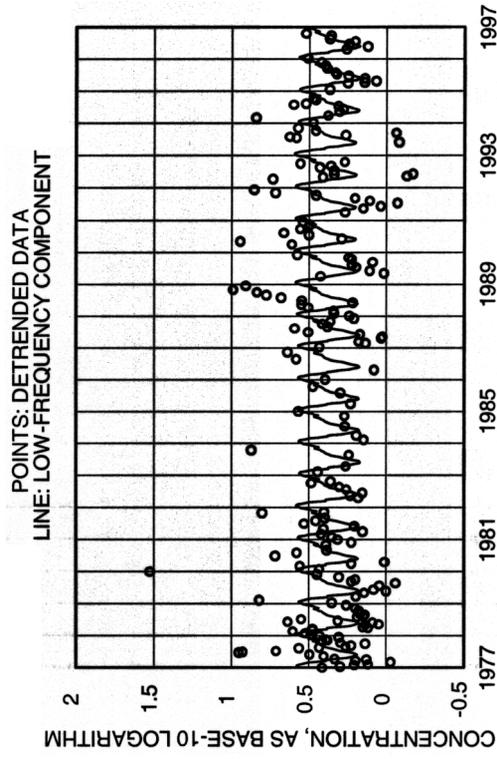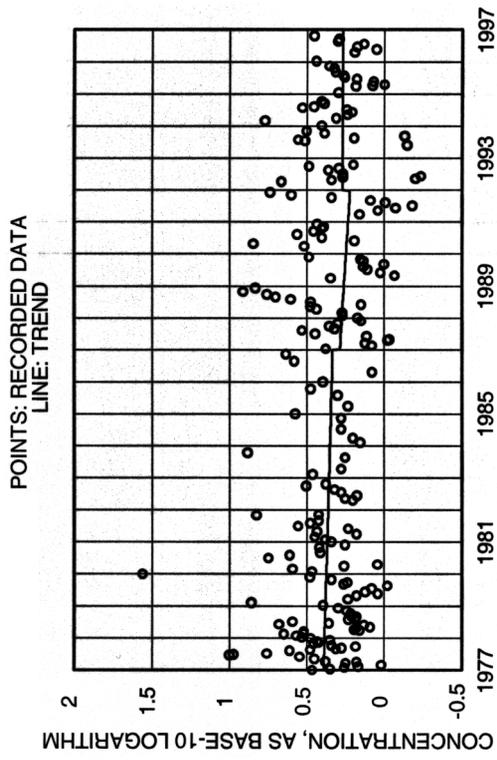
Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

