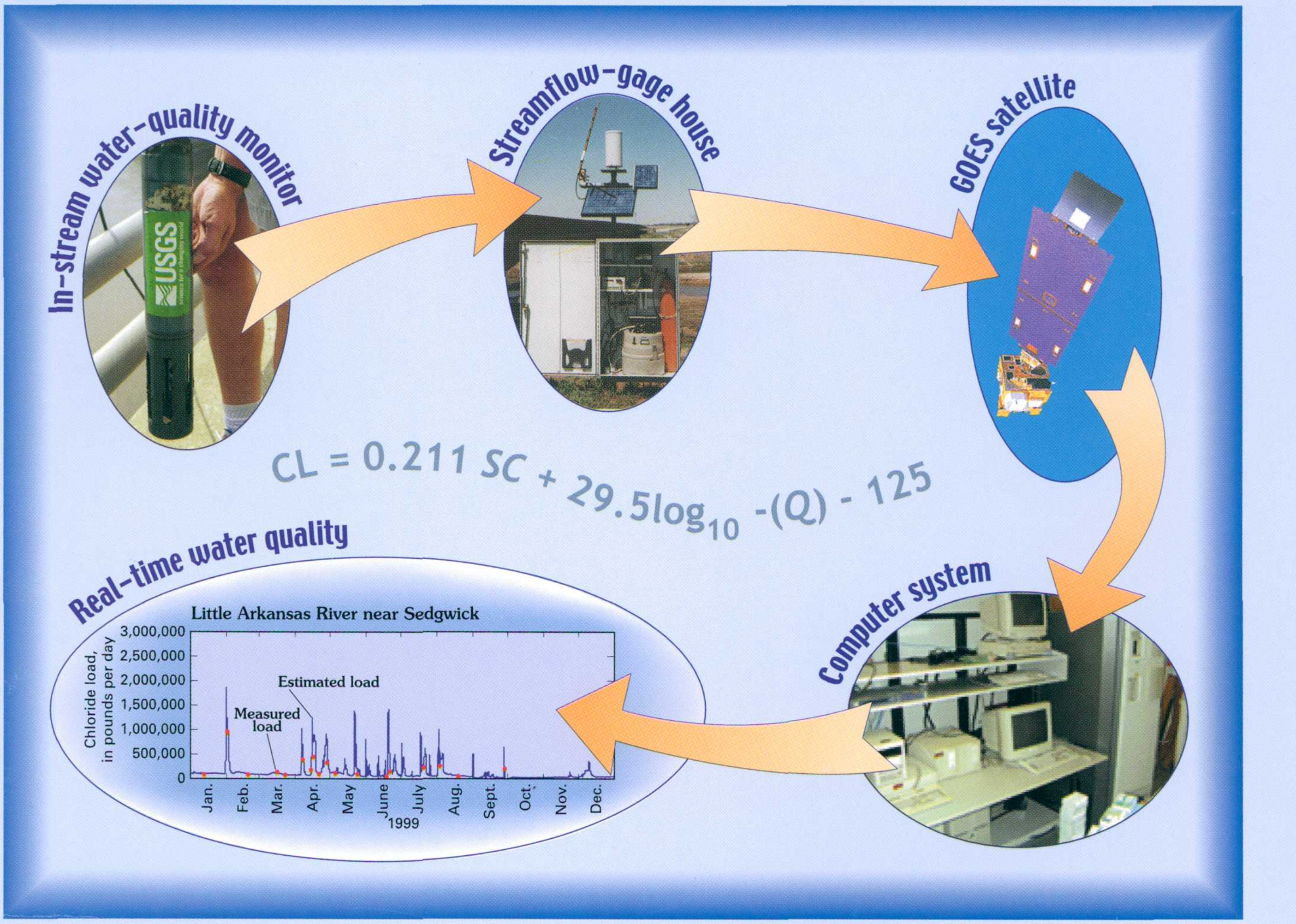
USGS
science for a changing world

Recd
8/00

Prepared in cooperation with the
CITY OF WICHITA, KANSAS, as part of the
Equus Beds Ground-Water Recharge Demonstration Project

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

Water-Resources Investigations Report 00–4126



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

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

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

By VICTORIA G. CHRISTENSEN, XIAODONG JIAN, and ANDREW C. ZIEGLER

Water-Resources Investigations Report 00–4126

Prepared in cooperation with the
CITY OF WICHITA, KANSAS, as part of the
Equus Beds Ground-Water Recharge Demonstration Project

Lawrence, Kansas
2000

U.S. Department of the Interior
Bruce Babbitt, Secretary

U.S. Geological Survey
Charles G. Groat, Director

For additional information write to:

District Chief
U.S. Geological Survey
4821 Quail Crest Place
Lawrence, KS 66049–3839

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Building 810, Federal Center
Box 25286
Denver, CO 80225–0286

CONTENTS

FIGURES

CONTENTS—Continued

FIGURES—Continued

TABLES

CONVERSION FACTORS, ABBREVIATIONS, and DEFINITIONS

Multiply	By	To obtain
acre	4,047	square meter
cubic foot per acre (ft^3/acre)	0.06997	cubic meter per hectare
cubic foot per second (ft^3/s)	0.02832	cubic meter per second
foot (ft)	0.3048	meter
mile (mi)	1.609	kilometer
pound (lb)	453.6	gram
pound per acre (lb/acre)	1.121	kilogram per hectare
pound per day (lb/d)	453.6	gram per day
square mile (mi^2)	2.590	square kilometer

Temperature can be converted to degrees Celsius (°C) or degrees Fahrenheit (°F) by the equations:

$$^{\circ}C=5/9\ (^{\circ}F - 32)$$
$$^{\circ}F=9/5\ (^{\circ}C) + 32.$$

Water year is the 12-month period from October 1 through September 30. The water year is designated by the calendar year in which it ends. For example, the period October 1, 1998, through September 30, 1999, is called the 1999 water year.

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

By Victoria G. Christensen, Xiaodong Jian, *and* Andrew C. Ziegler

Abstract

Water from the Little Arkansas River is used as source water for artificial recharge to the *Equus* Beds aquifer, which provides water for the city of Wichita in south-central Kansas. To assess the quality of the source water, continuous in-stream water-quality monitors were installed at two U.S. Geological Survey stream-gaging stations to provide real-time measurement of specific conductance, pH, water temperature, dissolved oxygen, and turbidity in the Little Arkansas River. In addition, periodic water samples were collected manually and analyzed for selected constituents, including alkalinity, dissolved solids, total suspended solids, chloride, sulfate, atrazine, and fecal coliform bacteria. However, these periodic samples do not provide real-time data on which to base aquifer-recharge operational decisions to prevent degradation of the *Equus* Beds aquifer.

Continuous and periodic monitoring enabled identification of seasonal trends in selected physical properties and chemical constituents and estimation of chemical mass transported in the Little Arkansas River. Identification of seasonal trends was especially important because high streamflows have a substantial effect on chemical loads and because concentration data from manually collected samples often were not available. Therefore, real-time water-quality monitoring of surrogates for the estimation of selected chemical constituents in streamflow can increase the accuracy of load and yield estimates and can decrease some manual data-collection activities.

Regression equations, which were based on physical properties and analysis of water samples collected from 1995 through 1998 throughout 95 percent of the stream's flow duration, were developed to estimate alkalinity, dissolved solids, total suspended solids, chloride, sulfate, atrazine, and fecal coliform bacteria concentrations. Error was evaluated for the first year of data collection and each subsequent year, and a decrease in error was observed as the number of samples increased. Generally, 2 years of data (35 to 55 samples) collected throughout 90 to 95 percent of the stream's flow duration were sufficient to define the relation between a constituent and its surrogate(s). Relations and resulting equations were site specific.

To test the regression equations developed from the first 3 years of data collection (1995–98), the equations were applied to the fourth year of data collection (1999) to calculate estimated constituent loads and the errors associated with these loads. Median relative percentage differences between measured constituent loads determined using the analysis of periodic, manual water samples and estimated constituent loads were less than 25 percent for alkalinity, dissolved solids, chloride, and sulfate. The percentage differences for total suspended solids, atrazine, and bacteria loads were more than 25 percent.

Even for those constituents with large relative percentage differences between the measured and

estimated loads, the estimation of constituent concentrations with regression analysis and real-time water-quality monitoring has numerous advantages over periodic manual sampling. The timely availability of bacteria and other constituent data may be important when considering recreation and the whole-body contact criteria established by the Kansas Department of Health and Environment for a specific water body. In addition, water suppliers would have timely information to use in adjusting water-treatment strategies; environmental changes could be assessed in time to prevent negative effects on fish or other aquatic life; and officials for the *Equus* Beds Ground-Water Recharge Demonstration project could use this information to prevent the possible degradation of the *Equus* Beds aquifer by choosing not to recharge when constituent concentrations in the source water are large.

Constituent loads calculated from the regression equations may be useful for calculating total maximum daily loads (TMDL's), which States are mandated to establish for stream segments that have been identified by the 1972 Clean Water Act as limited for specific uses because of water-quality concerns. Additionally, information on constituent loads and yields can be used to indicate in which subbasin to concentrate efforts with regard to land-resource best-management practices. These results have application anywhere constituent concentrations, loads, or transport are of concern.

INTRODUCTION

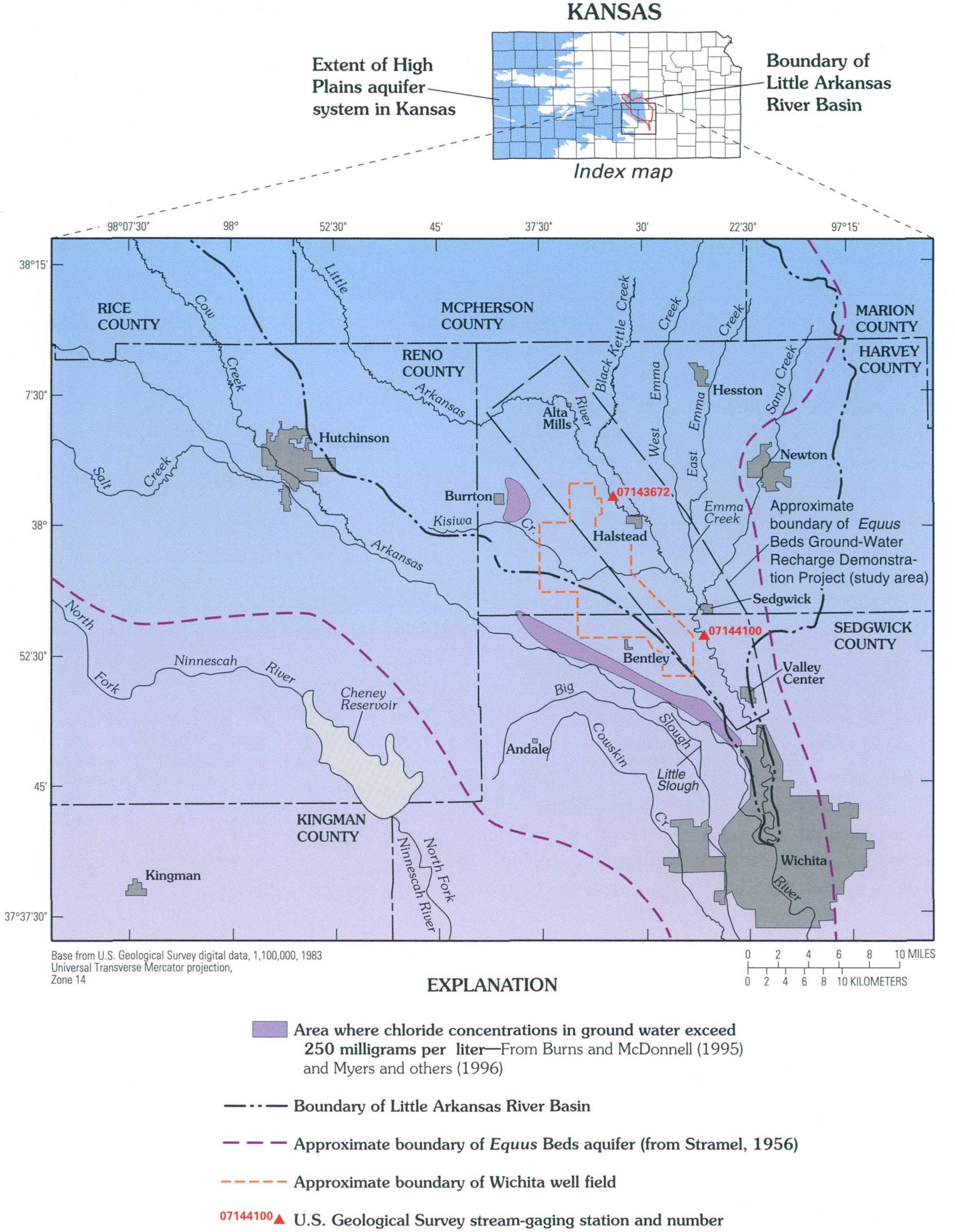
The Wichita well field (fig. 1), initiated in the 1940's and completed in the 1950's in the *Equus* Beds aquifer, is one of the primary sources of water for the city of Wichita and the surrounding area in south-central Kansas. Historical water use for municipal supply and irrigation caused water levels in the *Equus* Beds aquifer (easternmost part of High Plains aquifer in Kansas) to decline as much as 30 ft by 1993 (Aucott and others, 1998). Lower water levels not only represent a diminished water supply but also encourage saltwater intrusion as indicated in figure 1 by chloride concentrations greater than 250 milligrams per liter (mg/L) from the Burrton oil field to the northwest and from the Arkansas River to the southwest into the freshwater of the *Equus* Beds aquifer (Myers and others, 1996).

In 1995 a project was begun to investigate the feasibility of artificially recharging the *Equus* Beds aquifer with surface water from the Little Arkansas River as one alternative to meet future water-supply needs and to protect the aquifer from saltwater intrusion from natural and human-related sources. The *Equus* Beds Ground-Water Recharge Demonstration Project will use high flows from the Little Arkansas River to recharge the aquifer through various techniques, including recharge basins, a trench, and a recharge well.

The *Equus* Beds Ground-Water Recharge Demonstration Project is a collaborative effort among the city of Wichita, the Bureau of Reclamation, U.S. Department of the Interior, the U.S. Geological Survey (USGS), *Equus* Beds Groundwater Management District No. 2 (Halstead, Kansas), the U.S. Environmental Protection Agency (USEPA), the Kansas Department of Health and Environment (KDHE), the Kansas Water Office, and the Kansas Department of Agriculture, Division of Water Resources. Burns and McDonnell Engineering Consultants (Kansas City, Missouri) and Mid-Kansas Engineering Consultants (Wichita, Kansas) provide engineering expertise and project management. The maintenance and operation of the recharge facilities are performed by the city of Wichita.

The *Equus* Beds Ground-Water Recharge Demonstration Project is a part of the High Plains States Ground-Water Recharge Demonstration Program, which is a cooperative effort among the Bureau of Reclamation, USGS, and USEPA to study the potential for artificial recharge and its effects in 17 Western States. The USGS also has worked cooperatively with the city of Wichita for many years in evaluating the ground-water system's interaction with streams in the area to further the understanding of the entire hydrologic system and to provide information to improve local decisionmaking (Myers and others, 1996; Aucott and others, 1998; Christensen and Ziegler, 1998a; Ziegler and others, 1999).

During the recharge project, surface water at two USGS stream-gaging stations on the Little Arkansas River near Halstead and Sedgwick (fig. 1) is sampled to assess the quality of the source water. Real-time water-quality monitors at the two stations provide a



EXPLANATION

▨ **Area where chloride concentrations in ground water exceed 250 milligrams per liter**—From Burns and McDonnell (1995) and Myers and others (1996)

—·· — **Boundary of Little Arkansas River Basin**

— — **Approximate boundary of *Equus* Beds aquifer (from Stramel, 1956)**

- - - - **Approximate boundary of Wichita well field**

07144100▲ **U.S. Geological Survey stream-gaging station and number**

Figure 1. Location of Wichita well field, *Equus* Beds Ground-Water Recharge Demonstration Project in south-central Kansas, Little Arkansas River, and areas where chloride concentrations in ground water exceed 250 milligrams per liter.

continuous measure of specific conductance, pH, water temperature, dissolved oxygen, and turbidity. In addition, water samples are collected and analyzed at least monthly throughout 95 percent of the stream's flow duration (Ziegler and Combs, 1997).

During a recent study conducted as part of the *Equus* Beds Ground-Water Recharge Demonstration Project (Ziegler and others, 1999), chloride, atrazine, and fecal coliform bacteria were identified as constituents of concern. Alkalinity, dissolved solids, total suspended solids (TSS), and sulfate are important indicators of surface-water quality as well and potentially could have a detrimental effect on recharge. Although numerous surface-water samples are collected throughout the year as a part of the project for the analysis of chloride, atrazine, and other selected constituents, a continuous record is not available with which to make decisions regarding recharge operations. Regression equations that are based on surrogate physical properties measured in real time would be useful to estimate alkalinity, dissolved solids, TSS, chloride, sulfate, atrazine, and bacteria loads to support these decisions. A surrogate is a sensor measurement (for example, a physical property of water such as temperature) that is monitored continually in stream and may be used to estimate concentrations of a particular water-quality constituent for which continual data are not available.

In addition to the utility of surrogate relations to recharge operations, Section 303 (d) of the 1972 Clean Water Act requires States to establish total maximum daily loads (TMDL's). A TMDL is a calculation of the maximum amount of a constituent that a water body can receive and still meet water-quality standards. TMDL's may be based on surrogate relations, making a continuous record useful as it relates to TMDL's and water-quality standards in streams.

KDHE has included stream segments in the Little Arkansas River Basin on a list of water-quality-limited surface waters. The constituents that are water-quality limiting are dissolved oxygen, chloride, fluoride, sulfate, total ammonia, chlordane, and fecal coliform bacteria (Kansas Department of Health and Environment, 1999). Information on mass loading provided by real-time monitoring can be used to determine the effectiveness of implemented watershed management practices as a part of the TMDL process.

Purpose and Scope

This study is part of the *Equus* Beds Ground-Water Recharge Demonstration Project and is a cooperative effort between USGS and the city of Wichita. The purpose of the study is to develop regression equations to be used to estimate constituent concentrations, loads, and yields for the Little Arkansas River.

The purpose of this report is to present regression equations that use selected sensor measurements as surrogates to estimate concentrations, loads, and yields of alkalinity, dissolved solids, TSS, chloride, sulfate, atrazine, and fecal coliform bacteria in the Little Arkansas River. Separate sets of equations were developed for the Halstead and Sedgwick streamgaging stations to assist the city of Wichita in determining if water quality is adequate for recharge.

Regression equations are evaluated for the purpose of determining the decrease in error with each additional year of data collection. The purpose of this analysis is to determine the amount of data (number of samples) necessary to calculate a regression equation with a good statistical fit for each constituent. A comparison of the errors among the constituents and between the gaging stations is discussed.

Real-time monitoring of constituent properties in the source water potentially can improve the effectiveness of the current water-quality monitoring program for the *Equus* Beds Ground-Water Recharge Demonstration Project and is important to the maintenance of good quality water in the *Equus* Beds aquifer. With real-time monitoring and the application of regression equations, an undesirable level of a constituent in source water potentially could be identified almost immediately and action taken to either treat the water before recharge or the decision could be made not to recharge until water-quality conditions improved. To test the regression equations developed from the first 3 years of data collection (1995–98), the equations were applied to the fourth year of data collection (1999) to estimate constituent loads and yields and the errors associated with these loads and yields.

