

Prepared in cooperation with the  
Bureau of Reclamation, U.S. Department of the Interior

# **Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05**

Scientific Investigations Report 2006-5241

# **Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05**

By Karen R. Ryberg

Prepared in cooperation with the  
Bureau of Reclamation, U.S. Department of the Interior

Scientific Investigations Report 2006–5241

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

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

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

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

For product and ordering information:

World Wide Web: <http://www.usgs.gov/pubprod>

Telephone: 1-888-ASK-USGS

For more information on the USGS—the Federal source for science about the Earth, its natural and living resources, natural hazards, and the environment:

World Wide Web: <http://www.usgs.gov>

Telephone: 1-888-ASK-USGS

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

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

*Suggested citation:*

Ryberg, K.R., 2006, Continuous water-quality monitoring and regression analysis to estimate constituent concentrations and loads in the Red River of the North, Fargo, North Dakota, 2003-05: U.S. Geological Survey Scientific Investigations Report 2006-5241, 35 p.

# Contents

Abstract.....	1
Introduction.....	1
Purpose and Scope .....	2
Description of Study Area .....	2
Previous Studies .....	2
Methods.....	2
Streamflow Measurements .....	4
Continuous Water-Quality Monitoring.....	4
Manual Water-Quality Measurements, Sample Collection, and Analysis.....	4
Development of Regression Equations to Estimate Constituent Concentrations .....	7
Calculation of Measured and Estimated Constituent Loads .....	11
Results of Regression Analysis .....	12
Alkalinity.....	12
Dissolved Solids .....	13
Sulfate .....	13
Chloride.....	13
Total Nitrite Plus Nitrate .....	17
Total Nitrogen .....	17
Total Phosphorus .....	21
Suspended Sediment.....	21
Measured and Estimated Constituent Loads .....	24
Summary.....	24
References Cited.....	26

## Figures

1. Map showing location of the Fargo gaging station in the Red River of the North Basin, North Dakota.....	3
2. Aerial photograph showing location of gaging station and locations of water-quality sampling sites on the Red River of the North in Fargo, North Dakota .....	5
3–22. Graphs showing:	
3. Streamflow for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005 .....	6
4. Turbidity for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005.....	7
5. Specific conductance for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005 .....	8
6. Comparison of measured and estimated alkalinity concentrations in the Red River of the North at Fargo, North Dakota, 2003-05 .....	14
7. Comparison of measured and estimated dissolved-solids concentrations in the Red River of the North at Fargo, North Dakota, 2003-05 .....	15

8. Comparison of measured and estimated sulfate concentrations in the Red River of the North at Fargo, North Dakota, 2003-05 .....	16
9. Comparison of measured and estimated chloride concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.....	18
10. Comparison of measured and estimated total nitrite plus nitrate concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.....	19
11. Measured total nitrogen concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.....	20
12. Comparison of measured and estimated total nitrogen concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.....	22
13. Comparison of measured and estimated total phosphorus concentrations in the Red River of the North at Fargo, North Dakota, 2003-05 .....	23
14. Comparison of measured and estimated suspended-sediment concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.....	25
15. Comparison of measured and estimated alkalinity loads in the Red River of the North at Fargo, North Dakota, 2003-05.....	28
16. Comparison of measured and estimated dissolved-solids loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	29
17. Comparison of measured and estimated sulfate loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	30
18. Comparison of measured and estimated chloride loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	31
19. Comparison of measured and estimated total nitrite plus nitrate loads in the Red River of the North at Fargo, North Dakota, 2003-05.....	32
20. Comparison of measured and estimated total nitrogen loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	33
21. Comparison of measured and estimated total phosphorus loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	34
22. Comparison of measured and estimated suspended-sediment loads in the Red River of the North at Fargo, North Dakota, 2003-05 .....	35

## Tables

1. Summary statistics for manually measured water-quality physical properties and constituents for the Red River of the North at Fargo, North Dakota, 2003-05 .....	9
2. Constituents estimated using regression equations for the Red River of the North at Fargo, North Dakota, 2003-05 .....	9
3. Possible surrogate physical properties, or explanatory variables, used to develop regression equations to estimate constituent concentrations in the Red River of the North at Fargo, North Dakota, 2003-05 .....	10
4. Conversion factors used in calculation of measured and estimated loads .....	11
5. Regression equations for estimates of alkalinity, dissolved solids, chloride, sulfate, total nitrogen, total nitrite plus nitrate, total phosphorus, and suspended sediment in the Red River of the North at Fargo, North Dakota, 2003-05 .....	12

## Conversion Factors, Abbreviations, and Datum

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
yard (yd)	0.9144	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi <sup>2</sup> )	259.0	hectare (ha)
square mile (mi <sup>2</sup> )	2.590	square kilometer (km <sup>2</sup> )
Flow rate		
cubic foot per second (ft <sup>3</sup> /s)	0.02832	cubic meter per second (m <sup>3</sup> /s)
Mass		
pound, avoirdupois (lb)	0.4536	kilogram (kg)

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

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

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

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

A nanometer is one millionth of a millimeter.

Specific conductance is given in microsiemens per centimeter at 25 degrees Celsius ( $\mu\text{S}/\text{cm}$  at 25°C).

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

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

# Continuous Water-Quality Monitoring and Regression Analysis to Estimate Constituent Concentrations and Loads in the Red River of the North, Fargo, North Dakota, 2003-05

By Karen R. Ryberg

## Abstract

This report presents the results of a study by the U.S. Geological Survey, done in cooperation with the Bureau of Reclamation, U.S. Department of the Interior, to estimate water-quality constituent concentrations in the Red River of the North at Fargo, North Dakota. Regression analysis of water-quality data collected in 2003-05 was used to estimate concentrations and loads for alkalinity, dissolved solids, sulfate, chloride, total nitrite plus nitrate, total nitrogen, total phosphorus, and suspended sediment. The explanatory variables examined for regression relation were continuously monitored physical properties of water—streamflow, specific conductance, pH, water temperature, turbidity, and dissolved oxygen. For the conditions observed in 2003-05, streamflow was a significant explanatory variable for all estimated constituents except dissolved solids. pH, water temperature, and dissolved oxygen were not statistically significant explanatory variables for any of the constituents in this study. Specific conductance was a significant explanatory variable for alkalinity, dissolved solids, sulfate, and chloride. Turbidity was a significant explanatory variable for total phosphorus and suspended sediment. For the nutrients, total nitrite plus nitrate, total nitrogen, and total phosphorus, cosine and sine functions of time also were used to explain the seasonality in constituent concentrations.

The regression equations were evaluated using common measures of variability, including  $R^2$ , or the proportion of variability in the estimated constituent explained by the regression equation.  $R^2$  values ranged from 0.703 for total nitrogen concentration to 0.990 for dissolved-solids concentration. The regression equations also were evaluated by calculating the median relative percentage difference (*RPD*) between measured constituent concentration and the constituent concentration estimated by the regression equations. Median *RPDs* ranged from 1.1 for dissolved solids to 35.2 for total nitrite plus nitrate.

Regression equations also were used to estimate daily constituent loads. Load estimates can be used by water-quality

managers for comparison of current water-quality conditions to water-quality standards expressed as total maximum daily loads (TMDLs). TMDLs are a measure of the maximum amount of chemical constituents that a water body can receive and still meet established water-quality standards. The peak loads generally occurred in June and July when streamflow also peaked.

## Introduction

The U.S. Geological Survey (USGS) gaging station (05054000) at Fargo, North Dakota, was established in June 1901 to provide stage (gage-height) and streamflow values for the Red River of the North (Red River). The USGS began collecting water-quality samples at the gaging station in 1949. Since 1994, the gaging station has been equipped with a data-collection platform that uses satellite technology to transmit real-time stage and streamflow data 24 hours a day. In 1998, a continuous water-quality monitor began recording specific conductance (an indicator of dissolved chemical constituents in the water and the ability of water to conduct an electrical current) and water temperature. Continuous measurements of pH (a measure of hydrogen-ion activity), turbidity (an indicator of the amount of suspended sediment in the water), and dissolved oxygen concentration (an indicator of the balance between oxygen-consuming and oxygen-producing processes) began in October 2003. An acoustic doppler velocity meter (ADVM) was installed at the gaging station in 2004. The ADVM measures an instream index velocity that is related to mean channel velocity. Mean channel velocity, in conjunction with stage, allows for more accurate computation of real-time streamflows. For the study described in this report, the data collected at the USGS gaging station at Fargo during 2003-05 were analyzed in cooperation with the Bureau of Reclamation, U.S. Department of the Interior, to determine whether streamflow and the continuously recorded physical properties could act as surrogates to estimate water-quality constituents that are important indicators of surface-water quality. The methods

used in this study replicate those used in Kansas and may be replicated for other sites in North Dakota and the Nation to monitor water quality.

Estimation of water-quality constituents on the basis of surrogates provides several benefits. Although periodic Red River water samples are collected manually and analyzed, the delay associated with laboratory analysis does not permit immediate identification of undesirable levels of constituents. A relation between manually collected water-quality samples and real-time water-quality measurements allows immediate identification of potential water-quality problems. Examination of streamflow and physical properties of water that act as surrogates for constituents of interest also helps optimize visits for the collection of water-quality samples. For example, if turbidity is used as a surrogate for estimating suspended sediment, it is desirable to collect water-quality samples representative of the range of turbidity values possible for the Red River. In addition, the estimation of real-time water-quality information is useful in tracking changes in water quality as they occur. This information is important to North Dakota and Minnesota as well as Canada in assessing efforts to improve water quality in the upper Red River Basin. Availability of real-time information also could be used to assist the cities of Fargo, North Dakota, and Moorhead, Minnesota, in managing wastewater operations and in providing information that would allow citizens to assess the condition of the Red River relative to intended uses. Finally, real-time estimates of constituent concentrations that are based on surrogates monitored at the gaging station could more accurately estimate loads for comparison to total maximum daily loads (TMDLs). A TMDL is a calculation of the maximum amount of a constituent that a water body can receive and still meet water-quality standards. Section 303(d) of the 1972 Clean Water Act (U.S. Environmental Protection Agency, 2006) requires States to establish TMDLs.

### Purpose and Scope

This report presents the results of regression analysis of water-quality constituents for the Red River at Fargo, North Dakota, using manually collected water-quality data and continuously recorded streamflow and water-quality data from the USGS gaging station (05054000) at Fargo for 2003-05. The report provides regression equations that can be used to estimate concentrations and loads for alkalinity, dissolved solids, sulfate, chloride, total nitrite plus nitrate, total nitrogen, total phosphorus, and suspended sediment.

### Description of Study Area

The Red River (fig. 1) is formed by the confluence of the Bois de Sioux and Otter Tail Rivers at Wahpeton, North Dakota, and Breckenridge, Minnesota. The river then flows northward 394 mi to the international boundary and continues into Manitoba, Canada. The river from its confluence north to

the international boundary separates North Dakota and Minnesota. The drainage area of the Red River at Fargo is about 6,800 mi<sup>2</sup>, and the bed of the river slopes only about 200 ft over the reach between Wahpeton and the international boundary.

The Red River is a valuable resource and provides municipal and industrial water supplies, recreational opportunities, and wildlife habitat. The primary land use in the area is agriculture (Tornes, 2005). The July 1, 2003, population estimate for the city of Fargo was 91,048 and across the river the population estimate for Moorhead, Minnesota, was 32,849 (U.S. Census Bureau, 2006). Cass County, the county in which Fargo is located, had an estimated population of 126,769 in 2003 or 20 percent of North Dakota's total population, 633,400 (U.S. Census Bureau, 2006).

The Red River flows over lacustrine deposits of glacial Lake Agassiz through its entire length in North Dakota (Williams-Sether, 2004). Water quality in the Red River is affected by the lacustrine deposits, inflow from major tributaries in North Dakota and Minnesota, runoff from agricultural areas, ground-water discharge, industrial effluents, and wastewater discharges from cities along the river, including Fargo and Moorhead (Williams-Sether, 2004). Management strategies to control point- and nonpoint-source pollution from both agricultural and urban land uses are important issues in the Red River Basin, and the water quality of the Basin is an issue of international importance.

### Previous Studies

Many aspects of water quality in the Red River Basin have been studied. A bibliography of USGS reports related to the Red River Basin is available online at <http://nd.water.usgs.gov/pubs/key/redriver.html>. Notable among previous water-quality reports is Williams-Sether (2004), in which regression equations for estimating monthly mean concentrations of selected water-quality constituents were developed. Williams-Sether (2004) precedes continuous water-quality monitoring at the Fargo gaging station; therefore, the only continuously recorded variable available for regression analysis was streamflow. Regression analysis to estimate constituent concentrations and loads was used by Christensen and others (2000) for the Little Arkansas River in south-central Kansas and by Christensen (2001) for the Quivira National Wildlife Refuge, south-central Kansas.

### Methods

This section describes how streamflow was measured, water-quality samples were collected and analyzed, and data were interpreted for this study. "Although technological advances in turbidity measurement have produced a variety of instrument types to meet one or more ... differing objectives, turbidity instruments of different designs commonly do not



Figure 1. Location of the Fargo gaging station in the Red River of the North Basin, North Dakota.

yield identical or equivalent results” (Wilde, variously dated); therefore special attention is given to the equipment used to continuously record turbidity and to manually measure turbidity.

## Streamflow Measurements

Stage, or gage height, was measured to the nearest 0.01 ft at the Fargo gaging station with a nonsubmersible, pressure transducer. The data were electronically recorded and transmitted by satellite to a downlink site and then to the USGS office in Bismarck, North Dakota.

Methods used to determine streamflow are described in Buchanan and Somers (1969). Streamflow measurements were made approximately 15 times per year (S. Robinson, USGS, oral commun., 2006). A stage-streamflow relation was developed on the basis of streamflow measurements and the stage of the stream at the time of measurement (Kennedy, 1984) and this relation was used to compute a continuous record of streamflow (Kennedy, 1983). Stage and streamflow values were stored in the USGS National Water Information System (NWIS) database in 1-hour increments.

## Continuous Water-Quality Monitoring

Continuous monitoring of specific conductance and water temperature began at the Fargo gaging station in September 1998 using a Campbell Scientific CR247 conductivity and temperature sensor. In October 2003, a YSI Environmental YSI 6820 SONDE continuously recording water-quality monitor was installed at the gaging station. This water-quality monitor measures specific conductance, pH, water temperature, turbidity, and dissolved oxygen concentration. Initially, the monitor was installed inside the Fargo water-treatment pumping plant. In November 2004, the water-quality monitor was relocated to an in-stream location near the water-treatment plant intake structure. Specific conductance, pH, and turbidity values obtained at both locations were similar; however, water temperature and dissolved oxygen values differed between the two locations.

A YSI 6920 SONDE was used as a backup water-quality monitor, and beginning in 2005, the two models were used alternately to allow them to be brought back to the USGS field office for cleaning. A comparison of models 6820 and 6920 indicates that the two instruments are the same in terms of measurement range, resolution, and accuracy (YSI Environmental, 2001). The YSI 6920 model has an internal power supply, whereas the YSI 6820 does not. The sensors were calibrated according to methods presented in Wilde (variously dated).

