Flow-Adjusted Trends in Dissolved Selenium Load and Concentration in the Gunnison and Colorado Rivers near Grand Junction, Colorado, Water Years 1986–2008

Scientific Investigations Report 2012–5088

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

Flow-Adjusted Trends in Dissolved Selenium Load and Concentration in the Gunnison and Colorado Rivers near Grand Junction, Colorado, Water Years 1986–2008

By John W. Mayo and Kenneth J. Leib

Scientific Investigations Report 2012–5088

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

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2012

Contents

Figures

Tables

Conversion Factors

Multiply	By	To obtain
Length		
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
acre	4,047	square meter (m^2)
square mile (mi^2)	2.590	square kilometer (km^2)
Volume		
cubic foot (ft^3)	0.028317	cubic meter (m^3)
acre-foot (acre-ft)	1,233	cubic meter (m^3)
Flow		
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
Mass		
pound avoirdupois (lb avdp)	0.4536	kilogram (kg)
pound per day (lb/d)	0.9072	kilogram per day (kg/d)
pound per year (lb/yr)	0.9072	kilogram per year (kg/yr)

Water year in this report is defined as the period from October 1st of one year through September 30th of the following year and is named for the year of the ending date. The term "annual" in this report always refers to a water year.

Flow-Adjusted Trends in Dissolved Selenium Load and Concentration in the Gunnison and Colorado Rivers near Grand Junction, Colorado, Water Years 1986–2008

By John W. Mayo and Kenneth J. Leib

Abstract

As a result of elevated selenium concentrations, many western Colorado rivers and streams are on the U.S. Environmental Protection Agency 2010 Colorado 303(d) list, including the main stem of the Colorado River from the Gunnison River confluence to the Utah border. Selenium is a trace metal that bioaccumulates in aquatic food chains and can cause reproductive failure, deformities, and other adverse impacts in birds and fish, including several threatened and endangered fish species. Salinity in the upper Colorado River has been the focus of source-control efforts for many years. Although salinity loads and concentrations have been previously characterized at the U.S. Geological Survey (USGS) streamflow-gaging stations at the Gunnison River near Grand Junction, Colo., and at the Colorado River near the Colorado-Utah State line, trends in selenium load and concentration at these two stations have not been studied. The USGS, in cooperation with the Bureau of Reclamation and the Colorado River Water Conservation District, evaluated dissolved selenium (herein referred to as "selenium") load and concentration trends at these two sites to inform decision makers on the status and trends of selenium.

This report presents results of the evaluation of trends in selenium load and concentration for two USGS streamflow-gaging stations: the Gunnison River near Grand Junction, Colo. ("Gunnison River site"), USGS site 09152500, and the Colorado River near Colorado-Utah State line ("Colorado River site"), USGS site 09163500. Flow-adjusted selenium loads were estimated for the beginning water year (WY) of the study, 1986, and the ending WY of the study, 2008.

The difference between flow-adjusted selenium loads for WY 1986 and WY 2008 was selected as the method of analysis because flow adjustment removes the natural variations in load caused by changes in mean-daily streamflow, emphasizing human-caused changes in selenium load and concentration. Overall changes in human-caused effects in selenium loads and concentrations during the period of study are of primary interest to the cooperators. Selenium loads for each of the 2 water years were calculated by using normalized mean-daily streamflow, measured selenium concentration, standard linear regression techniques, and data previously collected at the two study sites. Mean-daily streamflow was normalized for each site by averaging the daily streamflow for each day of the year over the 23-year period of record. Thus, for the beginning and ending water years, estimations could be made of loads that would have occurred without the effect of year-to-year streamflow variation. The loads thus calculated are illustrative of the change in loads between water years 1986 and 2008, and are not the actual loads that occurred in those 2 water years.

The estimated 50th and 85th percentile selenium concentrations associated with the selenium loads were also calculated for WY 1986 and WY 2008 at each site. Time-trends in selenium concentration at the two sites were charted by using regression techniques for partial residuals for the entire study period (WY 1986 through WY 2008).

Annual selenium load for the Gunnison River site was estimated to be 23,196 pounds for WY 1986 and 16,560 pounds for WY 2008, a 28.6 percent decrease. Lower and upper 95-percent confidence levels for WY 1986 annual load were 22,360 and 24,032 pounds. Lower and upper 95-percent confidence levels for WY 2008 annual load were 15,724 and 17,396 pounds. Estimated 50th percentile daily selenium concentrations decreased from 6.41 to 4.57 micrograms/liter from WY 1986 to WY 2008, whereas estimated 85th percentile daily selenium concentrations decreased from 7.21 to 5.13 micrograms/liter from WY 1986 to WY 2008.

Annual selenium load for the Colorado River site was estimated to be 56,587 pounds for WY 1986 and 34,344 pounds for WY 2008, a 39.3 percent decrease. Lower and upper 95-percent confidence levels for WY 1986 annual load were 53,785 and 59,390 pounds. Lower and upper 95-percent confidence levels for WY 2008 annual load were 31,542 and 37,147 pounds. Estimated 50th percentile daily selenium concentrations decreased from 6.44 to 3.86 micrograms/liter from WY 1986 to WY 2008, whereas estimated 85th percentile daily selenium concentrations decreased from 7.94 to 4.72 micrograms/liter from WY 1986 to WY 2008.

Introduction

Selenium impairment of stream segments from nonpoint sources in western Colorado is of concern to local, State, and Federal governments, local water providers, and local land users. As a result of elevated selenium concentrations, many western Colorado rivers and streams are on the U.S. Environmental Protection Agency (EPA) 2010 Colorado 303(d) list (Colorado Department of Public Health and Environment, 2010), including the main stem of the Colorado River from the Gunnison River confluence to the Utah border (U.S. Environmental Protection Agency, 2011). The term "303(d) list" refers to the list of impaired and threatened streams, river segments, and lakes that all States are required to submit for EPA approval every 2 years. The States identify all waters where required pollution controls are not sufficient to attain or maintain applicable water-quality standards.

Selenium is a trace element that bioaccumulates in aquatic food chains and can cause reproductive failure, deformities, and other adverse impacts in birds and fish, which may include some threatened and endangered fish species native to the Colorado River (Hamilton, 1998; Lemly, 2002). The Colorado River along with portions of Colorado River tributaries in the Grand Valley of western Colorado located within the 100-year flood plain of the Colorado River are designated critical habitat for four fish species listed under the Endangered Species Act—the Colorado Pikeminnow, Razorback Sucker, Bonytail, and Humpback Chub (U.S. Fish & Wildlife Service, 2011).

Salinity in the upper Colorado River basin has been the focus of source-control efforts for many years (Kircher and others, 1984; Butler, 1996). Salinity is also referred to as total dissolved solids in water, or TDS. In response to the Salinity Control Act of 1974, the Bureau of Reclamation (USBR) and the Natural Resources Conservation Service have focused on salinity control since 1979 through the Colorado River Basin Salinity Control Program. The primary methods of salinity reduction are the lining of irrigation canals and laterals and assisting farmers to establish more efficient irrigation practices (Colorado River Salinity Control Forum, 2011). Starting in 1988, the National Irrigation Water Quality Program (NIWQP), a Federal-agency board, began investigations to determine whether selenium and other trace elements from irrigation drainage were having an adverse effect on water quality in the Western United States. The NIWQP investigations found high concentrations of selenium in water, biota, and sediment samples (Butler, 1996; Butler and others, 1996). These previous investigations determined that a relation exists between subbasin characteristics (such as selenium-rich shale outcrops, agricultural practices, and irrigation-water delivery-system design) and salinity and selenium loads in certain subbasins.

Although salinity loads and concentrations have been previously characterized for the U.S. Geological Survey (USGS) streamflow-gaging stations at Colorado River near Colorado-Utah State line and Gunnison River near Grand Junction, Colo. (Kircher and others, 1984; Butler, 1996; Vaill and Butler, 1999; Butler, 2001; Leib and Bauch, 2008), trends

in selenium at these two stations have not been studied. The Gunnison Basin and Grand Valley Selenium Task Forces have expressed a need to better understand selenium trends in the Gunnison and Colorado Rivers (Gunnison Basin & Grand Valley Selenium Task Forces, 2012).

The USGS, in cooperation with the USBR and the Colorado River Water Conservation District, evaluated the dissolved-selenium load and concentration trends at two streamflow-gaging stations in western Colorado to inform decision makers on the status and trends of selenium. For the purposes of this report, dissolved selenium load or concentration will be referred to as selenium load or concentration. This report presents results of the evaluation of flow-adjusted trends in selenium load and concentration for two USGS streamflow-gaging stations near Grand Junction, Colo. Flow-adjusted selenium loads were estimated for the beginning water year (WY) of the study, 1986, and the ending WY of the study, 2008. (A water year is the period from October 1st of one year through September 30th of the following year and is named for the year of the ending date. The term "annual" in this report always refers to a water year.)

The difference between flow-adjusted selenium loads for WY 1986 and WY 2008 was selected as the method of analysis because flow adjustment removes the natural variations in load caused by changes in mean-daily streamflow, emphasizing human-caused changes in selenium load and concentration. Overall changes in human-caused effects in selenium loads and concentrations during the period of study are of primary interest to the cooperators. Flow-adjusted selenium loads for each of the 2 water years were calculated by using normalized mean-daily streamflow, measured selenium concentration, standard linear regression techniques, and data previously collected at the two study sites. Mean-daily streamflow was normalized for each site by averaging the daily streamflow for each day of the year over the 23-year period of record. Thus, for the beginning and ending water years, estimations could be made of loads that would have occurred without the effect of year-to-year streamflow variation. The calculated loads would be illustrative of the change in loads between water years 1986 and 2008 and would not be the actual loads that occurred in those two water years.

The estimated 50th and 85th percentile selenium concentrations associated with the selenium loads were calculated for WY 1986 and WY 2008 for each site. The percentile values are presented in this report because regulatory agencies in Colorado make 303(d) selenium compliance decisions based on concentration percentile values. Also, time-trends in selenium concentration at the two sites were demonstrated by using regression techniques for partial residuals for the entire study period (WY 1986 through WY 2008).

Study Area

The study area (fig. 1) includes two sites: Gunnison River near Grand Junction, Colo. (herein referred to as the "Gunnison River site"), USGS site 09152500, and Colorado River near Colorado-Utah State line (herein referred to as the
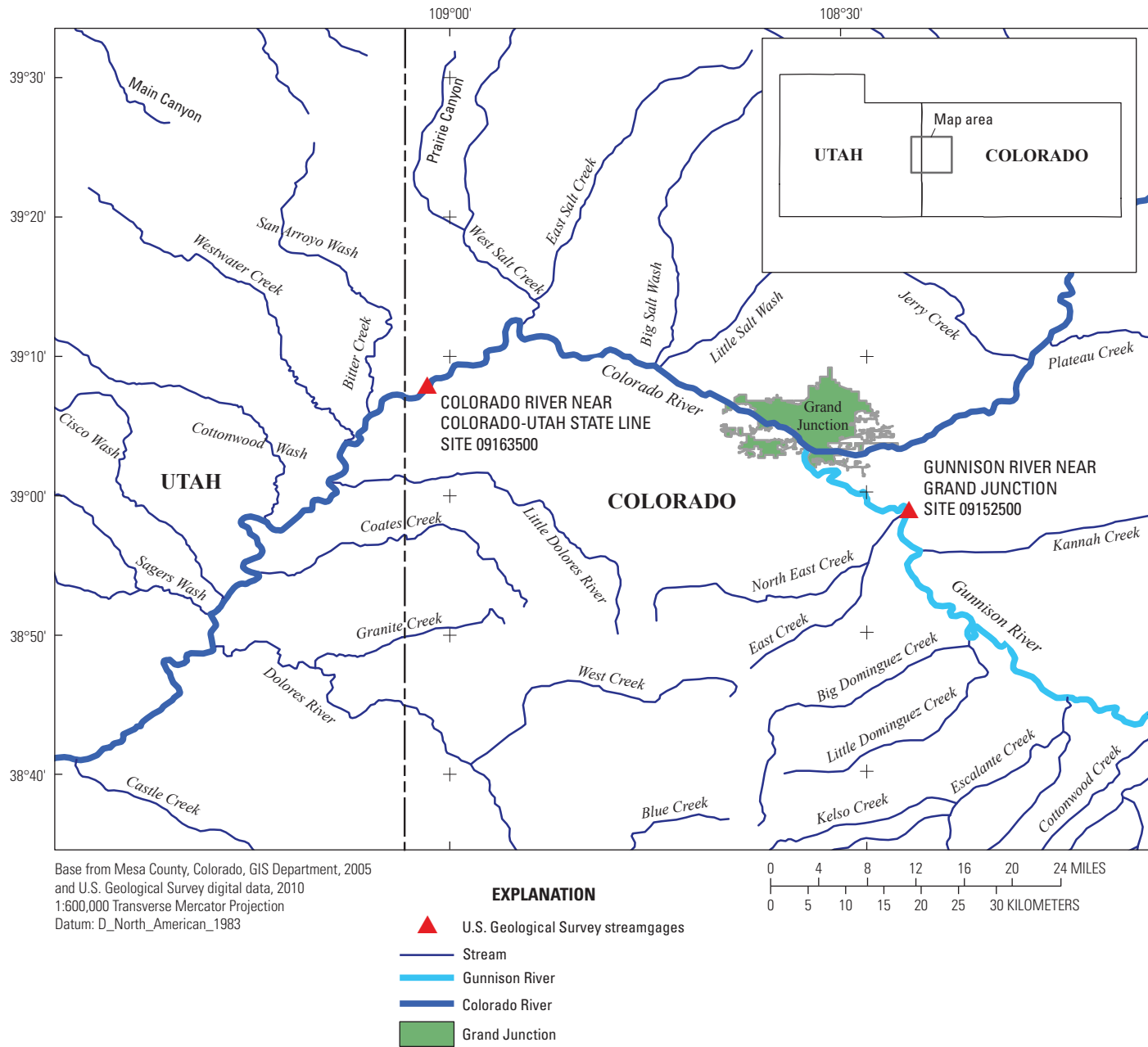
Figure 1. Location of the study sites: USGS streamflow-gaging stations 09152500, Gunnison River near Grand Junction, Colorado, and 09163500, Colorado River near Colorado-Utah State line.