The increasing public interest in TMDL's and water quality in general makes this study of regional, as well as national, importance because it shows how constituent loads may be calculated using real-time surrogate measurement of water-quality constituents. The methods used in this study may be used for other sites in Kansas and the Nation to provide input data for the development of TMDL's and to monitor future effectiveness of implemented land-resource best-management practices. Additionally, this approach can be used by water suppliers to anticipate treatment needs for water supplies.

Description of Study Area

The study area is located in south-central Kansas and includes parts of the Wichita well field and the Little Arkansas River between Alta Mills and Valley Center (fig. 1). The two data-collection sites established for this study are USGS stream-gaging stations on the Little Arkansas River at Highway 50 near Halstead (station 07143672) and Little Arkansas River near Sedgwick (station 07144100) (fig. 1). The Halstead stream-gaging station has a contributing drainage area of about 685 mi^2. The stream-gaging station near Sedgwick has a contributing drainage area of about 1,165 mi^2. Streamflow at both stream-gaging stations is affected by ground-water withdrawals, surface-water diversions, and return flow from irrigated areas (Putnam and others, 2000).

Land use in the Little Arkansas River Basin is primarily agricultural (fig. 2), with about 78 percent cropland and 19 percent grassland (Christensen and Ziegler, 1998a). The percentages of cropland in each subbasin are similar. The percentages of each crop type for the Halstead subbasin are about 63 percent wheat, 24 percent sorghum, 7 percent soybeans, and 5 percent corn. In the Sedgwick subbasin, the percentages of crop type are about 60 percent wheat, 25 percent sorghum, 8 percent soybeans, and 6 percent corn. Crop data were estimated from county data compiled by the Kansas Department of Agriculture and U.S. Department of Agriculture (1999). Agricultural chemicals applied to enhance crop production in the area include fertilizers (such as nitrate, ammonia, and phosphorus) and pesticides (primarily alachlor and atrazine) (Kansas Department of Agriculture and U.S. Department of Agriculture, 1999).

Rock formations are not well exposed in the study area because of a lack of topographic relief and the easily erodible nature of the rocks. The rocks that appear at the land surface include the Wellington Formation and the Ninnescah Shale of Permian age. Unconsolidated deposits of clay, silt, sand, and gravel also appear at the surface (Lane and Miller, 1965).

The *Equus* Beds aquifer, a part of the High Plains aquifer, consists of alluvial deposits of sand and gravel interbedded with clay or silt. In the study area, the general direction of ground-water movement in the *Equus* Beds aquifer is to the east (Aucott and others, 1998). However, in the vicinity of the Wichita well field and the Little Arkansas River, ground-water movement has been altered by pumping wells and a low-head dam on the river (fig. 2, inset map). The Little Arkansas River is primarily a gaining stream within the study area as indicated by higher water levels in wells adjacent to the stream than in the stream (Myers and others, 1996; Aucott and others, 1998). This is not the case, however, near Halstead where a low-head dam about 2 river mi downstream from the Halstead stream-gaging station (07143672, fig. 2) causes higher water levels in the stream than in the adjacent aquifer and results in stream-water recharge of the aquifer in this vicinity.

METHODS

For the *Equus* Beds Ground-Water Recharge Demonstration Project, surface-water quality and quantity are monitored frequently according to a monitoring plan established in consultation with State and Federal agencies. For additional information on the monitoring plan and other methods used during the project, refer to Ziegler and Combs (1997) and Ziegler and others (1999). Specific methods relevant to this study are presented in the following sections.

Streamflow Measurements

Stream water-surface elevation (stage) was measured to the nearest 0.01 ft at the Halstead and Sedgwick stream-gaging stations with nonsubmersible, pressure transducers (Buchanan and Somers, 1968). Stage data were electronically recorded and transmitted by satellite in 15-minute increments to a downlink site and then to the computer at the USGS office in Lawrence, Kansas.

Methods used to determine streamflow are described in Buchanan and Somers (1969). Streamflow measurements were made at least monthly. A stage-streamflow relation was developed on the basis of streamflow measurements and the stage of the stream at the time of measurement (Kennedy, 1983; 1984).

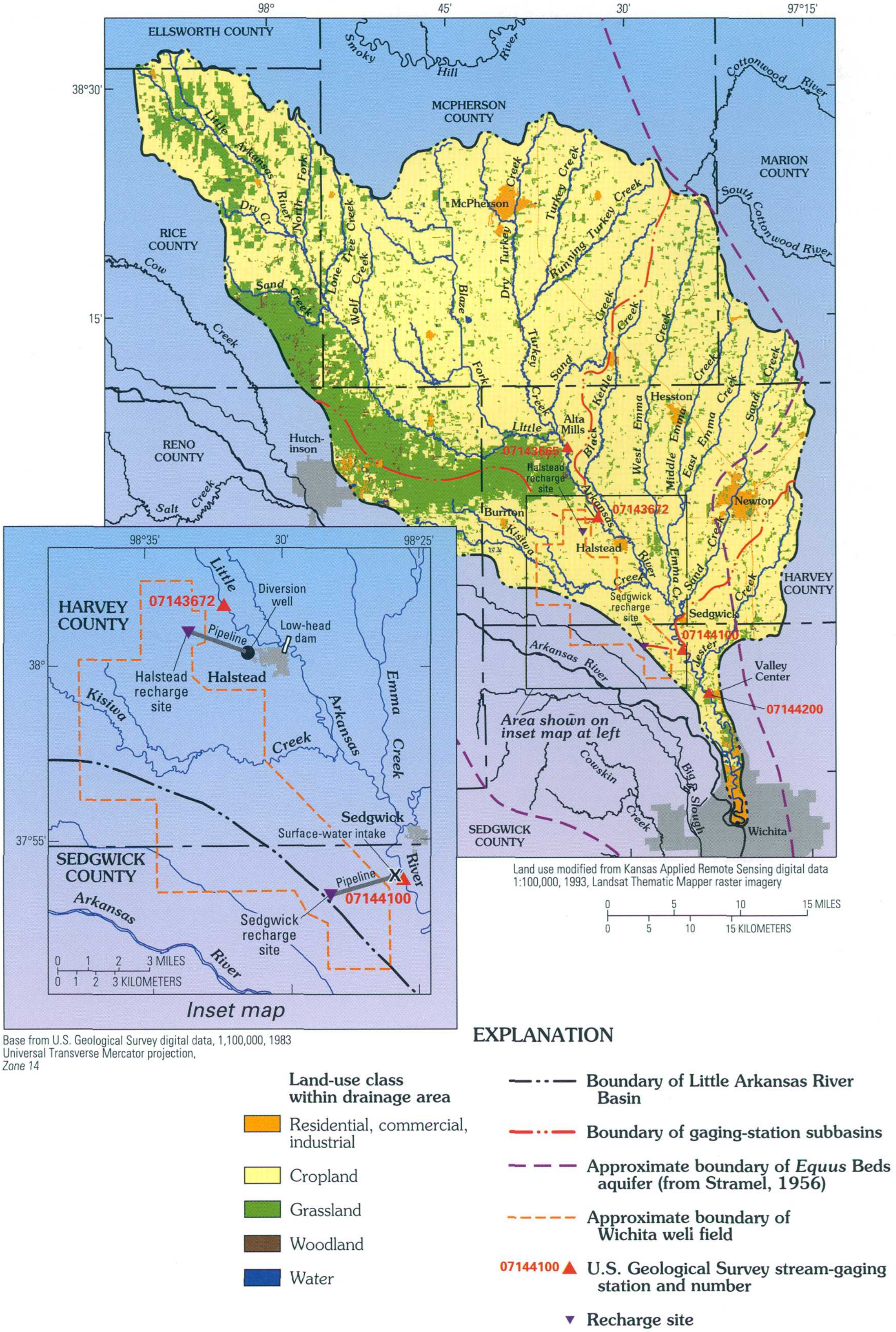


Figure 2. Land use in Little Arkansas River Basin and location of pertinent boundaries and sites.

Manual Sample Collection and Analysis

Water-quality samples were collected manually according to methods described in Ziegler and Combs (1997) and analyzed at the city of Wichita laboratory, the USGS National Water-Quality Laboratory (NWQL) in Denver, Colorado, and (or) the USGS laboratory in Lawrence, Kansas. Included were samples for the analysis of alkalinity, dissolved solids, TSS, chloride, sulfate, atrazine, triazine, and fecal coliform bacteria. In addition, water samples also were analyzed onsite at the time of collection for specific conductance, pH, water temperature, dissolved oxygen, and turbidity.

The collection of samples for most water-quality constituents was completed using depth- and width-integrating techniques; however, automated sampling equipment also was used for the collection of samples analyzed for triazine and atrazine concentrations. For the analysis of triazine compounds, the use of automated equipment for sample collection and enzyme-linked immunosorbent assay (ELISA) for herbicide analysis allowed for the low cost and efficient collection of about 2,500 samples from February 1995 through December 1998. A large number of samples was needed for triazine analysis because of substantial temporal variability. Gas chromatography/mass spectrometry (GC/MS) confirmation of selected analytical results indicated that at least 80 percent of the triazine herbicide concentration determined by ELISA was atrazine (Christensen and Ziegler, 1998b).

Because hydrologic conditions within a watershed may affect the variability of many chemical constituents, water samples also were collected manually throughout a range of streamflow conditions. Periodic monitoring enabled the estimation of seasonal trends in physical properties and chemical constituents. Figure 3*A* shows flow-duration curves for the Little Arkansas River at Alta Mills (station 07143665, fig. 2) and at Highway 50 near Halstead, Kansas (station 07143672, fig. 2). The stream-gaging station at Alta Mills is 8.7 river mi upstream from the Halstead stream-gaging station (Putnam and others, 2000). Because the period of record for the Halstead stream-gaging station is relatively short (1995–99), the flow-duration curve for the Alta Mills stream-gaging station provides an indication of the long-term flow duration (1973–99) for that segment of the Little Arkansas River. Figure 3*A* also shows the streamflows at which samples were collected. Figure 2*B* shows flow-duration curves for the Little Arkansas River near Sedgwick (station 07144100, fig. 2) and at Valley Center, Kansas (station 07144200, fig. 2). The stream-gaging station at Valley Center is 6.2 river mi downstream from the Sedgwick stream-gaging station (period of record 1995–99) and is more representative of long-term (1922–99) flow duration on that segment of the Little Arkansas River.

Real-Time Water-Quality Monitoring

Not only is there a desire to monitor the quality of water in the Little Arkansas River, but because the Little Arkansas River is the source for water used to recharge the *Equus* Beds aquifer, there also is a need to provide water-quality information immediately during recharge periods before the quality of the aquifer water is affected. In addition to collecting the periodic, manual water samples discussed previously, continuous, real-time water-quality monitoring was used to document physical properties of the water.

Real-time monitoring of specific conductance, pH, and temperature of the source water began in May 1998 at the Halstead stream-gaging station and in April 1998 at the Sedgwick stream-gaging station. Monitoring for dissolved oxygen and turbidity at both stream-gaging stations began in October 1998. The water-quality monitors installed at these two stream-gaging stations have multiple sensors that measure physical properties of water quality. The sensors measure specific conductance, pH, water temperature, dissolved oxygen, and turbidity. These devices are connected to data-collection platforms at each stream-gaging station, and data are transmitted by satellite to a downlink site and to the computer at the USGS office in Lawrence, Kansas (see report cover). The sensors were calibrated according to methods presented in Wilde and Radke (1998).

Development of Regression Equations to Estimate Constituent Concentrations

The concentrations of constituents in surface water often are strongly related to other constituent concentrations and factors such as hydrologic conditions, season, and location. It is possible to express one constituent concentration in terms of another constituent or constituents using simple regression equations (Helsel and Hirsch, 1992). The simplest regression equation can be expressed as:

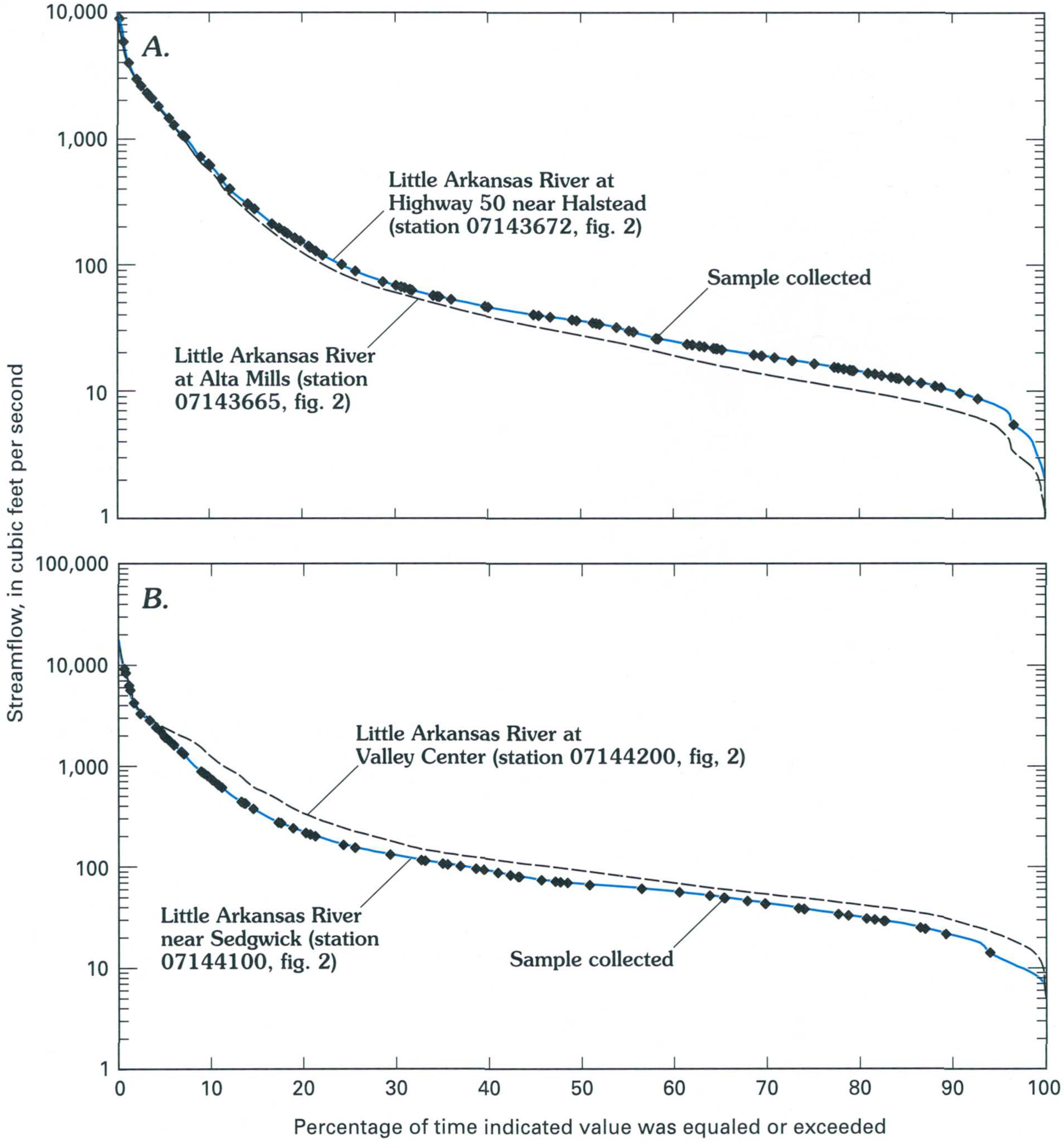


Figure 3. Flow-duration curves and samples collected from Little Arkansas River *(A)* at Alta Mills (1973–99), at Highway 50 near Halstead (1995–99), *(B)* near Sedgwick (1995–99), and at Valley Center (1922–99), Kansas. Samples collected represent instantaneous streamflow at which each sample was collected.

$$y_i = mx_i + b + e_i \qquad i = 1, 2, \ldots, n, \qquad (1)$$

where

y_i is the *i* th observation of the dependent variable;

m is the slope;

x_i is the *i* th observation of the independent variable;

b is the intercept;

e_i is the random error for the *i* th observation; and

n is the sample size.

The terms m and b represent the parameters that need to be estimated from the data set. The most common estimation technique is called least squares (Helsel and Hirsch, 1992). In least-squares estimation, the error term, e_i, is assumed to be normally distributed with a mean equal to zero and constant variance, σ^2.

To determine which independent variable or variables (x) to include in each regression equation, a stepwise procedure was used (Ott, 1993, p. 656). Each independent variable was added to the model one at a time to determine if there was a significant correlation. The possible independent variables included each of the sensor measurements (specific conductance, pH, water temperature, dissolved oxygen, and turbidity), streamflow (discharge), and time. Correlations were considered significant if the coefficient of determination, R^2, improved by at least 10 percent.

There are three measures commonly used to evaluate regression equations. These are the sum of squares of errors (*SSE*), the sum of squares of y about the mean (*SSy*), and the coefficient of determination (R^2). The least-squares estimators in equation 1 (m and b) are obtained by minimizing the *SSE*, which is calculated as follows:

$$SSE = \sum_{i=1}^{n} [y_i - E(y_i)]^2, \qquad (2)$$

where y_i represents the value of y at the i th data point, and $E(y_i)$ is the estimated value of y at the i th data point (where: $E(y_i) = mx_i + b$). *SSE* is a dimensional measure of fitting y on x and is a measure of the unexplained variability. The *SSy* represents the total variability (explained and unexplained) about the mean in y values and is calculated as follows:

$$SSy = \sum_{i=1}^{n} (y_i - \bar{y})^2, \qquad (3)$$

in which, $\bar{y}$ is the mean of y.

Both *SSE* and *SSy* are dimensional measures. Dimensionless measures often are required in practice for the purpose of comparing constituents with different dimensions (units of measure). A dimensionless measure of fitting y on x is the R^2, or the fraction of the variance explained by regression:

$$R^2 = 1.0 - (SSE/SSy). \qquad (4)$$

The larger the explained variability is compared to the unexplained variability, the better the equation fits the data, and this should lead to a more precise prediction of y (Ott, 1993). The R^2 ranges from 0 to 1 and often is called the multiple coefficient of determination in multiple linear regression. For the purpose of this study, if R^2 is greater than 0.50, the determination is considered acceptable.

Another measure often used to explain variability in regression equations is mean square error (*MSE*). *MSE* is calculated as follows:

$$MSE = \sqrt{\frac{SSE}{n-2}}. \qquad (5)$$

The *MSE* is presented for each equation to assess the variance between predicted and observed values.

Regression equations can be divided further into two groups, linear and nonlinear. Whether an equation is linear or nonlinear depends on parameters in the equation. A nonlinear regression equation is one in which at least one of the partial derivatives of y with respect to m and b is a function of at least one of the parameters, m or b (for example, $y = mx^c + b$). Linear regression is mathematically much easier to compute because the estimators of the parameters are from an explicit mathematical expression (for example, $y = mx + b$).

