![USGS logo] **USGS**
science for a changing world

Correlations of Turbidity to Suspended-Sediment Concentration in the Toutle River Basin, near Mount St. Helens, Washington, 2010–11

Open-File Report 2014–1204

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

Correlations of Turbidity to Suspended-Sediment Concentration in the Toutle River Basin, near Mount St. Helens, Washington, 2010–11

By Mark A. Uhrich, Jasna Kolasinac, Pamela L. Booth, Robert L. Fountain, Kurt R. Spicer, and Adam R. Mosbrucker

Open-File Report 2014–1204

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

U.S. Department of the Interior
SALLY JEWELL, Secretary

U.S. Geological Survey
Suzette M. Kimball, Acting Director

U.S. Geological Survey, Reston, Virginia: 2014

For more information on the USGS—the Federal source for science about the Earth,
its natural and living resources, natural hazards, and the environment—visit
http://www.usgs.gov or call 1–888–ASK–USGS

For an overview of USGS information products, including maps, imagery, and publications,
visit *http://www.usgs.gov/pubprod*

To order this and other USGS information products, visit *http://store.usgs.gov*

Contents

Figures

Tables

Conversion Factors and Datums

Conversion Factors

Inch/Pound to SI

Multiply	By	To obtain
Length		
inch (in.)	2.54	centimeter (cm)
inch (in.)	25.4	millimeter (mm)
foot (ft)	0.3048	meter (m)
mile (mi)	1.609	kilometer (km)
Area		
square mile (mi^2)	259.0	hectare (ha)
square mile (mi^2)	2.590	square kilometer (km^2)
Volume		
cubic yard (yd^3)	0.7646	cubic meter (m^3)
Flow rate		
cubic foot per second (ft^3/s)	0.02832	cubic meter per second (m^3/s)
Mass		
ton, short (2,000 lb)	0.9072	megagram (Mg)

SI to Inch/Pound

Multiply	By	To obtain
Length		
millimeter (mm)	0.03937	inch (in)

Concentrations of suspended sediment in water are given in milligrams per liter (mg/L).

Datums

Horizontal coordinate information is referenced to the North American Datum of 1983 (NAD 83).
Vertical coordinate information is referenced to the North American Vertical Datum of 1929 (NAVD 29).
Elevation, as used in this report, refers to distance above the vertical datum.

Correlations of Turbidity to Suspended-Sediment Concentration in the Toutle River Basin, near Mount St. Helens, Washington, 2010–11

By Mark A. Uhrich, Jasna Kolasinac[2], Pamela L. Booth[3], Robert L. Fountain[2], Kurt R. Spicer[1], and Adam R. Mosbrucker[1]

Abstract

Researchers at the U.S. Geological Survey, Cascades Volcano Observatory, investigated alternative methods for the traditional sample-based sediment record procedure in determining suspended-sediment concentration (SSC) and discharge. One such sediment-surrogate technique was developed using turbidity and discharge to estimate SSC for two gaging stations in the Toutle River Basin near Mount St. Helens, Washington. To provide context for the study, methods for collecting sediment data and monitoring turbidity are discussed. Statistical methods used include the development of ordinary least squares regression models for each gaging station. Issues of time-related autocorrelation also are evaluated. Addition of lagged explanatory variables was used to account for autocorrelation in the turbidity, discharge, and SSC data. Final regression model equations and plots are presented for the two gaging stations. The regression models support near-real-time estimates of SSC and improved suspended-sediment discharge records by incorporating continuous instream turbidity. Future use of such models may potentially lower the costs of sediment monitoring by reducing time it takes to collect and process samples and to derive a sediment-discharge record.

Introduction

Suspended-sediment transport throughout the Toutle River Basin has been monitored and studied since 1980–81, following the eruption of Mount St. Helens on May 18, 1980. This study used standard U.S. Geological Survey (USGS) methods to compute sediment-discharge for gaging stations in the basin (Porterfield, 1977), along with standard laboratory and field procedures (Guy, 1977; Edwards and Glysson, 1999). Streamflow and suspended-sediment concentration (SSC) have been measured, and suspended-sediment discharge (SSQ) has been computed, in several drainages in the Toutle River Basin (Dinehart, 1998). SSC data are collected by pump sample most days and by depth-integrated methods on infrequent days. This report uses data from two long-term gaging stations on the North Forth Toutle River and main-stem Toutle River. Daily, monthly, and annual SSC and SSQ data are available online.

---

[1]U.S. Geological Survey.
[2]Portland State University.
[3]University of Rhode Island.