"Colorado River site"), USGS site 09163500. Detailed information about these sites can be found on the USGS National Water Information System (NWIS) Web site at

   http://waterdata.usgs.gov/co/nwis/inventory/?site_no=09152500&amp (Gunnison River site) and *http://waterdata.usgs.gov/co/nwis/inventory/?site_no=09163500&amp* (Colorado River site).

Data Sources

   Daily streamflow and periodic selenium concentration data were retrieved from the USGS NWIS (*http://waterdata.usgs.gov/nwis*). The analyzed period of record for both sites was WY 1986 through WY 2008. Typically, three to five water samples were collected at each site per year and analyzed for selenium (dissolved fraction) concentration; the samples were filtered at collection time through a 0.45-µm filter as described in the USGS National Field Manual (U.S. Geological Survey, variously dated).

   The selenium samples were analyzed at the USGS National Water Quality Lab (NWQL). Prior to WY 2000, the NWQL established a Minimum Reporting Level (MRL) for each constituent as the less-than value (<) reported to customers. The MRL is the value reported when a constituent either is not detected or is detected at a concentration less than the MRL (Childress and others, 1999). If a measured value fell below the MRL, the entry into NWIS was shown at the MRL with a less-than symbol in the remark column

for that parameter. (For example, < 1.0 indicates that the value was not necessarily zero but was below the minimum reporting level of 1.0 µg/L.) This limits the false negative rate of reported values. There were three samples in the study between WY 1991 and WY 1997 that fell below the MRL of 1.0 µg/L (tables 8 and 9 in the Supplemental Data section at the back of the report).

   Starting in WY 2000, the NWQL established both a Long Term Method Detection Level (LT MDL) and a Laboratory Reporting Level (LRL), which is set at twice the LT MDL. The LT MDL is the lowest concentration of a constituent that is reported by the NWQL and represents that value at which the probability of a false positive is statistically limited to less than or equal to 1 percent. The LRL represents the value at which the probability of a false negative is less than or equal to 1 percent (Childress and others, 1999). Measured values that fell below the LRL but above the LT MDL were entered with their measured value and an "E" (for estimate) in the remark column in NWIS. Values that fell below the LT MDL were shown in NWIS as less than (<) the LRL value. In WY 2000, one study value was reported as 2.0 with a remark code of "E" (table 8 in the Supplemental Data section at the back of the report).

   All data were analyzed and quality assured according to standard USGS procedures and policies (U.S. Geological Survey, variously dated; Patricia Solberg, U.S. Geological Survey, written commun., 2010). The data are summarized in table 1, and shown in detail in tables 8 and 9 in the Supplemental Data section of the report.

Table 1.   Summary of U.S. Geological Survey National Water Information System (NWIS) records for study sites, water years 1986–2008.

[<, less than; µg/L, micrograms per liter]

Study site and number	Number of daily streamflow values	Number of dissolved selenium concentrations	Number of censored dissolved selenium concentrations (<1.0 µg/L)[1]	Number of estimated dissolved selenium concentrations[2]
Gunnison River (09152500)	8,401	171	1	1
Colorado River (09163500)	8,401	198	2	0

   [1]Censored values are automatically handled by the regression software.

   [2]Estimated concentration value had been set equal to 2.0 µg/L in NWIS. Estimated values are automatically handled by the regression software.

Organization of this Report

The results of this study are of interest to a broad section of the community. Therefore, the mathematical tools and principles used to arrive at the conclusions are discussed in considerable detail in order to meet the needs of such a broad audience. Development of a regression model to estimate load and concentration can best be described as a three-step process (Runkel and others, 2004):

1. Model Formulation. The section titled "Study Methods and Model Formulation" provides justification for the choice of model form and describes the technique for model building.

2. Model Calibration. The section titled "Regression Model Calibration" shows the selected models with estimated coefficients and describes the diagnostics used to validate the model's accuracy.

3. Load Estimation. The section titled "Flow-Adjusted Trends in Selenium Load and Concentration" gives the results of the model, which are the estimated flow-adjusted trends in selenium loads and concentrations.

Study Methods and Model Formulation

This section of the report discusses the technique of flow-adjusted trend analysis, the methods used for regression analysis, and the use of regression analysis software in the study. The concept of normalized streamflow is explained, and the estimation of load and concentration trends is shown.

General Approach of the Analysis

Regression analysis is a long-accepted and widely used method for analyzing trends in water-quality constituents (Kircher and others, 1984; Butler, 1996; Richards and Leib, 2011). Variables selected to estimate trends in water-quality constituents in these types of studies commonly include daily streamflow, time, and measured constituent (selenium) values. Various transformations are commonly used to enhance estimation accuracy (logarithmic (log) transformation, quadratic terms, decimal time, centered time, and sinusoidal transformations of time). In addition, seasonality variables such as irrigation season for a river with managed flow can be included to increase the accuracy of the estimation (Kircher and others, 1984). For this study, daily streamflow, decimal time, various transformations, and irrigation season were used in estimating trends in selenium load and concentration.

Flow-Adjusted Trend Analysis

Trends in loads and concentrations of water-quality constituents can be approached from two perspectives: nonflow-adjusted (which shows the overall influence from both human and natural factors) and flow-adjusted (which removes natural streamflow variability and emphasizes human-caused influences) (Sprague and others, 2006). Only flow-adjusted trend analyses were performed at the two sites in this study because the effect of selenium-control efforts over the study period was of primary interest to the cooperators.

Normalized Mean-Daily Streamflow

Daily streamflow values were averaged to produce a mean-daily streamflow (Q_n) for each day of the calendar year over the 23-year period of record. An averaging function available on the NWIS Web site (*http://waterdata.usgs.gov/co/nwis/dvstat/*) was used to calculate these normalized mean-daily streamflow values. For example, an average of all the January 1st daily streamflow values was calculated for January 1, 1986 through January 1, 2008. This creates a Q_n value for January 1st over the 23-year period. By calculating a similar Q_n for every day of the year, the year-to-year fluctuations in daily streamflow are removed when computing daily selenium loads.

Mean-daily streamflow (Q_n) was only used to compare the changes in selenium load and concentration between water years 1986 and 2008. It is important to remember that because the estimated loads and concentrations given for WY 1986 and WY 2008 were based on normalized streamflow, the results were only illustrative of the change in selenium loads and concentrations over the period of study. They were not the actual loads and concentrations that occurred in WY 1986 and WY 2008.

Regression Analysis

This section of the report discusses the principle of multiple linear regression, the use of regression analysis software in this study, estimation accuracy of the regression model, the calculation of percentile values for selenium concentration, and the indication of selenium load and concentration trends.

Multiple Linear Regression

Ordinary least squares (OLS) regression, commonly referred to as "linear regression," is an analytical tool that seeks to describe the relation between one or more variables of interest and a response variable. Simple linear regression models use one variable of interest, whereas multiple linear

regression models use more than one variable of interest (Helsel and Hirsch, 2002). Each variable of interest explains part of the variation in the response variable. Regression is performed to estimate values of the response variable based on knowledge of the variables of interest. For example, in this report the variables of interest were daily streamflow, time, and irrigation season, with the response variable being selenium load.

The general form of the multiple linear regression model is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k + \varepsilon \qquad (1)$$

where

y	is the response variable,
β_0	is the intercept on the y-axis,
β_1	is the slope coefficient for the first explanatory variable,
β_2	is the slope coefficient for the second explanatory variable,
β_k	is the slope coefficient of the k*th* explanatory variable,
$x_1...x_k$	are the variables of interest, and
ε	is the remaining unexplained variability in the data (the error).

Log-Linear Regression Models

Linear regression only works if there is a linear relation between the explanatory variables and the response variable. In some circumstances where the relation is not linear, it is possible to transform the explanatory and response variables mathematically so that the transformed relation becomes linear (Helsel and Hirsch, 2002). A common transformation that achieves this purpose is to take the natural logarithm (ln) of both sides of the model, as in this simplified selenium concentration example utilizing streamflow and time as the explanatory variables:

$$ln(C) = \beta_0 + \beta_1 ln(Q) + \beta_2 ln(T) + \varepsilon \qquad (2)$$

where

$ln(\ )$	is the natural logarithm function,
C	is concentration of selenium,
β_0	is the intercept on the y-axis,
β_1, β_2	are the slope coefficients for the two explanatory variables,
Q	is daily streamflow,
T	is time, and
ε	is the remaining unexplained variability in the data (the error).

The resulting log-linear model has been found to accurately estimate the relation between streamflow, time, and the concentration of constituents (selenium in this instance). Load estimates assuming the validity of a log-linear relation appear to be fairly

insensitive to modest amounts of model misspecification or nonnormality of residual errors (Cohn and others, 1992). Any bias that is introduced by the log transformation needs to be corrected when the results are transformed out of log space (Cohn and others, 1989), but this is automatically applied by the statistical software used for the regression analysis.

Regression Analysis Software

To build the regression model, the USGS software program S-LOADEST was selected because it is designed to calculate constituent loads using daily streamflow, time, seasonality, and other explanatory variables. S-LOADEST was derived from LOADEST (Runkel and others, 2004) and is provided by the USGS (David L. Lorenz, U.S. Geological Survey, electronic commun., January 12, 2009) in their internal distribution of the statistical software program Tibco Spotfire S+ (Tibco Software, Inc., 1988–2008). S-LOADEST was used to calculate daily selenium loads and concentrations from measured selenium-concentration calibration data spanning WY 1986 through WY 2008.

Automatic Variable Selection for Models

S-LOADEST can be used with a predefined/automatic model selection option or with a custom model selection option defined by the user. In the predefined option, S-LOADEST automatically selects the best regression model from among a set of nine predefined models, based on the lowest value of the Akaike Information Criterion (AIC). (The predefined models are listed in table 10 in the Supplemental Data at the end of the report). AIC is calculated for each of the nine models, and the lowest value of AIC determines the best model (Runkel and others, 2004).

The nine models use various combinations of daily streamflow, daily streamflow squared, time, time squared, and Fourier time-variable transformations. Compensation for differences in seasonal load is accomplished using Fourier variables. Fourier variables use sine and cosine terms to account for continual changes over the seasonal (annual) period.

Dummy variables (such as irrigation season) are used to account for abrupt seasonal changes (step changes) during the year. A dummy variable cannot be automatically included with the S-LOADEST predefined models; rather, it is added manually by the user as part of a custom S-LOADEST model.

The Adjusted Maximum Likelihood Estimation (AMLE) method of load estimation was selected in S-LOADEST because of the presence of censored selenium-concentration values (< 1.0 μg/L) in the calibration files. AMLE is an alternative regression method, similar to OLS regression, which is designed to correct for bias in the model coefficients caused by the inclusion of censored data (Runkel and others, 2004).

Data Centering and Decimal Time

S-LOADEST uses a "centering" technique to transform streamflow and decimal time (Runkel and others, 2004). The technique removes the effects of multicollinearity, which

arises when one of the explanatory variables is related to one or more of the other explanatory variables. Multicollinearity can be caused by natural phenomenon as well as mathematical artifacts, such as when one explanatory variable is a function of another explanatory variable. Multicollinearity is common in load-estimation models where quadratic terms of decimal time or log streamflow are included in the model (Cohn and others, 1992). Centering is automatically done by S-LOADEST for streamflow and decimal time using equation 3 for streamflow and equation 4 for decimal time:

$$\ln Q^* = \overline{\ln Q} + \frac{\sum_{i=1}^{N}\left(\ln Q_i - \overline{\ln Q}\right)^3}{2\sum_{i=1}^{N}\left(\ln Q_i - \overline{\ln Q}\right)^2} \quad \text{and} \quad \overline{\ln Q} = \frac{\sum_{i=1}^{N}\ln Q_i}{N} \quad (3)$$

where

$\ln Q^*$ is the natural logarithm of streamflow, centered value for the calibration dataset, in cubic feet per second;

$\overline{\ln Q}$ is the mean of the natural logarithm of streamflow in the dataset, in cubic feet per second;

$\ln Q_i$ is the natural logarithm of daily mean streamflow for day i, in cubic feet per second, and

N is the number of daily values in the dataset.

$$t^* = \bar{t} + \frac{\sum_{i=1}^{N}\left(t_i - \bar{t}\right)^3}{2\sum_{i=1}^{N}\left(t_i - \bar{t}\right)^2} \quad \text{and} \quad \bar{t} = \frac{\sum_{i=1}^{N}t_i}{N} \quad (4)$$

where

t^* is the time, centered value for the calibration dataset, in decimal years;

$\bar{t}$ is the mean of the time in the dataset, in decimal years;

t_i is time for day i, in decimal years, and

N is the number of daily values in the dataset.

S-LOADEST uses values of date and time that have been converted to decimal values. A decimal date consists of the integer value of the year with the day and time for that date added as a decimal value to the year. For example, July 16, 1987, at 12:00 pm as decimal time (expressed to 2 decimal places) would be 1987.54. The *dectime* term in S-LOADEST model equations is the difference between the decimal sample date and time and the decimal centered date and time for the study period in question. For the Gunnison River site, the center of decimal time for the study period WY 1986 through WY 2008 is 1997.44. The *dectime* value for July 16, 1987 at 12:00 pm would then be –9.90 (negative means that it is 9.90 years before the centered date).

Load and Concentration Estimation with Regression Models

To perform regression analysis in S-LOADEST, a calibration data set ("calibration file"), comprising rows of explanatory variables having a corresponding measured value of the response variable, is used with statistical analysis software to determine the intercept and slope coefficients of the explanatory variables. Then, by using the derived regression model with a set of estimation data ("estimate file") having rows of explanatory variables (without measured response variables), an estimated response variable can be calculated for each row of explanatory variables. The calibration data sets for this study are included in tables 8 and 9 of the Supplemental Data at the back of the report. The estimation data sets are simply the daily streamflow and irrigation-season code by date for each day of the study period at each site. For the irrigation season (April 1 through October 31), the irrigation-season code was set to 1, and for the non-irrigation season (November 1 through March 31), the irrigation-season code was set to 0.