Graphical plots were constructed to determine linearity. For certain equations, either the independent variable, dependent variable, or both were transformed to convert all equations presented herein to linear equations. Log transformations of variables can eliminate curvature in the data and simplify the analysis of the data (Ott, 1993, p. 454). For example, by taking the natural log of an independent variable in a nonlinear regression equation, it is possible to achieve a simpler linear equation. Outliers were identified graphically and generally were those points where either the dependent variable or independent variable was more than two standard deviations away from the mean. Each outlier was investigated to determine a possible reason for its occurrence. In all cases, only local outliers (for example, more than two standard deviations away from the dependent variable mean for a fixed range of the independent variable) were removed from the data set. Outliers that were not used in the development of each regression equation are discussed in the "Results..." section. Linear regression equations determined with the least-squares method were developed using the SAS System statistical software (SAS Institute Inc., 1996, release 6.12). Selected regression equations also were verified for differences in calculation accuracy using S-Plus 2000 statistical software (Math Soft, Inc., 1999).

To determine the effect of sample size on the error between measured and estimated concentrations, estimated concentrations were calculated with the given equations using the first year of data (for example, 1995). The estimated concentration then was calculated for each subsequent year using the cumulative data set (for example, all data from 1995 through 1996). The *SSE* for each year of cumulative data through December 1998 was calculated. The change in the *SSE* between 2 years of data collection was determined with the absolute value of the equation:

$$(A-B)/A \times 100, \qquad (6)$$

where

A is *SSE* for the first year of data; and

B is *SSE* for the subsequent year of cumulative data.

Calculation of Measured and Estimated Constituent Loads and Yields

Measured and estimated mean daily constituent loads were calculated by multiplying instantaneous constituent concentrations by instantaneous streamflow and an appropriate conversion factor (table 1). Annual loads were calculated by multiplying the mean daily loads by 365. Measured and estimated yields (in pounds per acre or billions of colonies per acre) were calculated by dividing the instantaneous constituent load by the stream-gaging-station contributing-subbasin area, in acres.

The relative percentage differences (*RPD*'s) between measured and estimated constituent loads were calculated for each instantaneous data set using the absolute value of the following equation:

$$RPD = (B-A)/A \times 100, \qquad (7)$$

Table 1. Conversion factors used in calculation of measured and estimated loads

[ft^3/s, cubic feet per second]

Multiply	By	By	To obtain
colonies per 100 milliliters (col/100 mL)	0.02447	streamflow, in ft^3/s	billion colonies per day
micrograms per liter (µg/L)	0.00539	streamflow, in ft^3/s	pounds per day
milligrams per liter (mg/L)	5.39	streamflow, in ft^3/s	pounds per day

where

RPD is the relative percentage difference;

A is the measured constituent load; and

B is the estimated constituent load.

RESULTS OF REGRESSION ANALYSIS AND EFFECT OF SAMPLE SIZE FOR SELECTED CONSTITUENTS

The relation between concentrations of each constituent of concern and surrogate physical property(s) was determined, and a regression equation was developed (table 2). *MSE* and R^2 were calculated for each equation. A discussion of each constituent, the associated regression equations for the Halstead and Sedgwick stream-gaging stations, and the effect of sample size follows.

Alkalinity

The alkalinity of water may be defined as the capacity for solutes in the water to react with and neutralize acid. Alkalinity is important in determining a stream's ability to neutralize acidic pollution from rainfall or wastewater. It is one of the best measures of the sensitivity of the stream to acid inputs (U.S. Environmental Protection Agency, 1999). Streams with low alkalinity are not well buffered and often are affected adversely by acid inputs. The pH levels in low-alkalinity streams can decrease to a point that eliminates acid-intolerant forms of aquatic life. Fish are particularly affected by low pH water.

Table 2. Regression equations for estimates of alkalinity, dissolved solids, total suspended solids, chloride, sulfate, atrazine, and fecal coliform bacteria in Little Arkansas River at Highway 50 near Halstead and near Sedgwick, Kansas, February 1995–December 1998

[*MSE*, mean square error; R^2, coefficient of determination; *SC*, specific conductance; *Q*, streamflow; *Turb*, turbidity]

Constituent	Equation	*MSE*	R^2
	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)		
Alkalinity (*Alk*)	$\log_{10}(Alk) = 0.651\log_{10}SC - 0.101\log_{10}(Q) + 0.487$	0.0667	0.946
Dissolved solids (*DS*)	$DS = 0.545SC + 33.3$	49.4	.976
Total suspended solids (*TSS*)	$\log_{10}(TSS) = 0.920\log_{10}(Turb) + 0.243$	.193	.911
Chloride (*Cl*)	$Cl = 0.255SC + 30.9\log_{10}(Q) - 140$	16.4	.985
Sulfate (SO_4)	$\log_{10}SO_4 = 0.911\log_{10}(SC) - 1.12$	.0891	.919
Triazine	$\log_{10}(Triazine) = (1.42)e^{\frac{-(month-6.24)^2}{3.75}} - 0.0000288Q - 0.000581SC - 0.104$	.334	.777
Fecal coliform bacteria (*Bact*)	$\log_{10}(Bact) = 0.490\cos\left(2\pi\left(\frac{month+2.06}{8.76}\right)\right) + 0.00106Turb + 0.417\log_{10}(Turb) + 1.65$	.609	.620
	Little Arkansas River near Sedgwick (station 07144100, fig. 2)		
Alkalinity (*Alk*)	$Alk = 0.141SC - 44.4\log_{10}(Q) + 167$	25.1	.912
Dissolved solids (*DS*)	$DS = 0.556SC + 18.3$	24.5	.857
Total suspended solids (*TSS*)	$\log_{10}(TSS) = 0.878\log_{10}(Turb) + 0.300$	.209	.889
Chloride (*Cl*)	$Cl = 0.211SC + 29.5\log_{10}(Q) - 125$	13.6	.954
Sulfate (SO_4)	$SO_4 = 0.0482SC - 5.50\log_{10}(Q) + 18.8$	5.57	.931
Triazine	$\log_{10}(Triazine) = (1.38)e^{-\frac{(month-5.94)^2}{3.92}} - 0.0000652Q - 0.000820SC + 0.0681$	.324	.736
Fecal coliform bacteria (*Bact*)	$\log_{10}(Bact) = -(0.286)\cos\left(2\pi\left(\frac{month-29.4}{-6.8}\right)\right) - 0.000422(Turb) + 1.26\log_{10}(Turb) + 0.519$	.784	.556

Several different solute species contribute to the alkalinity of water; however, in most natural water, alkalinity is produced by the dissolved carbon dioxide species, bicarbonate and carbonate. There are some noncarbonate contributors to alkalinity, including hydroxide, silicate, borate, and organic ligands. The alkalinity of natural water, however, can be assigned entirely to dissolved bicarbonate and carbonate (Hem, 1992), except for water with unusual chemical composition or high pH. A relation can be established between hardness (alkalinity as $CaCO_3$) and specific conductance (Hem, 1992), and because runoff can have a dilution effect on many constituents, streamflow also may have a relation to specific conductance and alkalinity.

The surrogates used for the determination of alkalinity in water from the Little Arkansas River at the Halstead stream-gaging station were specific conductance and streamflow. Ranges in the independent variable concentrations were 98.9 to 1,850 microsiemens per centimeter at 25 °C (μS/cm) for specific conductance and 5.6 to 9,220 ft^3/s for streamflow. For water at the Halstead stream-gaging station, there were no outliers recorded from February 1995 through December 1998. The multiple regression equation with logarithmic transformation was:

$$\log_{10}(ALK) = 0.651 \log_{10}(SC) - 0.101 \log_{10}(Q) + 0.487, \quad (8)$$

where *ALK* is the estimated alkalinity concentration, in milligrams per liter; *SC* is the specific conductance, in microsiemens per centimeter at 25 °C; and *Q* is discharge or streamflow, in cubic feet per second. The measured alkalinity concentration was compared to the estimated alkalinity concentration calculated on the basis of equation 8 (fig. 4*A*).

For the estimation of alkalinity in water at the Sedgwick stream-gaging station, the ranges in the indepedent variable concentrations were 76 to 1,350 μS/cm for specific conductance and 14.2 to 9,250 ft^3/s for streamflow. The multiple linear regression equation was:

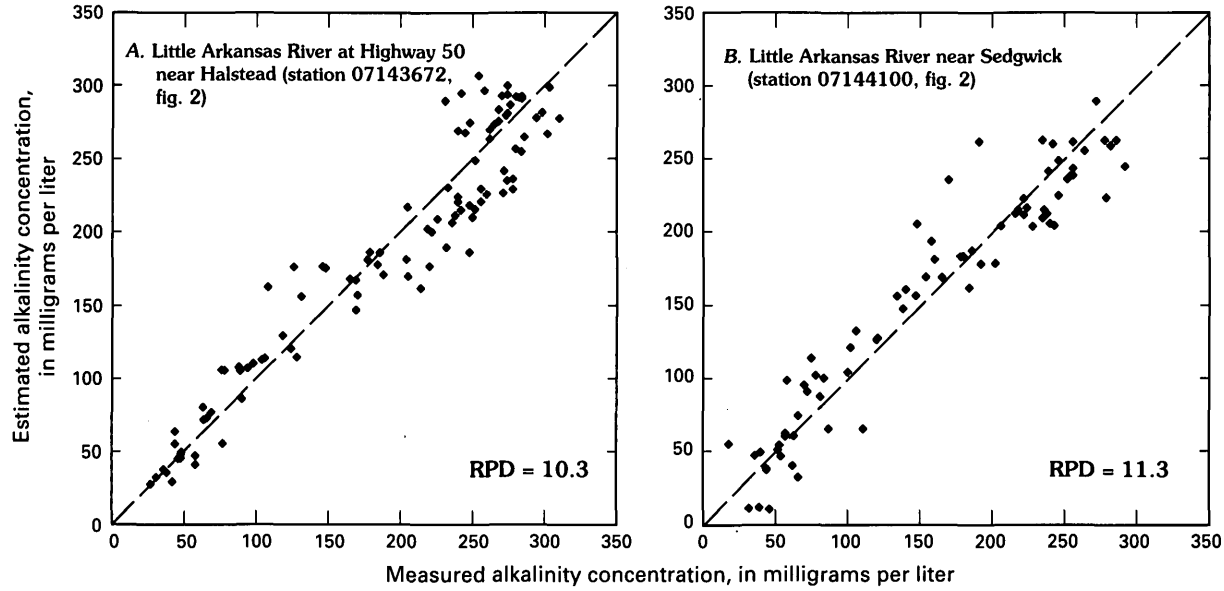
$$ALK = 0.141SC - 44.4\log_{10}(Q) + 167, \quad (9)$$

where *ALK* is the estimated alkalinity concentration, in milligrams per liter; *SC* is the specific conductance, in microsiemens per centimeter at 25 °C; and *Q* is streamflow, in cubic feet per second. There were no outlier concentrations recorded from February 1995 through December 1998. The comparison of measured and estimated alkalinity concentrations at the Sedgwick stream-gaging station is shown in figure 4*B*. The estimated alkalinity concentration has a better fit to measured alkalinity concentration for water at the Halstead stream-gaging station (fig. 4*A*) than for water at the Sedgwick stream-gaging station (fig. 4*B*). More manual samples were collected at the Halstead station (102 samples) than at the Sedgwick stream-gaging station (77 samples). The additional samples may account for the lower *RPD* associated with the Halstead equation than with the Sedgwick equation.

An evaluation of sample-size effect is given in table 3. The increase in sample size at the Halstead stream-gaging station from 1995 to 1996 improves the *SSE* by 9.06 percent. Between 1997 and 1998, the *SSE* changes by approximately the same amount (7.68 percent). At the Sedgwick stream-gaging station, the effect of an increase in sample size is apparent from 1995 to 1996, but after the initial 2 years of data collection, an increase in sample size does not appear to have much effect, with changes of 1.68 and 2.48 percent at the end of 1997 and 1998, respectively. An important note with respect to table 3 and subsequent tables is the magnitude of the difference in *SSE* between the Halstead and Sedgwick stream-gaging stations. Because the estimated alkalinity is transformed in equation 8, the *SSE* is affected and should not be compared to the *SSE* for equation 9. The purpose of table 3 is to note the percentage change in *SSE* with additional samples.

Dissolved Solids

Dissolved solids include both organic and inorganic material dissolved in a sample of water (Bates and Jackson, 1984). One of the physical properties that is related to dissolved solids is specific conductance. However, this is not a simple relation; it is dependent on the ions present in the water. Some water is characterized by well-defined relations between dissolved solids and specific conductance (Hem, 1992). At locations where this is true, the calculation of concentrations of certain constituents, such as chloride, is possible with a good level of accuracy; however, during high flows when a larger percentage of streamflow originates from precipitation than from ground water, the relation could change. Therefore, at some



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured alkalinity concentrations and instantaneous estimated alkalinity concentrations

Figure 4. Comparison of measured and estimated alkalinity concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

Table 3. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based alkalinity equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	19	0.939	0.508	--	20	0.902	52,300	--
1996	54	.944	.462	-9.06	38	.911	47,500	-9.18
1997	74	.942	.482	4.33	54	.910	48,300	1.68
1998	102	.946	.445	-7.68	77	.912	47,100	-2.48

locations, streamflow or time (season) are variables that could affect the relation.

Dissolved solids commonly are used as a general indicator of salinity or water quality. Although dissolved-solids concentration has an important effect on surface water, there is no USEPA primary drinking-water standard. The Secondary Maximum Contaminant Level (SMCL) for dissolved solids in drinking water is 500 mg/L (U.S. Environmental Protection Agency, 1995). SMCL's are not enforceable water-quality standards. Large dissolved-solids concentrations can make water cloudy and give it a bitter taste. In addition, water with large dissolved-solids concentrations can produce scaly deposits and cause staining, wear, or corrosion of pipes and fittings. Excessively large concentrations of dissolved solids are objectionable in drinking water because of possible physiological effects, unpalatable mineral tastes, and higher costs due to corrosion or the necessity for additional treatment (U.S. Environmental Protection Agency, 1986).

Dissolved solids in surface water from the Little Arkansas River at the Halstead stream-gaging station are highly correlated with specific conductance. Because of the high correlation between dissolved solids and specific conductance, the linear regression equation was developed between dissolved solids and specific conductance and does not include discharge or time (season). The range in concentration for the independent variable, specific conductance, was 98.9 to 3,550 µS/cm. The final regression equation was:

$$DS = 0.545SC + 33.3, \qquad (10)$$

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 concentration compares well with the estimated dissolved-solids concentration for water samples collected at the Halstead stream-gaging station from February 1995 through December 1998 (fig. 5*A*). There was one measurement with a specific conductance concentration of 3,550 µS/cm and a dissolved-solids concentration of 1,960 mg/L (fig. 5*A*). Although this data point is considered an outlier (both the dependent and independent variables are more than two standard deviations from the mean), the inclusion of the point did not affect the final regression equation and was not removed from the data set.

Dissolved solids also are highly correlated with specific conductance in water from the Sedgwick stream-gaging station; therefore, a linear regression equation was developed between dissolved solids and specific conductance. The range in concentration for specific conductance was 76 to 1,350 µS/cm. The final regression equation was:

$$DS = 0.556SC + 18.3, \qquad (11)$$

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 comparison of measured dissolved-solids concentrations to estimated dissolved-solids concentrations shows a good relation (fig. 5*B*). There were no outlier concentrations in water from the Sedgwick stream-gaging station from February 1995 through December 1998.

The relation between dissolved solids and specific conductance has a higher correlation in water from the Halstead stream-gaging station than in water from the Sedgwick stream-gaging station based on the R^2 (table 2). An analysis of the effect of sample size (table 4) shows that the change in *SSE* is less than

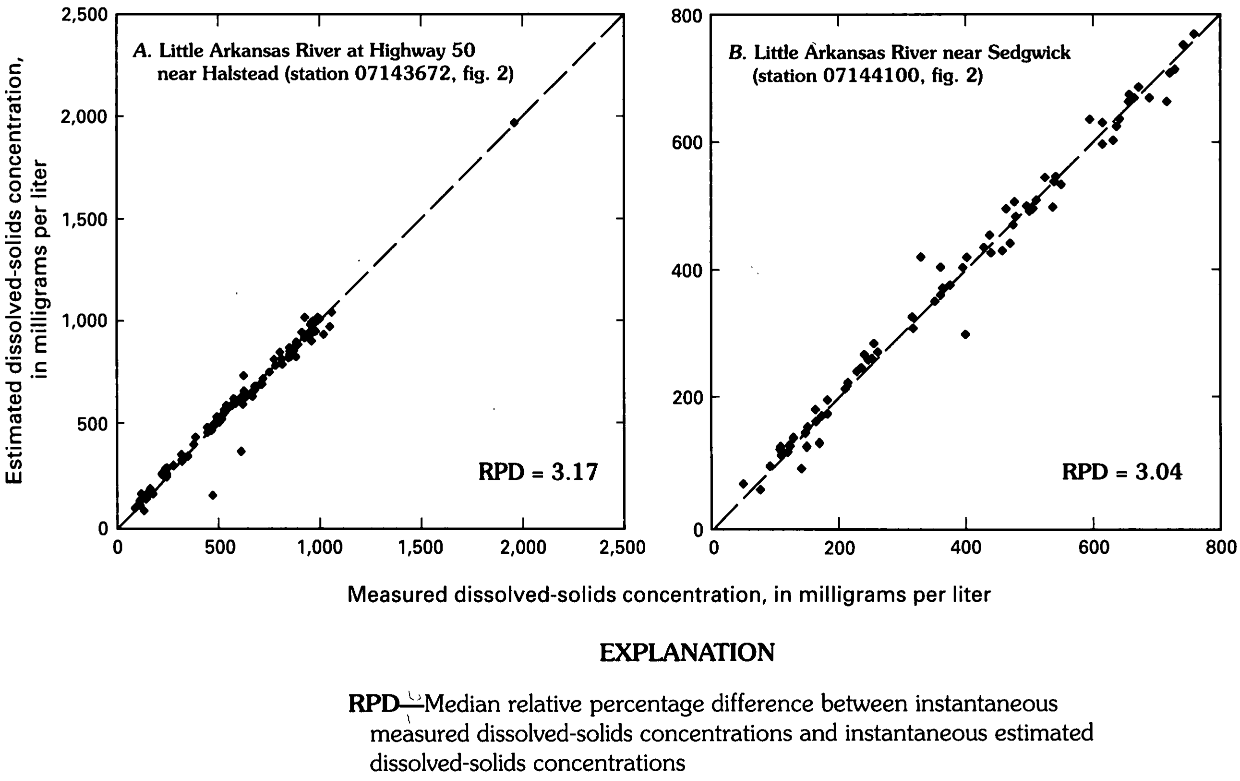


Figure 5. Comparison of measured and estimated dissolved-solids concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

Table 4. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based dissolved-solids equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	19	0.974	272,000	--	20	0.846	46,600	--
1996	54	.975	260,000	-4.41	38	.852	44,800	-3.86
1997	74	.976	248,000	-4.62	54	.855	43,700	-2.46
1998	103	.976	247,000	-.40	74	.857	43,100	-1.37

5 percent for each subsequent year of data collection for both stations. These small improvements in error indicate that 20 samples, collected throughout a range approximating historic hydrologic conditions, may be enough to define the relation between dissolved solids and specific conductance in water from these two stream-gaging stations.