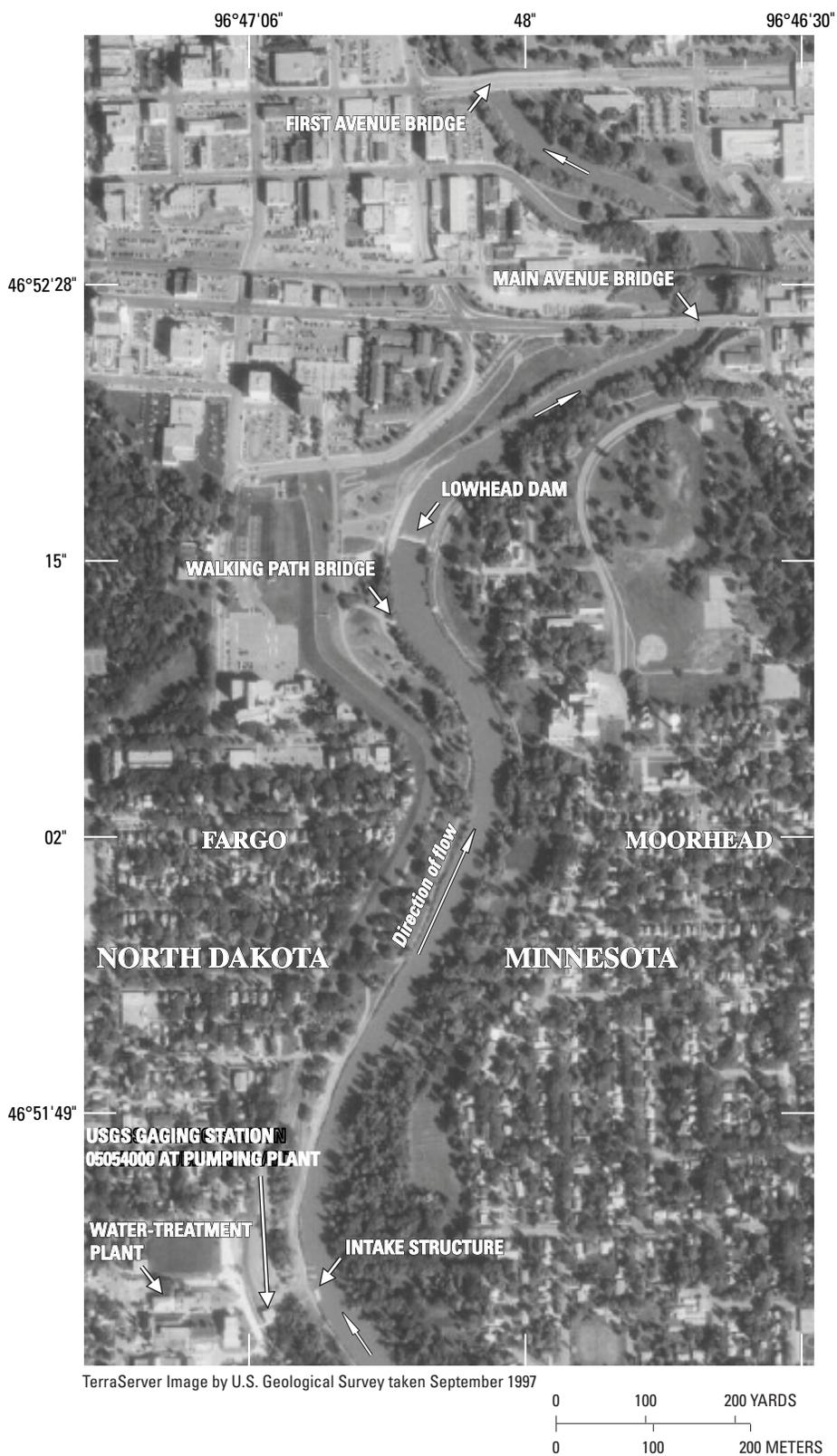
## Manual Water-Quality Measurements, Sample Collection, and Analysis

Water-quality samples were collected manually according to methods described in U.S. Geological Survey techniques manuals (variously dated). All samples were collected downstream (north) of the Red River at Fargo gaging station (05054000) (fig. 2) because it was difficult to access the river by boat at the gaging station. In 2003, samples were collected from the Main Avenue Bridge or the First Avenue Bridge. In 2004 and 2005, samples were collected from the First Avenue Bridge during high flow and from the walking-path bridge, between the gaging station and the Main Avenue Bridge, during low flow. Waste water from the cities of Fargo and Moorhead is returned to the Red River downstream from the water-quality sampling locations. Water-quality samples were analyzed by the North Dakota Department of Health Laboratory using published methods (North Dakota Department of Health, 2003b) and quality assurance procedures (North Dakota Department of Health, 2003a). Sediment samples were analyzed by the USGS Iowa Sediment Laboratory using procedures described in Guy (1969) and Matthes and others (1992). During the collection of water-quality samples, physical properties of water were measured, including specific conductance, pH, temperature, turbidity, and dissolved oxygen. In 2003, the physical properties also were measured at the gaging station. The comparison of the two sets of measurements verified that river conditions were similar at the water-quality sampling location and the gaging station.

Because hydrologic conditions within a watershed may affect variability of many chemical constituents, water-quality samples were collected manually throughout a range of streamflow conditions. Streamflow from January 1, 2003, to December 31, 2005, and the days and streamflow values for which water-quality samples were collected and physical properties were manually measured are shown in figure 3. It is important to note that water-quality samples for this study were not collected under ice conditions during the winter.

Efforts also were made to collect water-quality samples throughout a range of turbidity values because turbidity was anticipated to be an important surrogate variable for this study. Manually measured and continuously recorded turbidity data are shown in figure 4. The manually measured turbidity values and dates correspond with the collection of water-quality samples.

Turbidity was manually measured using two different instruments. In 2003, manual turbidity measurements were obtained using a Hach Environmental Hydrolab DataSonde 3. The DataSonde 3 method of determining turbidity was the USGS-approved method ISO 7027, the European



**Figure 2.** Aerial photograph showing location of gaging station and water-quality sampling sites on the Red River of the North in Fargo, North Dakota.

## 6 Continuous Water-Quality Monitoring and Regression Analysis, Red River of the North, 2003-05

drinking-water protocol (International Organization for Standardization, 1999). In 2004 and 2005 the manual turbidity measurements were made by removing the YSI 6820 or YSI 6920 water-quality monitor (which also uses ISO method 7027) from the gaging station, using it to make manual measurements, and then placing the monitor back in the gage.

Manual and continuously recorded turbidity measurements for 2003 and 2004 were stored in the USGS NWIS database under parameter code P61028 (turbidity, water, unfiltered, field, nephelometric turbidity units). Effective October 1, 2004, a change was made in USGS turbidity data-reporting procedures. The new procedures were a result of USGS collaboration with public and private sector entities and established “a suite of reporting units for storing and reporting turbidity data that is based on instrument design” (U.S.

Geological Survey, 2004). Manual and continuously recorded turbidity measurements made using these new procedures were stored under parameter code P63680 (turbidity, water, monochrome near infra-red LED light, 780-900 nm, detection angle 90 +/- 2.5 degrees, FNU). FNU is the acronym for the formazin nephelometric unit of measurement and, under the new procedures, is used with instruments that comply with ISO method 7027. Despite the differences in monitor, parameter code, and units of measurement, the turbidity measurements were made using the same method. Therefore, the manual and continuously recorded turbidity values for 2003 and 2004 were changed in the USGS database to parameter code P63680, and all turbidity data were treated as equivalent. Turbidity units are referred to as FNU throughout the remainder of the report.

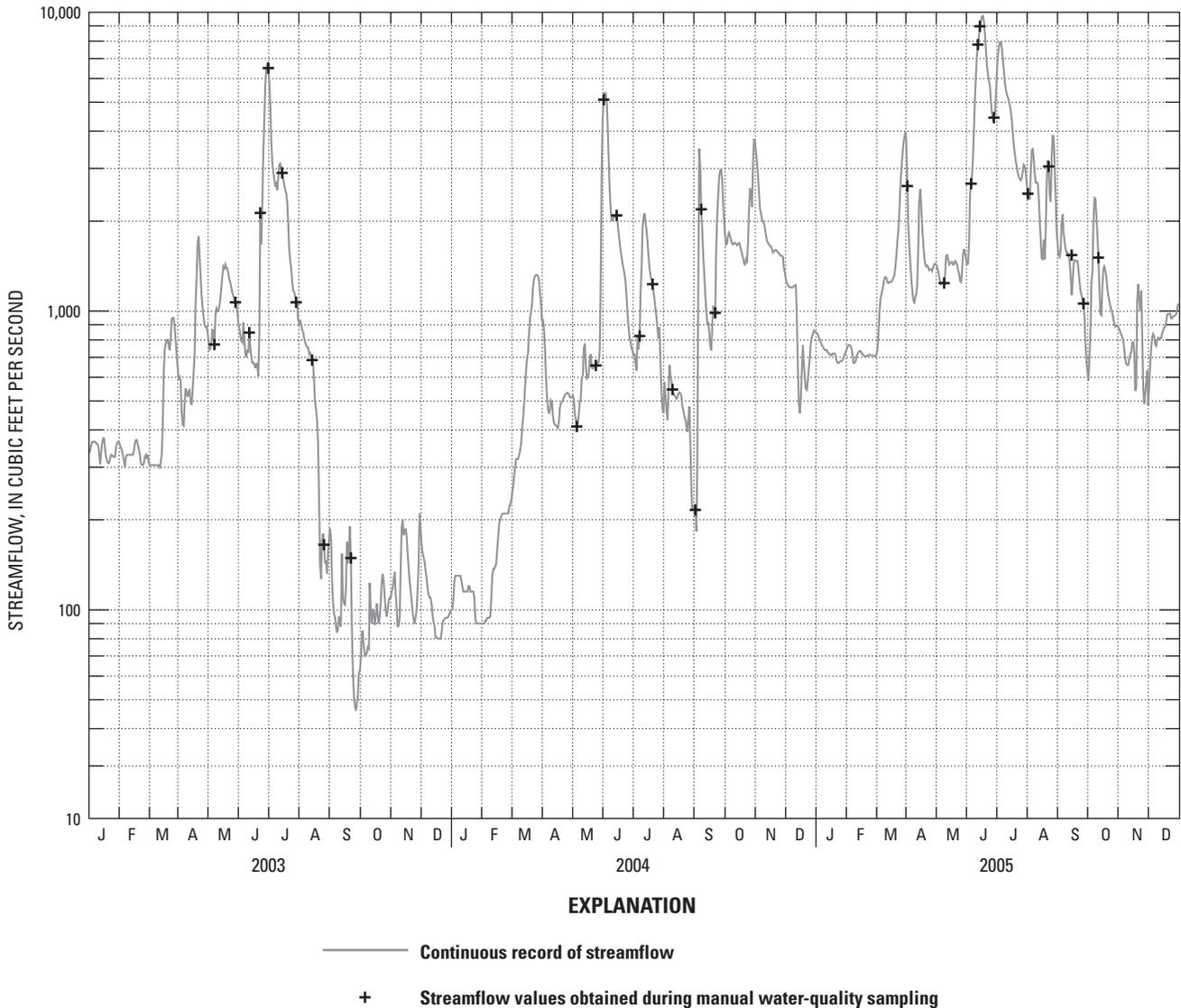


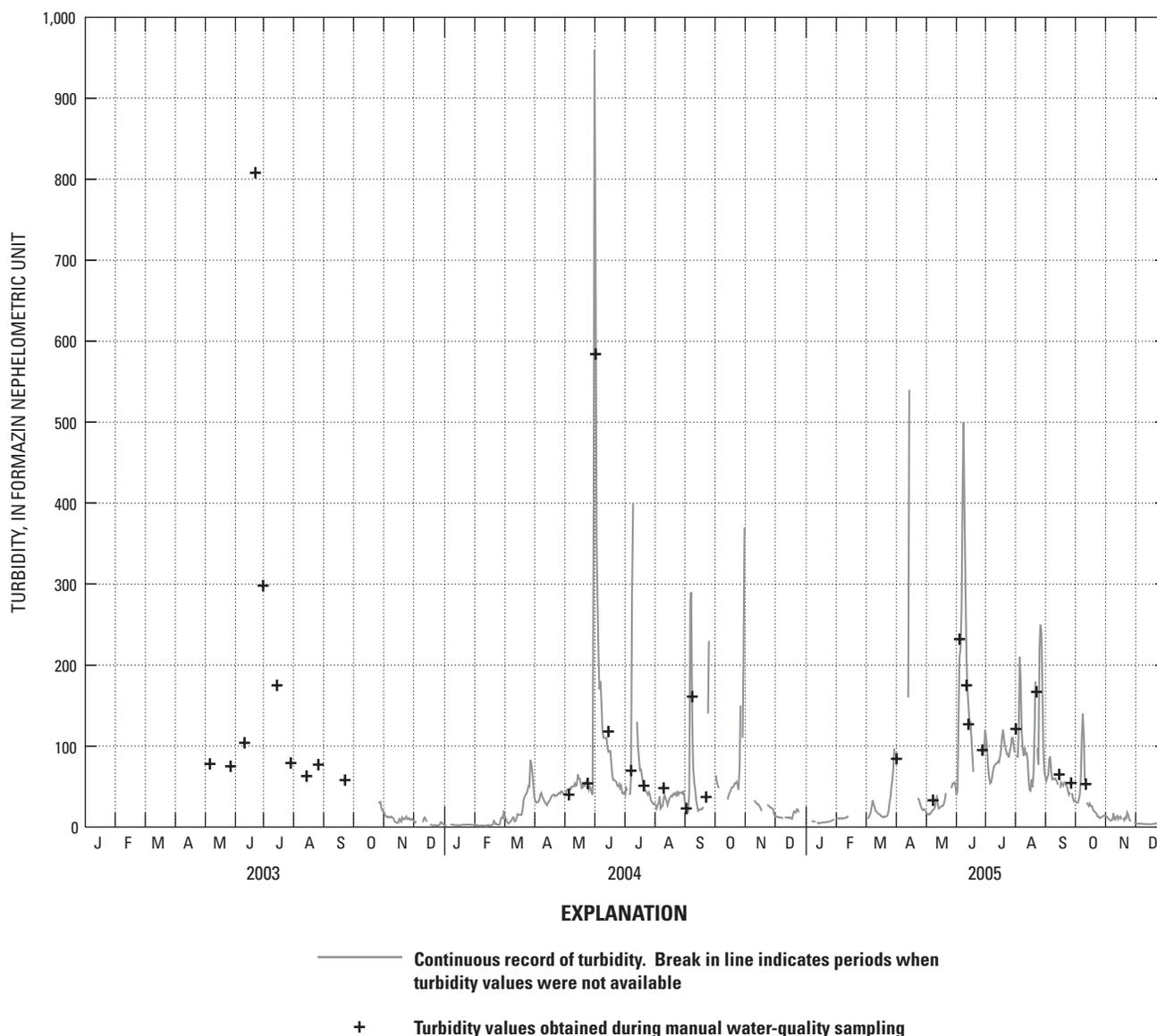
Figure 3. Streamflow for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005.

Manual and continuously recorded specific conductance data collected are shown in figure 5. The manual specific conductance values and dates correspond to the collection of water-quality samples, which were made over a broad range of specific conductance values.

Summary statistics of the manually measured physical properties and water-quality constituents are listed in table 1. The data are stored in the USGS NWIS database and are available online at [http://nwis.waterdata.usgs.gov/nd/nwis/qwdata?site\\_no=05054000](http://nwis.waterdata.usgs.gov/nd/nwis/qwdata?site_no=05054000).

### Development of Regression Equations to Estimate Constituent Concentrations

The water-quality constituents selected for estimation are listed in table 2. The rationale and methodology for expressing water-quality constituent concentrations in terms of other surrogate constituents or physical properties in a regression equation are explained in Helsel and Hirsch (1995). Helsel and Hirsch (1995) also detail the computations for regression estimation and identify measures commonly used to evaluate regression equations, including mean square error, standard deviation, and  $R^2$ . They explain the benefit of log transformation of variables, which was a technique used for some of the regression relations in this study.



**Figure 4.** Turbidity for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005.

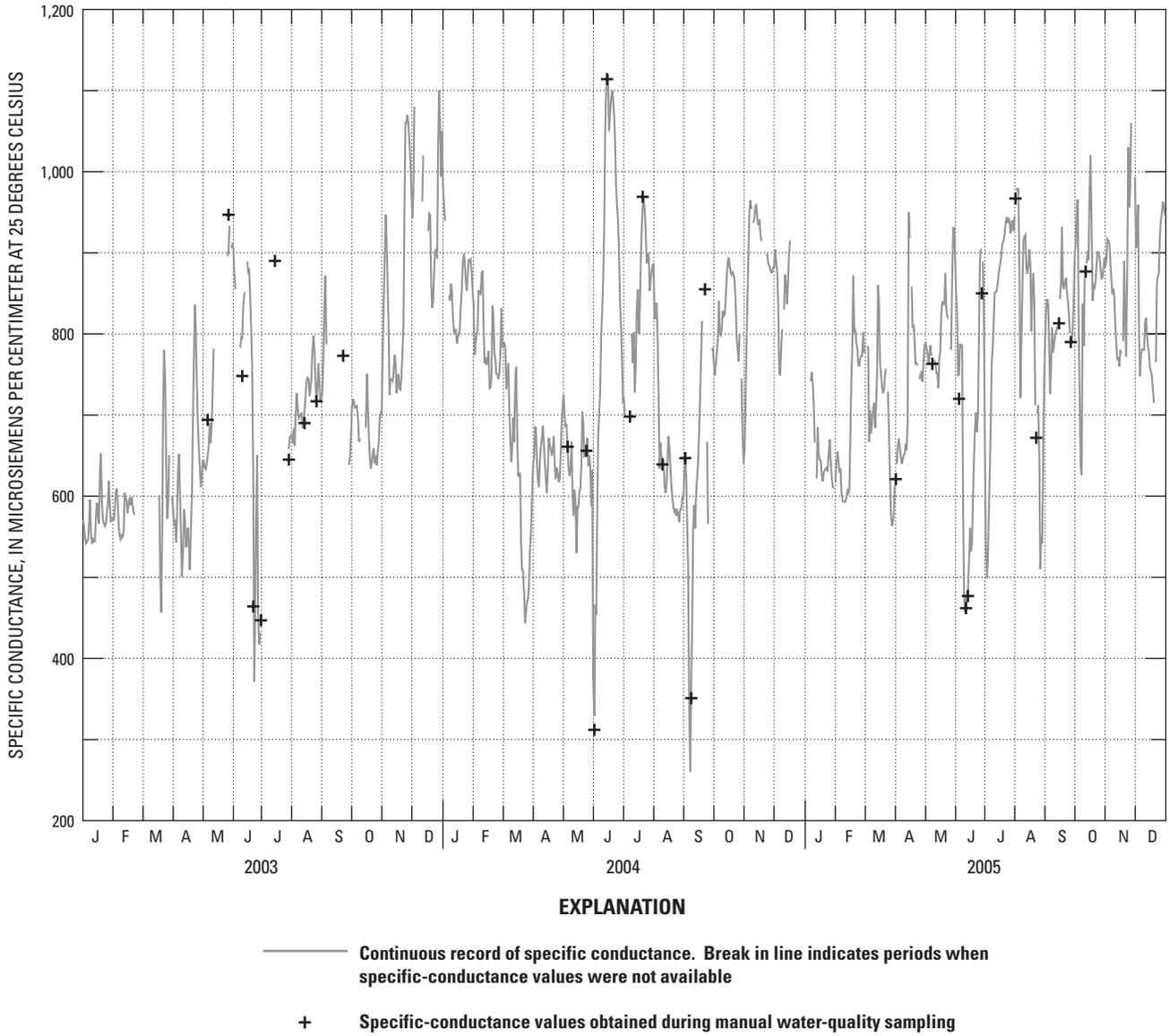


Figure 5. Specific conductance for the Red River of the North at Fargo, North Dakota, January 1, 2003, to December 31, 2005.