Estimation Accuracy

One measure of the accuracy of a regression model is evaluated by computing the difference between each measured value of the response variable and its corresponding estimated value. This difference is called the residual value. Residual values are calculated by the equation:

$$e_i = y_i - \hat{y}_i \quad (5)$$

where

e_i is the estimated residual for observation i,

y_i is the *ith* value of the actual response variable, and

$\hat{y}_i$ is the *ith* value of the estimated response variable.

In order to ensure that the regression model is valid for use in estimations, a number of criteria are required to be met for the residuals: the residuals are normally distributed, are independent, and have constant variance (Helsel and Hirsch, 2002).

An important indicator of the accuracy of the regression model is residual standard error (RSE), which is the standard deviation of the residual values, and also the square root of the estimated residual variance. RSE is a measure of the dispersion (variance) of the data around the regression line. Low values of RSE (closer to zero) are desirable (Helsel and Hirsch, 2002). Another measure of how well the explanatory variables estimate the response variable is the coefficient of determination, R^2, which indicates how much of the variance in the response variable is explained by the regression model (Helsel and Hirsch, 2002). Values of R^2 range from 0.0 to 1.0, with higher values (closer to 1.0) showing more of the variance being explained by the model. R^2 also can be expressed as a percentage from 0 to 100 (used in this report). R^2 can be misleading in a load model, however. Because flow is found

on both sides of the equation, a model for a stream with lower variability in flow will have a lower R^2 than one with a higher variability in streamflow. Another caution is that when censored values exist in the data, the value of R^2 reported by S-LOADEST is an approximation (David L. Lorenz, U.S. Geological Survey, written commun., October 25, 2011). Values for RSE and R^2 are shown in the model calibration section for both sites. Only three values out of 369 selenium samples were censored, which is less than one percent of the data used in the analysis.

Each model coefficient (β_0, β_1, ... β_k) has an associated p-value, which is a measure of the "attained significance level" of the coefficient (Helsel and Hirsch, 2002). If the p-value is less than a chosen value (for example, 0.05), then the coefficient (and hence the corresponding variable of interest) is statistically significant in the regression model. The p-values for each coefficient are shown in the model calibration section for both sites. Another indicator of the model's accuracy is the estimation confidence interval, which shows, for each estimated value, an upper and lower value for which there is some level of probability (for example, 95 percent) that the estimated value falls between the upper and lower values.

Percentile Values for Concentrations

The 50th and 85th percentile values of estimated selenium concentration were calculated for WY 1986 and WY 2008 from the estimated daily selenium concentrations. The percentile values are presented in this report because regulatory agencies in Colorado make 303(d) selenium compliance decisions based on percentile values of concentration. It is important to note that these percentile values were calculated using normalized flow values, and only were illustrative of the changes in 50th and 85th percentile values between the two water years, rather than being actual values of concentration percentiles for the two water years.

Load and Concentration Trend Indication

The sign of the coefficient for the time variable in the regression model indicates any multi-year trend in selenium load and concentration over the study period (David K. Mueller, U.S. Geological Survey, written commun., March 14, 2011). If the sign of the time coefficient is positive, then the trend in selenium load is upward. If the sign of the time coefficient is negative, then the trend in selenium load is downward. The selenium concentration trend will follow the same trend as for the selenium load, and the residuals will be the same (David L. Lorenz, U.S. Geological Survey, written commun., October 28, 2011).

In order to demonstrate a time-trend in selenium concentration, regressions for partial residuals can be used which remove the time variables of interest from the regression model. Removal of any one of the variables of interest shows the effect of that variable on the regression model. By calculating regression partial residuals and plotting these partial residuals over the study period, the trend is shown graphically. Using a smoothing technique called Locally Weighted Scatterplot Smoothing (LOWESS), a line can then be fitted to the partial residuals to show the trend in selenium concentration over the study period (Helsel and Hirsch, 2002).

Model Diagnostics

There are five requirements for successful use of linear regression analysis (Helsel and Hirsch, 2002). These requirements are

1. The model form is correct: y is linearly related to x.

2. Data used to fit the model are representative of data of interest.

3. Variance of the residuals is constant.

4. The residuals are independent.

5. The residuals are normally distributed.

Selenium load and concentration for this study were observed to be linearly related to streamflow when log transformations were performed. The data used for the selenium load and concentration model (streamflow, time, selenium concentration) have been routinely collected for many years by the USGS and are the variables that represent the data of interest. Thus requirements 1 and 2 are deemed to be met.

For requirement 4, the independence of the data samples can be assumed from the fact that over the study period, the average number of days between samples was 48.8 days for the Gunnison River site, and was 41.9 days for the Colorado River site. To ensure sample independence, the USGS generally collected a minimum of 4 samples a year, one for each season, with rotation of the months that the samples were taken from year to year. Sampling during different streamflow regimes typically was planned (Steve Anders, U.S. Geological Survey, oral commun., October 28, 2011). Sampling intervals of 2 weeks or longer are considered necessary to ensure sample independence (David L. Lorenz, U.S. Geological Survey, written commun., October 25, 2011).

Diagnostic plots generated by S-LOADEST enable the user to determine whether requirements 3 and 5 have been met. These plots are of three types (Helsel and Hirsch, 2002; Runkel and others, 2004):

1. Q-Normal Plot. This shows quantiles of standard normal distribution on the x-axis and normalized residuals on the y-axis. A one-to-one line is included in the plot. If the plotted normalized residuals generally fall along the one-to-one line, then the residuals can be characterized as coming from a normal distribution.

2. Residuals versus Log-Fitted Values Plots. S-LOADEST generates two plots of this type:

a. S-L Plot. This shows the log-fitted values (selenium load in this report) on the x-axis and the square root of the absolute residuals on the y-axis. A LOWESS smoothing line is fitted to the residuals. The scatter of the residuals indicates how well the estimated values match their corresponding measured values. If the scatter of the residuals is random throughout the plot about the LOWESS line, if the residual points do not fall into curves or the variance does not change along the line, and if the LOWESS line is generally horizontal, then the results indicate that the residuals are normal, residual variance is acceptable, and the design of the model is valid. If discernible patterns in the residuals are seen, or the LOWESS line is not generally horizontal, then the residuals are not normal and their variance is not random. This indicates problems with the regression model such as the incorrect choice of variables of interest or problems with the calibration data.

b. Residuals versus Log-Fitted Values Plot. The interpretation of this plot is the same as for the S-L plot. The only difference is that the y-axis variable is the residual rather than the square root of the residual used in the S-L plot.

3. Residuals versus Explanatory Variables Plot(s). S-LOADEST will output a separate plot for each category of explanatory variable (streamflow, time, transformations of time, irrigation season) in the regression model. These plots indicate how the estimated selenium load values are varying with each explanatory variable. The desired condition is to have random distribution of the residuals over all explanatory variables. If the residuals are not randomly distributed, then the explanatory variable is biasing the estimation.

The interpretations of these diagnostic plots are given in the model calibration section for each model and site.

Regression Model Calibration

This section discusses the four calibration steps used for each site. These steps include selecting the initial regression model, testing the addition of irrigation season to the model, estimating selenium loads, and demonstrating any trend in selenium concentration.

Calibration Process Steps

The detailed steps followed for each site to select the regression model and get estimations of selenium load, selenium concentration, and time-trend in concentration were as follows:

1. Select a base regression model of selenium load using daily streamflow, decimal time, and various transformations of streamflow (squared) and decimal time (squared, Fourier). Test all variables of interest for statistical significance (p-value < 0.05). In addition, test for the validity of the various model assumptions such as linearity, uniformity of variance, normality, and independence of the variables.

2. Add irrigation season as a variable (step) of interest in the regression model from step 1, and test for statistical significance and model assumptions after the addition of irrigation season.

3. Use the selected load regression model from steps 1 or 2 with normalized streamflow to estimate daily and annual selenium loads for WY 1986 and WY 2008. Derive daily mean selenium concentrations from estimated loads and daily flows for WY 1986 and WY 2008.

4. Examine the coefficient for *dectime* to determine if a time trend exists in load and concentration. Demonstrate graphically any trend in selenium concentration over time by removing the *dectime* terms from the selected load regression model (regression technique for partial residuals), deriving estimated concentrations from the estimated daily loads, and charting the concentration residuals with a fitted LOWESS trend line over the years of the study period.

Gunnison River Site Calibration Steps

Gunnison River Step 1—Select the Initial Selenium Load Regression Model

The Gunnison River site data used to generate the regression model were 171 paired NWIS records of daily streamflow in cubic feet per second and selenium concentration values in micrograms per liter. The data were collected from November 26, 1985 to August 13, 2008 (Supplemental Data, table 8, back of report).

Predefined regression model 8 (table 10) was selected by S-LOADEST as having the lowest AIC value for the input data for the Gunnison River site:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \varepsilon \qquad (6)$$

where

ln	is the natural logarithm;
load	is selenium load, in pounds per day;
β_0	is the intercept of the regression on the y-axis;
$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	are regression coefficients;
Q	is centered daily streamflow, in cubic feet per second;
dectime	is centered decimal time, in decimal years;
sin(2π·dectime),	
cos(2π·dectime)	are sine and cosine Fourier functions;
ε	is the remaining unexplained variability in the data (the error); *and*
π	is *pi*, approximately 3.141593.

Table 2 lists the coefficients, p-values, RSE, centered streamflow, and centered decimal time for equation 6. All terms of equation 6 had p-values < 0.05, with the exception of the cosine term. It is necessary, however, to retain the cosine term if the sine term is used. The negative coefficient value for *dectime* (–0.016) indicates that the selenium load trend is downward over time.

S-LOADEST generated several diagnostic plots to examine model diagnostics. The Q-Normal plot indicated that the residuals were in a normal distribution. The Residuals versus Log-Fitted Values plot showed random distribution of the residuals, and the LOWESS line showed a slight bow upward toward the right (fig. 2). This plot indicated that the residual variance was acceptable. The slight bow upward of the fit line indicated some underestimation of loads for higher load values, but was deemed acceptable. Three other plots of residuals versus explanatory variables (streamflow, decimal time, and proportion of year) also indicated that there was a normal distribution of the residuals.

S-LOADEST reported an R^2 value of 53.48 for selenium load in equation 6. The RSE for selenium load in equation 6 was 0.255 pounds of selenium per day, which was determined to be an acceptable amount of scatter about the regression line.

Gunnison River Step 2—Test the Addition of Irrigation Season to the Base Regression Model

As a test, a daily binary dummy variable for irrigation season was added to equation 6 in S-LOADEST to test whether this improved the accuracy of the load estimation. Irrigation season provided a step change that cannot be modeled by Fourier functions. Equation 7 was used as a custom model in S-LOADEST with the same calibration data set as in step 1:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \beta_6 \cdot season + \varepsilon \quad (7)$$

where

β_6	is a regression coefficient,
season	is irrigation season, and
all other terms	are the same as for equation 6.

Table 3 lists the coefficients, p-values, RSE, centered streamflow, and centered decimal time for equation 7. The p-value for the irrigation season variable was 0.425, which indicated that irrigation season did not make a statistically significant contribution to the regression model for the Gunnison River site. The p-values > 0.05 for $ln(Q)^2$ and *cos(2π·dectime)* were not important in this instance because the decision to use equation 7 depended only on the p-value for *season*. As such, equation 7 was rejected because irrigation season did not make a significant contribution to the model for this site. Equation 6 was the selected model to determine selenium loads in step 3.

Table 2.   Regression results for selenium load model (equation 6), Gunnison River site.

[ln, natural logarithm; sin, sine function; cos, cosine function; π, pi; Q, daily streamflow; dectime, decimal time; <, less than; RSE, residual standard error; ft³/s, cubic feet per second]