Total Suspended Solids

Total suspended solids (*TSS*) include both suspended sediment and organic material collected with the water sample. Total suspended sediment may consist of clay, silt, and sand (Bates and Jackson, 1984) for which particle-size distributions often are determined. The significance of organic material with respect to the determination of *TSS* concentration is minimal in most drainage basins (Guy, 1969), However, another physical property, the effective surface area of the sediment per unit weight, has more bearing on the physical-chemical nature of the material than does the particle size alone (Hem, 1992). Certain solute-sediment interactions make total suspended sediment, and thus *TSS*, an important water-quality factor.

TSS can cause problems for fish by clogging gills and for aquatic plants by reducing light penetration and, thus, limiting growth. In addition, *TSS* provides a medium for the accumulation and transport of other constituents such as phosphorus and bacteria.

Some physical properties that are related to *TSS* are water temperature and river stage. *TSS* suspension and deposition can be affected by water temperature (or viscosity). Stream stage is important because these data are used to compute streamflow, which can affect *TSS* suspension (Edwards and Glysson, 1988). Turbidity is an indicator of sediment and other solid material transported in a stream and, therefore, has a relation to *TSS*.

Water samples may be collected from a single point in the stream section, a vertical line between the surface and the streambed, or several vertical lines across the entire stream section. A sample collected using depth- and width-integrating techniques is considered more representative of the entire cross-sectional streamflow than would a sample collected at a single point in the stream.

Manual samples were collected using depth- and width-integrating techniques, and because a real-time water-quality monitor probes the turbidity at a single point in the stream, there was some difference between turbidity of water samples collected manually and turbidity measured using the real-time water-quality monitor. There was no adjustment applied to the equations to correct for this difference.

There is a linear relation between *TSS* and turbidity after logarithm transformation. Graphical plots of this relation at the Halstead stream-gaging station indicate that there are two outliers (one with a turbidity value 0.3 nephelometric turbidity unit (NTU) and *TSS* of 262 mg/L, and one with a turbidity of 1,780 NTU and *TSS* of 92 mg/L). To reduce the effects of these outliers, these two points were not used in the regression equation. The range in turbidity concentrations used to develop the regression equation was 4.69 to 1,150 NTU. The final linear regression equation for TSS with log transformation was:

$$\log_{10}(TSS) = 0.920\log_{10}(Turb) + 0.243, \qquad (12)$$

where *TSS* is estimated total suspended solids, in milligrams per liter, and *Turb* is turbidity, in nephelometric turbidity units. A graph presenting the comparison of measured and estimated *TSS* in water from the Halstead stream-gaging station is shown in figure 6*A*.

In water from the Sedgwick stream-gaging station (fig. 6*B*), there were three local outliers with turbidity values of 3.99, 5.01, and 1.44 NTU and *TSS* values of 342, 261, and 30.8 mg/L, respectively. These outliers were not used in the development of the regression equation. The range in turbidity concentrations used in the development of the regression equation was 3.63 to 1,030 NTU. The final linear regression equation for *TSS* after log transformation was:

$$\log_{10}(TSS) = 0.878\log_{10}(Turb) + 0.300, \qquad (13)$$

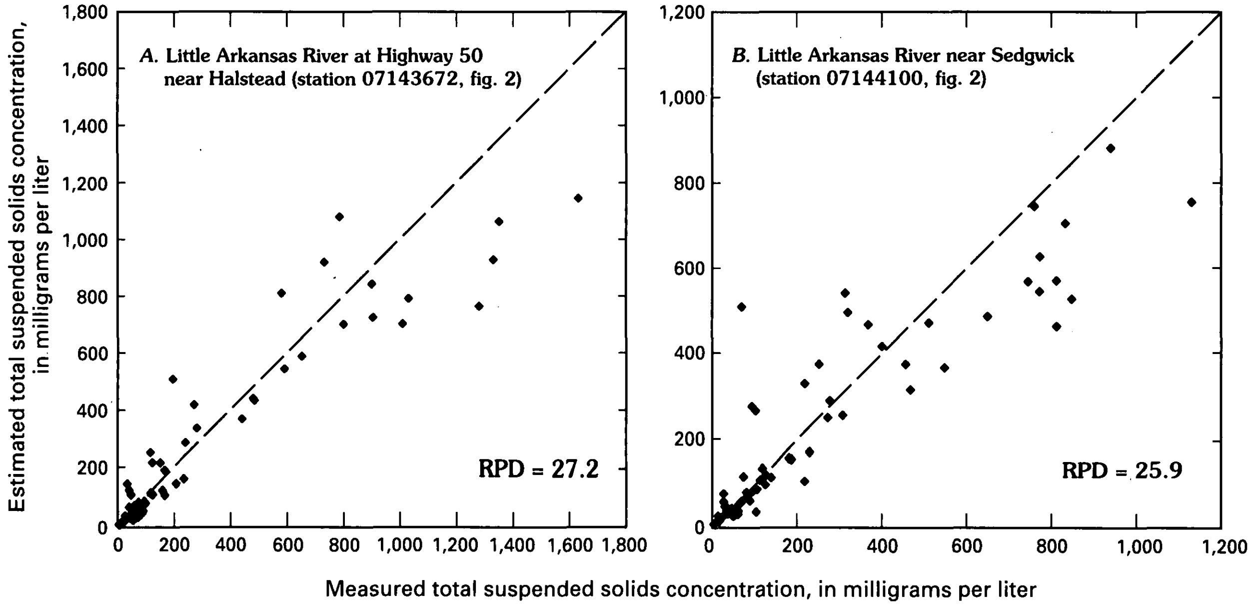
where *TSS* is estimated total suspended solids, in milligrams per liter, and *Turb* is turbidity, in nephelometric turbidity units.

The relation between *TSS* and turbidity is similar for water from the Halstead and Sedgwick stream-gaging stations. In addition, it appears there is very little improvement in *SSE* with the addition of samples (table 5). This would indicate that 20 samples, collected throughout a range approximating historic hydrologic conditions, may be enough to define the relation between *TSS* and turbidity.

Chloride

Chloride concentrations in water from the study area are naturally high, and chloride was identified by Ziegler and others (1999) as being one of the primary constituents of concern for the *Equus* Beds Ground-Water Recharge Demonstration Project. In addition to solution of natural salt, increased chloride concentrations in the area are due in part to oil-field brines resulting from past oil and gas activities (Albert, 1964). Additionally, water softening by municipalities such as McPherson and Newton (fig. 2) can contribute salts (Donald Whittemore, Kansas Geological Survey, oral commun., 1999).

Chloride is conservative (unchanging) in water; therefore, the circulation of the chloride ion is largely through physical processes rather than chemical or biochemical processes (Hem, 1992). Chloride is a charged ionic species that makes water conductive. As chloride concentrations increase, the conductivity of a solution increases (Hem, 1992); therefore, specific conductance and chloride are directly related. This relation between specific conductance and chloride in the Little Arkansas River was studied as early as 1960 with the installation of a continuously recording



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured total suspended solids concentrations and instantaneous estimated total suspended solids concentrations

Figure 6. Comparison of measured and estimated total suspended solids concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

Table 5. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based total suspended solids equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	19	0.907	2.79	--	19	0.881	3.24	--
1996	41	.908	2.78	-0.36	35	.879	3.28	1.23
1997	58	.909	2.75	-1.08	51	.885	3.12	-4.88
1998	74	.911	2.69	-2.18	71	.889	3.01	-3.53

specific-conductance monitor and the development of a relation between specific conductance and chloride (Albert, 1964).

Chloride concentrations vary seasonally in the Little Arkansas River and are larger in the winter when most of the streamflow is from ground-water discharge. The effects of chloride in ground-water discharge on water in the Little Arkansas River may be different between the Halstead and Sedgwick stream-gaging stations because of the low-head dam downstream from Halstead that causes the stream to recharge the aquifer in this location, reducing the ground-water percentage contribution to streamflow.

The multiple linear regression equation for chloride uses both specific conductance and streamflow to increase accuracy in the estimates of chloride concentrations in water at the Halstead stream-gaging station. The ranges in concentrations for the independent variables were 98.9 to 3,550 µS/cm for specific conductance and 5.61 to 9,220 ft^3/s for streamflow. The final regression equation was:

$$Cl = 0.255SC + 30.9\log_{10}(Q) - 140, \qquad (14)$$

where Cl is the estimated chloride concentration, in milligrams per liter, SC is specific conductance, in microsiemens per centimeter at 25 °C, and Q is streamflow, in cubic feet per second. There was one outlier with a specific conductance of 3,550 µS/cm and a chloride concentration of 932 mg/L. However, this was not a local outlier and did not have a substantial effect on the final regression equation; therefore, the outlier was not removed from the data set. Estimated chloride concentrations show a very good relation with measured chloride concentrations (fig. 7*A*).

The multiple linear regression equation developed for the estimation of chloride at the Sedgwick stream-gaging station also used both specific conductance and streamflow. The range in specific conductance concentrations was 93 to 1,350 µS/cm, and the range in streamflow was 14.2 to 9,250 ft^3/s. The final linear regression equation was:

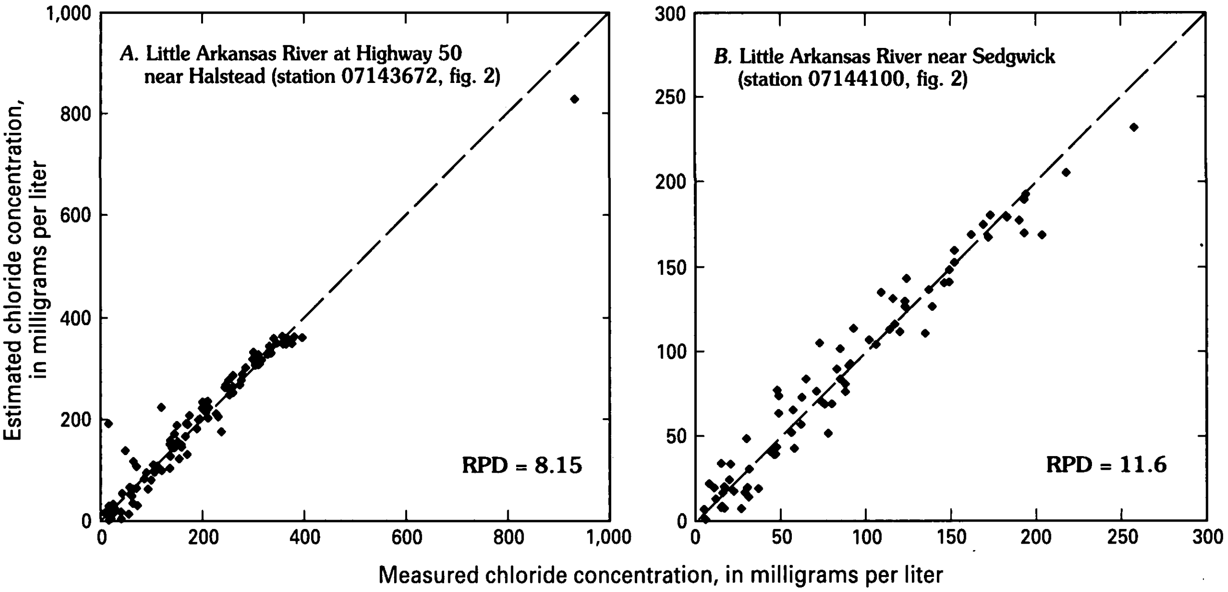
$$Cl = 0.211SC + 29.5\log_{10}(Q) - 125, \qquad (15)$$

where Cl is the estimated chloride concentration, in milligrams per liter, SC is specific conductance, in microsiemens per centimeter at 25 °C, and Q is streamflow, in cubic feet per second. There were no local outliers identified from February 1995 through December 1998. Estimated chloride concentrations compare well with measured chloride concentrations (fig. 7*B*).

The *SSE* between measured and estimated chloride concentrations improved significantly after the first year of data collection at both stream-gaging stations. Subsequent data collection had a minimal effect on improving the *SSE* (table 6) for the Sedgwick stream-gaging station; however, the *SSE* for water from the Halstead stream-gaging station continued to improve significantly. This may be evidence that chloride concentrations at the Halstead stream-gaging station are subject to more variability, and thus the collection of additional samples was necessary to define this variability.

Sulfate

Sulfate, like chloride, is a charged ionic species, and specific conductance is directly related to sulfate. The chemical behavior of sulfur is strongly related to the redox properties in the water (Hem, 1992); therefore, the pH of water is important for determining the sulfate concentration. However, pH was not strongly



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured chloride concentrations and instantaneous estimated chloride concentrations

Figure 7. Comparison of measured and estimated chloride concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

Table 6. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based chloride equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	19	0.938	115,000	--	20	0.921	22,700	--
1996	54	.974	48,100	-58.2	38	.950	14,300	-37.0
1997	74	.978	40,300	-16.2	54	.952	13,800	-3.50
1998	104	.985	27,500	-31.8	74	.954	13,400	-2.90

related to sulfate in water from the two stream-gaging stations in this study. Sulfate in water with saline base flows can have well-defined relations with specific conductance (Hem, 1992), except during high flows.

Specific conductance, which ranged from 98.9 to 3,550 µS/cm, was used as a surrogate for estimating sulfate concentrations in water from the Halstead stream-gaging station. The final linear regression equation for sulfate after logarithmic transformation was:

$$\log_{10}(SO_4) = 0.911\log_{10}(SC) - 1.12, \qquad (16)$$

where SO_4 is the estimated sulfate concentration, in milligrams per liter, and SC is specific conductance, in microsiemens per centimeter at 25 °C. There was one outlier, with a specific conductance of 3,550 µS/cm and a sulfate concentration of 95.6 mg/L. However, because this was not a local outlier, it did not affect the final regression equation substantially and was not removed from the data set.

At the Halstead stream-gaging station, streamflow is not strongly related to sulfate concentration because there is a low-head dam downstream from the gaging station that affects streamflow. The relation between sulfate and streamflow is poor because of mixing of water behind the dam. The relation is particularly poor at larger concentrations, which is evident in the scatter between measured and estimated sulfate concentrations for water from the Halstead stream-gaging station (fig. 8*A*). Log transformations did improve the linearity of the regression equation but did not have much effect on the scatter at large concentrations.

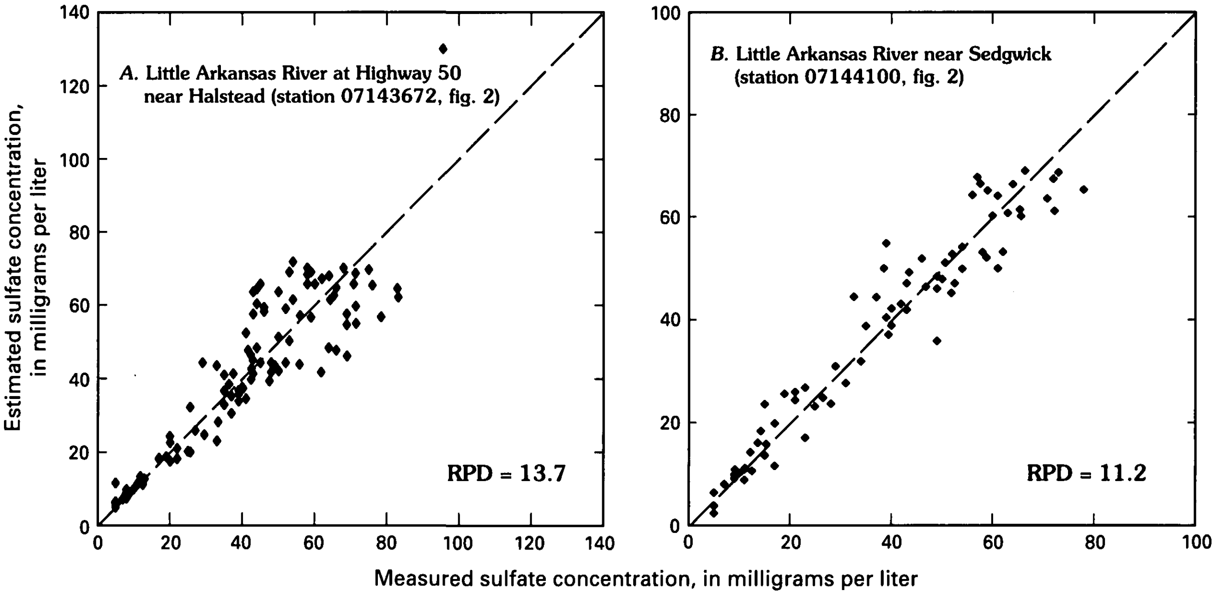
At the Sedgwick stream-gaging station, there was one outlier (with a sulfate concentration of 211 mg/L and a specific conductance of 1,350 μS/cm), which was not used in developing the regression equation. The ranges in the independent variables used in the development of the regression equation were 76 to 1,320 μS/cm for specific conductance and 14.2 to 9,250 ft^3/s for streamflow. The multiple linear regression equation was:

$$SO_4 = 0.0482SC - 5.50 \log_{10}(Q) + 18.8, \qquad (17)$$

where SO_4 is the estimated sulfate concentration, in milligrams per liter, SC is specific conductance, in microsiemens per centimeter at 25 °C, and Q is streamflow, in cubic feet per second.

It is important to note the difference in the form of the equations for the Halstead and Sedgwick stream-gaging stations in addition to the difference in the relation between measured and estimated sulfate concentrations (fig. 8). Unlike the Halstead stream-gaging station, streamflow at the Sedgwick stream-gaging station is not affected by a low-head dam. Perhaps there is a better relation between streamflow and sulfate at the Sedgwick stream-gaging station because of the less restricted streamflow; therefore, streamflow is included as a variable in the regression equation. A comparison of the measured and estimated sulfate concentrations is shown in figure 8*B*.

There was an improvement in error (*SSE*) for estimates of sulfate concentrations for both sites from 1995 to 1998. For the Halstead stream-gaging station, an increase in sample size from 1997 to 1998 continued to improve the *SSE* between measured and estimated sulfate concentrations (table 7). However, the improvement in error (*SSE*) for estimates of sulfate concentrations in water at the Sedgwick stream-gaging station with an additional year of data collection (1997–98) was not as substantial.