**Table 1.** Summary statistics for manually measured water-quality physical properties and constituents for the Red River of the North at Fargo, North Dakota, 2003-05.

[Streamflow in cubic feet per second, ft<sup>3</sup>/s; specific conductance in microsiemens per centimeter at 25 degrees Celsius,  $\mu$ S/cm at 25°C; pH in standard units; temperature in degrees Celsius, °C; turbidity in formazin nephelometric units, FNU; other physical properties and constituent concentrations in milligrams per liter, mg/L]

Physical property or constituent	Descriptive statistics				Percent of samples in which values were less than or equal to those shown				
	Sample size	Maximum	Minimum	Mean	0.95	0.75	(Median) 50 percent	0.25	0.05
Streamflow	31	9,010	165	2,190	7,170	2,720	1,230	812	193
Specific conductance	31	1,110	312	707	968	832	698	642	399
pH	31	8.50	7.30	8.00	8.35	8.20	8.00	7.80	7.53
Temperature, water	31	27.2	3.1	19.1	25.9	21.4	20.0	15.9	11.3
Turbidity	31	808.0	22.9	136.0	441.0	144.0	78.0	54.1	35.0
Dissolved oxygen	31	11.5	4.7	7.5	10.2	8.4	7.4	6.4	5.1
Alkalinity	13	239	127	192	238	228	209	162	127
Dissolved solids	13	610	264	430	591	497	439	370	266
Sulfate	13	271.0	91.1	167.0	260.0	190.0	167.0	113.0	93.7
Chloride	13	21.10	6.51	14.20	21.00	16.60	15.70	11.30	6.76
Nitrite plus nitrate	31	2.14	.02	.36	1.09	.41	.24	.06	.02
Nitrogen, total	31	3.26	.42	1.04	1.69	1.13	.95	.75	.57
Phosphorus, total	31	.70	.08	.28	.55	.40	.22	.16	.09
Suspended sediment	31	1,020.0	46.0	205.0	533.0	266.0	111.0	79.5	50.0

**Table 2.** Constituents estimated using regression equations for the Red River of the North at Fargo, North Dakota, 2003-05.

Constituent	Unit of measurement	U.S. Geological Survey parameter code
Alkalinity	Milligrams per liter as calcium carbonate	P90410
Dissolved solids	Milligrams per liter	P70301
Sulfate	Milligrams per liter	P00945
Chloride	Milligrams per liter	P00940
Total nitrite plus nitrate	Milligrams per liter as nitrogen	P00630
Total nitrogen	Milligrams per liter	P00600
Total phosphorus	Milligrams per liter	P00665
Suspended sediment	Milligrams per liter	P80154

Standard deviation,  $s$ , is the positive square root of the mean square error ( $MSE$ ), a common measure of variability in regression equations. Standard deviation is calculated as follows:

$$s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-p}} \quad (1)$$

where

- $SSE$  is the error sum of squares defined in Helsel and Hirsch (1995),
- $n$  is the number of observations used to develop the regression equation,
- and
- $p$  is the number of parameters estimated in the regression equation.

The standard deviation, like  $MSE$ , is an indicator of the variability of the probability distributions of the response, or explanatory, variable. However, the standard deviation is in the same units as the response variable, milligrams per liter (mg/L) for example, whereas  $MSE$  is in squared milligrams per liter (mg/L)<sup>2</sup>. Therefore, the standard deviation is easier to interpret in relation to the response variable.

$R^2$ , the multiple coefficient of determination, is calculated as follows:

$$R^2 = 1 - \frac{SSE}{SS_y} \quad (2)$$

where

- $SSE$  is the error sum of squares,
- and
- $SS_y$  is the sums of squares  $y$ , or total sums of squares, defined in Helsel and Hirsch (1995).

$R^2$  is a number, 0 through 1, that when multiplied by 100 is interpreted as the percentage of the variability in the response variable explained by the explanatory variables and the regression equation. Generally, the higher the  $R^2$ , the better the regression equation. However, this does not guarantee the regression equation is useful (Neter and others, 1996). For example, if estimates require extrapolation outside the observed response variables, the regression equation may not provide accurate estimates. In relation to this study, a regression equation with a high  $R^2$  does not imply that the equation will work well for estimating constituent concentrations and loads during the winter because winter constituent concentrations are outside the range of observed constituent concentrations.

As an indicator of the ability of the regression relations to estimate constituent concentrations, the measured concentrations were compared to the concentrations estimated by the regression relations by calculating relative percentage difference ( $RPD$ s) using the following equation:

$$RPD = \left| \frac{B - A}{A} \times 100 \right|, \quad (3)$$

where

- $B$  is the constituent concentration estimated from the regression equation,
- and
- $A$  is the measured constituent concentration.

$A$  is assumed to be correct and the  $RPD$  is the relative difference of  $B$  from  $A$ , expressed as a percentage.

The potential pool of surrogate physical properties, or explanatory variables, is listed in table 3. In addition to the properties in table 3, variables related to time were included in the pool of possible explanatory variables for nutrients in order to model the cyclical behavior of nutrients. These variables

**Table 3.** Possible surrogate physical properties, or explanatory variables, used to develop regression equations to estimate constituent concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

Physical properties	Unit of measurement	U.S. Geological Survey parameter code
Streamflow	Cubic feet per second	P00060
Specific conductance	Microsiemens per centimeter at 25 degrees Celsius	P00095
pH	Standard units	P00400
Temperature, water	Degrees Celsius	P00010
Turbidity	Nephelometric turbidity units Formazin nephelometric unit	P61028 P63680
Dissolved oxygen	Milligrams per liter	P00300

included  $t$ ,  $t^2$ ,  $\cos\left(\frac{2j\pi t}{365}\right)$ , and  $\sin\left(\frac{2j\pi t}{365}\right)$ , where  $t$  is

the Julian date, or the ordinal day of the year (92 for April 3, 2005),  $j$  is an integer 1 through 3,  $\cos$  is the cosine function, and  $\sin$  is the sine function. An increase in  $j$  decreases the period of the  $\cos$  and  $\sin$  functions. Larger  $j$  values may be used to model behavior that has multiple cycles per year (Helsel and Hirsch, 1995). A model may include two or more  $\cos/\sin$  pairs with different periods. This may indicate two different seasonal processes that affect the response variable. In selection of a regression model,  $\cos/\sin$  terms were required to

be used in pairs. For example, if  $\cos\left(\frac{2\pi t}{365}\right)$  was a statistically significant explanatory variable, the corresponding  $\sin$  term,

$\sin\left(\frac{2\pi t}{365}\right)$ , was included in the model. Including pairs of

$\cos/\sin$  terms may result in models where one member of the pair is significant and the other is not. However, using only one member of the  $\cos/\sin$  pair forces an arbitrary phase shift rather than a phase shift determined by the data (Helsel and Hirsch, 1995).

To select the best model for estimating a particular constituent, stepwise regression (Insightful Corporation, 2001b), an automatic search procedure that sequentially adds terms to and deletes terms from the model, was performed to determine the best subset of explanatory variables. These variables were

examined for scientific validity, and diagnostics were examined that in some cases resulted in logarithmic transformations of one or more of the explanatory variables or the response variable. In the case of transformation of one or more variables, stepwise regression was performed again to verify that the best subset of explanatory variables had been selected.

## Calculation of Measured and Estimated Constituent Loads

Daily load is the total mass of a constituent that is transported past a gaging station in 1 day. Measured constituent loads were calculated by multiplying measured constituent concentrations by streamflow at the time the constituent concentrations were measured and multiplying by the conversion factor listed in table 4. Estimated mean daily constituent loads were calculated by multiplying the constituent concentrations estimated using the regression equations by daily streamflow and by a conversion factor (table 4).

**Table 4.** Conversion factors used in calculation of measured and estimated loads.

[ft<sup>3</sup>/s, cubic feet per second]

Multiply	By	By	To obtain
milligrams per liter	streamflow, in ft <sup>3</sup> /s	5.39	pounds per day

## Results of Regression Analysis

Regression relations between constituents of concern and surrogate physical properties were examined, and a regression equation was developed for each constituent using one or more surrogate variables. The regression equations and each equation's associated standard deviation,  $R^2$ , and median *RPD* are listed in table 5. A discussion of each constituent and the associated regression equations follows.

**Table 5.** Regression equations for estimates of alkalinity, dissolved solids, chloride, sulfate, total nitrogen, total nitrite plus nitrate, total phosphorus, and suspended sediment in the Red River of the North at Fargo, North Dakota, 2003-05.

[n, numbers of samples used to develop regression equation;  $R^2$ , coefficient of multiple determination; median *RPD*, median relative percentage difference; *SC*, specific conductance;  $Q$ , streamflow;  $t$ , ordinal day of year; *Turb*, turbidity]

Constituent	n	Equation	Standard deviation	$R^2$	Median <i>RPD</i>
Alkalinity ( <i>Alk</i> )	13	$Alk = 0.165SC - 54.3\log_{10}(Q) + 261$	10.5	0.951	2.3
Dissolved solids ( <i>DS</i> )	13	$DS = 0.689SC - 52.0$	12.3	.990	1.1
Sulfate ( $SO_4$ )	13	$\log_{10}(SO_4) = 0.128\log_{10}(Q) + 1.67\log_{10}(SC) - 2.96$	0.034	.963	5.8
Chloride ( <i>Cl</i> )	13	$Cl = -9.55\log_{10}(Q) + 0.011SC + 38.8$	1.81	.895	10.7
Total nitrite plus nitrate ( $NO_2 + NO_3$ )	30	$\log_{10}(NO_2 + NO_3) = 0.818\log_{10}(Q) + 1.69\cos(2\pi t/365) - 0.134\sin(2\pi t/365) + 0.755\cos(4\pi t/365) + 0.016\sin(4\pi t/365) - 2.31$	.309	.714	35.2
Total nitrogen ( <i>N</i> )	30	$N = 0.422\log_{10}(Q) + 0.699\cos(2\pi t/365) - 0.318\sin(2\pi t/365) + 0.400\cos(4\pi t/365) - 0.202\sin(4\pi t/365) + 0.030$	.188	.703	14.7
Total phosphorus ( <i>P</i> )	30	$P = 0.111\log_{10}(Q) + 0.353\log_{10}(Turb) + 0.056\cos(2\pi t/365) - 0.047\sin(2\pi t/365) - 0.734$	.080	.771	17.2
Suspended sediment ( <i>SSC</i> )	30	$\log_{10}(SSC) = 0.213\log_{10}(Q) + 0.814\log_{10}(Turb) - 0.092$	.128	.873	19.3

### Alkalinity

Alkalinity is an important characteristic of natural and polluted water and is the capacity of the solutes in the water to react with and neutralize acid, such as inflow from acid rain or wastewater (Hem, 1985). Several different compounds contribute to the alkalinity of water, mainly bicarbonates, carbonates, and hydroxides. If water is unable to neutralize acid inputs, the affected stream becomes more acidic (pH below 7.0 standard units). Acidification is generally harmful at pH below 5.0 and results in a reduction in the number of aquatic species present in the water body (Allan, 1995). Most rivers, in areas not affected by pollution, have a pH of 6.5 to 8.5 standard units, although some water bodies are naturally acidic even without the effects of acid rain or other human-related acid sources (Hem, 1985). For the water-quality samples used in this study, the mean pH was 8.0, the minimum was 7.3, and the maximum was 8.5 standard units. Although high alkalinity may

result in drinking water with an unpleasant taste, it is not detrimental to humans (North Dakota Department of Health, 2006).

Hem (1985) showed a positive linear relation between specific conductance and hardness, which is measured in milligrams per liter as calcium carbonate, as is alkalinity, for the Gila River in Arizona. Christensen and others (2000) showed a positive relation between specific conductance and alkalinity for the Little Arkansas River in Kansas. Streamflow also may have a relation to specific conductance and alkalinity as runoff can have a dilution effect on water-quality constituents (Christensen, 2000).

The surrogate variables chosen to estimate alkalinity of the Red River at Fargo were specific conductance and streamflow. Ranges for these explanatory variables were 447.0 to 967.0  $\mu\text{S}/\text{cm}$  for specific conductance and 560 to 9,010  $\text{ft}^3/\text{s}$  for streamflow. The multiple regression equation was:

$$Alk = 0.165SC - 54.3\log_{10}(Q) + 261, \quad (4)$$

where

- $Alk$  is the estimated alkalinity concentration, in milligrams per liter,  
 $SC$  is specific conductance, in microsiemens per centimeter at 25°C,

and

- $\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second.

The measured alkalinity concentration was compared to the estimated alkalinity concentration calculated by equation 4 graphically and by calculating the *RPD* (fig. 6).

## Dissolved Solids

Dissolved-solids, also called total dissolved solids, concentration is used to evaluate water quality and to compare water from one location to the next. Dissolved-solids concentration may be determined in two ways: (1) by the weight of the dry residue remaining after evaporation of the water samples, or (2) by summing the concentrations of other dissolved constituents if the concentrations of major ions are known (Hem, 1985). According to Hem, the computed value for dissolved solids, which was used in this Red River at Fargo, North Dakota, study, “may give a more useful indication of total dissolved-ion concentration than the residue left by evaporation.” High concentrations of dissolved solids may have an adverse effect on taste and make the water appear cloudy (North Dakota Department of Health, 2006).

Specific conductance “provides general indication of the content of dissolved matter for water that is not too saline or too dilute” (Hem, 1985). Streamflow also may affect dissolved solids, but the relation is not a simple one (Allan, 1995). For the Red River at Fargo, specific conductance was a significant explanatory variable for estimating dissolved solids; streamflow was not. Regression diagnostics, Cooks’ distance (Neter and others, 1996), and jackknife after bootstrap (Insightful Corporation, 2001b) indicated that the specific conductance value of the October 12, 2005, water-quality sample had a large effect on the estimation of the regression parameters. This sample changed the intercept term for the regression equation but had little effect on  $R^2$  or the median *RPD*. Therefore, the sample was retained for the estimation of the regression parameters. The range for specific conductance was 447.0 to 967.0  $\mu\text{S}/\text{cm}$ . The multiple regression equation was:

$$DS = 0.689SC - 52.0, \quad (5)$$

where

- $DS$  is the estimated dissolved-solids concentration, in milligrams per liter,

and

- $SC$  is specific conductance, in microsiemens per centimeter at 25°C.

The measured dissolved-solids concentrations were compared to the estimated dissolved-solids concentrations calculated by equation 5 graphically and by calculating the *RPD* (fig. 7).

## Sulfate

Sulfate sources include the weathering of rocks, agricultural runoff, fuel combustion, municipal and industrial effluent, and precipitation (Hem, 1985; Allan, 1995). The presence of too much sulfate has three main undesirable effects in drinking water: (1) sulfate can have a laxative effect with excessive intake, (2) water with high sulfate concentrations can form hard scales in boilers, and (3) sulfate can negatively affect taste (North Dakota Department of Health, 2006).