Variable of interest	Coefficient	Coefficient value	p-value	RSE (pounds of selenium per day)	Centered streamflow (ft³/s)	Centered decimal time
intercept	β_0	3.988	<0.001	0.255	2,680	1997.44
ln(Q)	β_1	0.320	<0.001	0.255	2,680	1997.44
ln(Q)²	β_2	0.068	0.042	0.255	2,680	1997.44
dectime	β_3	–0.016	<0.001	0.255	2,680	1997.44
sin(2πdectime)	β_4	–0.223	<0.001	0.255	2,680	1997.44
cos(2πdectime)	β_5	0.047	0.122	0.255	2,680	1997.44
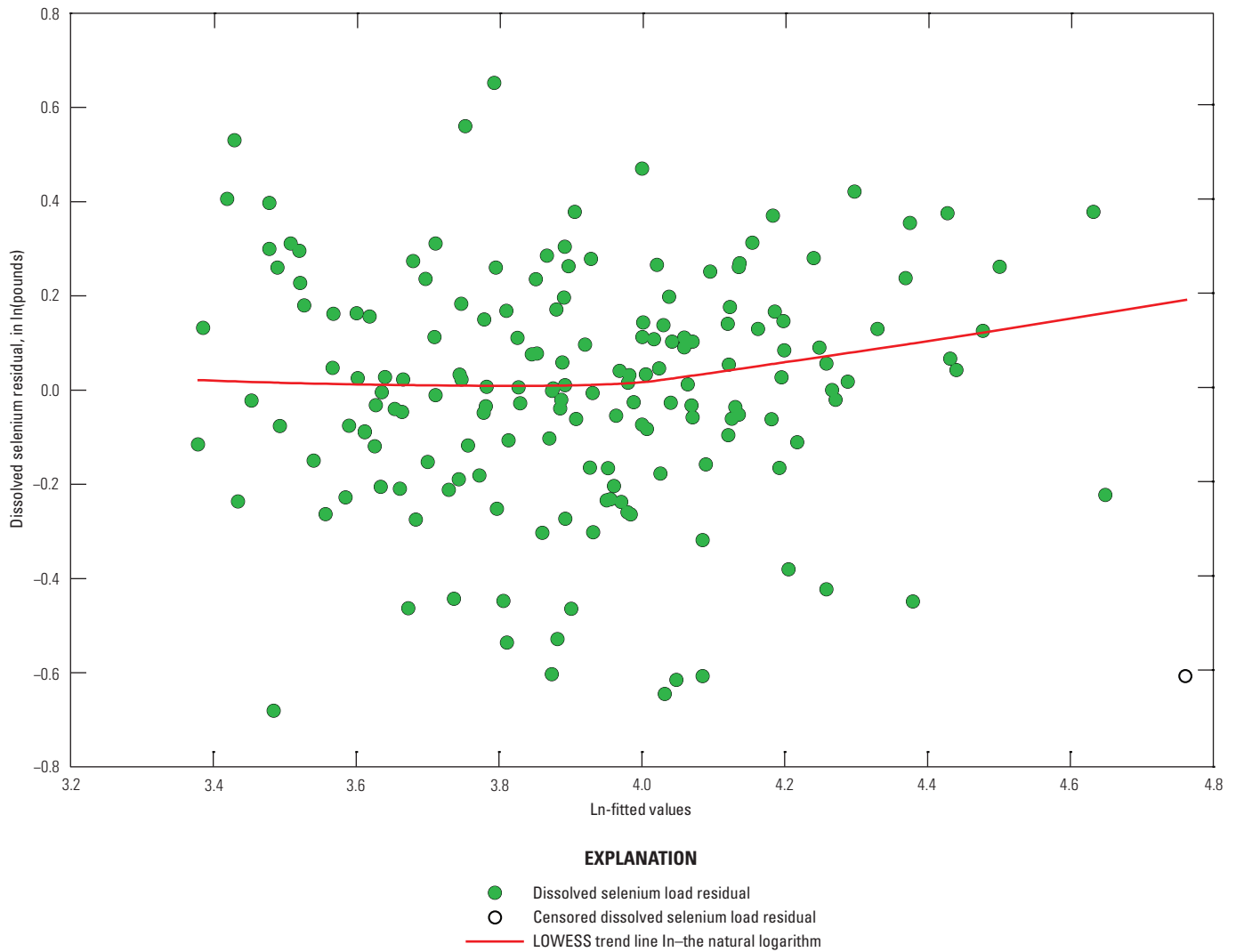
Figure 2. Dissloved selenium load residuals and LOWESS fit line using the step 1 load regression model (equation 6) for Gunnison River site, water years 1986–2008.

Table 3. Regression results for selenium load model (equation 7), Gunnison River site.

[ln, natural logarithm; sin, sine function; cos, cosine function; π, pi; Q, daily streamflow; dectime, decimal time; <, less than; RSE, residual standard error; ft³/s, cubic feet per second]

Variable of interest	Coefficient	Coefficient value	p-value	RSE (pounds of selenium per day)	Centered streamflow (ft³/s)	Centered decimal time
intercept	β_0	4.028	<0.001	0.256	2,680	1997.44
ln(Q)	β_1	0.325	<0.001	0.256	2,680	1997.44
ln(Q)²	β_2	0.065	0.051	0.256	2,680	1997.44
dectime	β_3	−0.015	<0.001	0.256	2,680	1997.44
sin(2πdectime)	β_4	−0.234	<0.001	0.256	2,680	1997.44
cos(2πdectime)	β_5	0.010	0.883	0.256	2,680	1997.44
irrigation season	β_6	−0.066	0.425	0.256	2,680	1997.44

Gunnison River Step 3—Estimate Selenium Loads for the First and Last Water Years of the Study Period

Equation 6 was used again in S-LOADEST, this time with an estimate file of normalized mean-daily streamflow (Q_n) from the 23-year period of record for only the first and last water years of the study period (WY 1986 and WY 2008). The model computed estimated daily and annual selenium loads and concentrations that illustrated the change between the two water years. S-LOADEST calculated the concentration from the estimated daily load value using equation 8 (David L. Lorenz, U.S. Geological Survey, electronic commun., January 12, 2009):

$$C = \frac{L}{(k \cdot Q_n)} \tag{8}$$

where

- C is selenium concentration, in micrograms per liter;
- L is selenium load, in pounds;
- k is a units conversion factor (0.005395); and
- Q_n is normalized mean-daily streamflow, in cubic feet per second.

The 50th and 85th percentile values of estimated selenium concentration were calculated for WY 1986 and WY 2008 from the estimated daily concentrations.

Gunnison River Step 4—Demonstrate Selenium Load and Concentration Trend over the Years of the Study

The β_3 coefficient for *dectime* in equation 6 had a value of –0.016 (table 2). The negative value indicated that the time-trend in selenium load and concentration from WY 1986 through WY 2008 was downward. This coefficient had a p-value <0.001, which meant that the time trend explanatory variable in equation 6 was statistically significant.

To demonstrate this downward time-trend of estimated selenium concentration over the study period graphically, the $\beta_3 \cdot dectime$ term was removed from equation 6, and a new load regression model was fitted in S-LOADEST:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \varepsilon \tag{9}$$

(The β_4 and β_5 terms are retained because these *dectime* terms repeat their cycle each year and do not contribute to a multi-year long-term trend.)

Variable removal was done to compute partial residuals that were plotted against time over the study period. Thus, equation 9 yielded estimated selenium load and concentration values for each day of the study period without the influence of decimal time in the regression. Equation 9 had an R^2 of 48.24 and a RSE of 0.268 pounds of selenium per day. The diagnostic plots indicated that there were no problems of residual normality or residual variance with the regression model.

The estimated daily selenium-concentration records from equation 9 were then paired by date (in Microsoft Access) with matching NWIS records of measured selenium concentration during the study period. This yielded pairs of measured and estimated values of selenium concentration by date. The residual values of measured selenium concentration minus estimated selenium concentration were calculated, and these residual values were plotted as a function of time over the study period (fig. 3). A LOWESS trend line for these residuals indicated a downward trend in selenium concentration over the study period.

Colorado River Site Calibration Steps

Colorado River Step 1—Select the Initial Selenium Load Regression Model

The Colorado River site data used to generate the regression model were 198 paired NWIS records of daily streamflow in cubic feet per second and selenium concentration in micrograms per liter. The data were collected between January 8, 1986 and August 12, 2008 (Supplemental Data, table 9, back of report).

Predefined regression model 9 was automatically selected by S-LOADEST for the Colorado River site:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \beta_4 \cdot dectime^2 + \beta_5 \cdot sin(2\pi \cdot dectime) + \beta_6 \cdot cos(2\pi \cdot dectime) + \varepsilon \tag{10}$$

where

- ln is the natural logarithm;
- $load$ is selenium load, in pounds per day;
- β_0 is the intercept of the regression on the y-axis;
- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6$ are regression coefficients;
- Q is centered daily streamflow, in cubic feet per second;
- $dectime$ is centered decimal time, in decimal years;
- $sin(2\pi \cdot dectime)$, $cos(2\pi \cdot dectime)$ are sine and cosine Fourier functions;
- ε is the remaining unexplained variability in the data (the error); and
- π is *pi*, approximately 3.141593.

The Q-Normal plot indicated that the residuals were in a normal distribution. The Residuals versus Log-fitted Values plot showed random distribution of the residuals, and the LOWESS line showed a generally horizontal fit (fig. 4). This plot indicated that the residual variance was acceptable. Three other plots of residuals versus explanatory variables (streamflow, decimal time, and proportion of year) also indicated that there was a normal distribution of the residuals.

EXPLANATION

● Dissolved selenium concentration partial residual
—— LOWESS trend line In–the natural logarithm
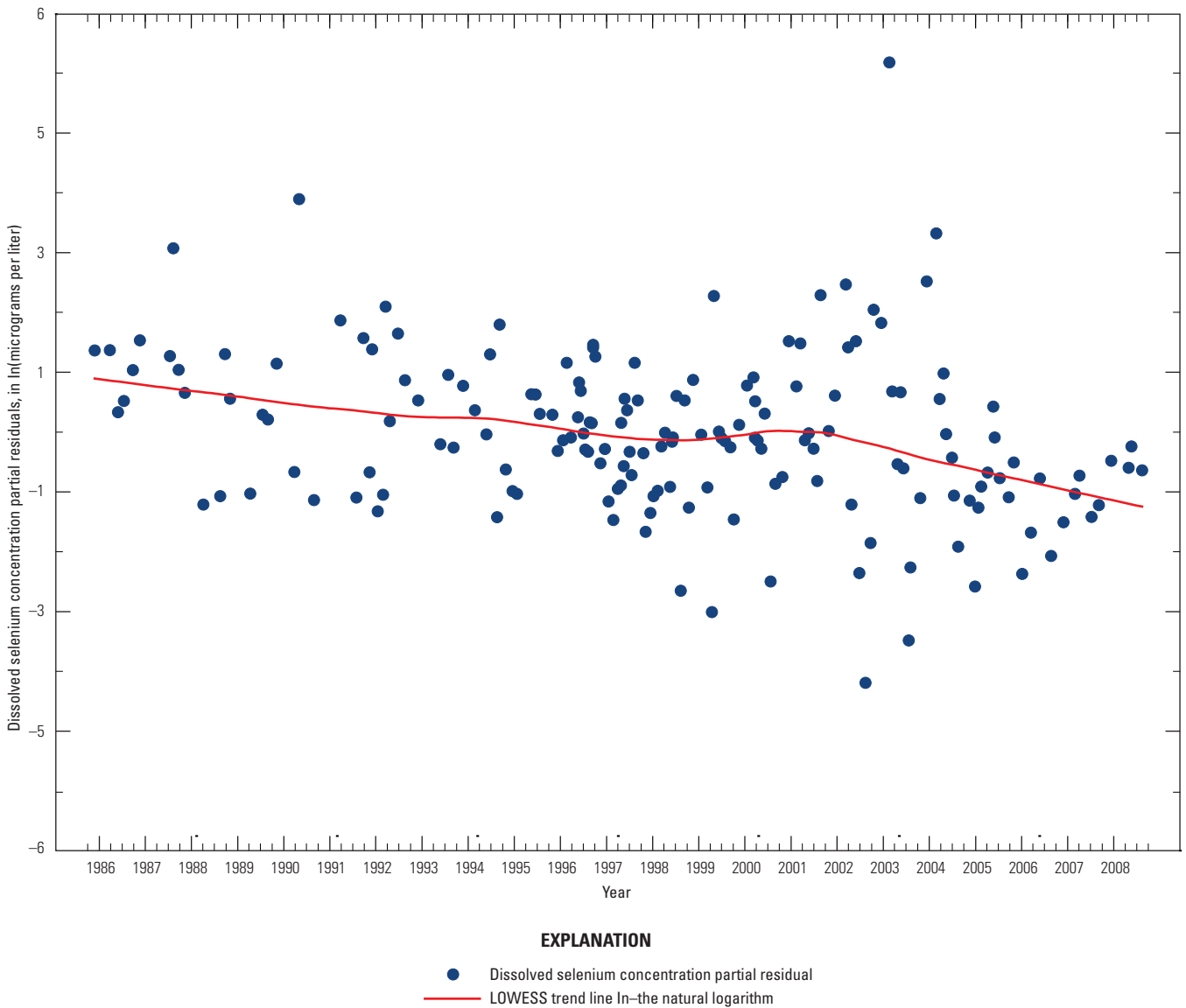
Figure 3. Dissloved selenium concentration partial residuals and LOWESS fit line using the step 4 regression model (equation 9) for Gunnison River site, water years 1986–2008.

Table 4 lists the coefficients, p-values, RSE, centered streamflow, and centered decimal time for equation 10. All terms of equation 10 had p-values <0.05 with the exception of the $ln(Q)^2$ term (0.145). The negative coefficient value for *dectime* (–0.021) indicated that the selenium load trend was downward over time.

The RSE for the load regression was 0.209 pounds of selenium per day, which was determined to be an acceptable amount of scatter about the regression line. The $ln(Q)^2$ term was dropped in subsequent steps because the p-value (0.145) was not significant, which yielded a new step 1 model in S-LOADEST:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot dectime + \beta_3 \cdot dectime^2 + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \varepsilon \quad (11)$$

Table 4.   Regression results for selenium load model (equation 10), Colorado River site.

[ln, natural logarithm; sin, sine function; cos, cosine function; π, pi; Q, daily streamflow; dectime, decimal time; <, less than; RSE, residual standard error; ft³/s, cubic feet per second]

Variable of interest	Coefficient	Coefficient value	p-value	RSE (pounds of selenium per day)	Centered streamflow (ft³/s)	Centered decimal time
intercept	β_0	4.593	<0.001	0.209	7,149	1997.76
ln(Q)	β_1	0.270	<0.001	0.209	7,149	1997.76
ln(Q)²	β_2	0.042	0.145	0.209	7,149	1997.76
dectime	β_3	−0.021	<0.001	0.209	7,149	1997.76
dectime²	β_4	0.002	<0.001	0.209	7,149	1997.76
sin(2πdectime)	β_5	−0.249	<0.001	0.209	7,149	1997.76
cos(2πdectime)	β_6	−0.085	<0.001	0.209	7,149	1997.76

EXPLANATION

● Dissolved selenium load residual
○ Censored dissolved selenium load residual
── LOWESS trend line ln–the natural logarithm

Figure 4.   Dissloved selenium load residuals and LOWESS fit line using the step 1 load regression model (equation 10) for Colorado River site, water years 1986–2008.

Colorado River Step 2—Test the Addition of Irrigation Season to the Base Regression Model

As a test, a daily binary dummy variable for irrigation season was added to equation 11 to test whether this improved the accuracy of the load estimation:

$$ln(load) = \beta_0 + \beta_1 \cdot lnQ + \beta_2 \cdot dectime + \beta_3 \cdot dectime^2 + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \beta_6 \cdot season + \varepsilon \quad (12)$$

where

β_6	is a regression coefficient,
season	is irrigation season, and
all other terms	are the same as for equation 10.