TOTAL PHOSPHORUS, IN MILLIGRAMS PER LITER
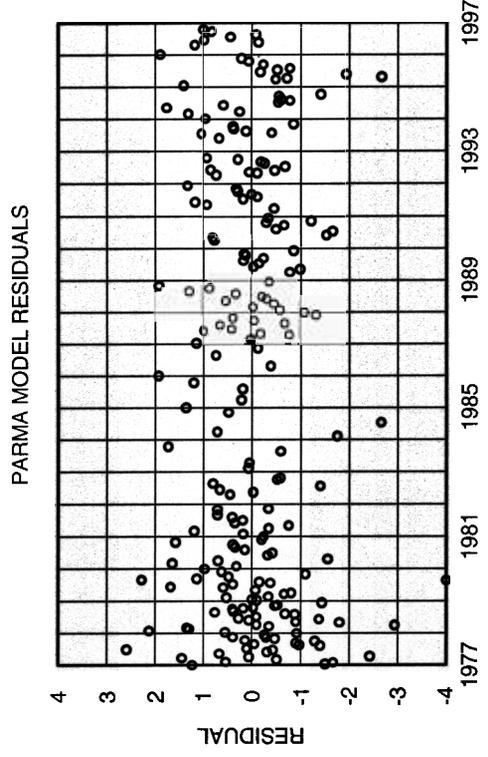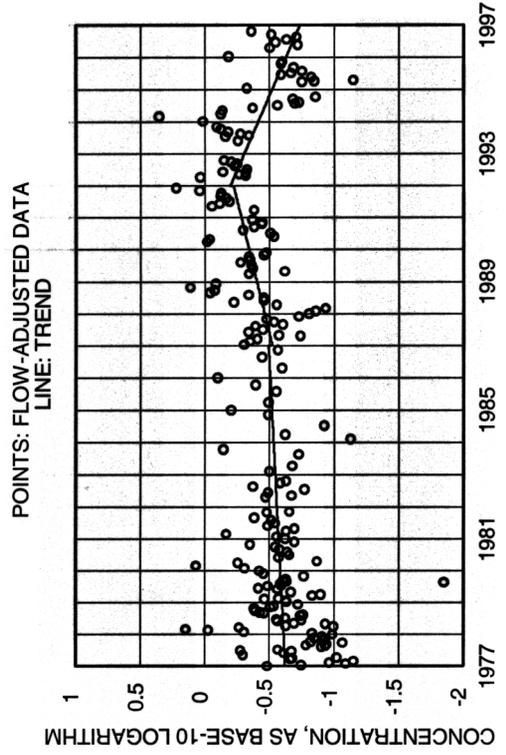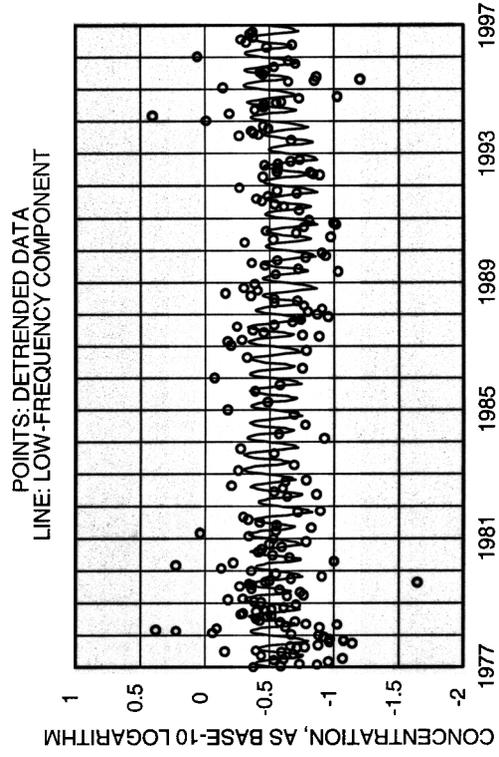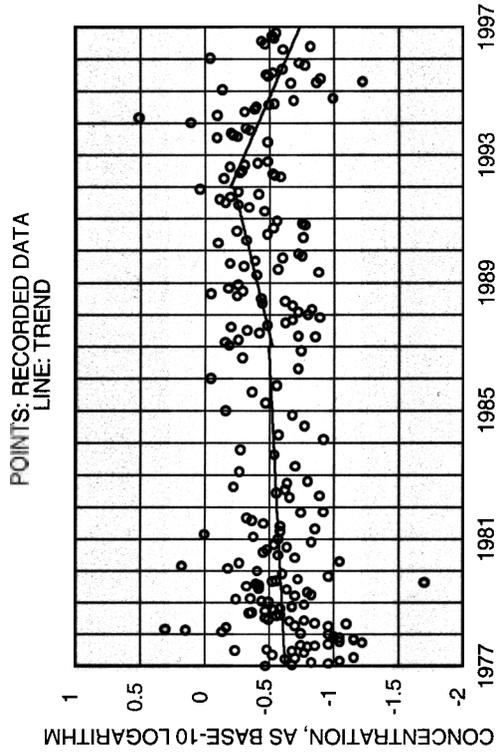
POINTS: RECORDED DATA
LINE: TREND