Sulfate is negatively charged and increased sulfate concentrations increase specific conductance (Hem, 1985). For the Red River at Fargo, both streamflow and specific conductance were statistically significant explanatory variables for estimating sulfate. The range for streamflow values was 560 to 9,010  $\text{ft}^3/\text{s}$ , and the range for specific conductance was 447.0 to 967.0  $\mu\text{S}/\text{cm}$ . The multiple regression equation was:

$$\log_{10}(SO_4) = 0.128\log_{10}(Q) + 1.67\log_{10}(SC) - 2.96, \quad (6)$$

where

- $\log_{10}(SO_4)$  is the base-10 logarithm of estimated sulfate concentration, in milligrams per liter,  
 $\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,

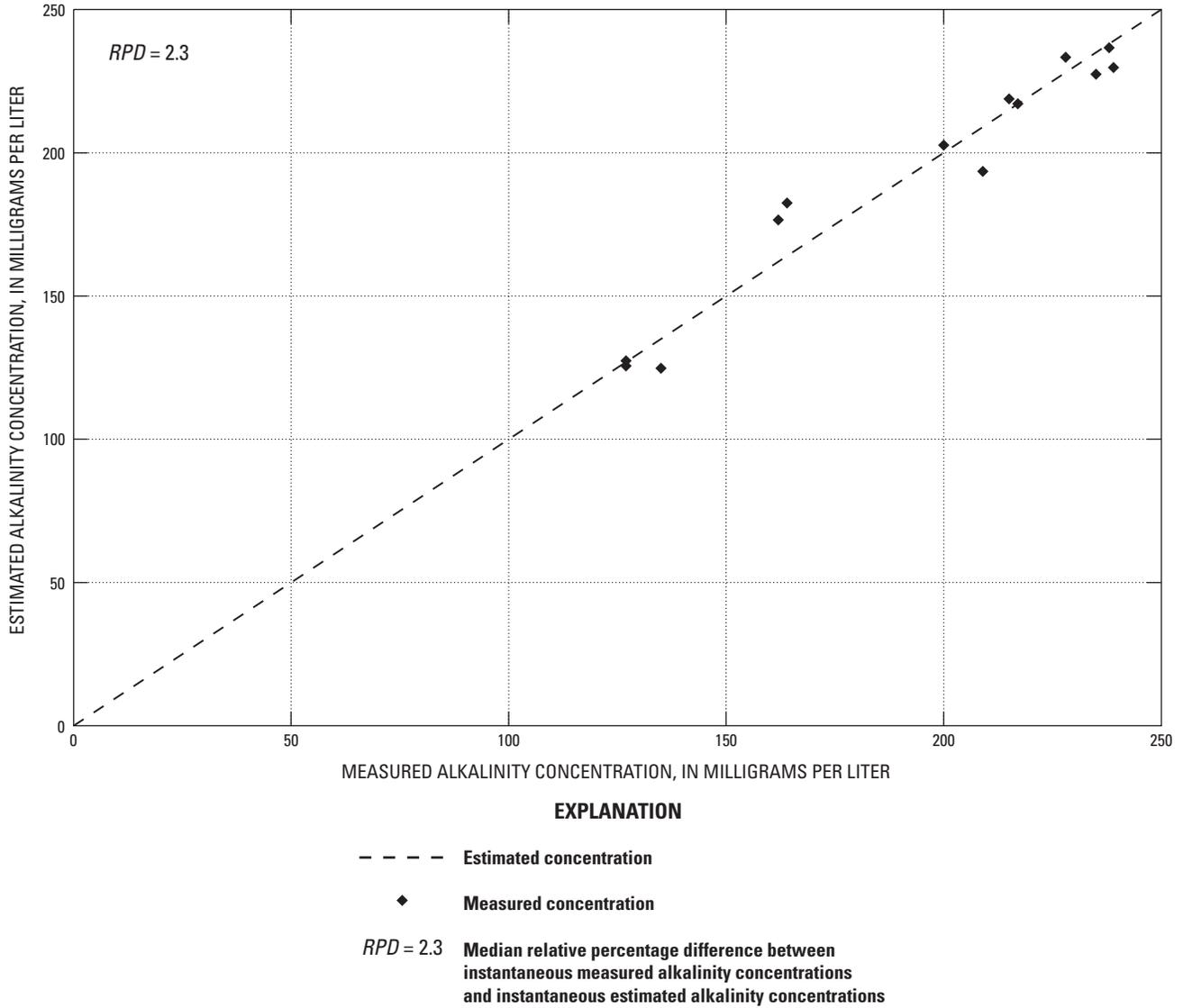
and

- $SC$  is specific conductance, in microsiemens per centimeter at 25°C.

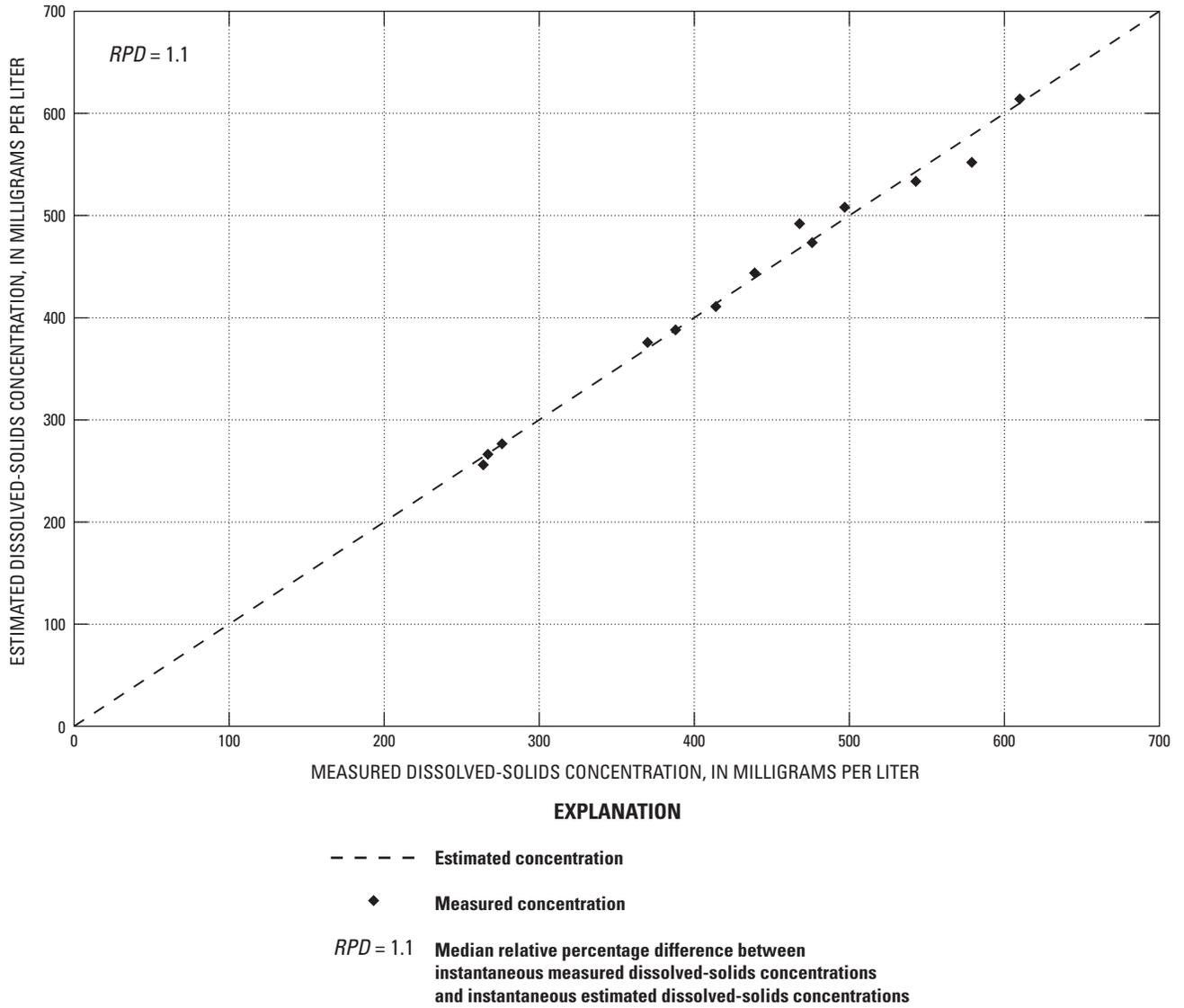
The measured sulfate concentrations were compared to the estimated sulfate concentrations calculated by equation 6 graphically and by calculating the *RPD* (fig. 8).

## Chloride

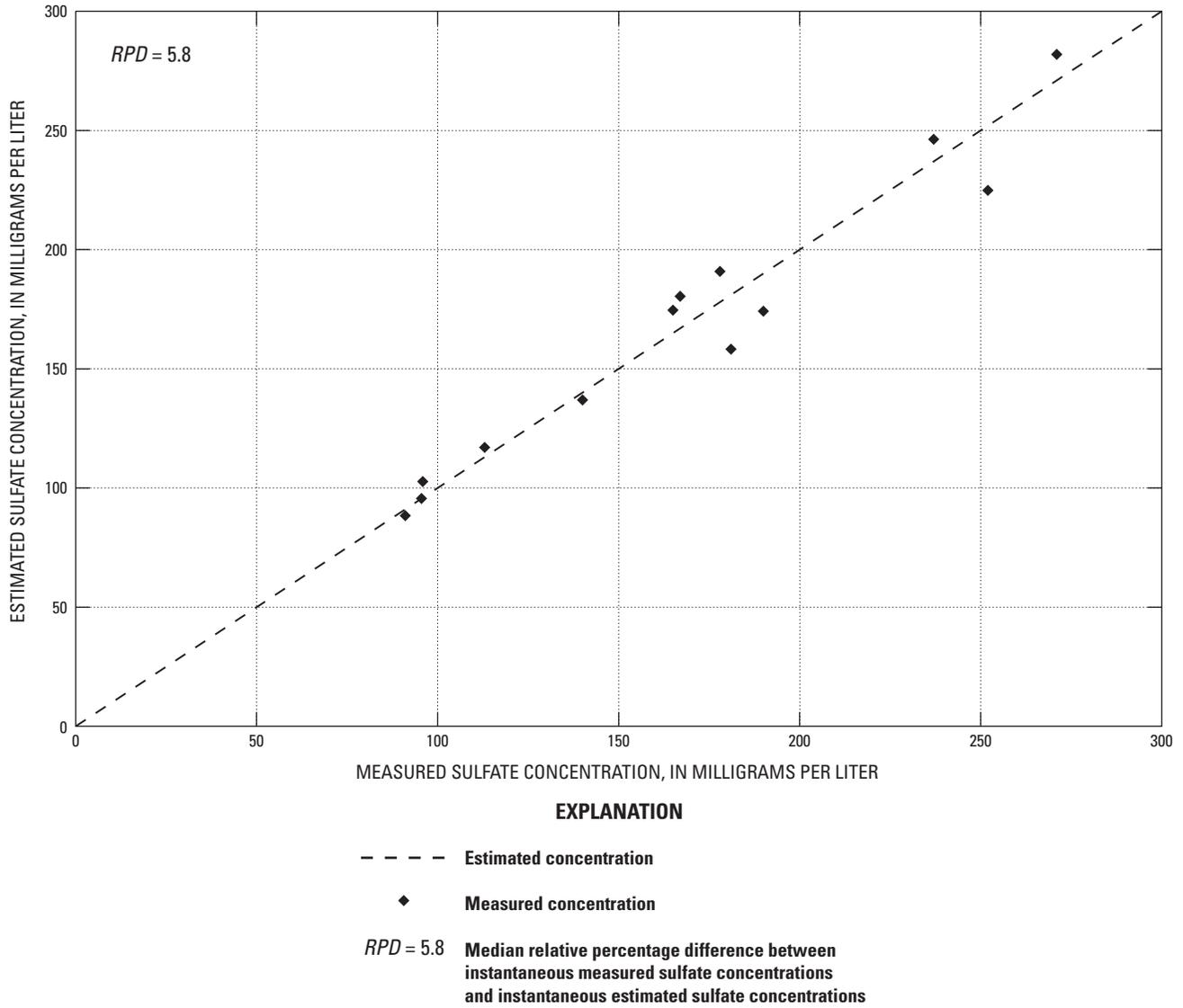
Chloride is present in virtually all rocks and is present in the tissues of all plants and animals. In surface water, concentrations of chloride range from less than 1 to more than 280,000  $\text{mg}/\text{L}$ , the extreme value representative of the Dead Sea (Feth, 1981), and in most streams the concentration is less than that of sulfate (Hem, 1985). Chloride is generally conservative (unchanging) in water; therefore, its circulation in the hydrologic cycle is mostly through physical processes (Hem, 1985). Chloride in streams may originate from the weathering of rocks, ground-water inflow, precipitation, and municipal and industrial effluent (Feth, 1981; Hem, 1985). The presence of too much chloride can have undesirable effects in drinking water; consumers may find the taste objectionable, and pipes in hot water systems may corrode (North Dakota Department of Health, 2006).



**Figure 6.** Comparison of measured and estimated alkalinity concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 7.** Comparison of measured and estimated dissolved-solids concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 8.** Comparison of measured and estimated sulfate concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

Chloride is a negatively charged ionic species that makes water conductive. Therefore, as chloride concentrations increase, the specific conductance increases. Chloride concentrations may be higher in the winter when a higher percentage of streamflow is from ground-water discharge, and sudden increases in chloride may indicate pollution in the stream (Hem, 1985). Feth (1981) found that western rivers in the United States generally have higher chloride concentrations in drought years.

For the Red River at Fargo, both streamflow and specific conductance were statistically significant explanatory variables for estimating chloride. The range for streamflow values was 560 to 9,010 ft<sup>3</sup>/s, and the range for specific conductance was 447 to 967  $\mu$ S/cm. The multiple regression equation was:

$$Cl = -9.55\log_{10}(Q) + 0.011SC + 38.8, \quad (7)$$

where

$Cl$  is the estimated chloride concentration, in milligrams per liter,

$\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,

and

$SC$  is specific conductance, in microsiemens per centimeter at 25°C.

The measured chloride concentrations were compared to the estimated chloride concentrations calculated by equation 7 graphically and by calculating the *RPD* (fig. 9).

## Total Nitrite Plus Nitrate

Nitrite is seldom present in high concentrations and is generally an indicator of pollution from sewage or other organic waste (Hem, 1985). Nitrate is an important plant nutrient produced during nitrification. Nitrification is a two-step process in which bacteria convert ammonia to nitrite, and then a second kind of bacteria convert nitrite to nitrate (Chapra, 1997).

Observed total nitrite plus nitrate was seasonal as was total nitrogen (fig. 10) and the same explanatory variables were used for nitrite plus nitrate as for total nitrogen. The range for streamflow values was 165 to 9,010 ft<sup>3</sup>/s. Regression diagnostics and the jackknife after bootstrap indicated that the June 6, 2005, total nitrite plus nitrate concentration had a large effect on the estimation of the regression parameters. This total nitrite plus nitrate measurement was very high compared to the other measured concentrations and was removed as an outlier (see discussion in "Total Nitrogen" section). The multiple regression equation was:

$$\begin{aligned} \log_{10}(NO_2 + NO_3) = & 0.818 \log_{10}(Q) + 1.69 \cos\left(2\pi \frac{t}{365}\right) \\ & - 0.134 \sin\left(2\pi \frac{t}{365}\right) + 0.755 \cos\left(4\pi \frac{t}{365}\right) \\ & + 0.016 \sin\left(4\pi \frac{t}{365}\right) - 2.31, \end{aligned} \quad (8)$$

where

$\log_{10}(NO_2 + NO_3)$  is the base-10 logarithm of the estimated total nitrite plus nitrate concentration, in milligrams per liter,

$\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,

and

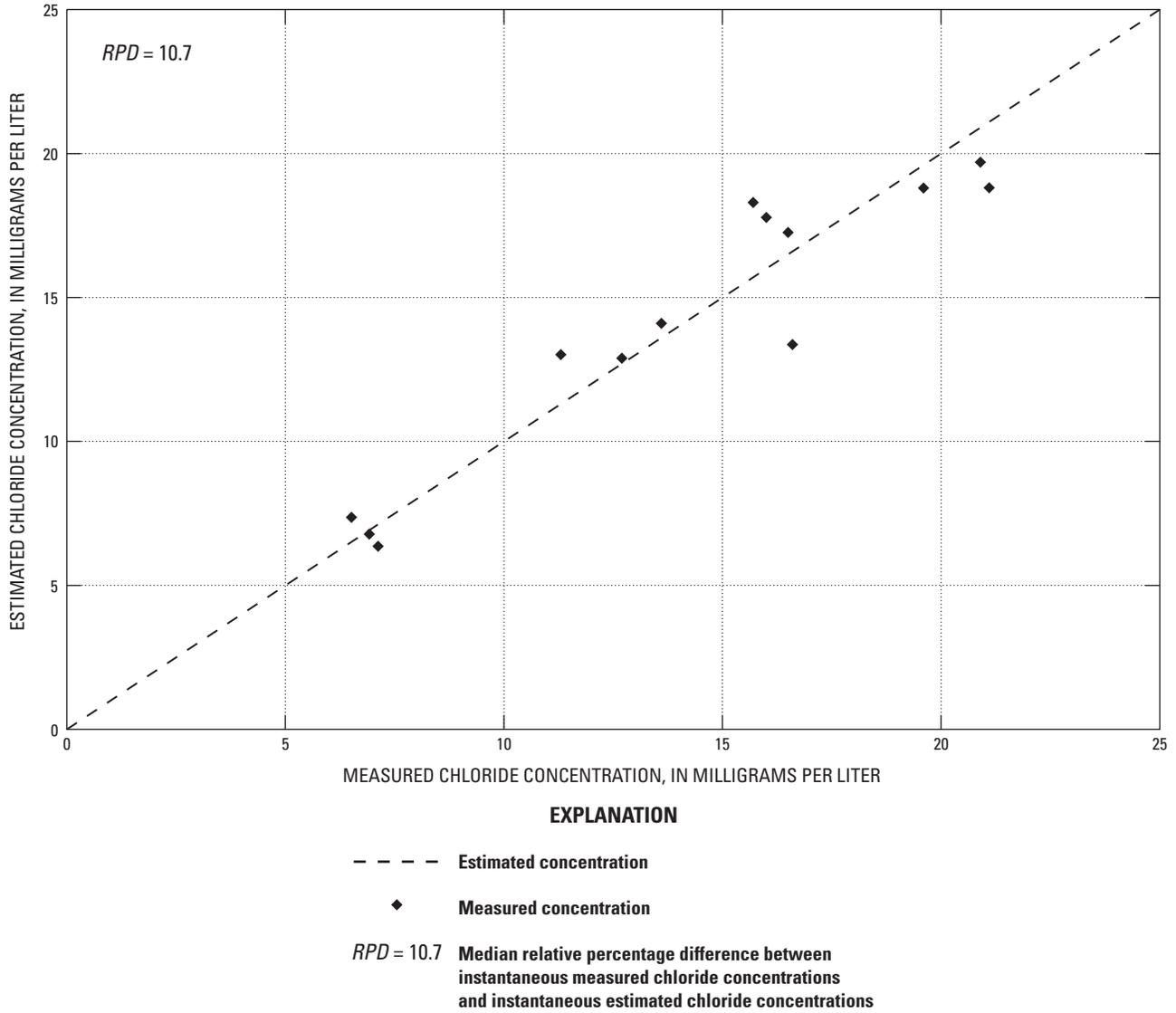
$t$  is the Julian date.