The regression model details for equation 12 are shown in table 5. The p-value for the irrigation season variable was 0.033, which indicated that irrigation season does make a statistically significant contribution to the regression model for the Colorado River site. Therefore, equation 12 was used to determine selenium loads in step 3.

The Q-Normal plot indicated that the residuals were in a normal distribution. The Residuals versus Log-fitted Values plot showed random distribution of the residuals, and the LOWESS line showed a generally horizontal fit (fig 5). This plot indicated that the residual variance was acceptable. Three other residual versus explanatory variable plots (streamflow, proportion of year, and season) also indicated that there was a normal distribution of the residuals.

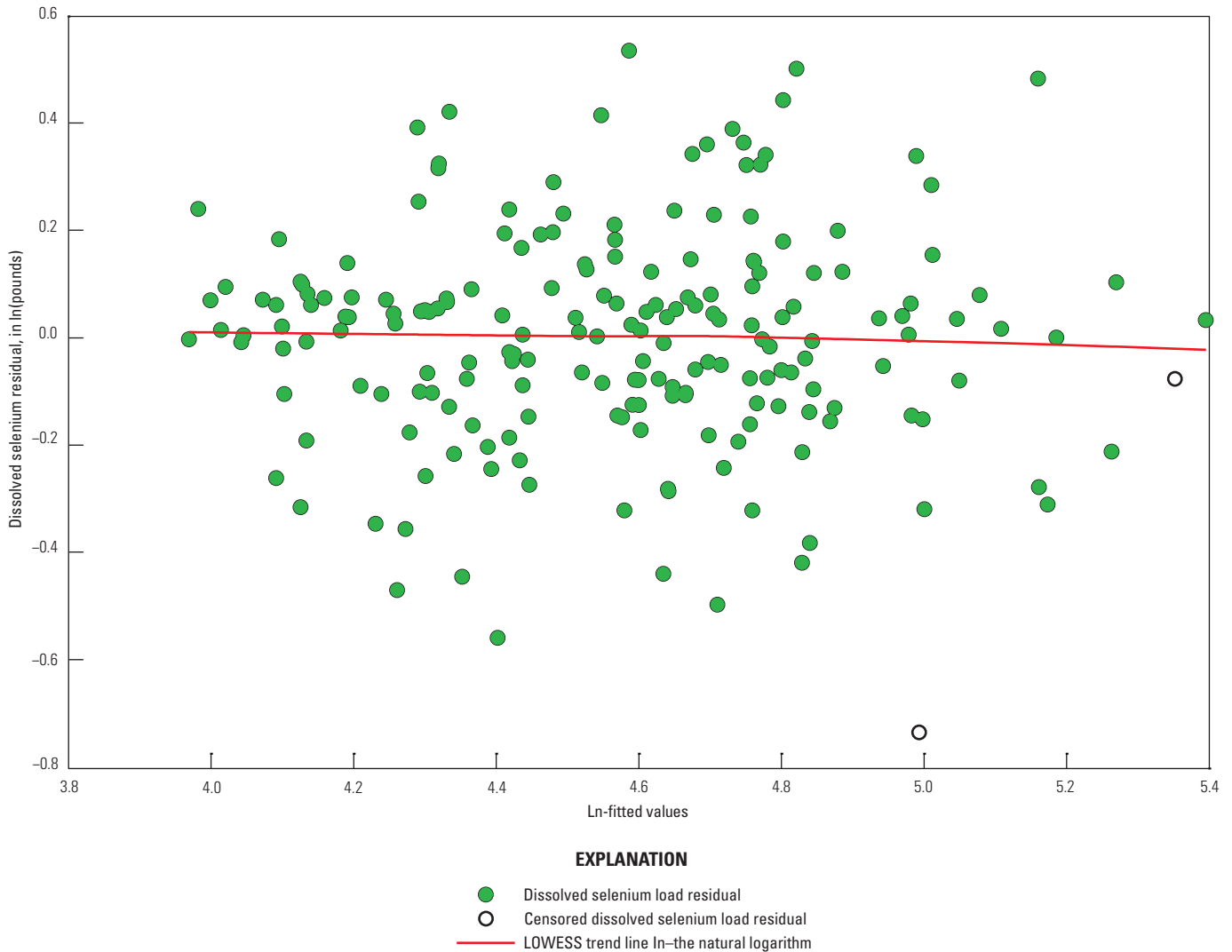S-LOADEST reported an R^2 value of 68.13 for selenium load in equation 12. The RSE for the load regression was 0.208 pounds of selenium per day, which was deemed to be an acceptable amount of scatter about the regression line.

Colorado River Step 3—Estimate Selenium Loads for the First and Last Water Years of the Study Period

Equation 12 was used again in S-LOADEST, this time with an estimate file of normalized mean-daily streamflow (Q_n) from the 23-year period of record for only the first and last water years of the study period (WY 1986 and WY 2008). This model computed estimated daily and annual selenium loads and concentrations that illustrated the change between the two water years. The 50th and 85th percentile values of estimated daily selenium concentration were also determined for WY 1986 and WY 2008.

Colorado River Step 4—Demonstrate Selenium Load and Concentration Trend over the Years of the Study

The β_2 coefficient for *dectime* in equation 12 was –0.021 (table 5). The negative value indicated that the time-trend in selenium load and concentration from WY1986 through WY 2008 was downward. This coefficient had a p-value <0.001, which meant that the time-trend explanatory variable in equation 12 was statistically significant.

To demonstrate this downward trend of estimated selenium concentration over the study period graphically, the $\beta_2 \cdot dectime$ and the $\beta_3 \cdot dectime^2$ terms were removed from equation 12 in S-LOADEST to yield a load regression for partial residuals:

$$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \beta_6 \cdot season + \varepsilon \quad (13)$$

Table 5. Regression results for selenium load model (equation 12), Colorado River site.

[ln, natural logarithm; sin, sine function; cos, cosine function; π, pi; Q, daily streamflow; dectime, decimal time; <, less than; RSE, residual standard error; ft³/s, cubic feet per second]

Variable of interest	Coefficient	Coefficient value	p-value	RSE (pounds of selenium per day)	Centered streamflow (ft³/s)	Centered decimal time
intercept	β_0	4.535	<0.001	0.208	7,149	1997.76
ln(Q)	β_1	0.261	<0.001	0.208	7,149	1997.76
dectime	β_2	–0.021	<0.001	0.208	7,149	1997.76
dectime²	β_3	0.002	<0.001	0.208	7,149	1997.76
sin(2πdectime)	β_4	–0.219	<0.001	0.208	7,149	1997.76
cos(2πdectime)	β_5	–0.015	0.712	0.208	7,149	1997.76
irrigation season	β_6	0.133	0.033	0.208	7,149	1997.76

EXPLANATION

- Dissolved selenium load residual
- Censored dissolved selenium load residual
- LOWESS trend line ln–the natural logarithm
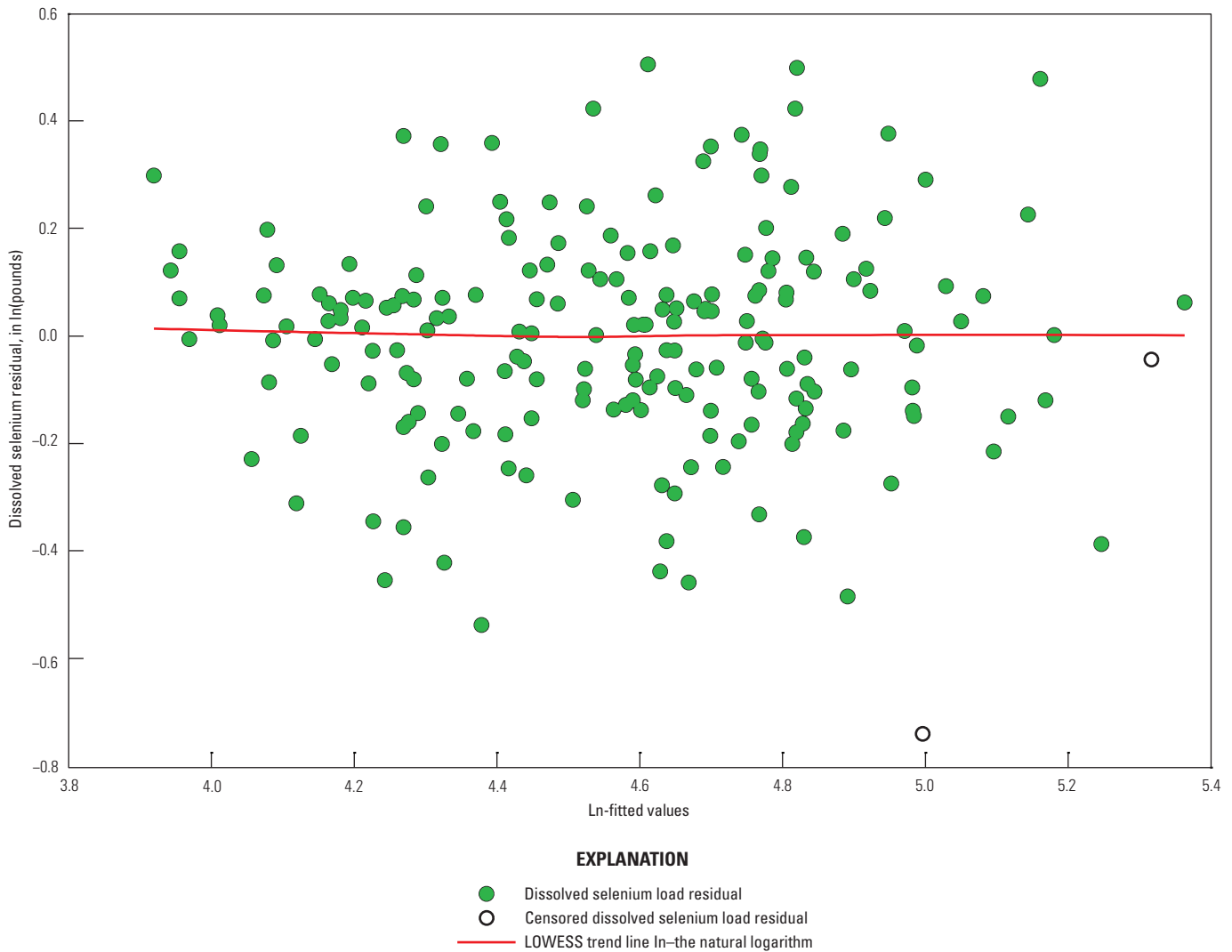
Figure 5. Dissloved selenium load residuals and LOWESS fit line using the step 2 load regression model (equation 12) for Colorado River site, water years 1986–2008.

Equation 13 yielded estimated selenium load and concentration values for each day of the study period, without the influence of decimal time in the regression. Equation 13 had an R^2 value of 55.01 and a RSE of 0.245 pounds of selenium per day. The diagnostic plots indicated no problems of residual normality or residual variance with the regression model.

The estimated daily selenium concentration records were paired by date (in Microsoft Access) with matching NWIS records of measured selenium concentration during the study period. This yielded pairs of measured and estimated values of selenium concentration by date. The residual values of measured selenium concentration minus estimated selenium concentration were calculated, and these residual values were plotted as a function of time over the study period (fig. 6). A LOWESS trend line for these residuals indicated a downward trend of selenium concentration over the study period.

Flow-Adjusted Trends in Selenium Load and Concentration

Changes in estimated selenium load for the first and last years of the study were calculated for the Gunnison River and Colorado River sites. Changes in estimated 50th and 85th percentile concentrations are shown, and trends in selenium concentration are discussed for the two sites.

Interpretation of the Estimates

Estimated selenium loads and concentrations for WY 1986 and WY 2008 are provided in tables 6 and 7. It is important to remember that the estimated loads and concentrations given for WY 1986 and WY 2008 in tables 6 and 7
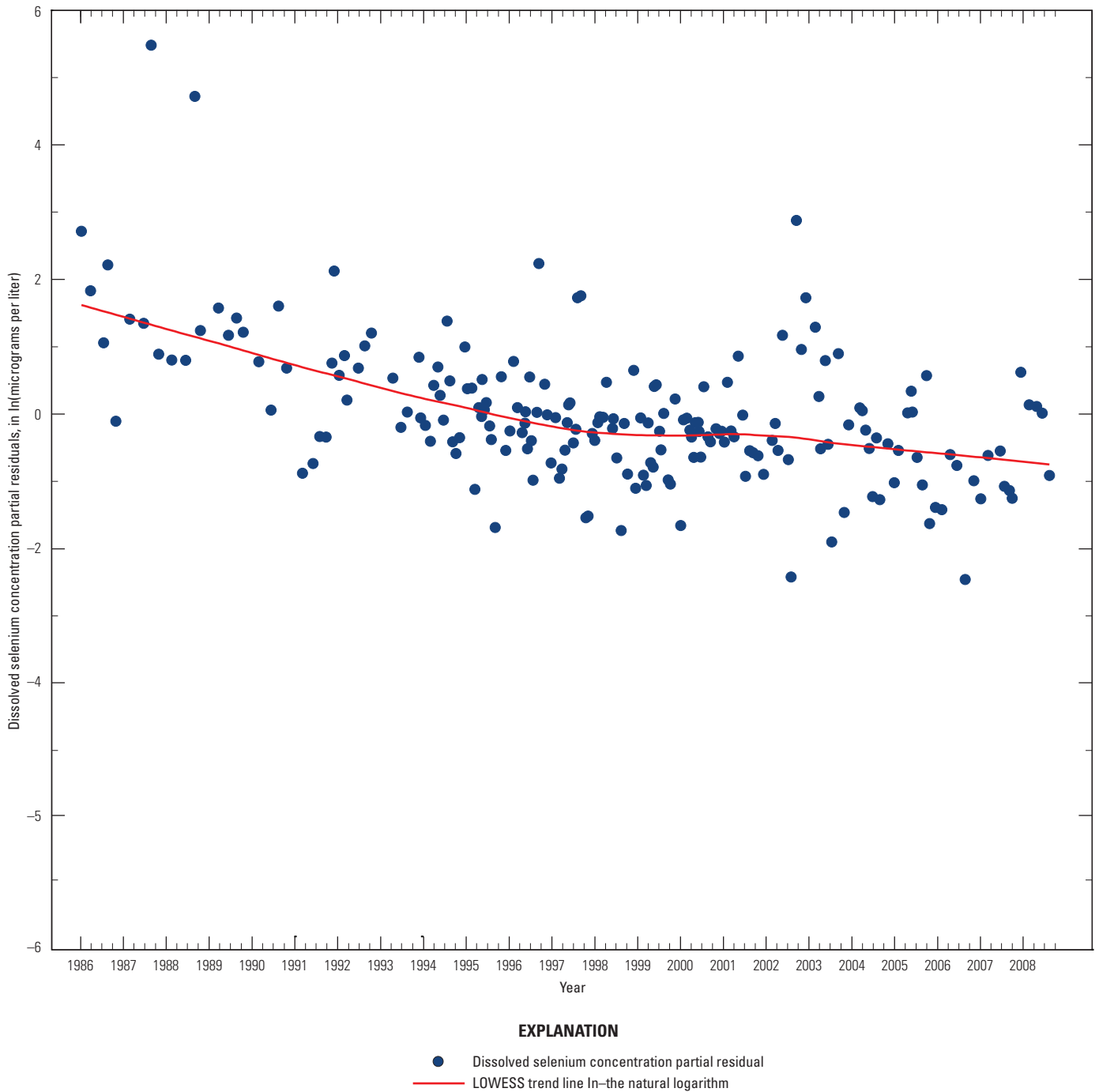
Figure 6. Dissloved selenium concentration partial residuals and LOWESS fit line using the step 4 regression model (equation 13) for Colorado River site, water years 1986–2008.