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

PARMA MODEL RESIDUALS

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED/TOTAL ARSENIC, IN MICROGRAMS PER LITER
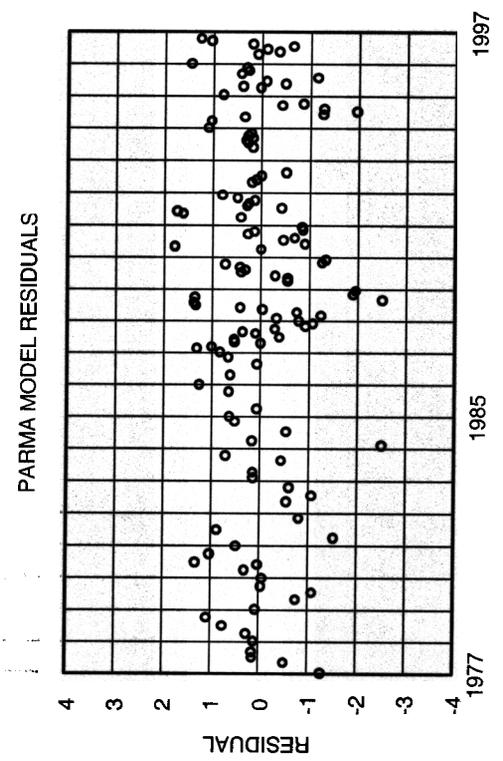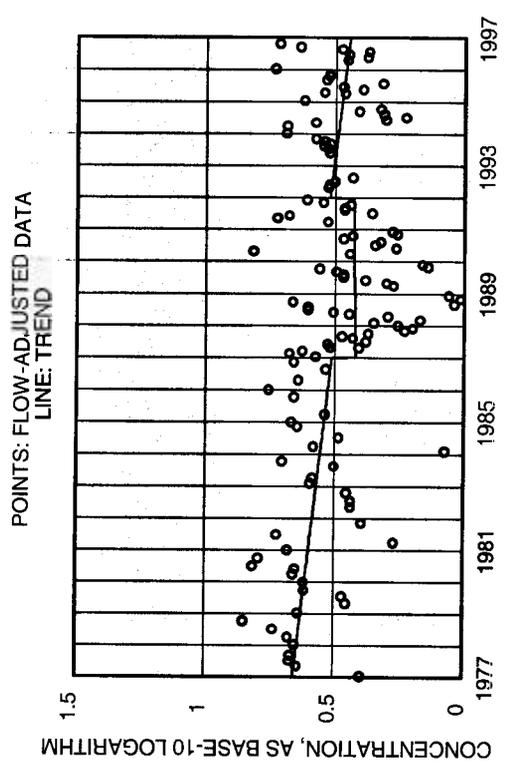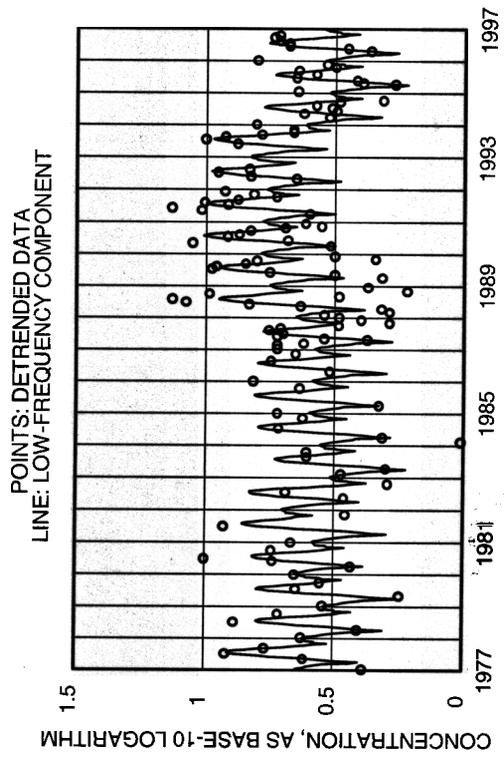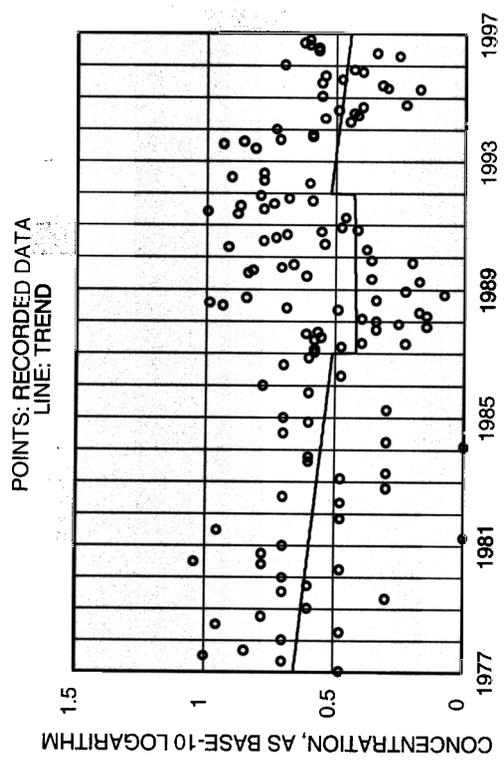
POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL RESIDUALS