In recent years, technology improvements have spawned efforts to develop innovative and improved methods of generating time-series records of SSC and SSQ. Traditionally, sample-based methods require lengthy evaluation and review before sediment records are finalized, although interactive software referred to as Graphical Constituent Loading Analysis System (GCLAS) has improved this processing (Koltun and others, 2006). Using recently approved methods (Rasmussen and others, 2009); this study examines turbidity as an alternative or surrogate for SSC with the intention of better defining SSQ, streamlining record computations, and possibly lowering costs. Additionally, land, water, fish, and wildlife resource planners need real-time estimations of SSC and SSQ to more effectively respond to changes and disturbances in basins under their management. These techniques, which compute SSC from turbidity and streamflow, coupled with a gaging-station telemetry system, potentially would allow delivery of near real-time SSC and SSQ data. Because real-time SSQ estimates are considered provisional owing to sensor and sampling uncertainty, regression-based SSQ records would be finalized annually following approval of turbidity and streamflow data. Use of a regression model to compute sediment records may improve accuracy by incorporating high-frequency measurements of explanatory variables, and also may lower costs by reducing record processing time and the number of samples collected and analyzed. The sediment-sample collection, turbidity monitoring, and regression analysis for this study were conducted in cooperation with the U.S. Army Corps of Engineers, Portland District.

Purpose and Scope

- The primary objective of this study is to test the feasibility and application of instream turbidity sensors at two sites in the Toutle River Basin and to demonstrate the use of these sensors as a surrogate for SSC, and document the results.
- Turbidity and streamflow data from April 2010 to September 2011 are used to generate regression models for estimating SSC. Such models can be updated as new turbidity, streamflow, and sampled SSC data become available.
- Regression equations are provided for both streamgages and could be used to provide near-real-time estimates of SSC and SSQ. The proof of concept is shown and regression-based estimates for the time-series data could be finalized if they were deemed beneficial. Future projections of SSC also could be made available as an online data series.
- Finally, we make a preliminary assessment as to whether using such a regression approach would provide a better-quality SSQ estimate and would reduce effort and expense compared to previous methods.

Study Area

The number and location of streamflow-gaging and sediment-monitoring stations in the Toutle River Basin have evolved since their establishment in 1980–81. Current (2014) gaging stations shown in figure 1 include North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington (NF Toutle-SRS, 14240525); and Toutle River at Tower Road near Silver Lake, Washington (Toutle-Tower, 14242580). A third gaging station, South Fork Toutle River at Toutle, Washington (SF Toutle, 14241500), was discontinued in 2013. For the 6 water years (WYs 2007–12)

the reported NF Toutle-SRS total SSQ was more than 18 million tons (units in short tons), constituting more than 67 percent of the total SSQ of nearly 27 million tons computed for Toutle-Tower. For the 20-year period, WYs 1993–2012, the reported total SSQ for Toutle-Tower was more than 60 million tons, an annual average of 3 million tons. For the 1.5-year (April 2010–September 2011) period of data in this report, the Toutle-Tower SSQ totaled nearly 2.9 million tons, slightly less than the yearly average (*http://wdr.water.usgs.gov/*).