were based on normalized streamflow, and are only illustrative of the change in selenium loads and concentrations over the period of study. Interpretation of the estimates was based on the percentage of change in load and concentration. The loads and concentrations shown in tables 6 and 7 were not the actual loads and concentrations that occurred in those years.

Gunnison River Site

Annual Selenium Loads and Selenium Concentration Percentiles for Gunnison River Site

Normalized mean-daily streamflow values were used with equation 6 (from Gunnison River methods step 3) in S-LOADEST to estimate annual selenium loads that would have been expected in WY 1986 and WY 2008 under conditions of long-term mean-daily streamflow.

Daily selenium concentrations were calculated by S-LOADEST as part of the daily load calculations. The 50th and 85th percentile selenium concentrations for WY 1986 and WY 2008 were derived from these daily selenium concentrations. These results, along with lower and upper 95-percent confidence levels are shown in table 6.

The flow-adjusted annual selenium load decreased from 23,196 lbs/yr in WY 1986 to 16,560 lbs/yr in WY 2008, a decrease of 6,636 lbs/yr or 28.6 percent. Lower and upper 95-percent confidence levels for WY 1986 annual load were 22,360 and 24,032 pounds, respectively. Lower and upper 95-percent confidence levels for WY 2008 annual load were 15,724 and 17,396 pounds, respectively. The 50th percentile flow-adjusted selenium concentration decreased from 6.41 µg/L in WY 1986 to 4.57 µg/L in WY 2008. The 85th percentile flow-adjusted selenium concentration decreased from 7.21 µg/L in WY 1986 to 5.13 µg/L in WY 2008.

Time-trend of Selenium Load and Concentration at Gunnison River Site

Model calibration step 4 for the Gunnison River site yielded a *dectime* coefficient that was negative and statistically significant ($\beta_3 = -0.016$, p-value <0.001, table 2). This indicated that the time-trend for selenium load, and therefore concentration, was downward over the study period. Figure 3 illustrates this generally downward trend in concentration over the study period. A slight upward bump in the trend line occurred from WY 1998 to 2001, after which the trend resumed downward. No analysis was done to attempt to explain this anomaly.

Colorado River Site

Annual Selenium Loads and Selenium Concentration Percentiles for Colorado River Site

Normalized mean-daily streamflow values were used with equation 12 (from Colorado River methods step 3) in the load calculation to estimate annual selenium loads for WY 1986 and WY 2008. Again, the annual loads derived were illustrative of the change in loads from WY 1986 to WY 2008 and were not actual loads for those two years.

Daily selenium concentrations were calculated by S-LOADEST as part of the daily load calculations. The 50th and 85th percentile concentrations were calculated from the estimated daily concentrations. These results, along with lower and upper 95-percent confidence levels, are shown in table 7.

The flow-adjusted annual selenium load decreased from 56,587 lbs/yr in WY 1986 to 34,344 lbs/yr in WY 2008, a decrease of 22,243 lbs/yr or 39.3 percent. Lower and upper 95-percent confidence levels for WY 1986 annual load were

Table 6. Estimated selenium loads and concentrations given normalized mean-daily streamflow for water years 1986 and 2008 for Gunnison River site.

[Water year, October 1st through the following September 30th; annual load, the total load for a water year; ft³/s, cubic feet per second; lbs, pounds; µg/L, micrograms per liter; %, percent; --, not applicable]

Water year	Average of mean-daily streamflow for 1986 to 2008 (ft³/s)	Estimated selenium annual load (lbs)	Lower 95% confidence level for estimated annual load (lbs)	Upper 95% confidence level for estimated annual load (lbs)	Estimated selenium annual load reduction %	50th percentile of estimated daily selenium concentration (µg/L)	85th percentile of estimated daily selenium concentration (µg/L)
1986	2,400	23,196	22,360	24,032	--	6.41	7.21
2008	2,400	16,560	15,724	17,396	28.6	4.57	5.13
Difference		6,636				1.84	2.08

Table 7.   Estimated selenium loads and concentrations given normalized mean-daily streamflow for water years 1986 and 2008 for Colorado River site.

[Water year, October 1st through the following September 30th; annual load, the total load for a water year; ft³/s, cubic feet per second; lbs, pounds; µg/L, micrograms per liter; %, percent; --, not applicable]

Water year	Average of mean daily streamflow for 1986 to 2008 (ft³/s)	Estimated selenium annual load (lbs)	Lower 95% confidence level for estimated annual load (lbs)	Upper 95% confidence level for estimated annual load (lbs)	Estimated selenium annual load reduction %	50th percentile of estimated daily selenium concentration (µg/L)	85th percentile of estimated daily selenium concentration (µg/L)
1986	5,908	56,587	53,785	59,390	--	6.44	7.94
2008	5,908	34,344	31,542	37,147	39.3	3.86	4.72
Difference		22,243				2.58	3.22

53,785 and 59,390 pounds, respectively. Lower and upper 95-percent confidence levels for WY 2008 annual load were 31,542 and 37,147 pounds, respectively. The 50th percentile flow-adjusted selenium concentration decreased from 6.44 µg/L in WY 1986 to 3.86 µg/L in WY 2008. The 85th percentile flow-adjusted selenium concentration decreased from 7.94 µg/L in WY 1986 to 4.72 µg/L in WY 2008.

Time-trend of Selenium Load and Concentration at Colorado River Site

Model calibration step 4 for the Colorado River site yielded a *dectime* coefficient that was negative and statistically significant ($\beta_2 = -0.021$, p-value <0.001, table 5). This indicated that the time-trend for selenium load, and therefore concentration, was downward over the study period. Figure 6 illustrates this general downward trend in concentration over the study period. A slight leveling off in the slope of the trend line occurred from WY 1998 to WY 2000, after which the trend resumed downward. No analysis was done to attempt to explain this anomaly.

Summary and Conclusions

As a result of elevated selenium concentrations, many western Colorado rivers and streams are on the U.S. Environmental Protection Agency 2010 Colorado 303(d) list, including the main stem of the Colorado River from the Gunnison River confluence to the Utah border. Selenium is a trace element that bioaccumulates in aquatic food chains and can cause reproductive failure, deformities, and other adverse impacts in birds and fish, including some threatened and endangered fish species. Salinity in the upper Colorado River has also been the focus of source-control efforts for many years. Although salinity loads and concentrations have been previously

characterized for the USGS streamflow-gaging stations at Colorado River near Colorado-Utah State line and at Gunnison River near Grand Junction, Colo., trends in selenium loads and concentrations for these two stations have not been studied. The USGS, in cooperation with the Bureau of Reclamation and the Colorado River Water Conservation District, evaluated the dissolved selenium (herein referred to as "selenium") load and concentration trends at two streamflow-gaging stations in western Colorado to inform decision makers on the status and trends of selenium.

This report presents results of the analysis of trends in selenium load and concentration for two USGS streamflow-gaging stations: Gunnison River near Grand Junction, Colo. ("Gunnison River site"), USGS site 09152500, and Colorado River near Colorado-Utah State line ("Colorado River site"), USGS site 09163500. Flow-adjusted selenium loads were estimated for the beginning water year (WY) of the study, 1986, and the ending WY of the study, 2008.

The difference between flow-adjusted selenium loads for WY 1986 and WY 2008 was selected as the method of analysis because flow adjustment removes the natural variations in load caused by changes in mean-daily streamflow, emphasizing human-caused changes in selenium load and concentration. Overall changes in human-caused effects in selenium loads and concentrations during the period of study are of primary interest to the cooperators. Selenium loads for each of the two water years were calculated by using normalized mean-daily streamflow, measured selenium concentration, standard linear regression techniques, and data previously collected at the two study sites. Mean-daily streamflow was normalized for each site by averaging the daily streamflow for each day of the year over the 23-year period of record. Thus, for the beginning and ending water years, estimates could be made of loads that would have occurred without the effect of year-to-year streamflow variation. The loads were illustrative of the change in loads between water years 1986 and 2008 and were not the actual loads that occurred in those two water years.

The estimated 50th and 85th percentile selenium concentrations associated with the selenium loads were also calculated for WY 1986 and WY 2008 at each site. Time-trends in selenium concentration at the two sites were charted by using regression techniques for partial residuals for the entire study period (WY 1986 through WY 2008).

A three-step process was chosen for this analysis based on model formulation, model calibration, and load estimation. Daily and annual selenium loads were calculated using multiple linear regression. The variables of interest used included daily streamflow, time, and irrigation season. Log transformation was used to linearize the relation between streamflow, time, and selenium concentration. The software package S-LOADEST was used to perform the regression analysis. The base regression model was automatically selected by S-LOADEST by minimizing the Akaike Information Criterion value for each of nine predefined models. Streamflow and decimal time were centered automatically by S-LOADEST. Calibration data sets composed of date, time, daily streamflow, measured selenium concentration, and irrigation season were used for each site. Residuals (actual load minus estimated load) were calculated by S-LOADEST for model testing. The residual standard error (RSE) and coefficient of determination (R^2) were calculated for each regression model.

The p-value for each variable of interest was examined, and if the p-value was greater than 0.05, the variable was not significant and was considered for exclusion from subsequent applications of the regression model. The 50th and 85th percentile values of estimated selenium concentration were calculated for use by regulatory agencies. The sign of the *dectime* coefficient (decimal date and time) in the regression model indicated whether the time-trend for the load and concentration was upward (positive sign on the coefficient) or downward (negative sign on the coefficient). Regressions for partial residuals were used with LOWESS smooth lines to graphically demonstrate the selenium load trend.

Various diagnostic plots were generated by S_LOADEST. These diagnostic plots were examined for normality, constant variance, and independence.

A four-step model calibration process was followed for both sites:

1. Select a base regression model of selenium load using daily streamflow, time, and transformations, and test for statistical significance of all variables of interest. Test for the validity of the various model assumptions.

2. Add irrigation season as a variable of interest in the regression model from step 1, and test for statistical significance of irrigation season.

3. Use the load regression model from steps 1 and 2 to estimate daily and annual selenium loads for WY 1986 and for WY 2008 and derive daily mean selenium concentrations from estimated loads and daily flows for WY 1986 and WY 2008.

4. Examine the coefficient for *dectime* to determine if a time-trend exists. Demonstrate graphically whether any trend in selenium concentration over time exists by removing the decimal time terms from the step 2 load regression model (regression analysis for partial residuals), deriving estimated concentrations from the estimated daily loads, and charting the concentration residuals with a fitted trend line over the years of the study period.

These steps were applied to both sites, resulting in a valid regression model being selected for each site. Annual selenium loads were estimated using the selected model for each site. In addition, 50th and 85th percentile selenium concentrations were calculated from the regression models. Time-trends in selenium concentration were examined and charted.

Annual selenium load for the Gunnison River site was estimated to be 23,196 pounds for WY 1986 and 16,560 pounds for WY 2008, a 28.6 percent decrease. Lower and upper 95-percent confidence levels for WY 1986 annual load were 22,360 and 24,032 pounds, respectively. Lower and upper 95-percent confidence levels for WY 2008 annual load were 15,724 and 17,396 pounds, respectively. Estimated 50th percentile daily selenium concentrations decreased from 6.41 to 4.57 micrograms/liter from WY 1986 to WY 2008, whereas estimated 85th percentile daily selenium concentrations decreased from 7.21 to 5.13 micrograms/liter from WY 1986 to WY 2008. It was determined that the selenium concentration for the Gunnison River site had a statistically significant downward trend over the study period.

Annual selenium load for the Colorado River site was estimated to be 56,587 pounds for WY 1986 and 34,344 pounds for WY 2008, a 39.3 percent decrease. Lower and upper 95-percent confidence levels for WY 1986 annual load were 53,785 and 59,390 pounds, respectively. Lower and upper 95-percent confidence levels for WY 2008 annual load were 31,542 and 37,147 pounds, respectively. Estimated 50th percentile daily selenium concentrations decreased from 6.44 to 3.86 micrograms/liter from WY 1986 to WY 2008, whereas estimated 85th percentile daily selenium concentrations decreased from 7.94 to 4.72 micrograms/liter from WY 1986 to WY 2008. It was determined that the selenium concentration for the Colorado River site had a statistically significant downward trend over the study period.