POINTS: FLOW-ADJUSTED DATA
LINE: TREND

74

Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

## DISSOLVED/TOTAL BORON, IN MICROGRAMS PER LITER

POINTS: DETRENDED DATA
LINE: LOW-FREQUENCY COMPONENT

POINTS: RECORDED DATA
LINE: TREND

PARMA MODEL RESIDUALS

S: FLOW-ADJUSTED DATA
LINE: TREND
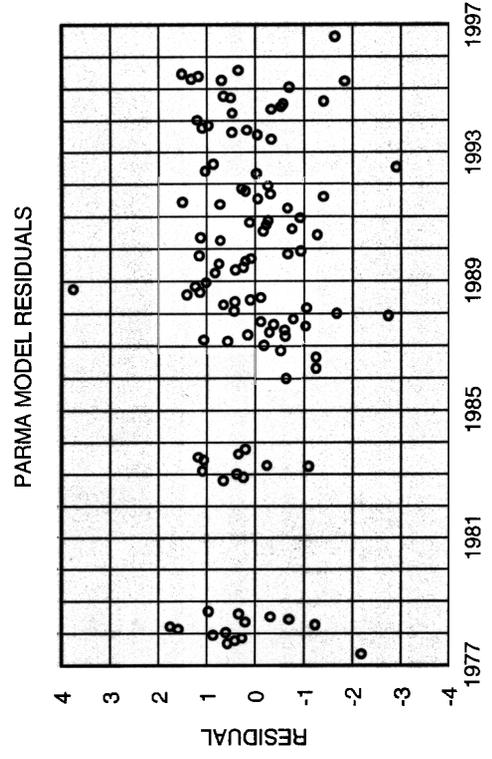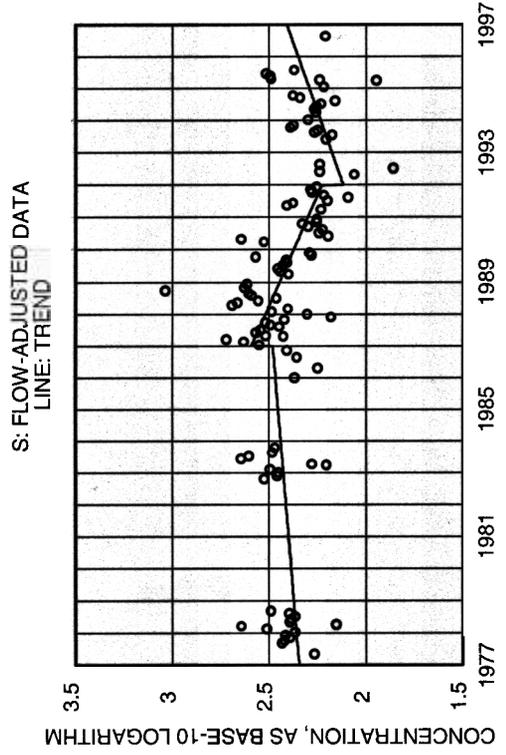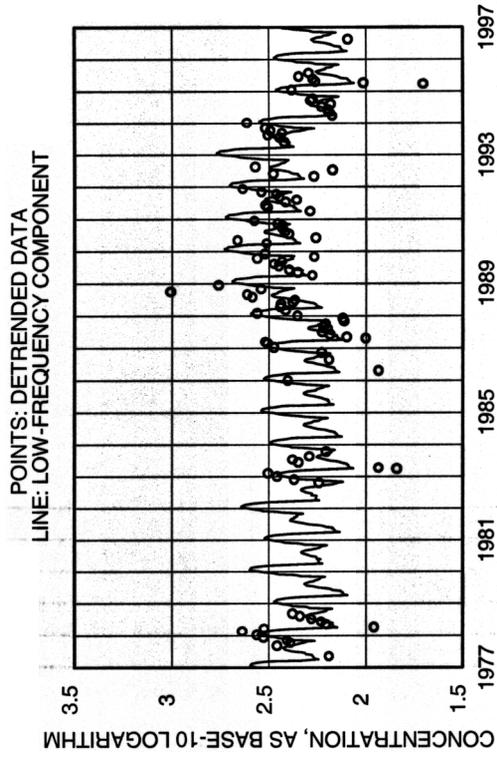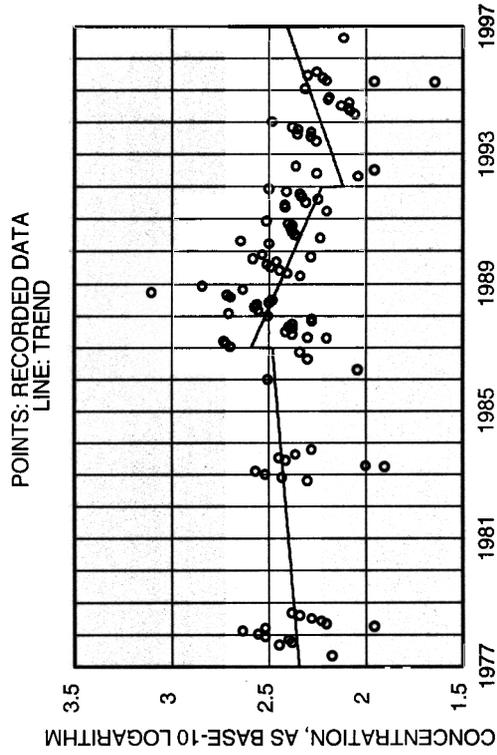
Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED IRON, IN MICROGRAMS PER LITER