The measured total nitrite plus nitrate concentrations were compared to the estimated total nitrite plus nitrate concentrations calculated by equation 8 graphically and by calculating the *RPD* (fig. 10). Generally, measured total nitrite plus nitrate concentrations greater than 0.5 mg/L deviated more from the estimated concentrations than the measured concentrations less than 0.5 mg/L.

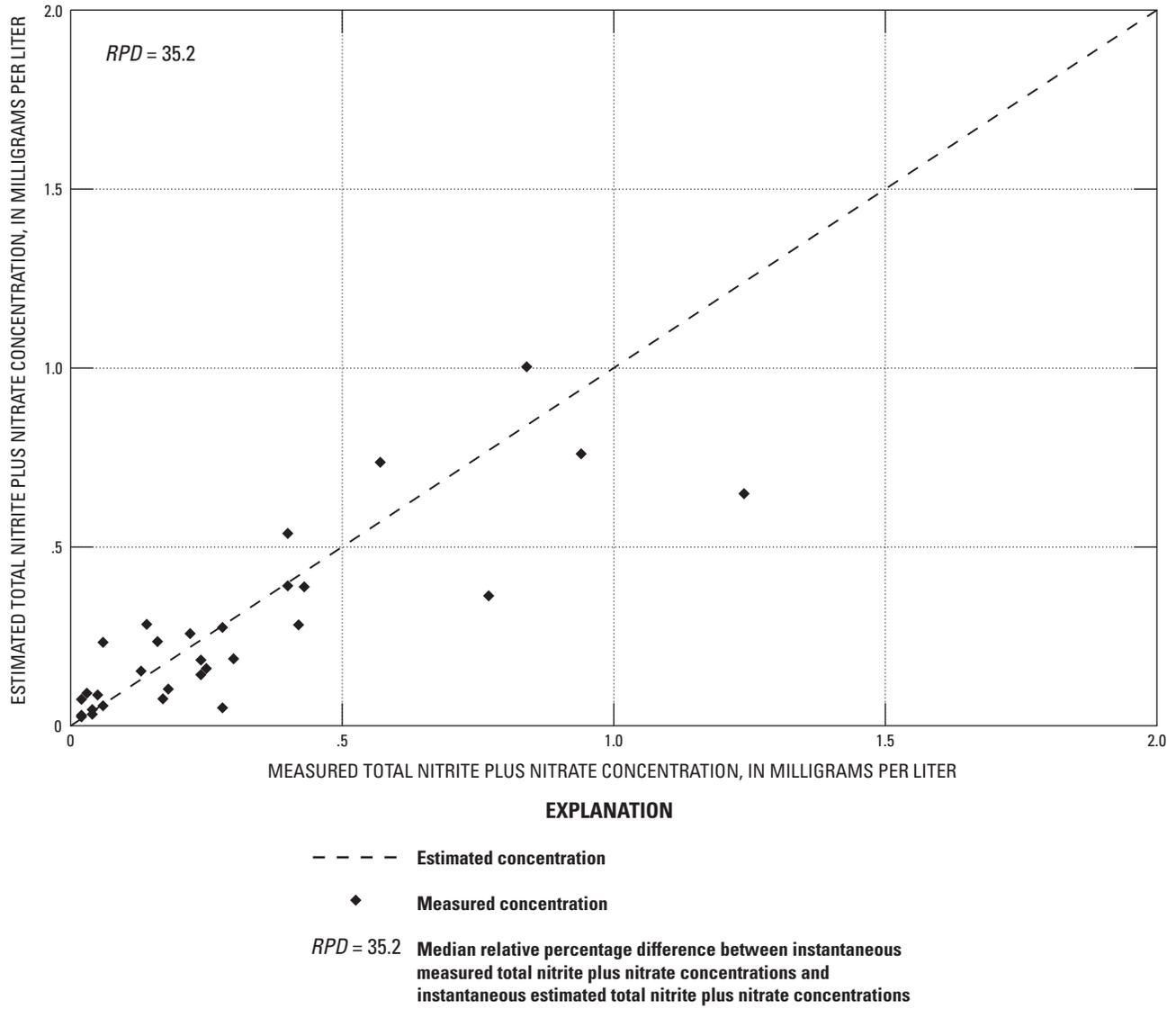
## Total Nitrogen

Nitrogen is of vital importance in plant and animal nutrition and occurs in water as nitrite or nitrate anions, ammonia cations, and in other forms such as cyanide. Total nitrogen concentration is related to land use and sources of nitrogen include precipitation, soil leaching, agricultural runoff, and waste disposal (Allan, 1995). Nitrogen levels may fluctuate seasonally, with high concentrations occurring during high flow when there is drainage from cultivated fields or feedlots (Hem, 1985) and with an autumn increase due to the decomposition of leaves (Allan, 1995). Figure 11 shows that total nitrogen concentrations in the Red River at Fargo do appear to be seasonal with the peak occurring during high streamflow (fig. 3), followed by a decrease, and then an autumn increase that corresponds to an autumn increase in streamflow and may correspond with autumn-shed leaf decomposition.

For the Red River at Fargo, streamflow and cosine (cos) and sine (sin) terms were chosen as explanatory variables for estimating nitrogen. The range for streamflow values was 165 to 9,010 ft<sup>3</sup>/s. Regression diagnostics and the jackknife after bootstrap indicated that the June 6, 2005, total nitrogen concentration of 2.73 mg/L (fig. 11) had a large effect on the estimation of the regression parameters. The field notes for June 6, 2005, indicated that the Red River at the water-quality



**Figure 9.** Comparison of measured and estimated chloride concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 10.** Comparison of measured and estimated total nitrite plus nitrate concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

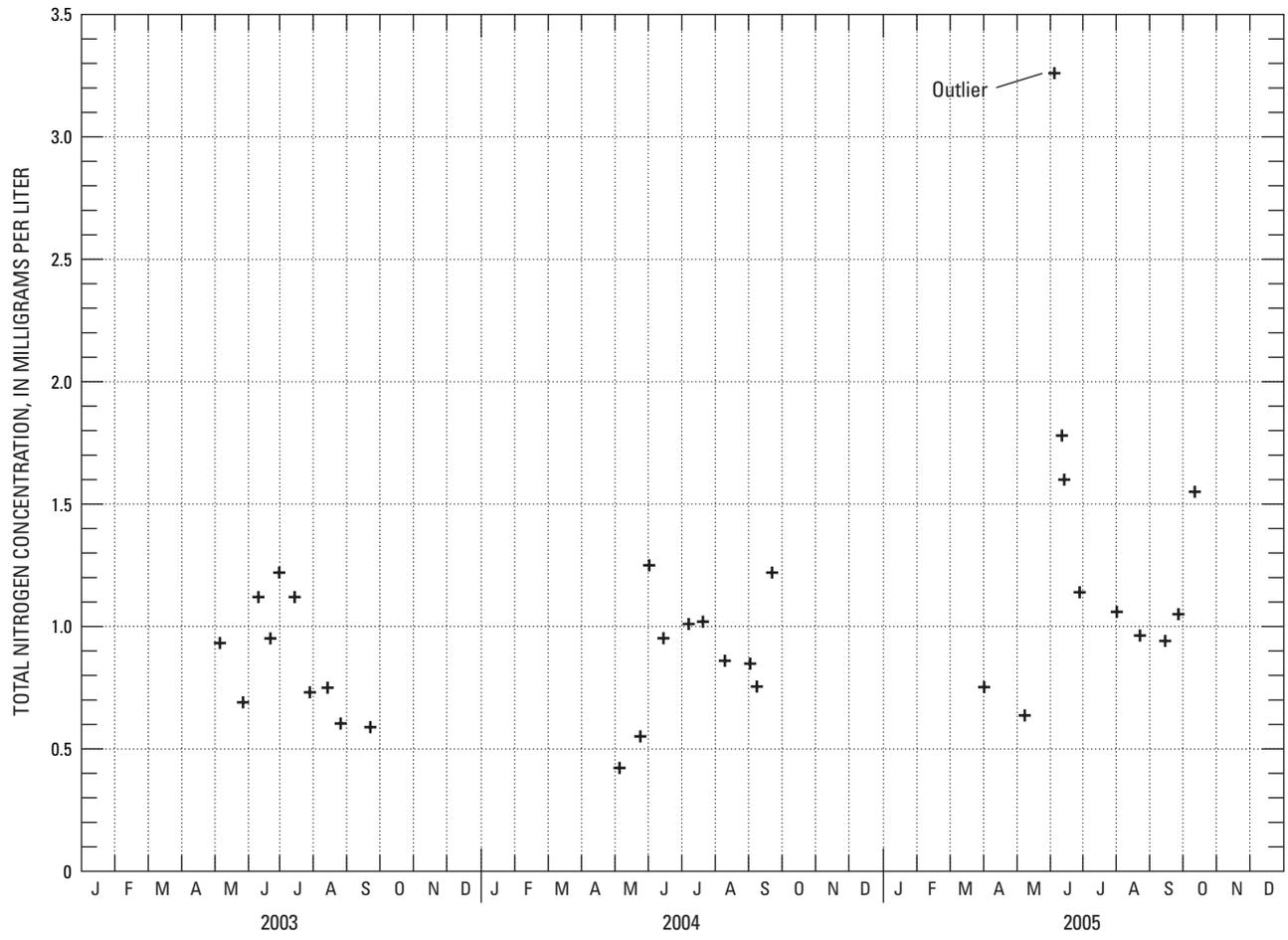


Figure 11. Measured total nitrogen concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

sampling site (the walking-path bridge) was “very turbid” with “lots of sediment suspended” (R. Nustad, USGS, unpub. data, 2005). The total nitrogen observed under those conditions was very high compared to the other observations (fig. 11) and was removed as an outlier. The multiple regression equation was:

$$N = 0.422 \log_{10}(Q) + 0.699 \cos\left(2\pi \frac{t}{365}\right) - 0.318 \sin\left(2\pi \frac{t}{365}\right) + 0.400 \cos\left(4\pi \frac{t}{365}\right) - 0.202 \sin\left(4\pi \frac{t}{365}\right) + 0.30, \quad (9)$$

where

$N$  is the estimated total nitrogen concentration, in milligrams per liter,  
 $\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,

and

$t$  is the Julian date.

The measured nitrogen concentrations were compared to the estimated nitrogen concentrations calculated by equation 9 graphically and by calculating the *RPD* (fig. 12).

## Total Phosphorus

Phosphorus, like nitrogen, is an essential nutrient for plant growth. Increased phosphorus, or nitrogen, in a water body may be accompanied by an undesirable high rate of plant growth, such as algal blooms. Sources of phosphorus include soil erosion and domestic and industrial sewage effluent (Hem, 1985). Like nitrogen, phosphorus may vary seasonally with streamflow and the growing season.

For the Red River at Fargo, streamflow, turbidity (an estimator of suspended sediment, which can transport constituents such as phosphorus), and cosine and sine terms were chosen as explanatory variables for estimating total phosphorus. Regression diagnostics and the jackknife after bootstrap indicated that the June 23, 2003, turbidity value (fig. 4) had a large effect on the estimation of the regression parameters. The field notes for June 23, 2003, indicated that it had rained the night before and while turbidity was measured it was raining and a small storm drain was discharging above the water-quality sampling site (First Avenue Bridge) (R. Nustad and C. Laveau, USGS, unpub. data, 2003). The measured turbidity value of 808 FNU was very high in comparison to other turbidity values, and because of its effect on the regression relation, it was removed as an outlier. The range for streamflow values used to develop the regression equation was 165 to 9,010 ft<sup>3</sup>/s, and the range for turbidity values was 22.9 to 584 FNU. The multiple regression equation was:

$$P = 0.111 \log_{10}(Q) + 0.353 \log_{10}(Turb) + 0.056 \cos\left(2\pi \frac{t}{365}\right) - 0.047 \sin\left(2\pi \frac{t}{365}\right) - 0.734, \quad (10)$$

where

$P$  is the estimated total phosphorus concentration, in milligrams per liter,  
 $\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,  
 $\log_{10}(Turb)$  is the base-10 logarithm of turbidity, in FNUs,

and

$t$  is the Julian date.

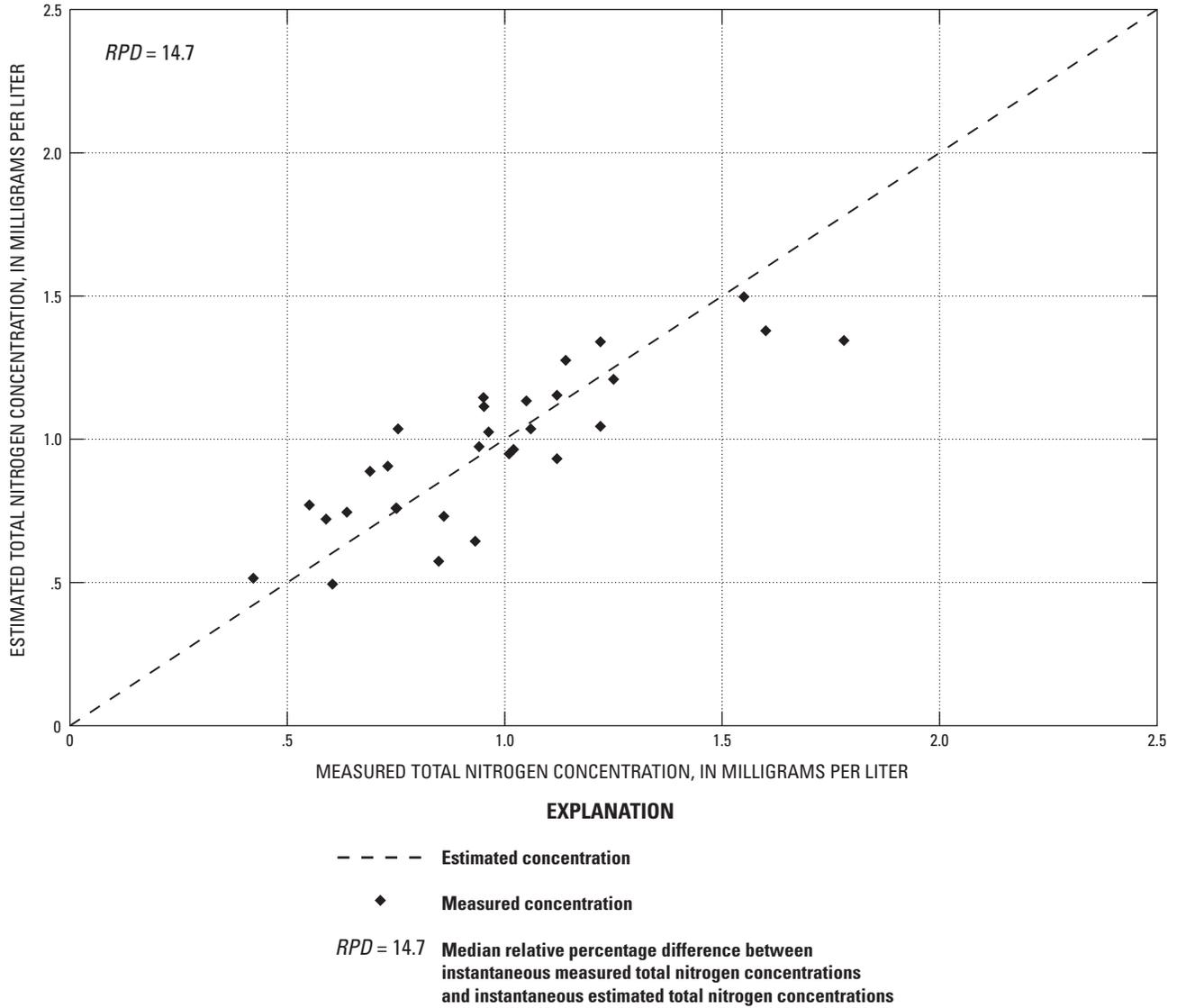
The measured total phosphorus concentrations were compared to the estimated total phosphorus concentrations calculated by equation 10 graphically and by calculating the *RPD* (fig. 13).