Acknowledgments

References Cited

Butler, D.L., 1996, Trend analysis of selected water-quality data associated with salinity control projects in the Grand Valley, in the lower Gunnison basin, and at Meeker dome, western Colorado: U.S. Geological Survey Water-Resources Investigations Report 95–4274, 38 p.

Butler, D.L., Wright, W.G., Stewart, K.C., Osmundson, B.C., Krueger, R.P., and Crabtree, D.W., 1996, Detailed study of selenium and other constituents in water, bottom sediment, soil, alfalfa, and biota associated with irrigation drainage in the Uncompahgre Project area and in the Grand Valley, west-central Colorado, 1991–93: U.S. Geological Survey Water-Resources Investigations Report 96–4138, 136 p.

Butler, D.L., 2001, Effects of piping irrigation laterals on selenium and salt loads, Montrose Arroyo basin, western Colorado: U.S. Geological Survey Water-Resources Investigations Report 01–4204, 14 p.

Childress, C.J.O., Foreman, W.T., Connor, B.F., and Maloney, T.J., 1999, New reporting procedures based on long-term method detection levels and some considerations for interpretations of water-quality data provided by the U.S. Geological Survey National Water Quality Laboratory: U.S. Geological Survey Open-File Report 99–193, 19 p.

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

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

Colorado Department of Public Health and Environment, 2010, Colorado's section 303(d) list of impaired waters and monitoring and evaluation list, effective April 30, 2010: Colorado Department of Public Health and Environment database accessed January 14, 2011, at *http://www.cdphe.state.co.us/regulations/wqccregs/100293wqlimitedsegtmdlsnew.pdf*.

Colorado River Salinity Control Forum, 2011, 2011 review, water quality standards for salinity: Colorado River Basin Salinity Control Forum, Bountiful, Utah, website accessed April 13, 2012 at *http://www.coloradoriversalinity.org/docs/2011%20REVIEW-October.pdf*.

Gunnison Basin & Grand Valley Selenium Task Forces, 2012, Current projects: Gunnison Basin & Grand Valley Selenium Task Forces website accessed February 27, 2012, at *http://www.seleniumtaskforce.org/projects.html*

Hamilton, S.J., 1998, Selenium effects on endangered fish in the Colorado River basin, *in* Frankenberger, W.T., and Engderg, R.A., eds., Environmental chemistry of selenium: New York, Marcel Dekker, Inc., 713 p.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, 510 p.

Kircher, J.E., Dinicola, R.S., and Middelburg, R.M., 1984, Trend analysis of salt load and evaluation of the frequency of water-quality measurements for the Gunnison, the Colorado, and the Dolores Rivers in Colorado and Utah: U.S. Geological Survey Water-Resources Investigations Report 84–4048, 69 p.

Leib, K.J., and Bauch, N.J., 2008, Salinity trends in the upper Colorado River basin upstream from the Grand Valley salinity control unit, Colorado, 1986–2003: U.S. Geological Survey Water-Resources Investigations Report 07–5288, 21 p.

Lemly, A.D., 2002, Selenium assessment in aquatic ecosystems—A guide for hazard evaluation and water quality criteria: New York, Springer-Verlag, 162 p.

Richards, R.J., and Leib, K.J., 2011, Characterization of hydrology and salinity in the Dolores project area, McElmo Creek Region, southwest Colorado, 1978–2006: U.S. Geological Survey Scientific Investigations Report 2010–5218, 38 p.

Runkel, R.L., Crawford, C.G., and Cohn, T.A., 2004, Load estimator (LOADEST)—A FORTRAN program for estimating constituent loads in streams and rivers: U.S. Geological Survey Techniques and Methods, book 4, chap. A5, 69 p.

Sprague, L.A., Clark, M.L., Rus, D.L., Zelt, R.B., Flynn, J.L., and Davis, J.V., 2006, Nutrient and suspended-sediment trends in the Missouri River basin, 1993–2003: U.S. Geological Survey Scientific Investigations Report 2006–5231, 80 p.

TIBCO Software Inc., 1988–2008, Spotfire S+ 8.1 for Windows: Somerville, Mass., accessed March 2, 2012, at *http://spotfire.tibco.com/products/s-plus/statistical-analysis-software.aspx*.

U.S. Environmental Protection Agency, 2011, What is a 303(d) list of impaired waters?: United States Environmental Protection Agency, accessed May 4, 2011, at *http://water.epa.gov/lawsregs/lawsguidance/cwa/tmdl/overview.cfm*.

U.S. Fish & Wildlife Service, 2011, Grand Junction Fish & Wildlife—Endangered Fish: U.S. Fish & Wildlife Service database accessed March 3, 2011, at *http://www.fws.gov/grandjunctionfishandwildlife/endangeredfish.html*.

U.S. Geological Survey, variously dated, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1–A9, accessed January 5, 2011, at *http://water.usgs.gov/owq/FieldManual/*.

Vaill, J.E., and Butler, D.L., 1999, Streamflow and dissolved-solids trends, through 1996, in the Colorado River basin upstream from Lake Powell—Colorado, Utah, and Wyoming: U.S. Geological Survey Water-Resources Investigations Report 99–4097, 47 p.

Supplemental Data

Table 8. Gunnison River site regression model calibration data.

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; E, laboratory remark for estimated dissolved selenium concentration in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
11/26/1985	2,670	12:00	none	6	0
03/26/1986	3,370	12:00	none	4	0
05/30/1986	8,200	12:00	none	2	1
07/16/1986	5,550	12:00	none	3	1
09/24/1986	3,700	12:00	none	5	1
11/20/1986	4,340	12:00	none	5	0
07/16/1987	2,000	12:00	none	6	1
08/11/1987	2,200	12:00	none	8	1
09/22/1987	1,980	12:00	none	7	1
11/10/1987	2,290	12:00	none	6	0
04/05/1988	2,440	12:00	none	2	1
08/16/1988	762	12:00	none	10	1
09/22/1988	1,680	12:00	none	8	1
11/02/1988	932	12:00	none	11	0
04/12/1989	2,650	12:00	none	2	1
07/18/1989	686	12:00	none	11	1
08/29/1989	938	12:00	none	10	1
11/07/1989	1,340	12:00	none	9	0
03/27/1990	760	12:00	none	7	0
05/02/1990	1,730	12:00	none	8	1
08/28/1990	794	12:00	none	10	1
03/26/1991	1,010	12:00	none	8	0
07/31/1991	1,530	12:00	none	5	1
09/24/1991	1,460	12:00	none	9	1
11/13/1991	2,080	12:00	none	5	0
12/03/1991	1,950	12:00	none	7	0
01/16/1992	1,300	12:00	none	5	0
02/26/1992	1,120	12:00	none	5	0
03/18/1992	1,380	12:00	none	7	0
04/21/1992	2,970	12:00	none	3	1
06/24/1992	1,980	12:00	none	6	1
08/19/1992	1,680	12:00	none	7	1
12/01/1992	1,120	12:00	none	9	0
05/26/1993	15,500	12:00	none	1	1
07/27/1993	1,950	12:00	none	6	1
09/08/1993	2,280	12:00	none	5	1

Table 8.   Gunnison River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; E, laboratory remark for estimated dissolved selenium concentration in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
11/23/1993	2,260	12:00	none	6	0
02/24/1994	1,000	12:00	none	7	0
05/25/1994	5,560	12:00	none	2	1
06/23/1994	2,490	12:00	none	5	1
08/17/1994	1,280	12:00	none	6	1
09/07/1994	1,790	12:00	none	8	1
10/27/1994	1,730	12:00	none	6	1
12/19/1994	1,330	12:00	none	6	0
01/24/1995	1,080	12:00	none	6	0
05/17/1995	11,300	12:00	none	2	1
06/19/1995	13,900	12:00	none	2	1
07/21/1995	11,800	12:00	<	1	1
10/30/1995	2,840	12:00	none	5	1
12/12/1995	2,750	12:00	none	4	0
01/23/1996	1,660	12:00	none	5	0
02/21/1996	2,140	12:00	none	5	0
03/26/1996	2,620	12:00	none	3	0
05/20/1996	7,140	12:00	none	2	1
05/29/1996	2,760	12:00	none	4	1
06/11/1996	2,760	12:00	none	4	1
07/03/1996	3,640	12:00	none	3	1
07/16/1996	2,310	12:00	none	4	1
08/07/1996	1,240	12:00	none	7	1
08/21/1996	1,460	12:00	none	7	1
09/05/1996	1,550	12:00	none	7	1
09/16/1996	2,170	12:00	none	7	1
09/17/1996	2,150	12:00	none	7	1
10/04/1996	2,840	12:00	none	6	1
11/13/1996	2,160	12:00	none	5	0
12/20/1996	1,950	12:00	none	5	0
01/17/1997	2,330	12:00	none	3	0
02/24/1997	2,450	12:00	none	2	0
03/31/1997	5,480	12:00	none	1	0
04/25/1997	5,740	12:00	none	1	1
04/28/1997	6,020	12:00	none	2	1
05/16/1997	8,580	12:00	none	1	1
05/24/1997	10,500	12:00	none	2	1
06/12/1997	9,280	12:00	none	2	1
07/03/1997	5,630	12:00	none	2	1
07/18/1997	2,900	12:00	none	3	1

Table 8.   Gunnison River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; E, laboratory remark for estimated dissolved selenium concentration in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
08/12/1997	3,230	12:00	none	5	1
09/05/1997	2,880	12:00	none	5	1
10/20/1997	3,250	12:00	none	4	1
11/07/1997	2,830	12:00	none	3	0
12/15/1997	2,670	12:00	none	3	0
01/08/1998	2,550	12:00	none	3	0
02/11/1998	2,140	12:00	none	3	0
03/12/1998	1,720	12:00	none	4	0
04/08/1998	2,680	12:00	none	3	1
05/18/1998	6,000	12:00	none	1	1
06/03/1998	5,250	12:00	none	2	1
06/11/1998	3,060	12:00	none	3	1
07/09/1998	2,130	12:00	none	5	1
08/11/1998	1,440	12:00	none	4	1
09/11/1998	1,710	12:00	none	7	1
10/16/1998	1,890	12:00	none	5	1
11/18/1998	1,830	12:00	none	7	0
01/20/1999	1,350	12:00	none	6	0
03/10/1999	1,100	12:00	none	5	0
04/16/1999	1,020	12:00	none	3	1
04/30/1999	3,160	12:00	none	5	1
06/10/1999	3,200	12:00	none	3	1
06/30/1999	3,420	12:00	none	3	1
07/29/1999	2,630	12:00	none	4	1
09/09/1999	3,160	12:00	none	4	1
10/06/1999	4,720	12:00	E	2	1
11/17/1999	1,950	12:00	none	6	0
01/19/2000	1,430	12:00	none	6.6	0
03/09/2000	1,110	12:00	none	6.8	0
03/21/2000	1,700	12:00	none	4.1	0
03/22/2000	1,700	12:00	none	4.7	0
04/13/2000	3,280	12:00	none	2.5	1
05/10/2000	4,340	12:00	none	2	1
06/06/2000	2,480	12:00	none	3.8	1
07/24/2000	1,440	12:00	none	3.7	1
08/29/2000	1,920	12:00	none	4.9	1
10/25/2000	1,800	12:00	none	5.7	1
12/14/2000	1,420	12:00	none	8.3	0
02/13/2001	1,030	12:00	none	7.5	0
03/15/2001	1,220	12:00	none	6.9	0

Table 8.   Gunnison River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; E, laboratory remark for estimated dissolved selenium concentration in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
04/20/2001	2,200	12:00	none	3.3	1
05/22/2001	3,600	12:00	none	2.6	1
06/28/2001	1,300	12:00	none	5.7	1
07/27/2001	1,310	12:00	none	5.9	1
08/23/2001	2,160	12:00	none	7.5	1
10/26/2001	1,642	12:00	none	6.9	1
12/14/2001	1,179	12:00	none	8.4	0
03/11/2002	1,027	12:00	none	8.7	0
03/29/2002	1,192	12:00	none	6.8	0
04/25/2002	692	12:00	none	7	1
05/31/2002	927	12:00	none	8.5	1
06/26/2002	832	12:00	none	6	1
08/13/2002	832	12:00	none	6	1
09/24/2002	1,003	12:00	none	8	1
10/18/2002	1,062	12:00	none	11.6	1
12/17/2002	717	12:00	none	13.2	0
02/20/2003	593	12:00	none	16.4	0
03/13/2003	651	12:00	none	9.6	0
04/25/2003	1,169	12:00	none	4.9	1
05/20/2003	4,638	12:00	none	2.9	1
06/11/2003	1,866	12:00	none	3.7	1
07/22/2003	998	12:00	none	4.6	1
08/04/2003	1,041	12:00	none	6	1
10/22/2003	1,068	12:00	none	8.4	1
12/11/2003	761	12:00	none	13.6	0
02/26/2004	708	12:00	none	12	0
03/24/2004	1,606	12:00	none	4.9	0
04/23/2004	1,594	12:00	none	5.3	1
05/13/2004	2,920	12:00	none	2.9	1
06/28/2004	997	12:00	none	6.9	1
07/14/2004	993	12:00	none	6.8	1
08/17/2004	1,022	12:00	none	6.9	1
11/16/2004	1,084	12:00	none	7.9	0
12/28/2004	900	12:00	none	6.5	0
01/24/2005	1,207	12:00	none	5.2	0
02/15/2005	1,556	12:00	none	4	0
04/08/2005	3218	12:00	none	2	1
05/23/2005	11,541	12:00	none	1.8	1
06/02/2005	6,033	12:00	none	1.9	1
07/12/2005	1,630	12:00	none	4.6	1

Table 8. Gunnison River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; E, laboratory remark for estimated dissolved selenium concentration in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
09/22/2005	1,647	12:00	none	5.7	1
10/31/2005	1,559	12:00	none	6.6	1
01/05/2006	1,003	12:00	none	5.7	0
03/17/2006	1,194	12:00	none	3.8	0
05/26/2006	3,539	12:00	none	1.9	1
08/24/2006	1,921	12:00	none	3.6	1
11/30/2006	2,336	12:00	none	3.5	0
03/01/2007	1,412	12:00	none	4	0
04/04/2007	2,809	12:00	none	2.2	1
07/09/2007	1,283	12:00	none	4.9	1
09/05/2007	1,771	12:00	none	5	1
12/13/2007	1,826	12:00	none	5.2	0
04/28/2008	7,049	12:00	none	1.1	1
05/20/2008	12,010	12:00	none	1.1	1
08/13/2008	2,098	12:00	none	4.5	1

[1]The number of decimal places shown for dissolved selenium concentration is determined at the U.S. Geological Survey laboratory and depends on the accuracy of the particular analytical method used at the time. The number of decimal places shown in table 8 is the same as those reported in the U.S. Geological Survey database. In general, the more recent the sample, the greater the number of decimal places shown.