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured sulfate concentrations and instantaneous estimated sulfate concentrations

Figure 8. Comparison of measured and estimated sulfate concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

Table 7. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based sulfate equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	19	0.916	0.833	--	20	0.924	2,570	--
1996	54	.908	.906	8.76	37	.927	2,490	-3.11
1997	74	.913	.865	-4.53	53	.930	2,370	-4.82
1998	103	.919	.802	-7.28	77	.931	2,330	-1.69

Atrazine

Atrazine, one of the triazine herbicides, was identified by Ziegler and others (1999) as one of the primary constituents of concern for the *Equus* Beds Ground-Water Recharge Demonstration Project. In the study reported by Ziegler and others (1999), atrazine was shown to have an inverse relation with chloride and, therefore, with specific conductance. This occurred because chloride concentrations in the Little Arkansas River, which are associated with ground water in the area, were large during the winter months when the percentage of streamflow coming from ground water was high. During these months, atrazine concentrations were small (atrazine concentrations in ground water are very small or nondetectable in this area; Ziegler and others, 1999).The largest concentrations of atrazine in the Little Arkansas River occurred during the spring and summer when atrazine is applied to crops and when rainfall is most abundant (Christensen and Ziegler, 1998a). Atrazine concentrations, therefore, are highly seasonal, which may account for the relation between time and atrazine in the regression equation. The regression equation for estimating triazine concentration in water at the Halstead stream-gaging station uses month (time), streamflow, and specific conductance as surrogates. The ranges were 11.3 to 6,030 ft^3/s for streamflow and 123 to 1,850 μS/cm for specific conductance. The final regression equation was:

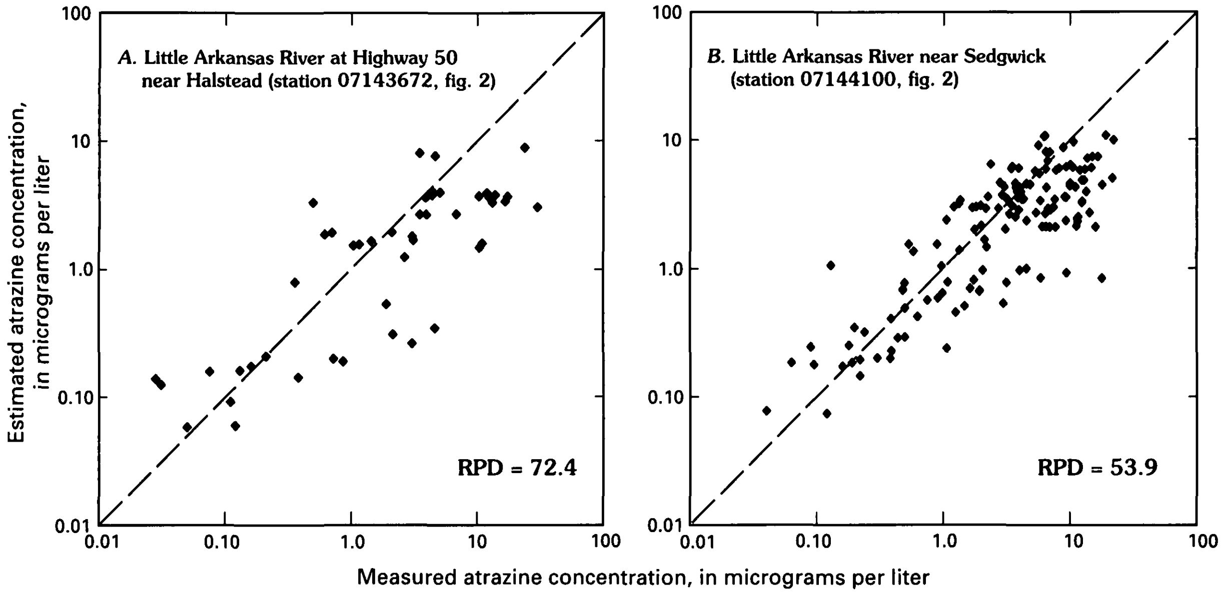
$$\log_{10}(triazine) = 1.42e^{\frac{-(month - 6.24)^2}{3.75}} - 0.0000288Q - 0.000581SC - 0.104, \qquad (18)$$

where *triazine* is in micrograms per liter, *month* is a number from 1 to 12, Q is streamflow, in cubic feet per second, and SC is specific conductance, in microsiemens per centimeter at 25 °C. About 700 samples were collected at the Halstead stream-gaging station for triazine herbicide analysis by ELISA, and 381 of these samples (table 7) were collected simultaneously with discharge and measurement of physical properties. No local outliers were identified from February 1995 through December 1998. Therefore, all 381 measurements were used in the development of the regression equation. Gas chromatography/mass spectrometry (GC/MS) confirmation indicated that atrazine and triazine concentrations were related to each other with an R^2 of 0.85 as shown in equation 19 (Ziegler and others, 1999):

$$atrazine = 0.80\ triazine. \qquad (19)$$

Results of the regression equation for estimating triazine concentrations in water at the Halstead stream-gaging station were converted to estimated atrazine concentrations using equation 19. Estimated atrazine concentrations did not compare as well with measured atrazine concentrations (fig. 9*A*) as other constituents discussed previously. This is due in part to the additional manipulations necessary to convert triazine to atrazine and the error involved in each of these steps that makes the final estimation less accurate than those of most other constituents presented in this report. In addition, atrazine concentration can vary by an order of magnitude within hours; therefore, timing visits for the collection of manual samples during peak concentration is difficult.

The regression equation for estimating triazine concentrations in water at the Sedgwick stream-gaging station uses the same surrogates as the equation for the Halstead stream-gaging station. The ranges were 14 to



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured atrazine concentrations and instantaneous estimated atrazine concentrations

Figure 9. Comparison of measured and estimated atrazine concentrations in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

6,300 ft^3/s for streamflow and 93 to 1,400 μS/cm for specific conductance. The final regression equation was:

$$\log_{10}(triazine) = 1.38e^{-\frac{(month - 5.94)^2}{3.92}} -$$

$$0.0000652Q - 0.000820SC + 0.0681\,, \qquad (20)$$

where *triazine* is in micrograms per liter, *month* is a number from 1 to 12, *Q* is streamflow, in cubic feet per second, and *SC* is specific conductance, in microsiemens per centimeter at 25 °C. Although streamflow is included in equation 20, it only improves the R^2 slightly (from 0.602 to 0.736).

About 1,300 water samples from the Little Arkansas River at the Sedgwick stream-gaging station were analyzed for triazine herbicide concentrations by ELISA, and 922 of these samples (table 8) were collected simultaneously with measurement of streamflow and physical properties. No local outliers were identified. GC/MS confirmation showed that the relation between the triazine herbicide concentration determined by ELISA and the atrazine concentration determined by GC/MS (Ziegler and others, 1999) is about the same as that used for the Halstead stream-gaging station (equation 19).

Triazine concentrations then were converted to atrazine concentrations using equation 19. As was the case with atrazine in water at the Halstead stream-gaging station, the measured and estimated atrazine concentrations for water at the Sedgwick stream-gaging station had relatively low correlation (fig. 9*B*). This is due in part because the concentration of atrazine in the stream can change by an order of magnitude within hours and timing visits for the collection of peak concentrations is difficult. In addition, atrazine is not strongly correlated with any of the physical properties that are currently being monitored in real time.

At the Halstead stream-gaging station, the addition of a second year of data (1996) made a substantial improvement in the error associated with the equation (table 8). At Halstead, more than 300 samples were collected in the second year, which accounts for much of the difference. Additional samples in 1997 and 1998 added only 54 analyses, and thus the improvement in *SSE* was minimal. At the Sedgwick stream-gaging station, both the second and third year of data collection added significantly to the improvement of

Table 8. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based atrazine equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)	Cumulative number of samples	R^2	*SSE*	Change in *SSE* (%)
1995	20	0.661	64.4	--	21	0.602	146	--
1996	327	.773	43.1	-33.1	311	.681	117	-19.9
1997	349	.775	42.7	-.93	649	.733	98.0	-16.2
1998	381	.777	42.4	-.70	922	.736	96.8	-1.22

the *SSE*. This is because there was a substantial number of samples collected in the second and third years. Although 1998 added another 273 samples to the development of the equation, there was very little improvement in the *SSE* after the third year. This would indicate that perhaps about 650 samples were enough to define a surrogate relation for atrazine.

Fecal Coliform Bacteria

Total coliform bacteria was identified by Ziegler and others (1999) as a constituent of concern in the Little Arkansas River during high flows or when water-treatment processes do not remove all the bacteria in the recharge water. The bacteria equations presented in this report are based on fecal coliform bacteria, one of the bacteria included in a total coliform analysis. The presence of fecal coliform bacteria in surface water indicates fecal contamination and possibly the presence of other organisms that could cause disease. Fecal coliform bacteria analyses were chosen for inclusion in this report because current (2000) State of Kansas water-quality criteria [2,000 colonies per 100 milliliters of water (col/100 mL) for noncontact recreation, 200 col/100 mL for contact recreation, and less than 1 col/100 mL for drinking water] are based on fecal coliform bacteria densities (Kansas Department of Health and Environment, 1999). Because runoff from a watershed may transport fecal coliform bacteria to streams, there may be a relation between bacteria densities and streamflow or, possibly, to time of year because runoff characteristics may vary with season. However, only month (time) and turbidity were significantly related to fecal coliform bacteria. The range in concentration for turbidity was 0.3 to 1,780 NTU.

For water at the Halstead stream-gaging station, the multiple regression equation used to estimate fecal coliform bacteria densities with logarithmic transformation was:

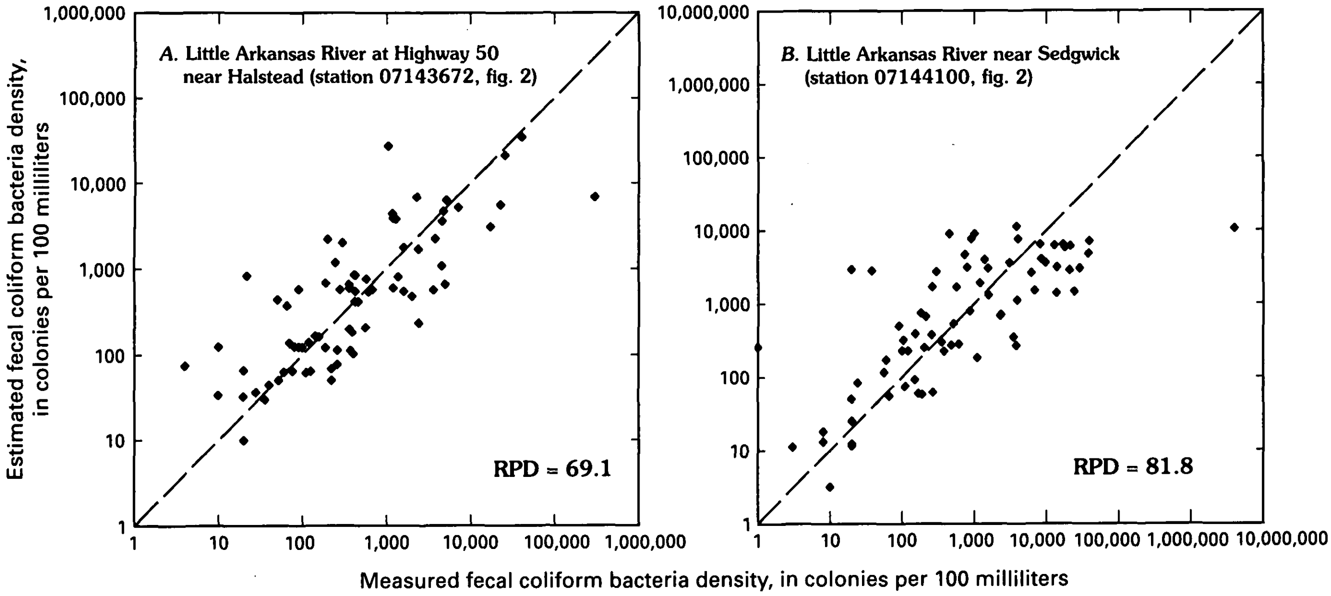
$$\log_{10}(Bact) = (0.490)\cos\left(2\pi\left(\frac{month + 2.06}{8.76}\right)\right) + 0.00106Turb + 0.417\log_{10}(Turb) + 1.65\,, \qquad (21)$$

where *Bact* is fecal coliform bacteria density, in colonies per 100 milliliters of water, *month* is a number from 1 to 12, and *Turb* is turbidity, in nephelometric turbidity units. A comparison of measured and estimated fecal coliform bacteria densities is shown in figure 10*A*.

For water at the Sedgwick stream-gaging station, the range in turbidity concentrations was 1.44 to 1,030 NTU. The multiple regression equation used to estimate fecal coliform bacteria densities with logarithmic transformation was:

$$\log_{10}(Bact) = -0.286\cos\left(2\pi\left(\frac{month - 29.4}{-6.8}\right)\right) - 0.000422(Turb) + 1.26\log_{10}(Turb) + 0.519\,, \qquad (22)$$

where *Bact* is fecal coliform bacteria density, in colonies per 100 milliliters of water, *month* is a number from 1 to 12, and *Turb* is turbidity, in nephelometric turbidity units. A comparison of measured and estimated bacteria densities is shown in figure 10*B*. For both the Halstead and Sedgwick stream-gaging stations there was one measurement for each that was an extreme outlier that was not removed from the data set. Both measurements occurred on March 20, 1997, and after an investigation verified these large densities, it was determined that there was not sufficient justification for their removal from the data set.



EXPLANATION

RPD—Median relative percentage difference between instantaneous measured fecal coliform bacteria densities and instantaneous estimated fecal coliform bacteria densities

Figure 10. Comparison of measured and estimated fecal coliform bacteria densities in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, February 1995–December 1998.

For estimates of bacteria densities in water at both the Halstead and Sedgwick stream-gaging stations, the second year of data collection offered a significant improvement in *SSE* (table 9). The *SSE* did not substantially improve for either stream-gaging station after 1996. This does not necessarily mean that 2 years of data collection are enough to define the relation. The R^2 is still not as good as those for other constituents and equations. This indicates that an ideal surrogate for fecal coliform bacteria has not been identified, but even with the errors associated with these estimates, this approach provides water-quality information not obtainable by collecting a few samples per month.

MEASURED AND ESTIMATED CONSTITUENT LOADS AND YIELDS

Although concentrations are useful for evaluating the Little Arkansas River with respect to current (2000) water-quality criteria, constituent loads indicate the chemical mass transported by the Little Arkansas River during a given period of time. Load estimates can be used to compute a yield for the contributing drainage area of a particular site. Yield is an area-normalized load calculated for the purpose of site comparison.

Real-time water-quality monitoring and instantaneous, measured water-quality data from January 1 through December 31, 1999, were not used to calibrate the seven regression equations presented in the previous sections but were used to test the equations. The relative percentage differences (RPD's) between instantaneous measured and estimated constituent loads at the Halstead and Sedgwick stream-gaging stations varied considerably (table 10).

In general, the majority of the estimated mean daily alkalinity load in the Little Arkansas River at both stream-gaging stations occurred in the late spring and summer (fig. 11) when rainfall was most abundant. The estimated alkalinity load compares well with the measured alkalinity load, and the median RPD was similar for both stream-gaging stations (a 19.6-percent difference at the Halstead stream-gaging station and a 17.1-percent difference at the Sedgwick stream-gaging station, table 10). The estimated alkalinity load for 1999 was larger for the Sedgwick stream-gaging station, but the 1999 estimated alkalinity yield was larger for the Halstead stream-gaging station (table 11). The smaller estimated alkalinity yield at the Sedgwick stream-gaging station may be because of the relatively

Table 9. Sample-size effect on improving sum of squares of error (*SSE*) for surrogate-based fecal coliform bacteria equation

[R^2, correlation coefficient; *SSE*, sum of squares of error; %, percent; --, not determined]

	Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)				Little Arkansas River near Sedgwick (station 07144100, fig. 2)			
Calendar year	**Cumulative number of samples**	R^2	***SSE***	**Change in *SSE* (%)**	**Cumulative number of samples**	R^2	***SSE***	**Change in *SSE* (%)**
1995	20	-0.574	75.5	--	18	0.043	94.1	--
1996	42	.578	30.1	-60.1	36	.567	42.6	-54.7
1997	58	.606	28.1	-6.64	50	.593	40.0	-6.10
1998	75	.620	27.1	-3.56	73	.556	43.6	9.00

small contribution of the intervening drainage area (fig. 12*A* and *B*) and because alkalinity is not conservative and, thus, subject to loss between the Halstead and Sedgwick stream-gaging stations.

Measured and estimated dissolved-solids load also followed a seasonal pattern (fig. 13) and had the smallest median RPD's (9.53 percent at the Halstead stream-gaging station and 4.09 percent at the Sedgwick stream-gaging station, table 10) and, thus, the closest surrogate relation of any of the constituents presented in this report. The 1999 estimated dissolved-solids load was larger for the Sedgwick stream-gaging station, yet the 1999 estimated dissolved-solids yield was larger for the Halstead stream-gaing station (table 11).

Changes in *TSS* loads generally are affected by changes in streamflow (Topping and others, 2000), and thus, measured and estimated *TSS* loads at the Halstead and Sedgwick stream-gaging stations showed a seasonal pattern (fig. 14). The RPD's between measured and estimated loads were substantial (table 10), and there was a large difference between estimated *TSS* loads at the Halstead and Sedgwick stream-gaging stations (table 11). Turbidity (the surrogate for *TSS*) may be altered significantly between the Halstead and Sedgwick stream-gaging stations due to the inflow of Kisiwa, Emma, and Sand Creeks (fig. 1). This difference in turbidity affected the regression equations and the resulting RPD's at each site. The 1999 estimated *TSS* load was larger at the Sedgwick stream-gaging station, whereas the estimated *TSS* yield was larger at the Halstead stream-gaging station. The relatively small contribution of *TSS* from the intervening drainage area (table 11) may result from storage of solids and sediment in the streambed between the Halstead and Sedgwick stream-gaging stations, or the inflow of Kisiwa, Emma, and Sand Creeks (fig. 1) may have a dilution effect on *TSS* at the Sedgwick stream-gaging station.

Table 10. Median relative percentage differences (RPD's) between instantaneous measured and estimated constituent loads for selected constituents in Little Arkansas River at Highway 50 near Halstead and near Sedgwick, Kansas, 1999

	RPD's for constituent loads	
Constituent	**Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)**	**Little Arkansas River near Sedgwick (station 0744100, fig. 2)**
Alkalinity	19.6	17.1
Dissolved solids	9.53	4.09
Total suspended solids	66.4	34.0
Chloride	10.8	19.4
Sulfate	20.1	15.0
Atrazine	53.6	87.1
Fecal coliform bacteria	242	83.7

Although chloride concentrations generally are larger during the winter (Albert, 1964; Ziegler and others, 1999), the large streamflows during the summer of 1999 caused most of the transport of chloride to occur from April through September (fig. 15). Measured and estimated chloride loads compared well (table 10); however, the RPD was substantially less for the Halstead stream-gaging station (10.8 percent) than for the Sedgwick stream-gaging station (19.4 percent).