## Suspended Sediment

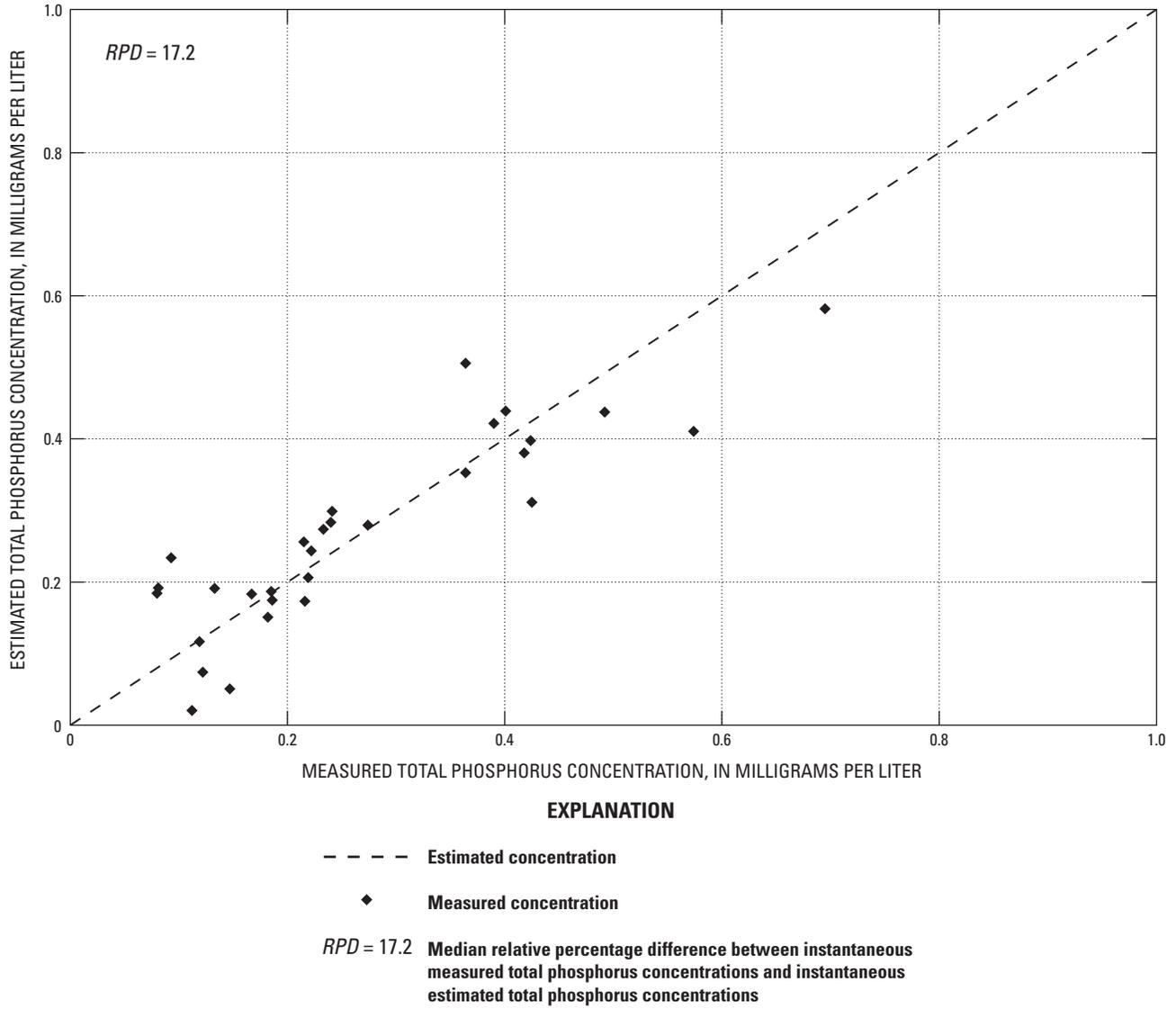
The terms suspended-sediment concentration (SSC) and total suspended solids (TSS) are often used interchangeably. However, SSC and TSS represent different laboratory analytical methods used to quantify concentrations of suspended solids in surface water. The methods are described in Guy (1969) and Gray and others (2000). Gray and others (2000) evaluated 3,235 paired SSC and TSS values. They stated, “The method for determining TSS, which was originally designed for analyses of wastewater samples, is shown to be fundamentally unreliable for the analysis of natural-water samples. In contrast, the method for determining SSC produces relatively reliable results for samples of natural water, regardless of the amount or percentage of sand-size material in the samples. SSC and TSS data collected from natural waters are not comparable and should not be used interchangeably.” Therefore, SSC was used in this study.

SSC is an important indicator of water quality for several reasons. Suspended sediments may act as a transport for other constituents, including toxic substances, and some fish can be affected by sediment deposition on spawning beds. “An increased amount of fine sediments within the streambed reduces its permeability to water movement, affecting the delivery and removal of gases, nutrients and metabolites” for stream biota, and potentially restricting movement of biota (Allan, 1995).

SSC increases with streamflow, unless the supply of sediment is depleted (Allan, 1995). SSC is also related to turbidity because turbidity includes suspended material. Turbidity is a field measurement that is an indicator of “suspended and



**Figure 12.** Comparison of measured and estimated total nitrogen concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 13.** Comparison of measured and estimated total phosphorus concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

dissolved matter, such as clay, silt, finely divided organic matter, plankton and other microscopic organisms, organic acids, and dyes” (Wilde, variously dated). For the Red River at Fargo, both streamflow and turbidity were statistically significant explanatory variables for estimating SSC.

As was the case with total phosphorus, regression diagnostics and jackknife after bootstrap indicated that the June 23, 2003, turbidity value (fig. 4) had a large effect on the estimation of the regression parameters. This turbidity value of 808 FNU was very high in comparison to other turbidity values, and because of its extreme effect on the regression relation, it was removed as an outlier. The range for streamflow values was 165 to 9,010 ft<sup>3</sup>/s, and the range for turbidity values was 22.9 to 584 FNU. The multiple regression equation was:

$$\log_{10}(SSC) = 0.213\log_{10}(Q) + 0.814\log_{10}(Turb) - 0.092, \quad (11)$$

where

$\log_{10}(SSC)$  is the base-10 logarithm of estimated suspended-sediment concentration, in milligrams per liter,

$\log_{10}(Q)$  is the base-10 logarithm of streamflow, in cubic feet per second,

and

$\log_{10}(Turb)$  is the base-10 logarithm of turbidity, in FNUs.

The measured suspended-sediment concentrations were compared to the estimated suspended-sediment concentrations calculated by equation 11 graphically and by calculating the *RPD* (fig. 14).

## Measured and Estimated Constituent Loads

Because the explanatory variables used in the regression equations (table 5) are related to time or are continuously monitored properties, the regression equations developed in this study may be used to estimate daily concentration loads in the Red River at Fargo. Measured daily load was computed for each constituent, and the daily load was estimated using the regression equations. Comparisons of measured and estimated daily loads are shown for each constituent in figures 15-22. No water-quality samples were collected under ice conditions, and the regression equations may not be appropriate for ice conditions. Therefore, loads were not estimated from November 1 to March 31. Breaks in the lines representing estimated daily loads indicate ice conditions or time periods in which one or more of the continuously monitored physical property values was unavailable.

The peak loads generally occurred in June and July when streamflow also peaked (fig. 3). In 2004, high streamflow occurred in autumn (September and October). These high flows caused an increase in estimated loads in September and October 2004, especially for the nutrients (figs. 19-21).

The estimated total nitrite plus nitrate nutrient loads were not graphed after October 24 (fig. 19) because the high streamflow in fall 2004 produced an extreme increase in the fall 2004 load that was not supported by measured concentrations during that time. Further water-quality sampling may result in a refined regression equation that better estimates loads under such conditions or further analysis may indicate that a separate regression equation is needed for fall and winter estimation.

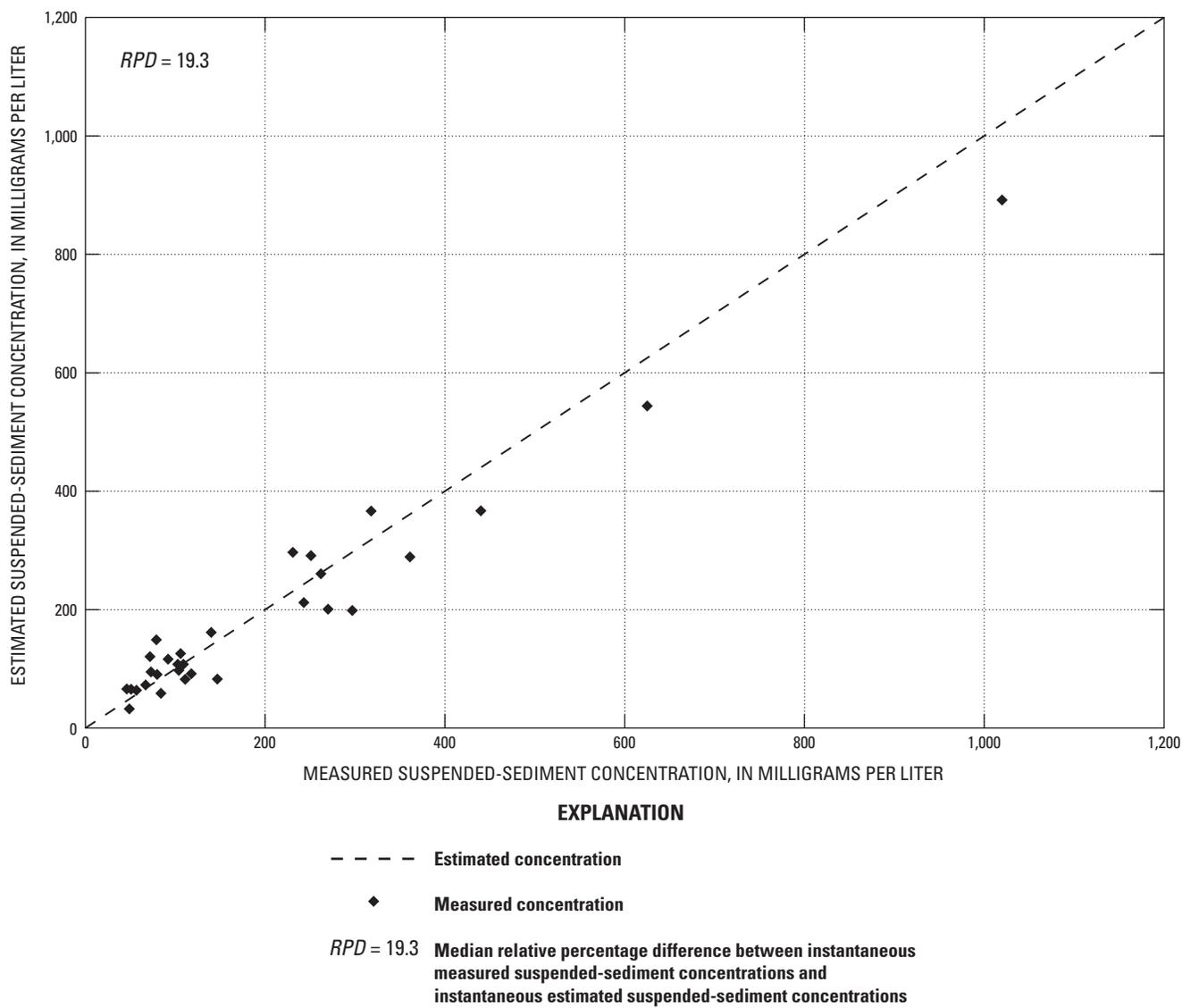
## Summary

The U.S. Geological Survey gaging station at Fargo, North Dakota, was established in June 1901 to provide stage and streamflow values for the Red River. The USGS began collecting water-quality samples at the gaging station in 1949. During more recent years, 2004 to the present (2006), the gaging station has provided real-time stage and streamflow data, instream velocity, and water-quality data (the physical properties of specific conductance, pH, water temperature, turbidity, and dissolved oxygen concentration).

As shown in a previous study, physical properties measured by water-quality monitors may be used as surrogates for concentrations of important water-quality constituents. Because the surrogate physical properties used in this study are continuously recorded, the regression equations may be used to track changes in water quality as they occur. Potential benefits include providing the cities of Fargo, North Dakota, and Moorhead, Minnesota, with real-time information for the management of wastewater and providing the public with water-quality information that may affect their use of the Red River.

A regression analysis was performed on the streamflow and water-quality data collected in 2003 through 2005 from the Red River at Fargo, North Dakota, to determine the relation between physical properties of water and water-quality constituent concentrations. The explanatory variables considered in the regression analysis were streamflow, specific conductance, pH, water temperature, turbidity, and dissolved oxygen concentration. The response variables were alkalinity, dissolved solids, sulfate, chloride, total nitrite plus nitrate, total nitrogen, total phosphorus, and suspended sediment. For total nitrite plus nitrate, total nitrogen, and total phosphorus, cosine and sine functions of time also were used to explain the seasonality in concentration.

Streamflow, or the volume of water passing a particular point per unit time, is a fundamental measurement of streams. For the conditions observed in 2003-05, the base-10 logarithm of streamflow was a statistically significant estimator of all of the constituents in this study except dissolved solids. Specific conductance, or water's ability to conduct an electric current, was a statistically significant estimator of alkalinity, dissolved solids, sulfate, and chloride. Turbidity, caused by the presence of suspended and dissolved matter in water, includes suspended sediment, which may act as a transport mechanism



**Figure 14.** Comparison of measured and estimated suspended-sediment concentrations in the Red River of the North at Fargo, North Dakota, 2003-05.

for water-quality constituents. Turbidity was a statistically significant explanatory variable for total phosphorus and suspended sediment. pH, water temperature, and dissolved oxygen were not statistically significant explanatory variables for any constituent concentrations in this study.

As an indicator of the ability of the regression equations to explain the variability in the constituent concentrations,  $R^2$ , the coefficient of multiple determination, was reported for each equation.  $R^2$  was 0.951 for alkalinity, 0.990 for dissolved solids, 0.963 for sulfate, 0.895 for chloride, 0.714 for total nitrite plus nitrate, 0.703 for total nitrogen, 0.771 for total phosphorus, and 0.873 for suspended sediment. Also, the median relative percentage difference (median *RPD*) was calculated for each equation. The median *RPD* compares the measured concentrations to the concentrations estimated by the regression equations and smaller median *RPDs* are desirable. The median *RPD* was 2.3 for alkalinity, 1.1 for dissolved solids, 5.8 for sulfate, 10.7 for chloride, 35.2 for total nitrite plus nitrate, 14.7 for total nitrogen, 17.2 for total phosphorus, and 19.3 for suspended sediment.

The regression equations presented in this report are site specific and apply only to the Red River at Fargo, North Dakota. An additional constraint is that the equations were developed with water-quality samples collected only during the months of March through October. The regression equations may not apply to the quality of water under ice. Future sampling may result in the refinement of these equations or in the development of different regression equations for spring/summer and fall/winter for some constituents.

In addition to constituent concentrations, constituent loads also were estimated using the regression equations. These load estimates provide a visual summary of the seasonality of constituent loads and make a continuous record available for comparison to water-quality standards expressed as TMDLs. The peak loads generally occurred in June and July when streamflow also peaked. The increasing population of the Fargo, North Dakota, region, increasing public interest in TMDLs, and the international interest in water quality of the Red River make this study of regional and international importance because it shows how constituent loads may be calculated with surrogate variables continuously recorded at the Fargo gaging station. The methods used in this study replicate those used in Kansas and may be replicated for other sites in North Dakota and the Nation to monitor water quality.

## References Cited

- Allan, J.D., 1995, Stream ecology—structure and function of running waters: London, Chapman & Hall, 388 p.
- Buchanan, T.J., and Somers, W.P., 1969, Discharge measurements at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A8, 65 p., accessed June 2, 2006, at <http://pubs.usgs.gov/twri/twri3a8/>
- Chapra, S.C., 1997, Surface water-quality modeling: New York, McGraw-Hill, 844 p.
- Christensen, V.G., 2001, Characterization of surface-water quality based on real-time monitoring and regression analysis, Quivira National Wildlife Refuge, south-central Kansas, December 1998 through June 2001: U.S. Geological Survey Water-Resources Investigations Report 01-4248, 28 p., accessed March 2, 2006, at <http://pubs.water.usgs.gov/wri014248>
- Christensen, V.G., Jian, Xiaodong, and Ziegler, A.C., 2000, Regression analysis and real-time water-quality monitoring to estimate constituent concentrations, loads, and yields in the Little Arkansas River, south-central Kansas, 1995-99: U.S. Geological Survey Water-Resources Investigations Report 00-4126, 36 p., accessed March 2, 2006, at <http://pubs.water.usgs.gov/wri004126>
- Feth, J.H., 1981, Chloride in natural continental water—a review: U.S. Geological Survey Water-Supply Paper 2176, 30 p., accessed June 9, 2006, at <http://pubs.er.usgs.gov/usgspubs/wsp/wsp2176>
- Gray, J.R., Glysson, G.D., Turcios, L.M., and Schwarz, G.E., 2000, Comparability of suspended-sediment concentration and total suspended solids data: U.S. Geological Survey Water-Resources Investigations Report 00-4191, 14 p., accessed June 2, 2006, at <http://pubs.er.usgs.gov/usgspubs/wri/wri004191>
- Guy, H.P., 1969, Laboratory theory and methods for sediment analysis: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. C1, 58 p., accessed June 2, 2006, at <http://pubs.usgs.gov/twri/twri5c1/>