Table 9. Colorado River site regression model calibration data.

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
01/08/1986	4,960	12:00	none	6	0
03/25/1986	6,750	12:00	none	4	0
07/15/1986	14,100	12:00	none	3	1
08/19/1986	4,740	12:00	none	7	1
10/29/1986	6,000	12:00	none	4	1
02/25/1987	5,530	12:00	none	4	0
06/23/1987	8,060	12:00	none	4	1
08/25/1987	5,240	12:00	none	10	1
10/27/1987	4,450	12:00	none	6	1
02/17/1988	4,240	12:00	none	4	0
06/14/1988	9,950	12:00	none	3	1
08/30/1988	3,360	12:00	none	11	1

Table 9. Colorado River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
10/18/1988	3,080	12:00	none	8	1
03/21/1989	3,580	12:00	none	5	0
06/13/1989	6,940	12:00	none	4	1
08/22/1989	3,870	12:00	none	7	1
10/18/1989	3,830	12:00	none	7	1
02/27/1990	2,070	12:00	none	6	0
06/12/1990	11,700	12:00	none	2	1
08/14/1990	2,120	12:00	none	10	1
10/24/1990	3,340	12:00	none	7	1
03/07/1991	3,070	12:00	none	3	0
06/04/1991	13,100	12:00	<	1	1
07/31/1991	3,750	12:00	none	5	1
09/24/1991	3,440	12:00	none	6	1
11/13/1991	4,690	12:00	none	5	0
12/03/1991	3,510	12:00	none	7	0
01/15/1992	3,150	12:00	none	5	0
02/26/1992	2,880	12:00	none	5	0
03/19/1992	3,110	12:00	none	4	0
06/24/1992	5,950	12:00	none	4	1
08/19/1992	3,470	12:00	none	7	1
10/14/1992	3,080	12:00	none	8	1
04/14/1993	6,630	12:00	none	3	1
06/24/1993	24,500	12:00	none	1	1
08/18/1993	4,480	12:00	none	5	1
11/24/1993	4,590	12:00	none	5	0
12/07/1993	4,440	12:00	none	4	0
01/19/1994	3,350	12:00	none	4	0
03/02/1994	3,730	12:00	none	3	0
03/30/1994	3,360	12:00	none	4	0
05/04/1994	4,640	12:00	none	4	1
05/24/1994	12,500	12:00	none	2	1
06/22/1994	6,470	12:00	none	3	1
07/22/1994	2,650	12:00	none	8	1
08/16/1994	3,050	12:00	none	7	1
09/08/1994	3,330	12:00	none	6	1
10/04/1994	4,100	12:00	none	5	1
11/07/1994	3,470	12:00	none	5	0
12/20/1994	3,080	12:00	none	6	0
01/12/1995	3,010	12:00	none	5	0
02/17/1995	2,540	12:00	none	5	0

Table 9. Colorado River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
03/16/1995	4,100	12:00	none	2	0
04/18/1995	5,320	12:00	none	3	1
05/11/1995	9,350	12:00	none	2	1
05/16/1995	14,800	12:00	none	2	1
06/06/1995	31,200	12:00	none	1	1
06/20/1995	40,000	12:00	none	1	1
07/20/1995	29,000	12:00	none	1	1
08/04/1995	11,800	12:00	none	2	1
09/06/1995	3,920	12:00	none	4	1
10/27/1995	5,410	12:00	none	5	1
12/05/1995	5,420	12:00	none	3	0
01/09/1996	5,000	12:00	none	3	0
02/08/1996	4,360	12:00	none	4	0
03/12/1996	4,550	12:00	none	3	0
04/23/1996	7,570	12:00	none	2	1
05/14/1996	21,400	12:00	none	1	1
05/21/1996	27,800	12:00	none	1	1
06/06/1996	16,000	12:00	none	1	1
06/26/1996	19,100	12:00	none	2	1
07/09/1996	10,200	12:00	none	2	1
07/23/1996	5,430	12:00	none	3	1
08/27/1996	3,570	12:00	none	6	1
09/10/1996	3,880	12:00	none	8	1
10/31/1996	5,150	12:00	none	5	1
11/21/1996	4,890	12:00	none	4	0
12/24/1996	4,540	12:00	none	3	0
01/30/1997	4,890	12:00	none	3	0
03/06/1997	4,500	12:00	none	2	0
03/27/1997	8,640	12:00	none	1	0
04/23/1997	13,100	12:00	none	1	1
05/13/1997	21,400	12:00	none	1	1
05/23/1997	32,500	12:00	none	1	1
06/04/1997	36,300	12:00	<	1	1
07/01/1997	20,000	12:00	none	1	1
07/24/1997	7,290	12:00	none	3	1
08/05/1997	7,600	12:00	none	5	1
09/03/1997	5,860	12:00	none	6	1
10/16/1997	7,620	12:00	none	2	1
11/06/1997	6,260	12:00	none	2	0
12/11/1997	5,810	12:00	none	3	0

Table 9.   Colorado River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
12/30/1997	4,990	12:00	none	3	0
01/27/1998	4,790	12:00	none	3	0
02/13/1998	4,620	12:00	none	3	0
03/13/1998	4,250	12:00	none	3	0
04/10/1998	6,360	12:00	none	3	1
06/02/1998	21,300	12:00	none	1	1
06/10/1998	10,600	12:00	none	2	1
07/07/1998	8,730	12:00	none	2	1
08/14/1998	3,610	12:00	none	4	1
09/10/1998	3,550	12:00	none	6	1
10/08/1998	4,920	12:00	none	4	1
11/30/1998	4,180	12:00	none	5	0
12/14/1998	3,100	12:00	none	4	0
01/28/1999	3,320	12:00	none	4	0
02/19/1999	3,180	12:00	none	3	0
03/15/1999	2,840	12:00	none	3	0
04/01/1999	3,180	12:00	none	4	1
04/23/1999	3,800	12:00	none	3	1
05/13/1999	6,080	12:00	none	2	1
05/24/1999	14,000	12:00	none	2	1
06/09/1999	15,500	12:00	none	2	1
07/06/1999	10,900	12:00	none	2	1
07/19/1999	6,270	12:00	none	3	1
08/12/1999	8,870	12:00	none	3	1
09/20/1999	6,580	12:00	none	3	1
10/08/1999	6,420	12:00	none	3	1
11/16/1999	3,920	12:00	none	5	0
01/04/2000	3,320	12:00	none	2.8	0
01/27/2000	2,900	12:00	none	4.4	0
02/24/2000	2,870	12:00	none	4.1	0
03/22/2000	3,700	12:00	none	3.1	0
04/05/2000	3,530	12:00	none	3.5	1
04/25/2000	5,570	12:00	none	2.2	1
05/08/2000	12,700	12:00	none	1.5	1
05/31/2000	17,000	12:00	none	1.3	1
06/07/2000	10,500	12:00	none	1.8	1
06/23/2000	7,010	12:00	none	2.3	1
07/18/2000	4,200	12:00	none	5.1	1
08/23/2000	4,020	12:00	none	5.1	1
09/14/2000	3,290	12:00	none	6.1	1

Table 9. Colorado River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
10/30/2000	3,770	12:00	none	5.5	1
11/28/2000	3,050	12:00	none	5.2	0
12/18/2000	2,690	12:00	none	5.3	0
01/11/2001	2,560	12:00	none	4.8	0
02/06/2001	2,570	12:00	none	5.2	0
03/08/2001	2,530	12:00	none	4.2	0
04/03/2001	2,710	12:00	none	4.3	1
05/08/2001	5,820	12:00	none	3.7	1
06/14/2001	6,710	12:00	none	2.9	1
07/09/2001	2,810	12:00	none	5.1	1
08/14/2001	3,600	12:00	none	5.2	1
09/10/2001	3,690	12:00	none	5.4	1
10/23/2001	3,273	12:00	none	5.8	1
12/10/2001	2,417	12:00	none	5.3	0
02/21/2002	2,375	12:00	none	4.4	0
03/19/2002	2,212	12:00	none	4.7	0
04/11/2002	2,997	12:00	none	3.8	1
05/22/2002	3,921	12:00	none	5.1	1
07/09/2002	1,899	12:00	none	7.3	1
08/01/2002	1,333	12:00	none	8.8	1
09/17/2002	2,384	12:00	none	11.1	1
10/29/2002	2,572	12:00	none	8.5	1
12/04/2002	2,350	12:00	none	8.2	0
02/25/2003	1,751	12:00	none	7.2	0
03/27/2003	2,268	12:00	none	5	0
04/10/2003	1,415	12:00	none	6.9	1
05/20/2003	12,416	12:00	none	2.5	1
06/11/2003	9,458	12:00	none	1.8	1
07/15/2003	2,731	12:00	none	4.4	1
09/10/2003	6,025	12:00	none	5.1	1
10/29/2003	3,189	12:00	none	5	1
12/05/2003	2,565	12:00	none	5.9	0
03/09/2004	2,081	12:00	none	5.2	0
03/30/2004	3,841	12:00	none	3.3	0
04/27/2004	2,919	12:00	none	4.3	1
05/28/2004	6,154	12:00	none	2.4	1
06/25/2004	3,543	12:00	none	3.6	1
07/28/2004	2,966	12:00	none	5.9	1
08/26/2004	2,557	12:00	none	6.3	1
11/04/2004	3,426	12:00	none	5	0

Table 9. Colorado River site regression model calibration data.—Continued

[Streamflow, daily streamflow; ft³/s, cubic feet per second; time, assumed time for regression model calculation, not the time of sample collection; R01145, U.S. Geological Survey laboratory remarks for database parameter code P01145; P01145, U.S. Geological Survey database parameter code for dissolved selenium concentrations; <, laboratory remark indicating dissolved selenium concentration is less than value in P01145 column; Season, irrigation season code, April 1 through October 31 = 1, November 1 through March 31 = 0]

Date	Streamflow (ft³/s)	Time	Laboratory remarks (R01145)	Dissolved selenium concentration in µg/L (P01145)[1]	Season
12/29/2004	2,754	12:00	none	4.2	0
02/03/2005	2,674	12:00	none	4.1	0
04/21/2005	11,441	12:00	none	1.7	1
05/20/2005	17,115	12:00	none	1.7	1
06/01/2005	17,834	12:00	none	1.4	1
07/11/2005	7,366	12:00	none	2.4	1
08/25/2005	3,739	12:00	none	4.7	1
09/29/2005	6,711	12:00	none	4.5	1
10/26/2005	4,323	12:00	none	3.6	1
12/15/2005	3,273	12:00	none	3.5	0
02/07/2006	3,094	12:00	none	2.7	0
04/19/2006	10,338	12:00	none	1.2	1
06/16/2006	9,659	12:00	none	1.5	1
08/24/2006	3,720	12:00	none	3.3	1
11/07/2006	4,301	12:00	none	3.6	0
01/04/2007	3,908	12:00	none	2.7	0
03/08/2007	3,252	12:00	none	3.1	0
06/20/2007	10,600	12:00	none	1.6	1
07/24/2007	3,882	12:00	none	4	1
09/06/2007	3,528	12:00	none	5	1
10/01/2007	4,041	12:00	none	4.4	1
12/11/2007	3,902	12:00	none	5	0
02/20/2008	4,262	12:00	none	3.3	0
04/25/2008	13,746	12:00	none	1.6	1
06/11/2008	26,189	12:00	none	1.1	1
08/12/2008	6,112	12:00	none	3	1

[1]The number of decimal places shown for dissolved selenium concentration is determined at the U.S. Geological Survey laboratory and depends on the accuracy of the particular analytical method used at the time. The number of decimal places shown in table 9 is the same as those reported in the U.S. Geological Survey database. In general, the more recent the sample, the greater the number of decimal places shown.

Table 10. Predefined regression models used by S-LOADEST.

[ln, natural logarithm; $\beta_0 - \beta_6$, regression model coefficients; sin, sine function; cos, cosine function; π, pi; Q, daily streamflow; dectime, decimal time; ε, error term]

Model number	Regression model
1	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \varepsilon$
2	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \varepsilon$
3	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot dectime + \varepsilon$
4	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot sin(2\pi \cdot dectime) + \beta_3 \cdot cos(2\pi \cdot dectime) + \varepsilon$
5	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \varepsilon$
6	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot sin(2\pi \cdot dectime) + \beta_4 \cdot cos(2\pi \cdot dectime) + \varepsilon$
7	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot dectime + \beta_3 \cdot sin(2\pi \cdot dectime) + \beta_4 \cdot cos(2\pi \cdot dectime) + \varepsilon$
8	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \beta_4 \cdot sin(2\pi \cdot dectime) + \beta_5 \cdot cos(2\pi \cdot dectime) + \varepsilon$
9	$ln(load) = \beta_0 + \beta_1 \cdot ln(Q) + \beta_2 \cdot ln(Q)^2 + \beta_3 \cdot dectime + \beta_4 \cdot dectime^2 + \beta_5 \cdot sin(2\pi \cdot dectime) + \beta_6 \cdot cos(2\pi \cdot dectime) + \varepsilon$