The use of specific conductance measurements as a surrogate for chloride concentrations in the Little Arkansas River was studied by Albert (1964) from 1960–61. The relation between these two constituents does not appear to have changed substantially in the past 40 years. In fact, the mean daily chloride load calculated for the 1960 water year for the Little Arkansas

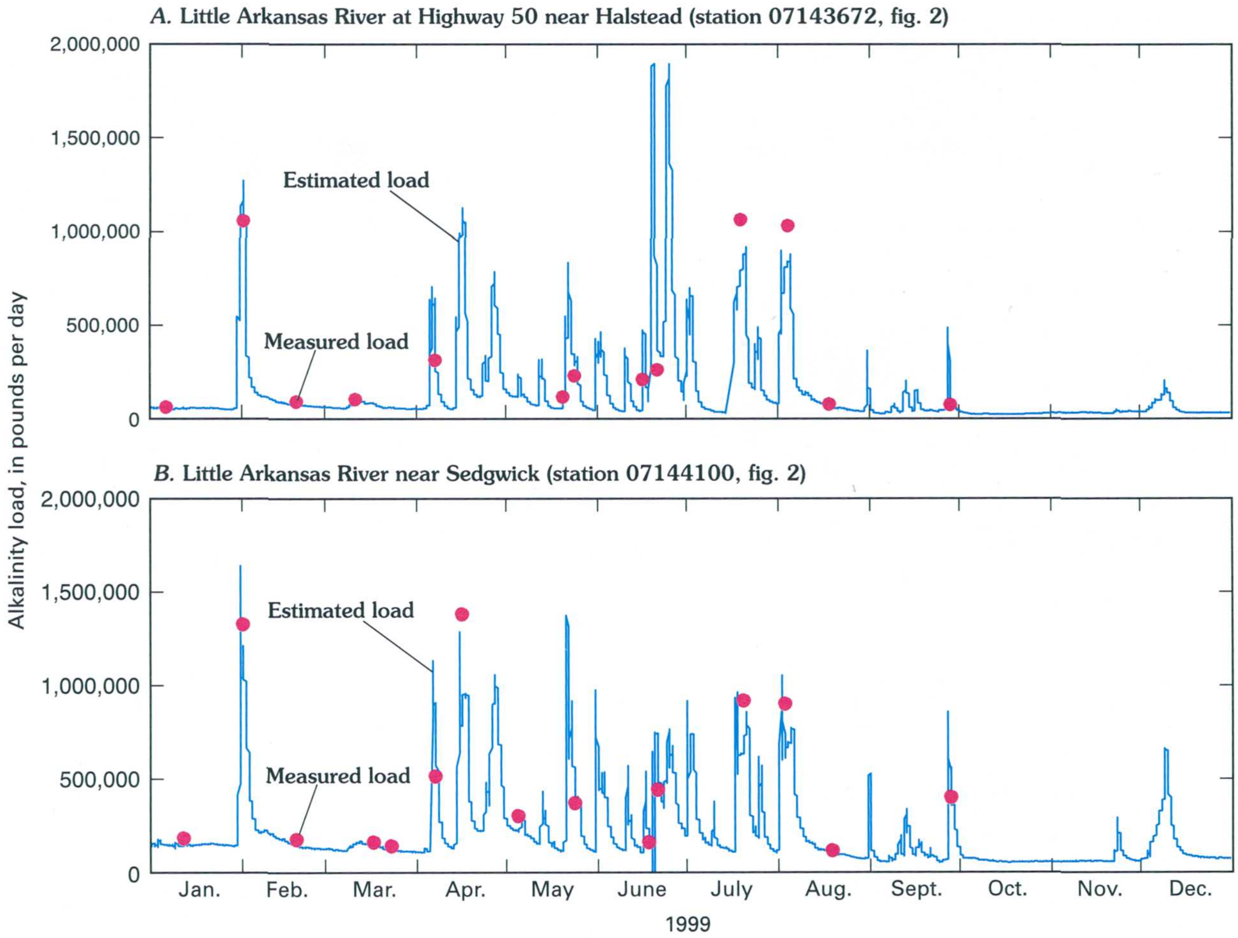


Figure 11. Comparison of measured and estimated alkalinity loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

Table 11. Estimated loads and yields of selected constituents in subbasins of Little Arkansas River, Kansas, 1999

[With the exception of bacteria, all loads are expressed in pounds per year, and yields are expressed in pounds per acre per year. Bacteria is expressed as billions of colonies per year and billions of colonies per acre per year for loads and yields, respectively]

	Subbasin areas (fig. 12)					
	Upstream from Little Arkansas River at Highway 50 near Halstead (drainage area = 439,000 acres)		Contribution of intervening drainage area (between Halstead and Sedgwick streamgaging stations) (drainage area = 303,000 acres)		Upstream from Little Arkansas River near Sedgwick (drainage area = 742,000 acres)	
Constituent	**Estimated load**	**Estimated yield**	**Estimated load**	**Estimated yield**	**Estimated load**	**Estimated yield**
Alkalinity	53,300,000	121	24,100,000	79.4	77,400,000	104
Dissolved solids	183,000,000	417	39,100,000	129	222,000,000	299
Total suspended solids	460,000,000	1,050	153,000,000	505	613,000,000	826
Chloride	44,500,000	101	730,000	2.41	45,300,000	61.0
Sulfate	12,700,000	28.9	6,100,000	20.1	18,800,000	25.3
Atrazine	1,290	.003	1,040	.003	2,330	.003
Fecal coliform bacteria	30,000,000	68.3	2,410,000	7.94	32,400,000	43.7

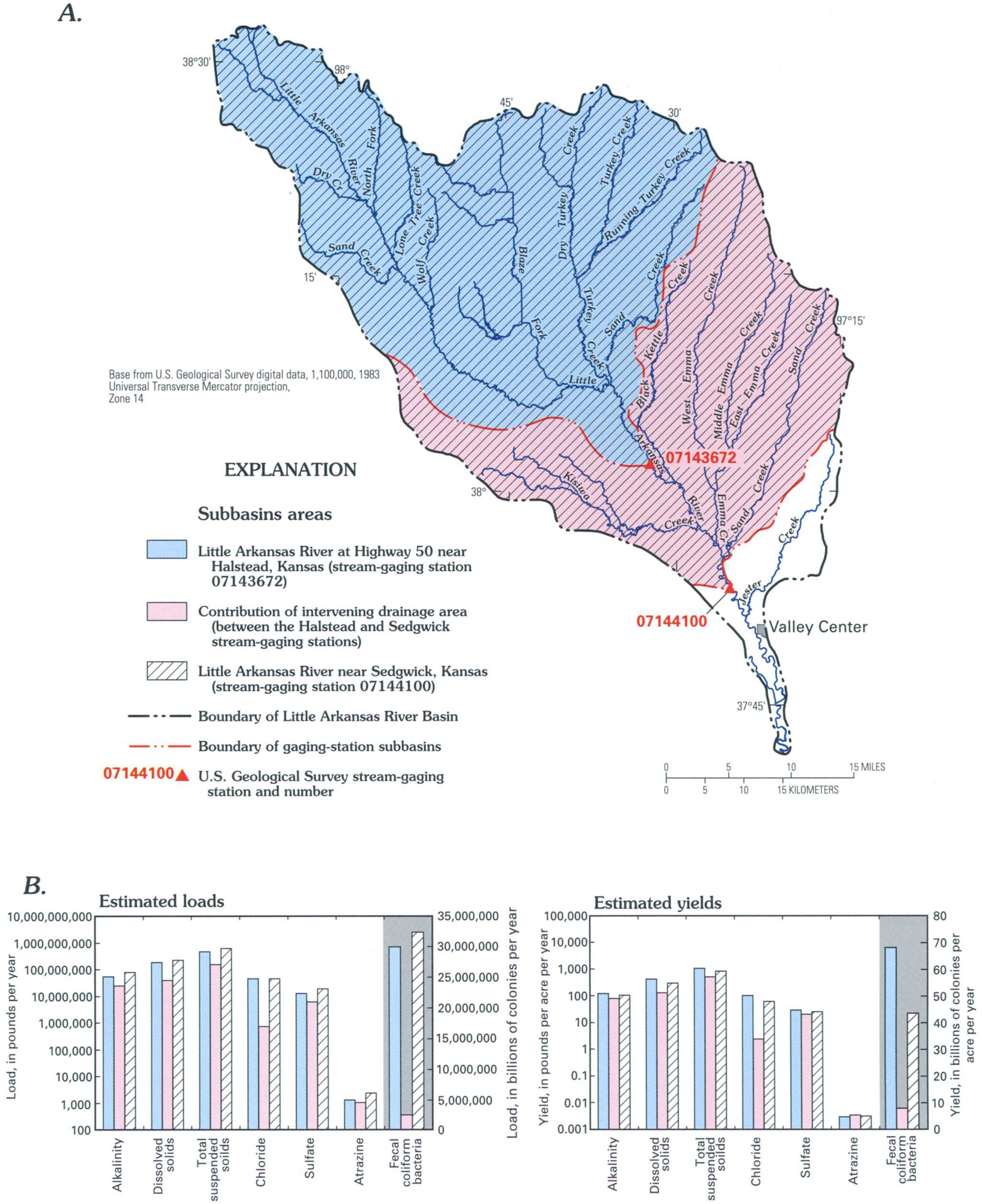


Figure 12. *(A)* Subbasin and intervening drainage areas between Halstead and Sedgwick stream-gaging stations and *(B)* estimated loads and yields for selected constituents in Little Arkansas River at Highway 50 near Halstead and near Sedgwick, Kansas, 1999.

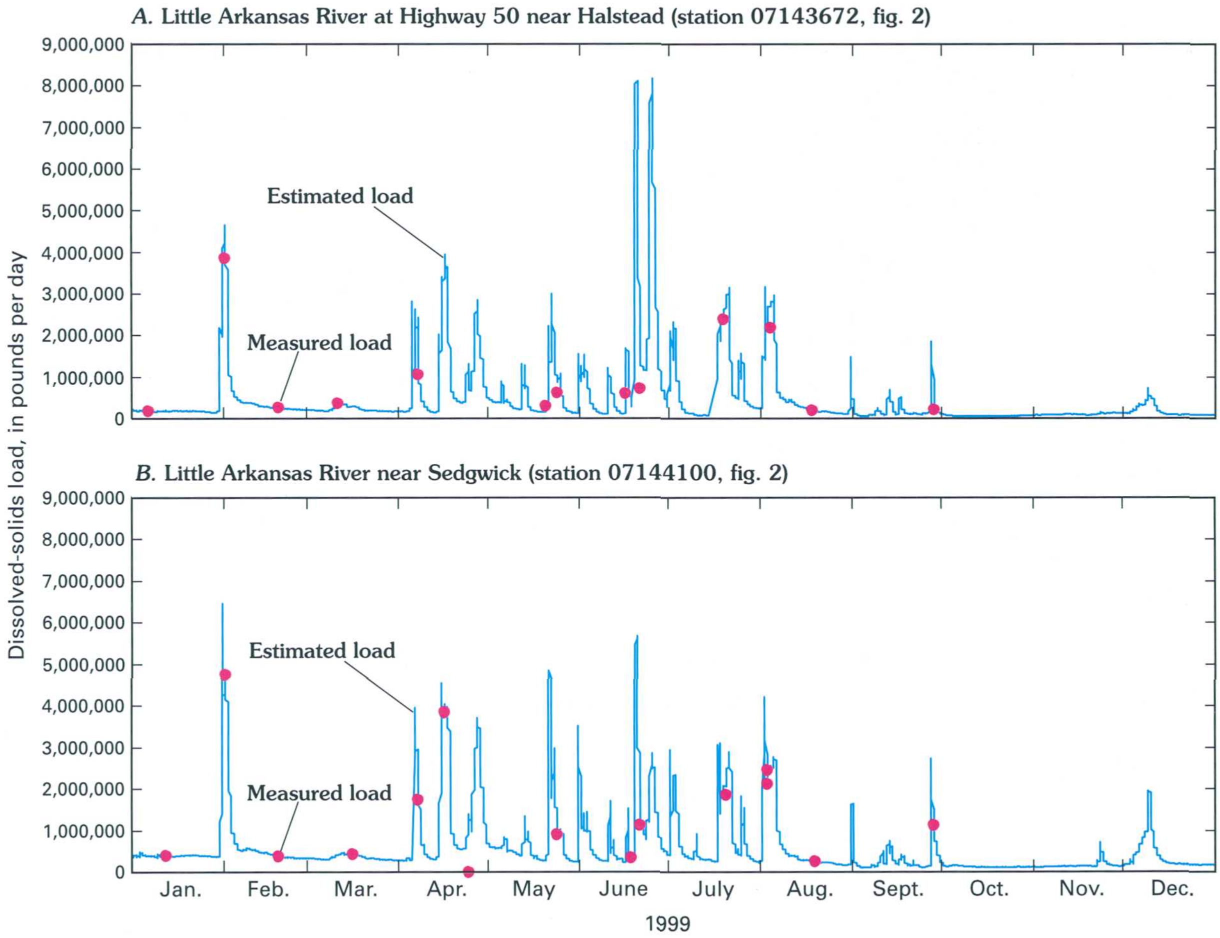


Figure 13. Comparison of measured and estimated dissolved-solids loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

River at Valley Center was similar to the estimated mean daily chloride load calculated for the 1999 calendar year for the Little Arkansas River at Sedgwick (126,000 and 124,000 lb, respectively). Streamflow for the two study periods also was similar. During the 1960 water year, the annual mean streamflow of the Little Arkansas River at Valley Center, Kansas (station 07144200, fig. 2) was 534 ft^3/s; whereas during the 1999 calendar year, the annual mean streamflow of the Little Arkansas River near Sedgwick (station 07144100, fig. 2) was 528 ft^3/s. The stream-gaging stations near Sedgwick and at Valley Center are 6.2 river mi apart. Considering the length of time between the two studies and the distance between the stream-gaging stations, the results from these studies may have important implications for the establishment of chloride TMDL's in the Little Arkansas River Basin.

The 1999 estimated chloride yields were 101 lb/acre for the Halstead stream-gaging station subbasin area and 61.0 lb/acre for the Sedgwick subbasin area (table 11). The estimated chloride yield for the intervening drainage area was very small (2.41 lb/acre). Because chloride is considered a conservative constituent, the smaller contribution of the intervening drainage area may be due to dilution of the river water caused by precipitation runoff. This indicates that the subbasin area upstream from the Halstead stream-gaging station may be a more substantial source for chloride.

Because the real-time surrogate measurement for sulfate also is specific conductance, the estimated sulfate loads were similar in distribution to those for chloride (fig. 16). The measured and estimated sulfate loads also were similar for the Halstead and Sedgwick stream-gaging stations, with median RPD's of 20.1 and 15.0 percent, respectively (table 10). Estimated sulfate yields were similar for the two stream-gaging stations at 28.9 and 25.3 lb/acre (table 11).

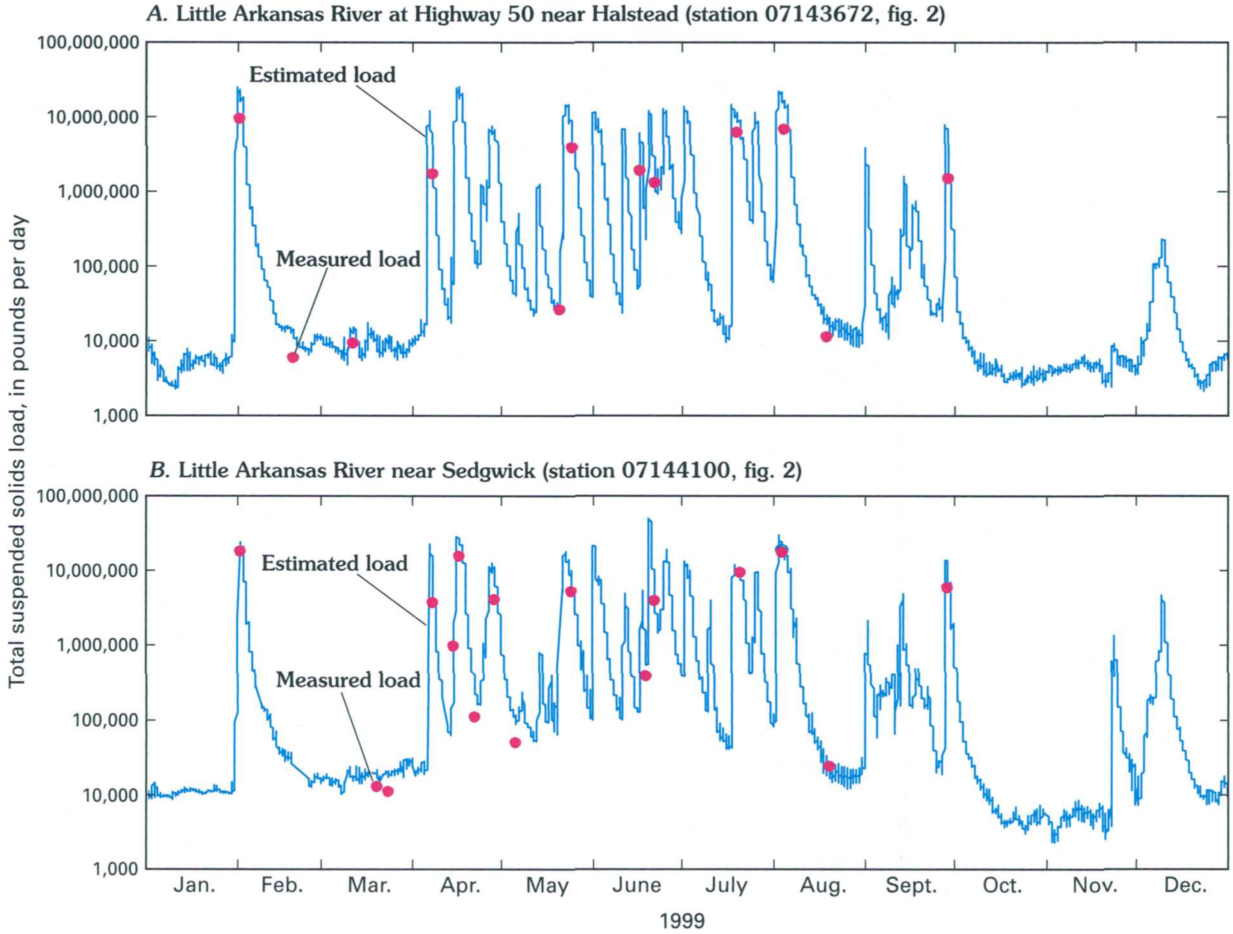


Figure 14. Comparison of measured and estimated total suspended solids loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

Atrazine concentrations generally are larger during late spring and summer (Christensen and Ziegler, 1998a) when streamflows also are larger. In fact, in a previous study (Christensen and Ziegler, 1998a), it was determined that 90 percent of the atrazine load occurred during about 15 days each year at the Halstead stream-gaging station and during about 40 days each year at the Sedgwick stream-gaging station. Generally, atrazine and streamflow increases occur at the same time and produce the largest loads during late spring and summer at the Halstead stream-gaging station (fig. 17*A*), and during the spring at the Sedgwick stream-gaging station (fig. 17*B*). Measured atrazine loads sometimes were an order of magnitude larger than estimated atrazine loads (fig. 17). The RPD was larger for measured and estimated atrazine loads at the Sedgwick stream-gaging station than at Halstead stream-gaging station (table 10).

Part of the variability in atrazine concentrations and loads probably is due to factors unrelated or weakly related to month, streamflow, and specific conductance. Examples of such factors include application rates, cropping practices, precipitation characteristics, soil characteristics, and topography. Also, 3 years of data collection may not have been enough to define atrazine fluctuations. The largest variance between measured and estimated atrazine loads occurred during early spring of 1999 (fig. 17). It is possible that atrazine was applied earlier than in the past 3 years due to exceptionally good weather in the spring of 1999 that encouraged farmers to plant (and apply pesticides) much earlier than in the past, which caused larger-than-expected measured atrazine loads in the early spring of 1999.