The Toutle-Tower gaging station, at 160-ft in elevation, is about 7 river miles (RMs) upstream of the confluence of the Toutle and Cowlitz Rivers and has a drainage area of 496 mi^2. The NF Toutle-SRS gaging station, at RM 12 of the North Fork Toutle River, drains 175 mi^2, and is about 30 RMs upstream of the Toutle-Tower gaging station (fig. 1). The NF Toutle-SRS station, at 700-ft elevation, is less than 2 RMs downstream of the Sediment Retention Structure (SRS). The SRS was completed in 1989 to retain and avert sediment eroded from the Mount St. Helens debris-avalanche deposit from being transported to the lower basin and eventually the Cowlitz and Columbia Rivers. Through 2012, the SRS has trapped about 115 million yd^3, representing about 3.5 percent of the total sand and gravel deposited after the 1980 eruption (Major and Spicer, 2003; Gibson and others, 2010). Nonetheless, a large volume of fluvial sediment passing the SRS is deposited downstream and is aggrading channel beds, thereby increasing the threat of flood inundation to the surrounding communities, as well as posing a hazard to river navigation and economically important commerce, drinking-water supplies, and migrating fish.
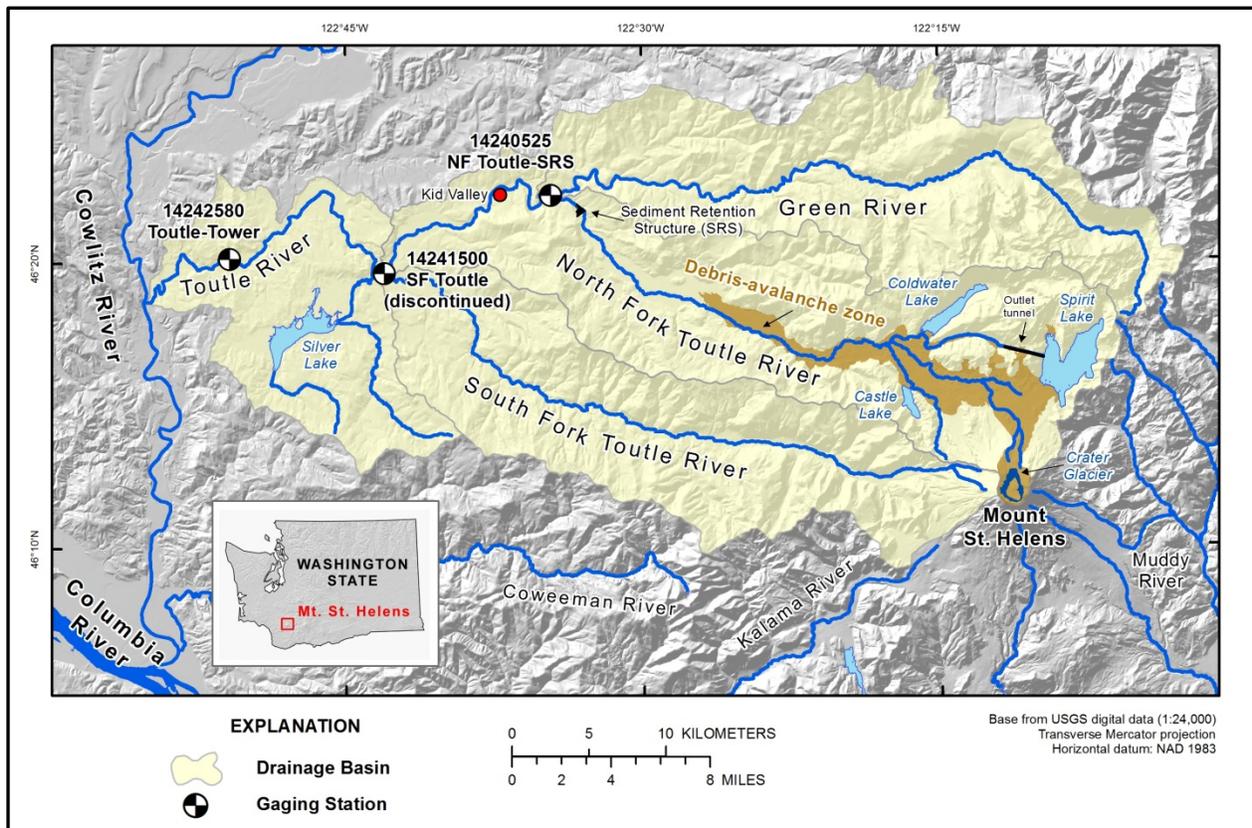
Figure 1. Map showing Toutle River Basin study area, drainage basins, and U.S. Geological Survey (USGS) gaging station locations, near Mount St. Helens, Washington.

Data Collection and Analysis Methods

To achieve the objectives in the section, "Purpose and Scope," we installed instream turbidity sensors at two gaging stations, NF Toutle-SRS and Toutle-Tower. Fifteen-minute unit-value turbidity and discharge data and periodic suspended-sediment samples were collected at both gaging stations (U.S. Geological Survey, 2010, 2011).

Matched pairs of turbidity and discharge with SSC were used as the explanatory and response variables, respectively, in a multi-linear regression using ordinary least squares (OLS) methods. Separate regression models were generated for each station. The resulting equations can be used to estimate 15-minute unit values of SSC from associated turbidity and water discharge unit values. Finally, the regression results, including accompanying uncertainty estimates, can be compared with previous sample-based sediment records computed for these stations in the Toutle River Basin in order to evaluate the relative utility of the traditional and surrogate methods.

Several established USGS methods were used to collect and process the suspended-sediment samples and to check, review, and publish the turbidity data.

Suspended-Sediment Sampling

This study started in April and May 2010 for the Toutle-Tower and NF Toutle-SRS streamgages, respectively, when turbidity calibrations and data collection began (figs. 2 and 3). Suspended-sediment samples were collected routinely in WYs 2010 and 2011, with an emphasis on storm, high-streamflow, and high-turbidity events.

Figure 2. Photographs showing suspended-sediment sampling, March 12, 2010 (large photograph), and servicing sensors, April 20, 2012 (inset), at North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington. Photographs taken by Kurt Spicer, USGS, Cascades Volcano Observatory.