- Helsel, D.R., and Hirsch, R.M., 1995, *Statistical methods in water resources*: New York, Elsevier Science B.V., 529 p.
- Hem, J.D., 1985, *Study and interpretation of the chemical characteristics of natural water*, (3rd ed.): U.S. Geological Survey Water-Supply Paper 2254, 264 p., accessed March 7, 2006, at <http://pubs.er.usgs.gov/usgspubs/wsp/wsp2254>
- Insightful Corporation, 2001a, Resampling techniques—bootstrap and jackknife, in *S-PLUS 6 for Windows guide to statistics*, volume 2: Seattle, Insightful Corporation, p. 537-562.
- Insightful Corporation, 2001b, *S-Plus 6 for Windows user's guide*: Seattle, Insightful Corporation, 688 p.
- International Organization for Standardization, 1999, *Water quality—Determination of turbidity*: Geneva, Switzerland, International Organization for Standardization, ISO 7027, 10 p.
- Kennedy, E.J., 1983, Computation of continuous records of streamflow: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A13, 53 p., accessed June 2, 2006, at <http://pubs.usgs.gov/twri/twri3-a13/>
- Kennedy, E.J., 1984, Discharge ratings at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A10, 59 p., accessed June 2, 2006, at <http://pubs.usgs.gov/twri/twri3-a10/>
- Matthes, W.J., Sholar, C.J., and George, J.R., 1992, Quality-assurance plan for analysis of fluvial sediment: U.S. Geological Survey Open-File Report 91-467, 31 p., accessed June 2, 2006, at <http://pubs.er.usgs.gov/usgspubs/ofr/ofr91467>
- Neter, John, Kutner, M.H., Nachtsheim, C.J., and Wasserman, William, 1996, *Applied linear statistical models*, (4th ed.): Boston, WCB/McGraw-Hill, 1408 p.
- North Dakota Department of Health, 2003a, North Dakota Department of Health Chemistry Division quality assurance plan: Bismarck, North Dakota, North Dakota Department of Health, [variously paged].
- North Dakota Department of Health, 2003b, North Dakota Department of Health environmental methods manual: Bismarck, North Dakota, North Dakota Department of Health, [variously paged].
- North Dakota Department of Health, 2006, How minerals affect water supplies: Miscellaneous publication, about 3 p., accessed June 9, 2006, at <http://www.health.state.nd.us/WQ/GW/pubs/mineral.htm>
- Tornes, L.H., 2005, Water quality of streams in the Red River of the North Basin, Minnesota, North Dakota, and South Dakota, 1970-2001: U.S. Geological Survey Scientific Investigations Report 2005-5095, 81 p., accessed July 19, 2006, at <http://pubs.water.usgs.gov/sir20055095/>
- U.S. Census Bureau, 2006, State and county quickfacts: accessed August 28, 2006, at <http://quickfacts.census.gov/qfd/states/>
- U.S. Environmental Protection Agency, 2006, Clean water act section 303: accessed July 19, 2006, at <http://www.epa.gov/waterscience/standards/303.htm>
- U.S. Geological Survey, [variously dated], National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A9, [variously paged], accessed June 2, 2006, at <http://pubs.water.usgs.gov/twri9A>
- U.S. Geological Survey, [2004?], U.S. Geological Survey implements new turbidity data-reporting procedures: U.S. Geological Survey Information Sheet, 1 p., accessed June 5, 2006, at <http://water.usgs.gov/owq/turbidity/TurbidityInfoSheet.pdf>
- Wilde, F.D., ed., [variously dated], Field measurements, in National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water Resource Investigations, book 9, chap. A6, [variously paged], accessed June 5, 2006, at [http://water.usgs.gov/owq/FieldManual/Chapter6/Ch6\\_contents.html](http://water.usgs.gov/owq/FieldManual/Chapter6/Ch6_contents.html)
- Williams-Sether, T.J., 2004, Regression equations for estimating concentrations of selected water-quality constituents for selected gaging stations in the Red River of the North Basin, North Dakota, Minnesota, and South Dakota: U.S. Geological Survey Water-Resources Investigations Report 03-4291, 33 p., accessed April 26, 2006, at <http://pubs.water.usgs.gov/wri034291/>
- YSI Environmental, 2001, Comparison – YSI 6820 SONDE: accessed June 5, 2006, at <http://www.yei.com/extranet/EPGKL.nsf/447554deba0f52f2852569f500696b21/575fa0ee78a627e5852569e7005bf7e6!OpenDocument>

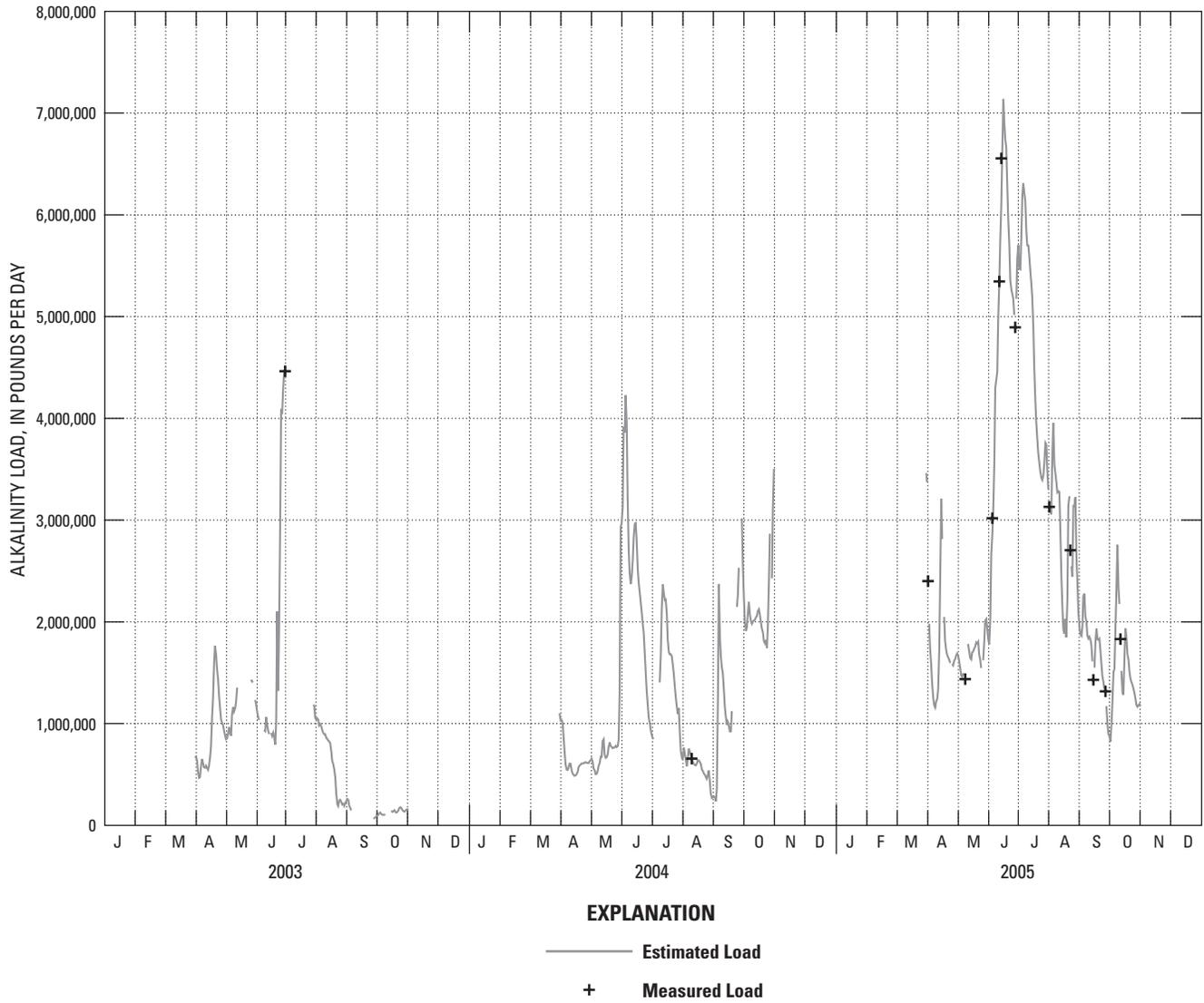
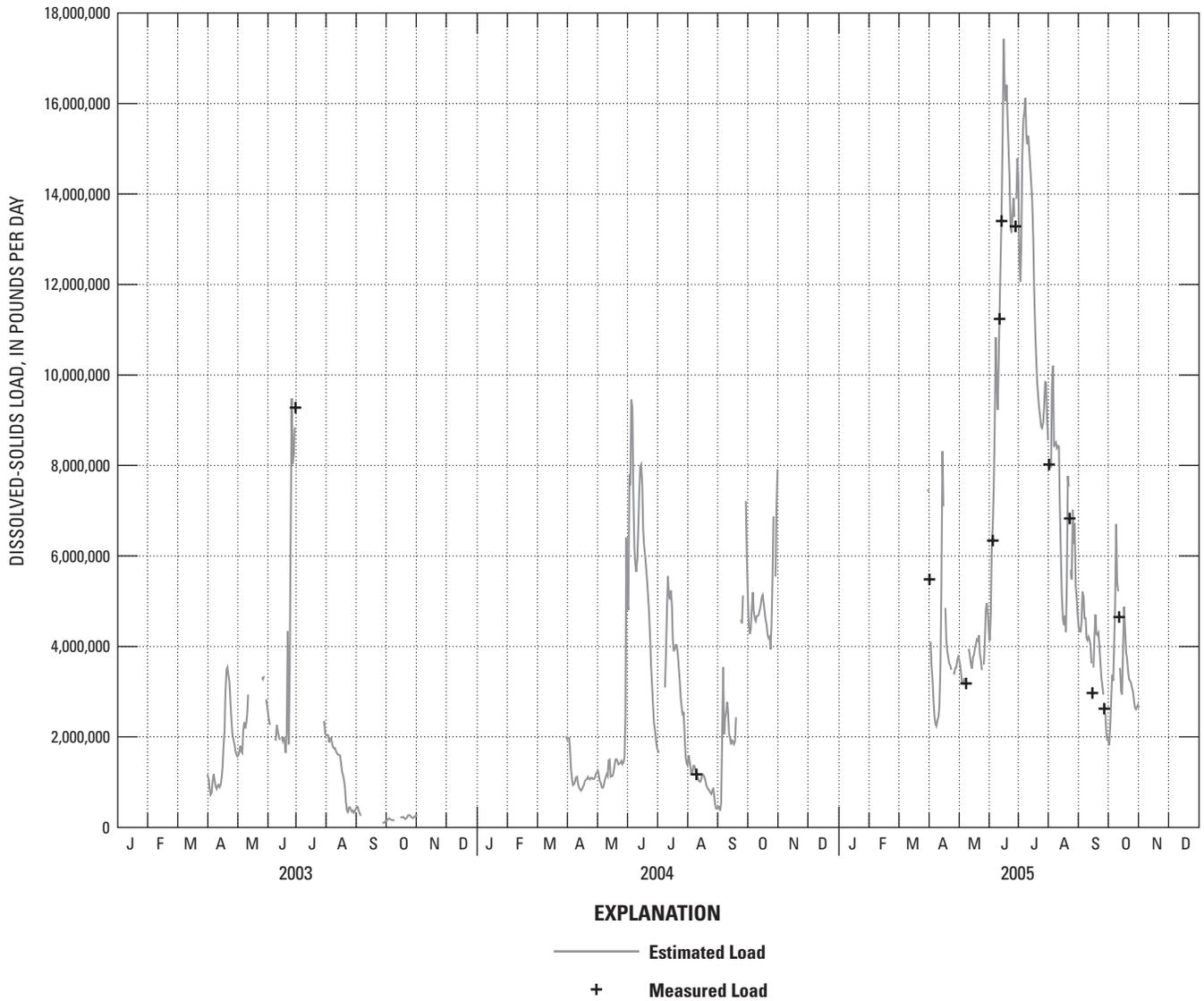


Figure 15. Comparison of measured and estimated alkalinity loads in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 16.** Comparison of measured and estimated dissolved-solids loads in the Red River of the North at Fargo, North Dakota, 2003-05.

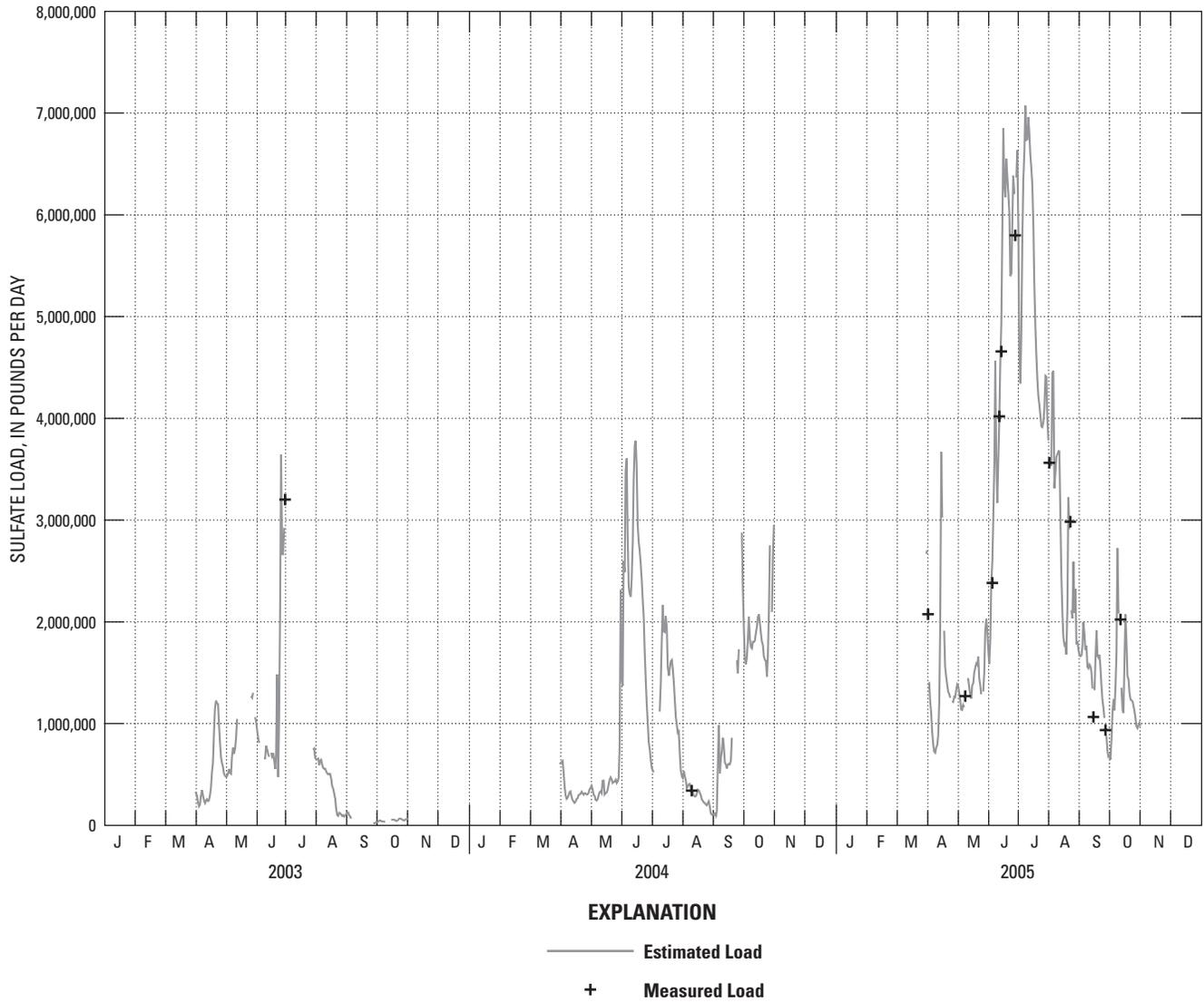


Figure 17. Comparison of measured and estimated sulfate loads in the Red River of the North at Fargo, North Dakota, 2003-05.