The estimated atrazine load at the Sedgwick stream-gaging station was larger in 1999 than in previous years (Christensen and Ziegler, 1998a). In 1996, however, the larger annual atrazine yield occurred at the Halstead stream-gaging station (Christensen and Ziegler, 1998a). Estimated atrazine yields were the

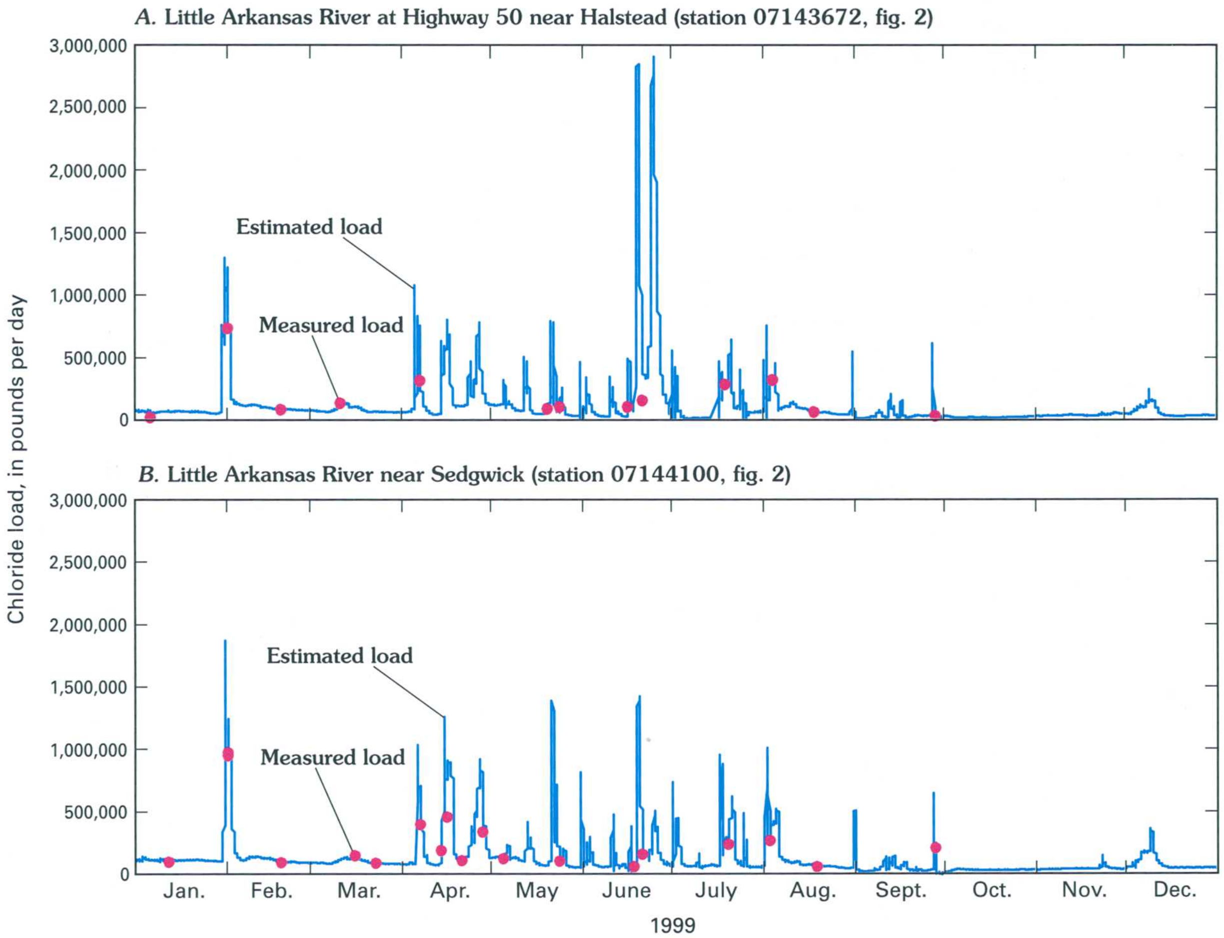


Figure 15. Comparison of measured and estimated chloride loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

same for the Halstead and Sedgwick stream-gaging stations and for the intervening drainage area between the two stream-gaging stations during 1999 (table 11). The changes in estimated atrazine loads and yields indicate that atrazine use and transport may fluctuate in the subbasins.

The regression equation for fecal coliform bacteria (table 2) cannot accurately estimate extreme fecal coliform bacteria densities that sometimes occur. The estimates from the regression equation are similar for smaller bacteria densities that may result from nonpoint sources, such as livestock production. However, if point sources, such as wastewater discharge or animal feedlot runoff, are contributing to the bacteria density, the regression equation cannot estimate this accurately. The estimated fecal coliform bacteria load for 1999 was largest for the Sedgwick stream-gaging station, but the 1999 estimated yield was largest for the Halstead stream-gaging station (table 11).

Although manual samples were collected during a wide range of streamflows, some of the differences between the measured and estimated bacteria loads (fig. 18) could be due to factors unrelated or weakly related to the surrogates used in the regression equations. Examples of such factors include manure application, feedlot runoff, sewage-treatment-plant discharges, precipitation characteristics, soil characteristics, and topography.

Despite some large differences, the regression-estimated mean daily loads of fecal coliform bacteria may be more reflective of actual loads than loads calculated from periodic measured data because of the continual nature of the real-time data used to develop the regression equation. There were few or no gaps in the sampling of hydrologic conditions and data with the real-time monitors, except in the case of water-quality monitor malfunction. On the other hand, the measured loads were based on a single discrete

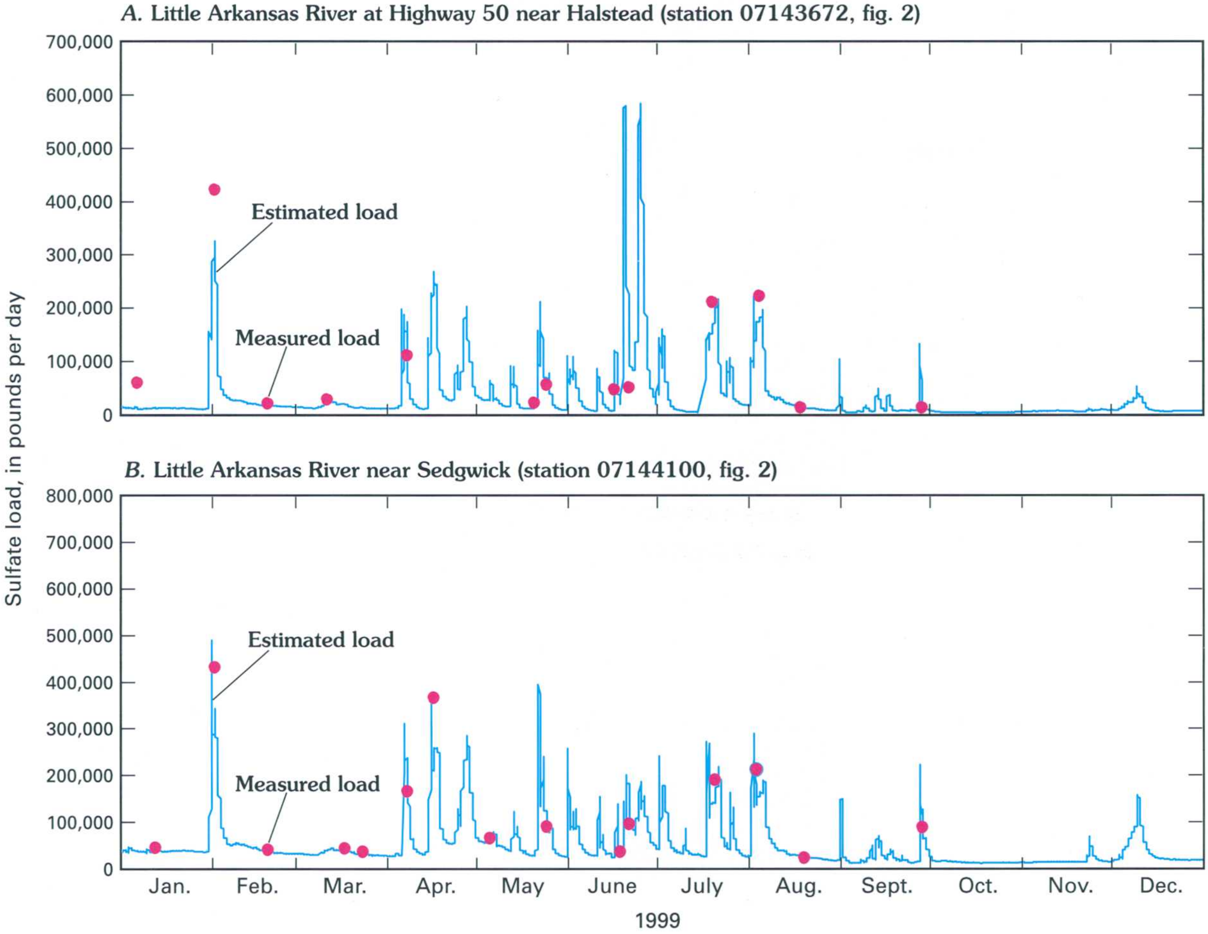


Figure 16. Comparison of measured and estimated sulfate loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

sample, and peaks in both concentrations and streamflow may have been missed. Despite large RPD's between measured and estimated fecal coliform bacteria loads for the extreme fecal coliform densities, surrogate estimates can be useful because the extreme bacteria densities were substantially higher than water-quality criteria established by KDHE.

Annual water yields for the subbasins and intervening drainage areas between the Halstead and Sedgwick stream-gaging stations were 68.7 ft^3/acre at the Halstead stream-gaging station, 51.2 ft^3/acre for the intervening drainage area, and 61.6 ft^3/acre for the Sedgwick stream-gaging station (annual water yield is the annual mean stream discharge in cubic feet divided by the subbasin area in acres). Yields of water at the Sedgwick stream-gaging station and intervening drainage area probably were smaller because of the more clay-rich soils in the northern subbasin that reduce infiltration and increase runoff and more sandy soils in the subbasin between Halstead and Sedgwick that allow for more infiltration and decrease runoff. Part of the differences in constituent yields between the two subbasins for the same period of time is due to the differences in water yield and also could be due to factors unrelated or weakly related to the surrogates used in the regression equations. Examples include point sources of nutrient, bacteria, or other constituents from sources such as wastewater discharge or animal feedlots.

BENEFITS OF REAL-TIME WATER-QUALITY MONITORING

The physical properties measured by water-quality monitors can be related to concentrations of dissolved ions. A notable example is the water's ability to conduct electricity (specific conductance), which is affected by the concentration of chloride ions (Hem, 1992). The pH and temperature of water can be useful

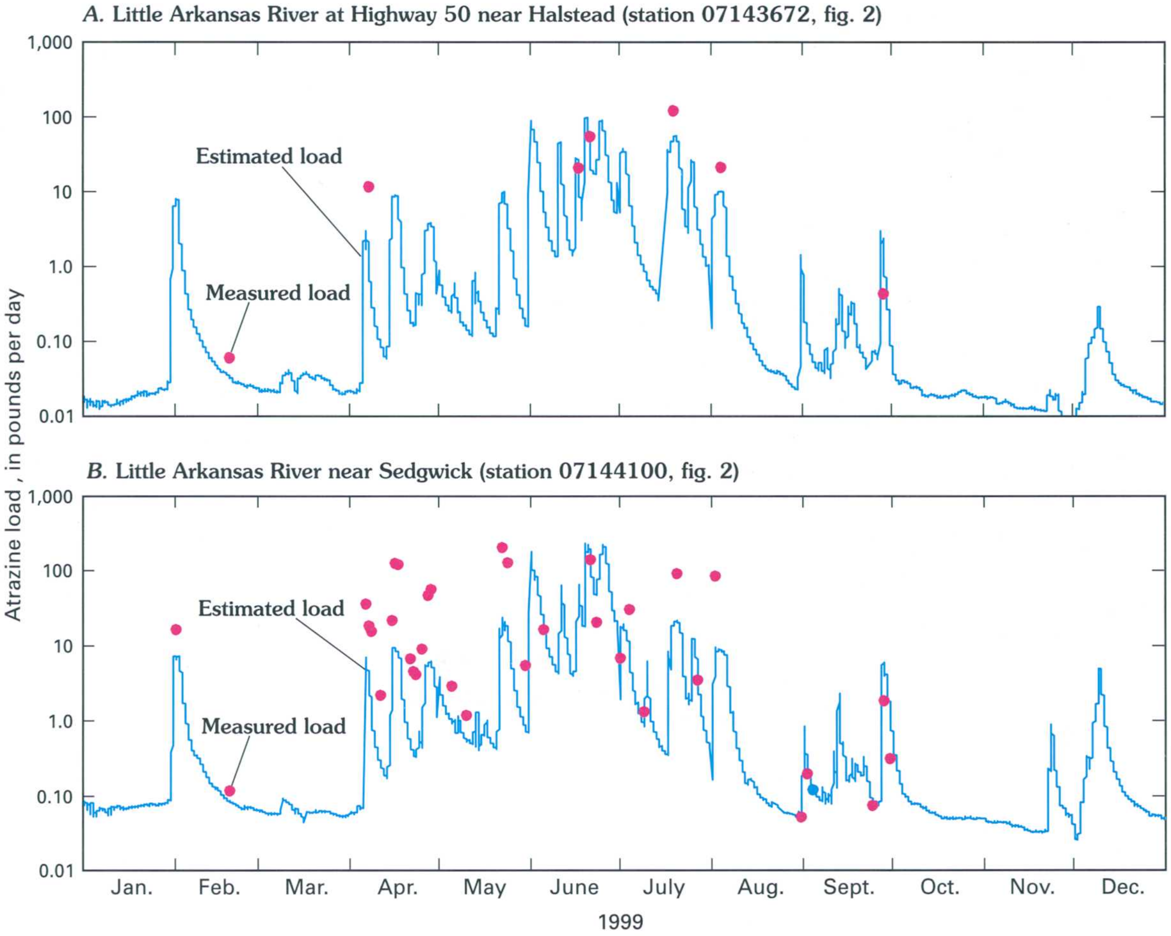


Figure 17. Comparison of measured and estimated atrazine loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

indicators of chemical equilibrium. The USEPA has established an acceptable range for pH in drinking water of 6.5 to 8.5 standard units (U.S. Environmental Protection Agency, 1999). Turbidity can be an indicator of the amount of sediment and related constituents transported by a stream. The acceptable range for turbidity in drinking water is 0.5 to 1.0 NTU (U.S. Environmental Protection Agency, 1999). Dissolved oxygen often is used to evaluate the biochemistry of water because oxygen is required for the survival of fish and other aquatic life (Hem, 1992). Dissolved oxygen also can be an indicator of the quality of the source water for recharge. KDHE has adopted a minimum criteria of 5.0 mg/L for dissolved oxygen in surface water (Kansas Department of Health and the Environment, 1997).

Continuous and periodic monitoring enable identification of seasonal trends in selected physical properties and chemical constituents and estimation of chemical mass transported in the Little Arkansas River. For example, real-time dissolved oxygen measurements provide information on the daily exposures of biota in the streams—in contrast to manual samples that are typically collected only during daylight hours. Identification of seasonal trends in water-quality constituents is especially important because high flows have such a substantial effect on chemical loads, and concentration data from manual samples often are not available. Therefore, real-time water-quality monitoring of surrogates for the estimation of water-quality constituents in streamflow may increase the accuracy of load estimations and decrease the need for some manual data-collection activities. Surrogates monitored on a continuous basis provide resource managers with real-time information on stream physical properties as shown in figures 19 and 20. Figure 19 shows specific conductance of water and streamflow at the Halstead and Sedgwick stream-gaging stations for 1999. Figure 20 shows turbidity of the water and streamflow at each stream-gaging station. Because the

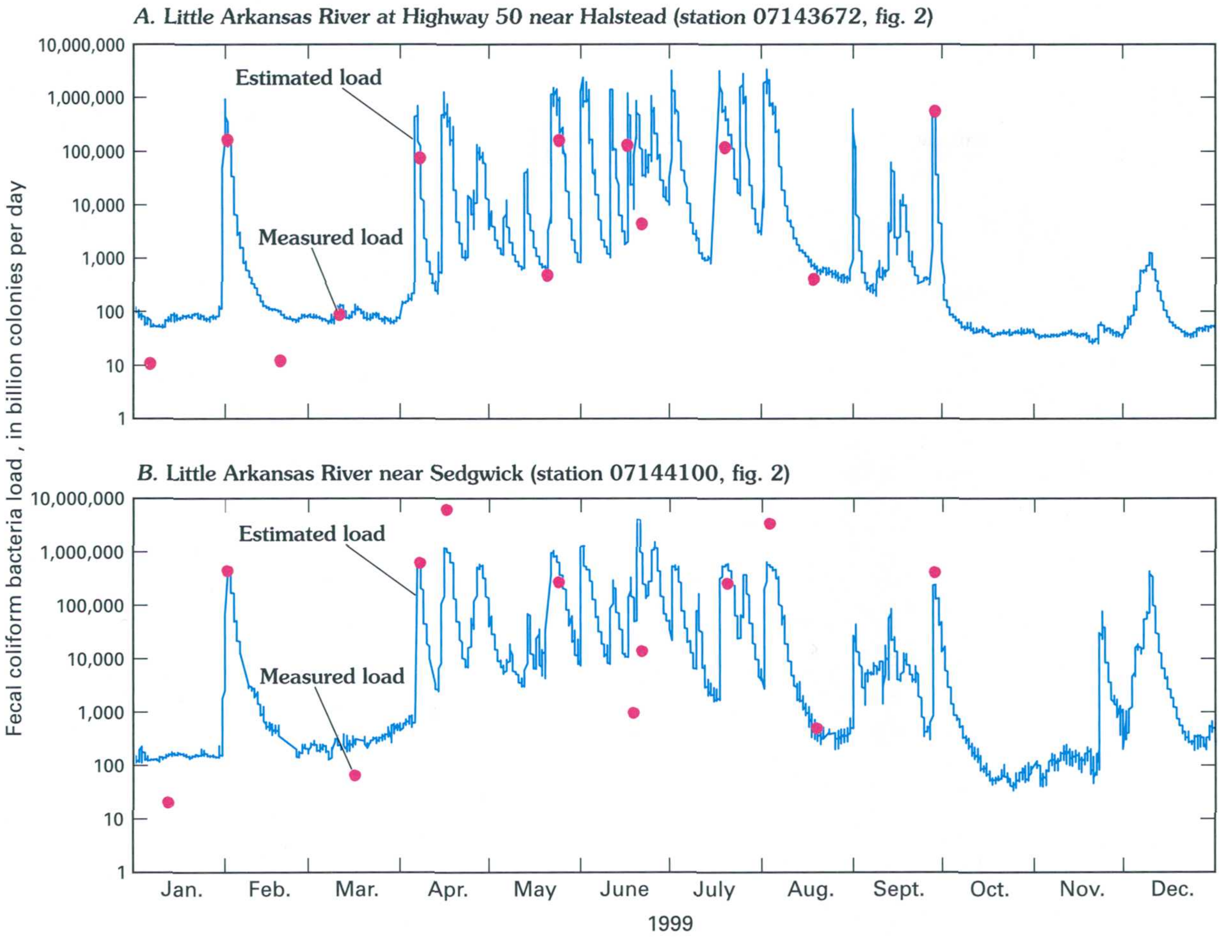


Figure 18. Comparison of measured and estimated fecal coliform bacteria loads in Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

other sensor measurements (pH, water temperature, and dissolved oxygen) were not significantly related to the water-quality constituents that were estimated in the previous section, they are not included here.