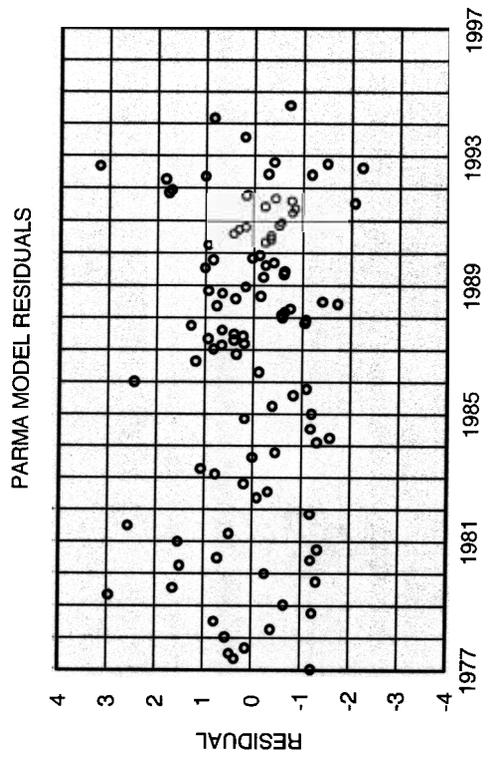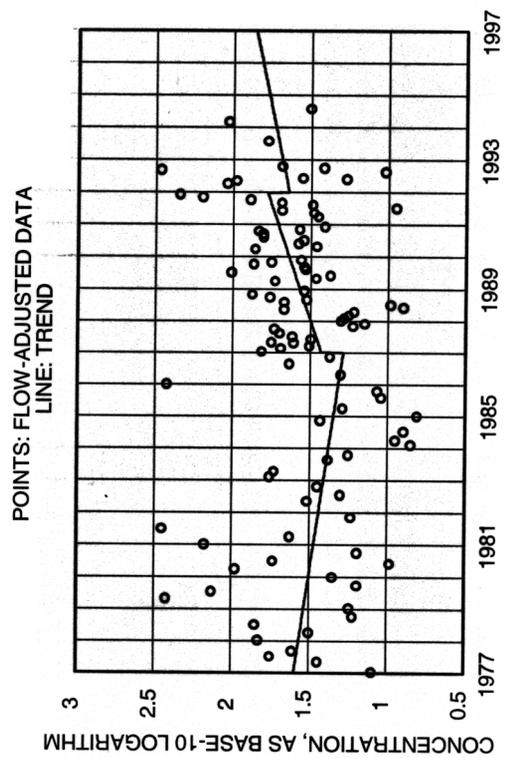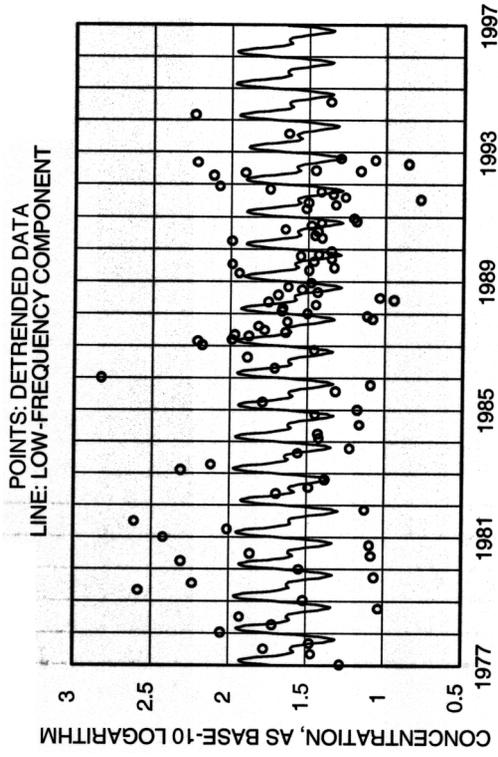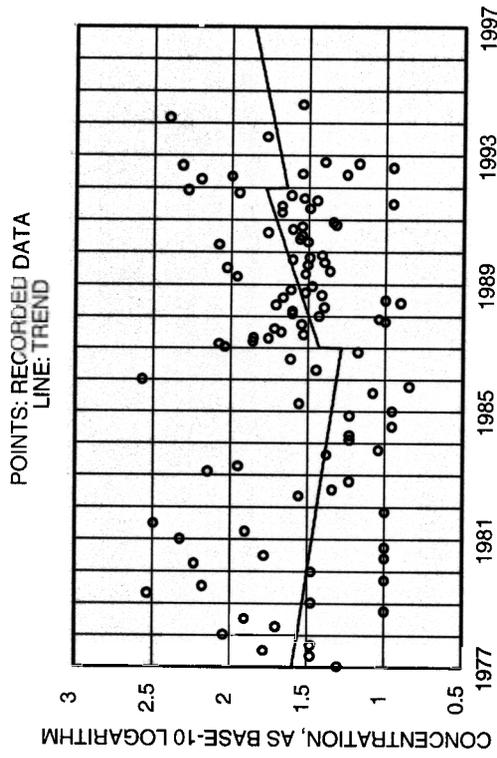
Souris River near Westhope, North Dakota, and Souris River near Coulter, Manitoba

DISSOLVED MANGANESE, IN MICROGRAMS PER LITER