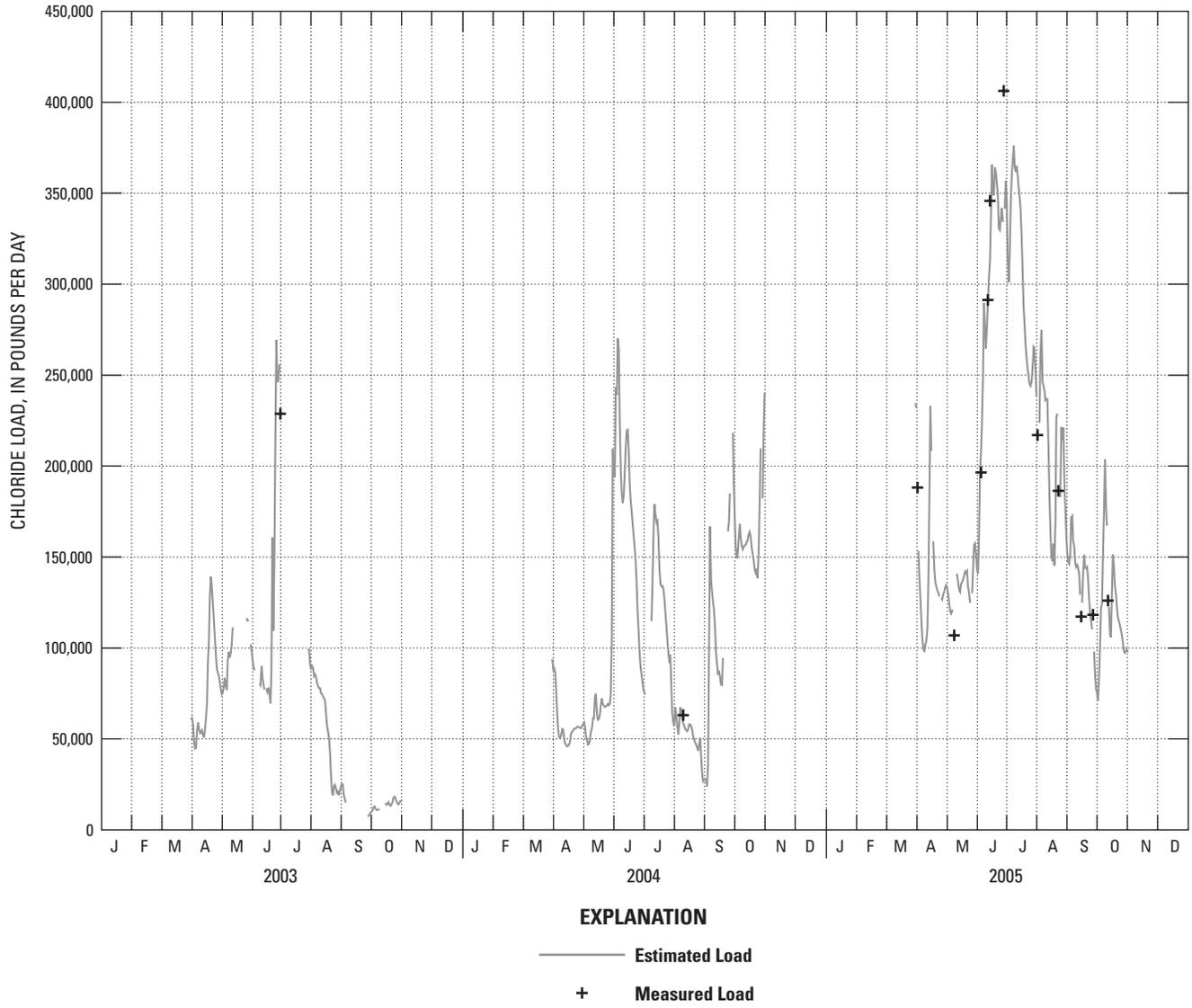
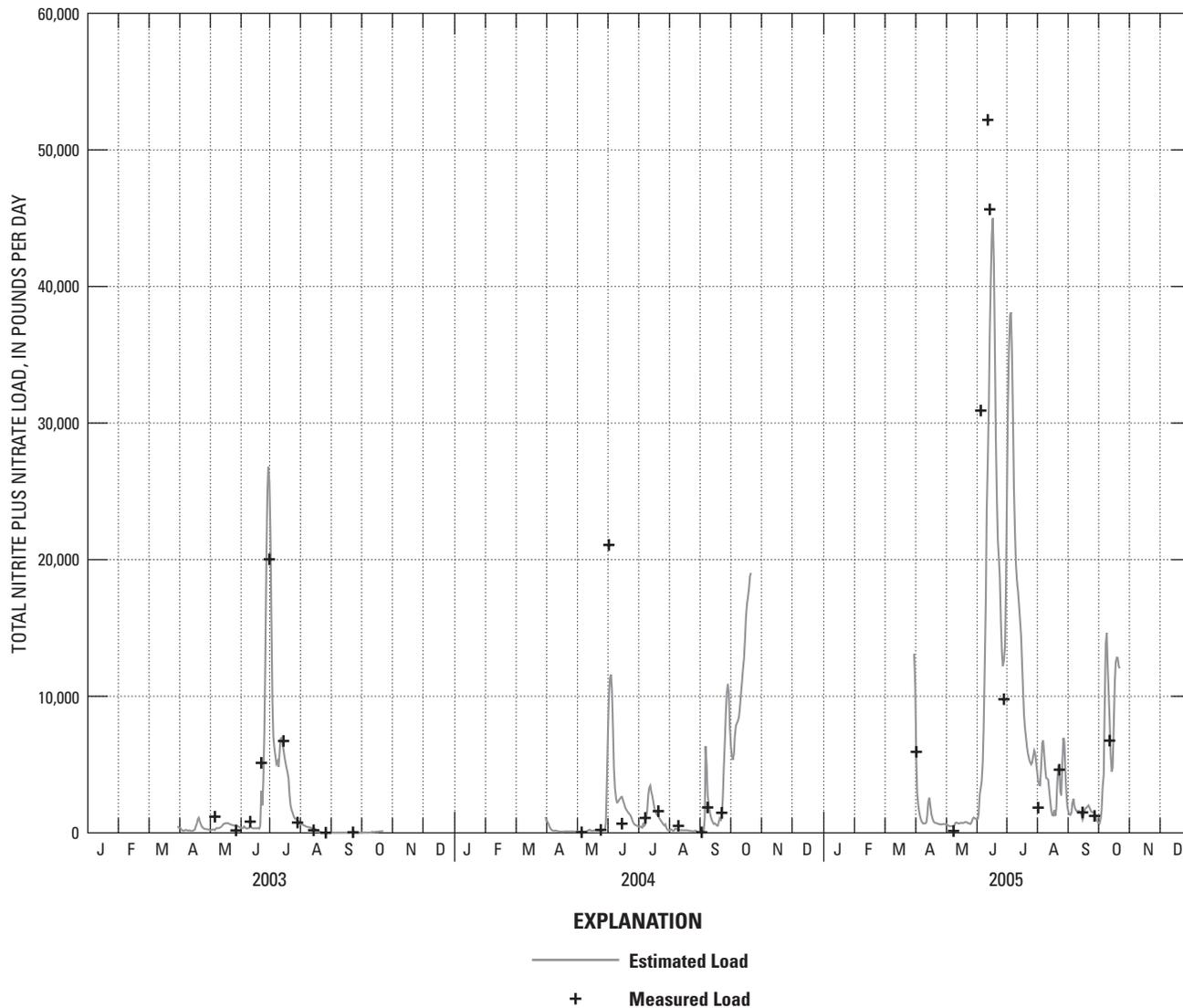
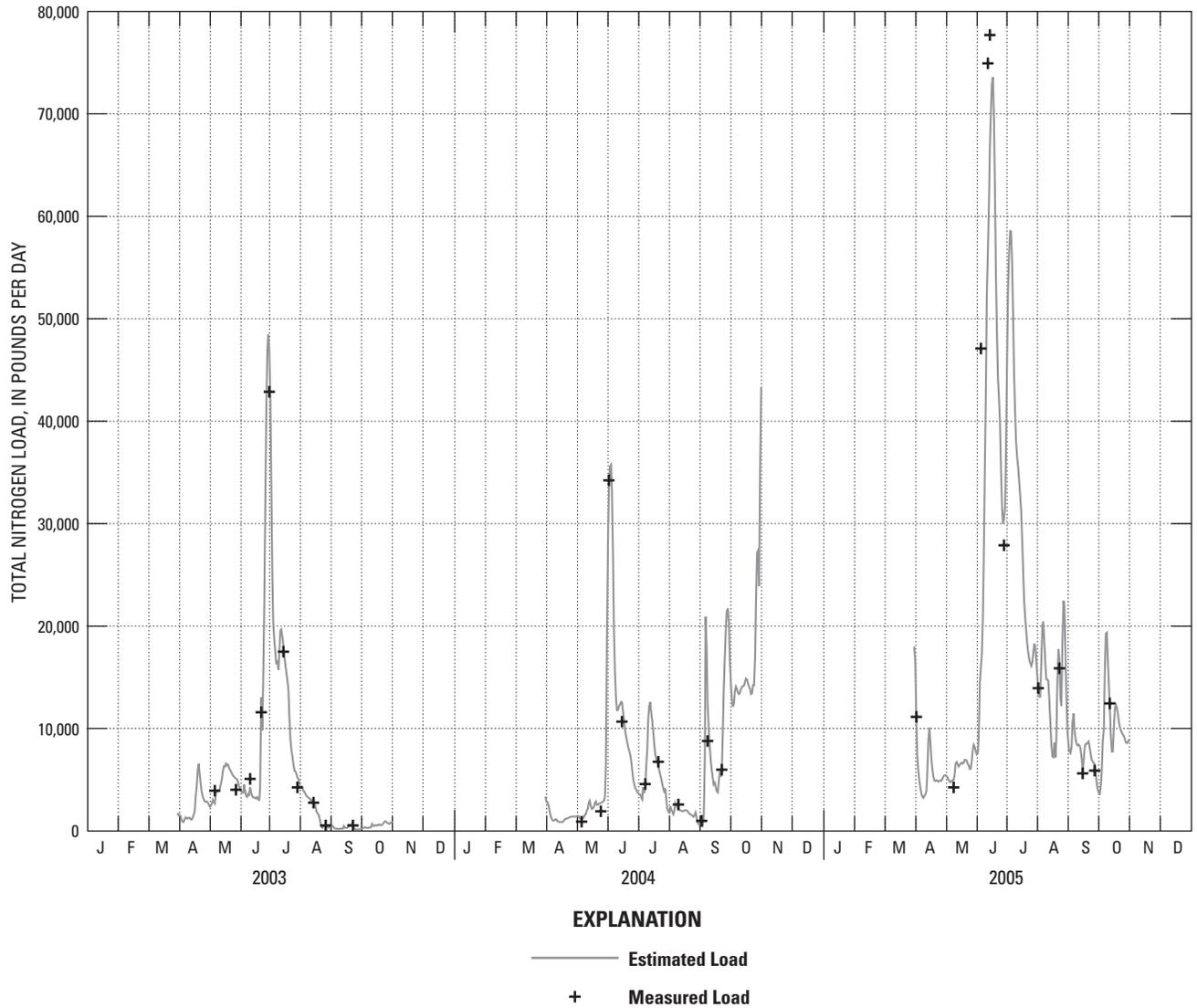


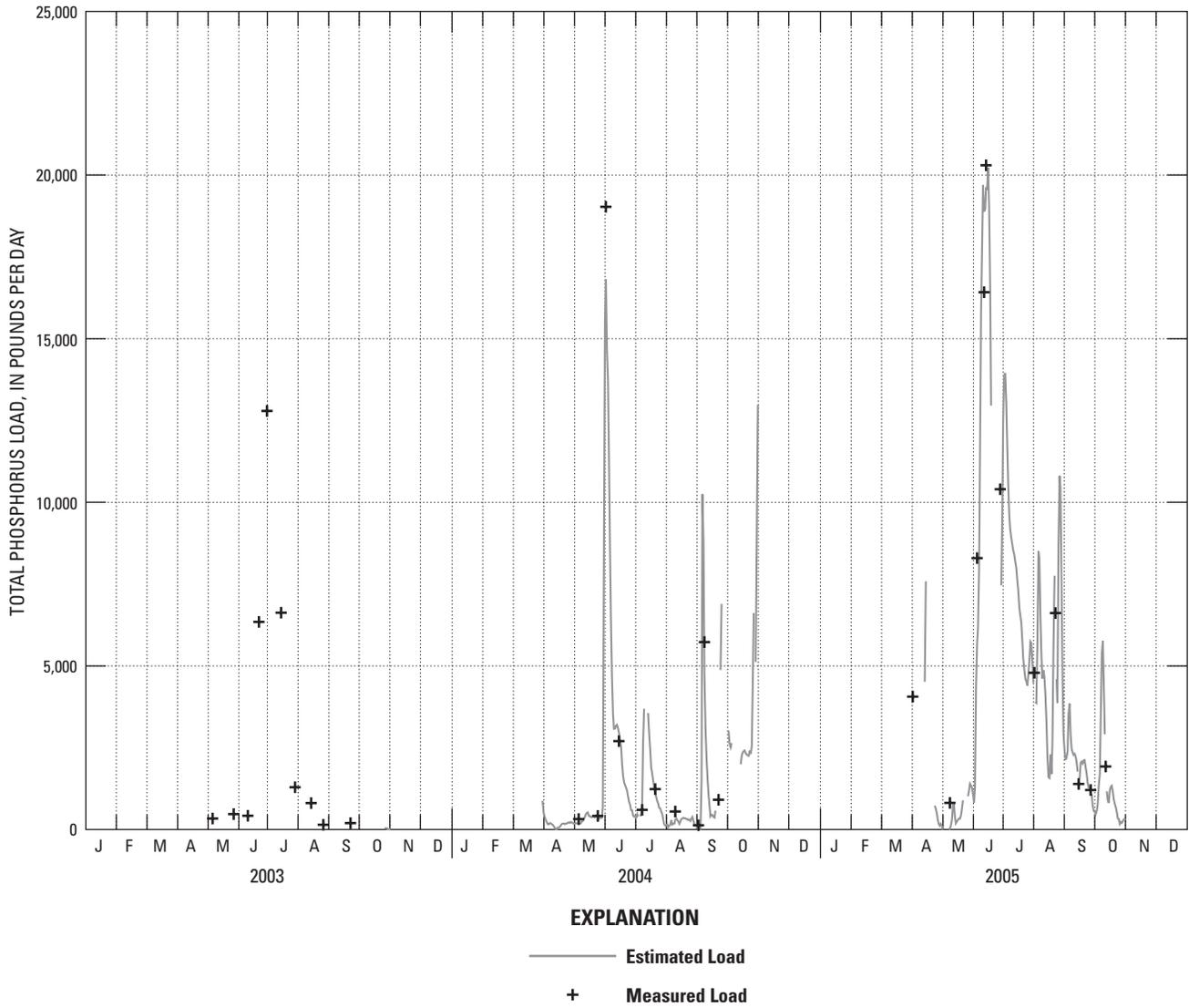
Figure 18. Comparison of measured and estimated chloride loads in the Red River of the North at Fargo, North Dakota, 2003-05.



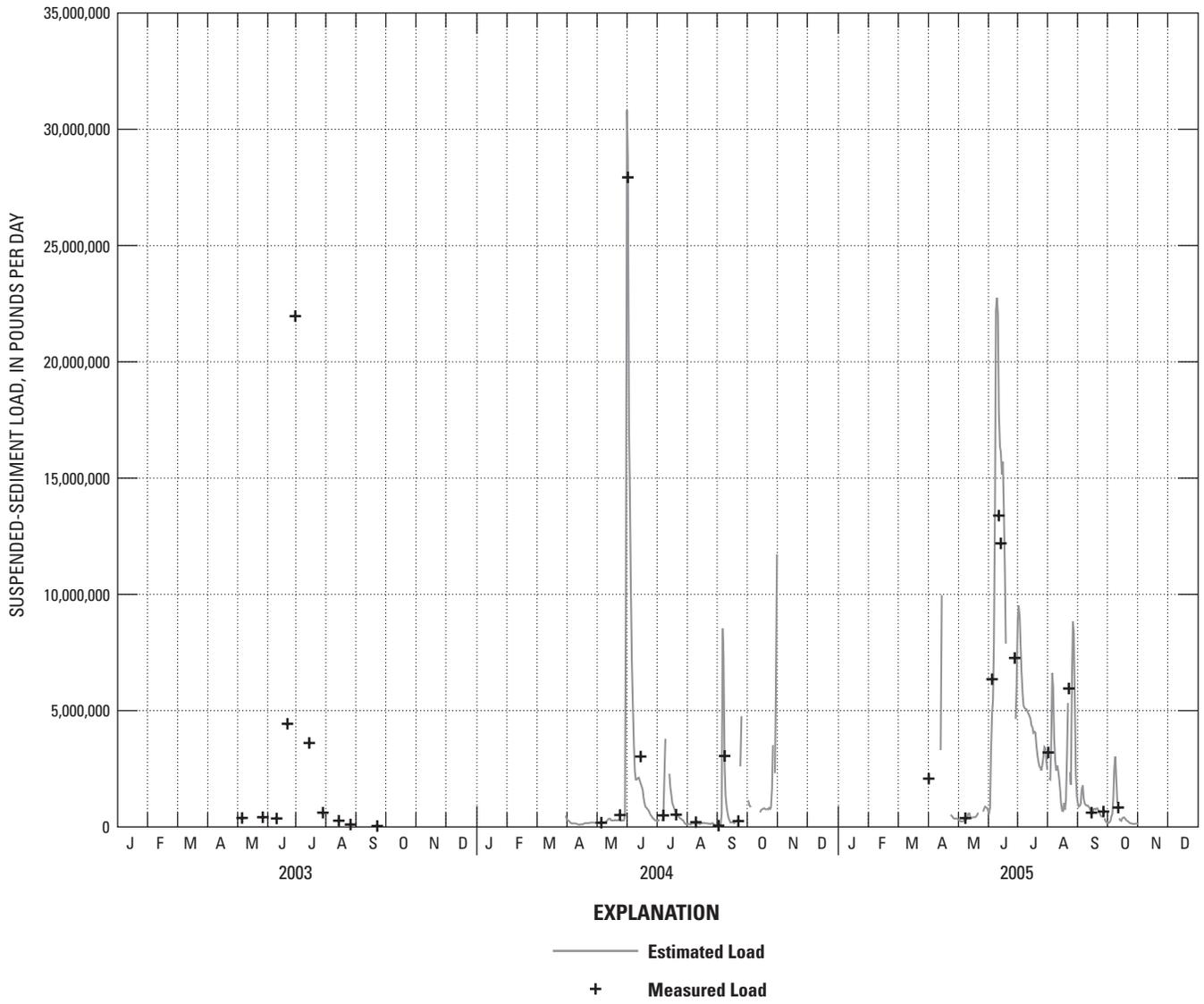
**Figure 19.** Comparison of measured and estimated total nitrite plus nitrate loads in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 20.** Comparison of measured and estimated total nitrogen loads in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 21.** Comparison of measured and estimated total phosphorus loads in the Red River of the North at Fargo, North Dakota, 2003-05.



**Figure 22.** Comparison of measured and estimated suspended-sediment loads in the Red River of the North at Fargo, North Dakota, 2003-05.

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

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

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