Even for those constituents with large RPD's between the measured and estimated loads, the estimation of constituent concentrations and loads with regression analysis and real-time water-quality monitoring has numerous advantages over periodic manual sampling. The timely availability of bacteria and other constituent data may be important when considering recreation and whole-body contact criteria for a water body; water suppliers would have timely information to use in adjusting water-treatment strategies; environmental changes could be assessed in time to prevent negative effects on fish or other aquatic life; and officials for the *Equus* Beds Ground-Water Recharge Demonstration Project could use this information to prevent the possible degradation of the *Equus* Beds aquifer by choosing not to recharge when constituent concentrations in the recharge water are large.

Information from the regression equations and real-time monitoring may be used to optimize visits to the sampling sites, and resource managers may use the regression-estimated concentrations to adjust management strategies rapidly when large concentrations of constituents may affect the quality of the water supply. In addition, information on loads and yields may be an indication in which subbasin to concentrate efforts with regard to land-resource best-management practices. Constituent loads were more substantial at the Sedgwick stream-gaging station due to its downstream location and thus higher streamflows. However, for all constituents, the Halstead subbasin had the largest yields (fig. 12).

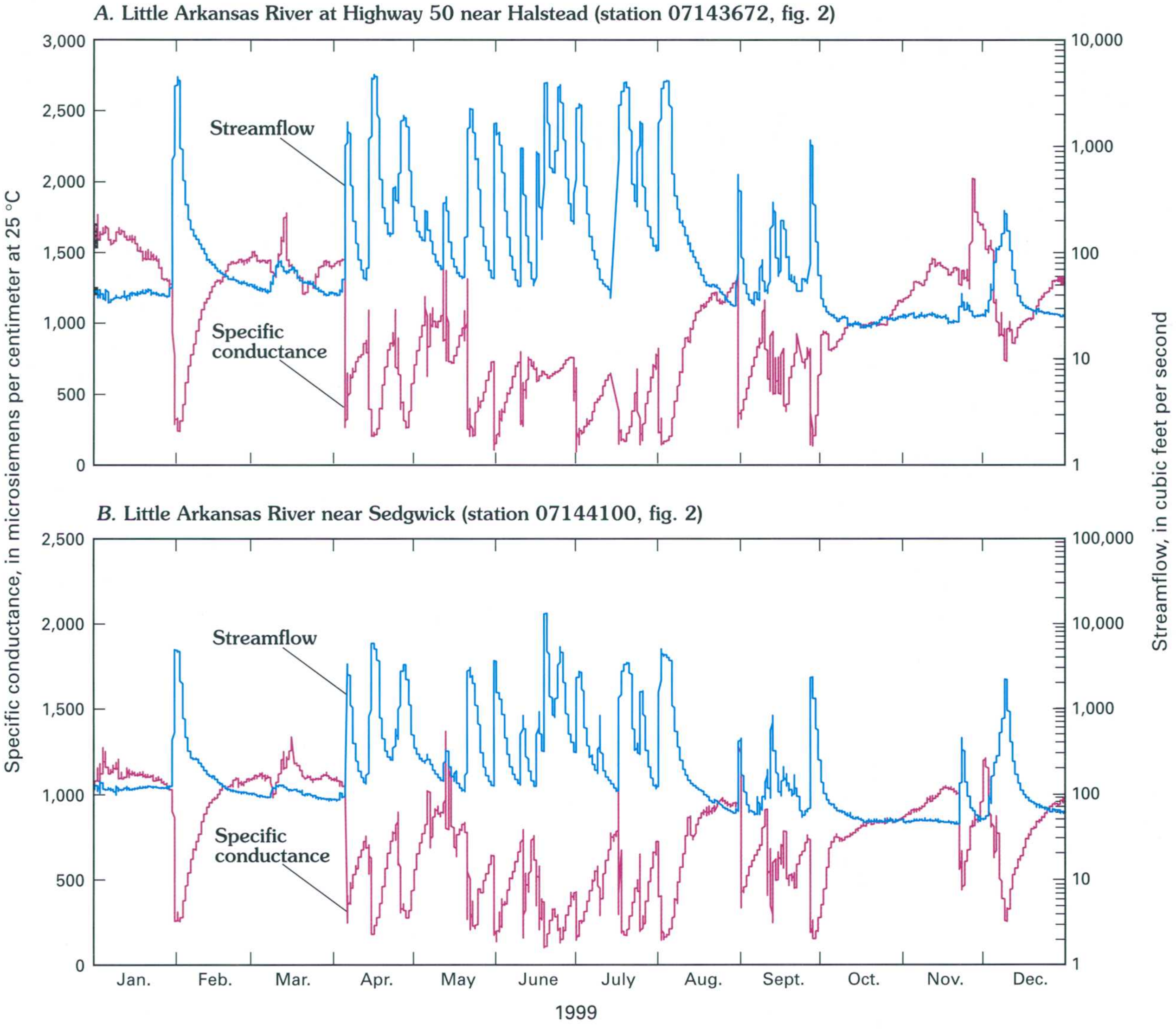
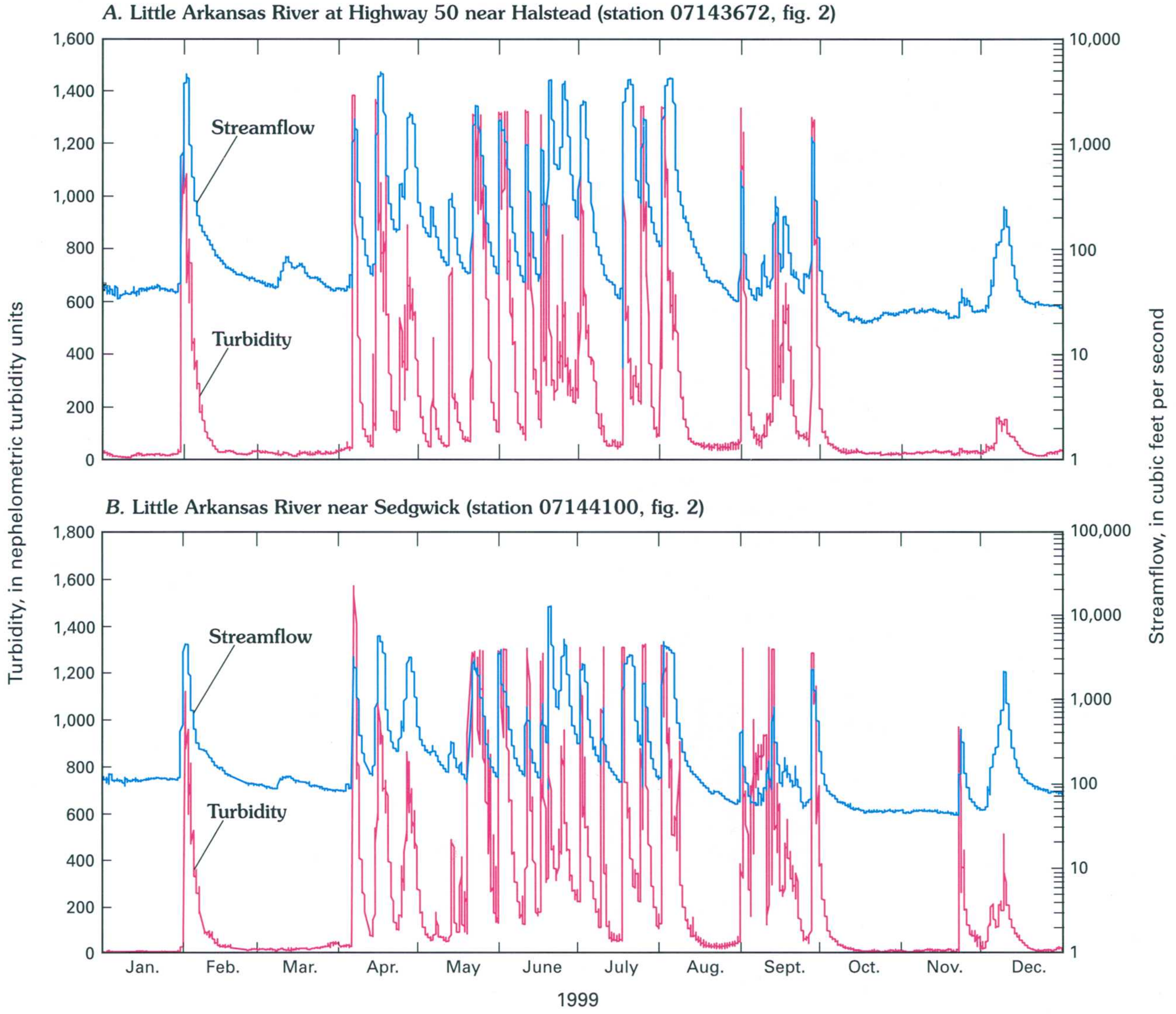


Figure 19. Comparison of specific conductance of water and streamflow for Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

CONCLUSIONS

The Wichita well field, located in the *Equus* Beds aquifer of south-central Kansas, is one of the primary sources of water for the city of Wichita. The *Equus* Beds Ground-Water Recharge Demonstration Project was initiated not only to increase ground-water levels in the *Equus* Beds aquifer but also to prevent saltwater intrusion into this important freshwater aquifer. During the project, surface water from two USGS streamgaging stations on the Little Arkansas River is used as the source for recharge water.

Because the quality of the water from the Little Arkansas River is likely to vary during different hydrologic conditions and different times of the year, it is important to know when recharge should not occur. The management of the *Equus* Beds Ground-Water Recharge Demonstration Project requires timely water-quality information to ensure that the quality of the receiving aquifer water is not degraded by poor quality source water. By developing regression equations using parameters that can be measured in real time to estimate water-quality conditions, it may be possible to avert possible contamination of the *Equus* Beds aquifer. The surrogate relations developed in this report may be used by resource managers to understand the changing quality of water in the Little Arkansas River and to quickly adjust water-management operations in the event that a significant


A. Little Arkansas River at Highway 50 near Halstead (station 07143672, fig. 2)

B. Little Arkansas River near Sedgwick (station 07144100, fig. 2)

Figure 20. Comparison of turbidity of water and streamflow for Little Arkansas River *(A)* at Highway 50 near Halstead and *(B)* near Sedgwick, Kansas, 1999.

change in constituent concentrations would affect the quality of the water used to recharge the *Equus* Beds aquifer.

Alkalinity, dissolved solids, chloride, and sulfate loads estimated on the basis of the regression equations developed in this report had median relative percentage differences of less than 25 percent. The estimated total suspended solids, atrazine, and fecal coliform bacteria loads, however, had median errors greater than 25 percent. The accuracy of total suspended solids, atrazine, and bacteria loads that are based on surrogate relations need to be evaluated further because of these large errors.

Even for those constituents with large relative percentage differences between the measured and estimated concentrations and loads, the estimation of constituent concentrations and loads with regression analysis and real-time water-quality monitoring has numerous advantages over periodic manual sampling. The timely availability of bacteria and other constituent data may be important when considering recreation and whole-body contact criteria for a water body; water suppliers would have timely information to use in adjusting water-treatment strategies; environmental changes could be assessed in time to prevent negative effects on fish or other aquatic life; and officials for the *Equus* Beds Ground-Water Recharge

Demonstration Project could use this information to prevent the possible degradation of the *Equus* Beds aquifer by choosing not to recharge when constituent concentrations in the source water are large.

For most of the constituents in this report, 2 years of data collection were sufficient (about 35 to 55 samples) to define the relation between the constituent and its surrogate(s) when samples were collected throughout 90 to 95 percent of the stream's flow duration— two very wet or dry years in a row may not define a relation because samples will not have been collected during varying conditions.

Continuous and periodic monitoring enabled identification of seasonal trends in selected physical properties and chemical constituents and estimation of chemical loads transported in the Little Arkansas River. For example, real-time dissolved oxygen measurements provided information on the daily exposures of biota in the streams—in contrast to manual samples that were collected only during daylight hours. Identification of seasonal trends is especially important because high flows have a substantial effect on chemical loads and because concentration data from manually collected samples often were not available. Therefore, real-time water-quality monitoring of surrogates for the estimation of constituents in streamflow probably increased the accuracy of load estimations and decreased the need for some manual data collection.

The stream-gaging network used in this study provides real-time streamflow data in 15-minute increments that allow water-resource and hazards managers to utilize the data immediately. In the past, this immediate availability has not been the case with water-quality constituents. However, with the installation of real-time water-quality monitors that measure physical properties of water and the development of regression equations that relate these physical properties to constituents of concern, water-resource managers have the necessary information to make immediate decisions regarding their water supply.

Information from the regression equations and real-time monitoring can be used to optimize visits to the sampling sites, and resource managers can use the real-time estimates of concentration to adjust management strategies rapidly when high concentrations of constituents may affect the quality of the water supply. In addition, information on constituent loads and yields may be an indication in which subbasin to concentrate efforts with regard to land-resource best-management practices. For example, constituent loads alone were more substantial at the Sedgwick gaging station due to its downstream location and higher streamflows. However, when compared to yields, it appeared that for all constituents the Halstead subbasin had the largest yields.

In addition, the regression equations also may be useful for calculating total maximum daily loads (TMDL's), which the State of Kansas is mandated to establish for stream segments that have been identified by section 303 (d) of the 1972 Clean Water Act as limited for specific uses because of water-quality concerns. With the development of surrogate relations between real-time sensor measurements (physical properties) and periodic collection of samples for analysis of chemical constituents, a more accurate representation of actual daily loads is possible. This is due in part to the continuous nature of data collection. The methods used in this report to develop regression equations can be used by water-resource managers to develop relations for constituents and surrogates at other sites where there may be concerns about water quality. With annual load and yield estimates, water-quality trends and the effectiveness of land-resource best-management strategies to improve water quality can be evaluated.

The increasing public interest in TMDL's and water quality in general makes this study of regional and national importance because it shows how constituent loads may be calculated with surrogates of water-quality constituents. The methods used in this study may be replicated for other sites in Kansas and the Nation to provide input data for the development of TMDL's and to monitor future effectiveness of implemented best-management practices. Also, by comparing annual loads, an evaluation of long-term trends is possible. However, it is important to note that the regression equations presented in this report are site specific.

REFERENCES CITED

Albert, C.D., 1964, Brine in surface water of the Little Arkansas River Basin, Kansas: U.S. Geological Survey Bulletin 1–5, 14 p.

Aucott, W.R., Myers, N.C., and Dague, B.J., 1998, Status of ground-water levels and storage in the Wichita well field area, south-central Kansas, 1997: U.S. Geological Survey Water-Resources Investigations Report 98–4095, 15 p.

Bates, R.L., and Jackson, J.A., eds., 1984, Dictionary of geological terms: New York, Doubleday, various pagination.

Buchanan, T.J., and Somers, W.P., 1968, Stage measurements at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A7, 28 p.

———1969, Discharge measurements at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations Report, book 3, chap. A8, 65 p.

Burns and McDonnell, 1995, *Equus* Beds Groundwater Recharge Demonstration environmental assessment: Kansas City, Missouri, Burns and McDonnell Engineering Consultants, 120 p.

Christensen, V.G., and Ziegler, A.C., 1998a, Atrazine in source water intended for artificial ground-water recharge, south-central Kansas: U.S. Geological Survey Fact Sheet 074–98, 4 p.

———1998b, Characterization of atrazine in source water—a case study from Kansas, *in* Proceedings from the Water Quality Technical Conference: Denver, Colorado, American Water Works Association, 9 p.

Edwards, T.K., and Glysson, G.D., 1988, Field methods for measurement of fluvial sediment: U.S. Geological Survey Open-File Report 86–531, 118 p.

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.

Helsel, D.R., and Hirsch, R.M., 1992, Statistical methods in water resources: New York, Elsevier, 522 p.

Hem, J.D., 1992, Study and interpretation of chemical characteristics of natural water (3d ed.): U.S. Geological Survey Water-Supply Paper 2254, 263 p.

Kansas Department of Agriculture and U.S. Department of Agriculture, 1999, Kansas farm facts: Topeka, Kansas, 114 p.

Kansas Department of Health and Environment, 1997, Kansas farm facts: Topeka, Kansas, 118 p.

———1999, List of water quality limited stream segments impacted predominantly by nonpoint and point sources: accessed March 2, 2000, at URL http://www.kdhe.state.ks.us/befs/303d/table1.htm

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.

———1984, Discharge ratings at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A10, 59 p.

Lane, C.W., and Miller, D.E., 1965, Geohydrology of Sedgwick County, Kansas: Kansas Geological Survey Bulletin 176, p. 14.

MathSoft, Inc., 1999, S-Plus 2000, modern statistics and advanced graphics, users guide: Cambridge, Massachusetts, 558 p.

Myers, N.C., Hargadine, G.D., and Gillespie, J.B., 1996, Hydrologic and chemical interactions of the Arkansas River and the *Equus* Beds aquifer between Hutchinson and Wichita, south-central Kansas: U.S. Geological Survey Water-Resources Investigations Report 95–4191, 100 p.

Ott, R.L., 1993, An introduction to statistical methods and data analysis: Belmont, California, Duxbury Press, 1051 p.

Putnam, J.E., Lacock, D.L., Schneider, D.R., and Carlson, M.D., 2000, Water resources data, Kansas, water year 1999: U.S. Geological Survey Water-Data Report KS–99–1, p. 288 and 302.

SAS Institute Inc., 1996, SAS/STAT user's guide, version 6.12: Cary, North Carolina.

Stramel, G.J., 1956, Progress report on the ground-water hydrology of the *Equus* Beds area: Kansas Geological Survey Bulletin 119, part 1, 59 p.

Topping, D.J., Rubin, D.M., Nelson, J.M., Kinzel, P.J., III, and Corson, I.C., 2000, Colorado River sediment transport, part 2. Systematic bed-elevation and grain-size effects of sandy supply limitation: Water Resources Research, February 2000, v. 36, no. 2, p. 543.

U.S. Environmental Protection Agency, 1986, Quality criteria for water, 1986: U.S. Environmental Protection Agency Report 440/5–86–001, unnumbered pages.

———1995, Drinking water regulations and health advisories: U.S. Environmental Protection Agency, 16 p.

———1999, Current drinking water standards: accessed October 12, 1999, at URL http://www.epa.gov/OST/Tools/destds.html

Wilde, F.D., and Radke, D.B., eds., 1998, 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, p. 3–20.

Ziegler, A.C., Christensen, V.G., and Ross, H.C., 1999, Preliminary effects of artificial recharge on ground-water quality, 1995–98—*Equus* Beds Ground-Water Recharge Demonstration Project, south-central Kansas: U.S. Geological Survey Water-Resources Investigations Report 99–4250, 74 p.

Ziegler, A.C., and Combs, L.J., 1997, Baseline data-collection and quality-control protocols and procedures for the *Equus* Beds Ground-Water Recharge Demonstration Project near Wichita, Kansas, 1995–96: U.S. Geological Survey Open-File Report 97–235, 57 p.

Back cover photograph: Little Arkansas River near Halstead during high flow, November 1999
(photograph taken by Trudy Bennett, USGS, Wichita, Kansas).

Victoria G. Christensen and others—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—USGS/WRIR 00–4126

Printed on recycled paper