Figure 3. Photographs showing suspended-sediment sampling at Toutle River at Tower Road near Silver Lake, Washington, December 13, 2012. Photographs taken by Kurt Spicer, USGS, Cascades Volcano Observatory.

Cross-Sectional, Depth-Integrated Sediment Samples

Manual suspended-sediment samples were collected at both gaging stations using standard USGS depth-integrated, equal-discharge-increment (EDI) and equal-width-increment (EWI) methods, (figs. 2 and 3; Edwards and Glysson, 1999). These sampling procedures have been used consistently at the NF Toutle-SRS and Toutle-Tower gaging stations since sampling began in the early 1980s. EDI and EWI sampling methods are the accepted procedures for providing representative cross-sectional SSCs.

Two sets of manual EDI/EWI samples (sets "A" and "B") usually were collected nearly simultaneously for each sampling visit and can be used independently or averaged to produce a single concentration and to better capture sample uncertainty (Topping and others, 2011). Manual data in this study used sample A and B sets so that each individual concentration could be used to better populate and define the regression model.

Point Samples

Automatic pumping samplers on the bank at each site were used to augment the EDI/EWI cross-sectional samples for periods between the manual collections. A single pump sample per day usually was collected in addition to multiple samples during high-flow events. Autosamplers draw water from a single point in depth and cross section, and, therefore, differ from the EDI/EWI methods that capture spatial variability throughout the stream width and water column. Autosampler concentrations nearest in

5

time to EDI/EWI samples were evaluated to determine if an adjustment or shift in the autosampler concentration was necessary. These point-sample correction adjustments or coefficients are used to shift autosampler concentrations to better reflect the mean cross-section concentration defined by manual EDI/EWI samples.

Autosampler concentrations typically are less than or equal to manual-sample concentrations (Glysson, 2008). To establish a correction coefficient, a pumping sample normally is manually triggered before and after an EDI/EWI cross-section sample set. Generally, if the pump and EDI/EWI sample concentrations agree to within 5 percent, no correction is applied and the coefficient is 1.0. If the difference is greater than 5 percent, the autosampler concentrations are adjusted to the manual concentration with a shift usually greater than 1.0. The shift is applied across time, either by relation to flow or by linear proration, until the next measured pumping sample coefficient. The corrections are defined by a manual cross-sectional sample or by a particular streamflow or turbidity event that may have altered the pumping efficiency or indicated a change in stream-channel dynamics (Guy, 1977; Guy, 1978; Porterfield, 1977; Bent and others, 2000). Finally, as in any sample collection program, there is a delay in acquiring the concentration data because of shipment time and laboratory processing, so that pump and manual sample results are not available in real time.

Turbidity Measurement and Data Processing

Turbidity data were collected and processed using established USGS procedures for continuous water-quality monitoring (Wagner and others, 2006). Continuous turbidity data were collected at NF Toutle-SRS and Toutle-Tower using a DTS-12 sensor™, manufactured by Forest Technology Systems, Victoria, Canada (*http://www.ftsenvironmental.com/products/sensors/dts12/*). The sensor has a large optical face that allows for a relatively wide water column area to be measured by the lens and detector. The probe has a large and durable wiper that virtually eliminates the need for cleaning corrections because debris buildup on the optics is removed at each reading. The head of the sensor is angled at 45 degrees to lessen the formation of air bubbles, which can interfere with the optics and cause false readings. The sensor head must be oriented facing down and into the main water body for correct turbidity readings. The DTS-12 sensor™ turbidity readings are reported in Formazin Nephelometric Units (FNU) (Anderson, 2005).

Suspended-sediment concentrations in the Toutle River Basin typically range from 10–50 milligrams per liter (mg/L) during extended periods of low flow, to 10,000–20,000 mg/L during storm runoff. Such sediment-laden waters can negatively affect instream electronic instrumentation. The DTS-12 sensors™ have worked consistently through these harsh conditions, requiring only routine cleanings with calibration checks every 3 to 6 months. The DTS-12 sensor™ takes 20 readings per second over 5 seconds and provides several parameters for those 100 readings. These parameters include mean, median, minimum, and maximum turbidity, and water temperature. Two variance parameters also are included to help with quality assurance for the other parameters. Near real-time median turbidity readings are reported on the USGS National Water Information System Web site (*http://waterdata.usgs.gov/wa/nwis/current/?type=flow*) in 15-minute intervals, and are used in the regression analyses. Daily median, minimum, and maximum turbidity for the two gaging stations are published in the Washington Annual Data Report (U.S. Geological Survey, 2010, 2011). Approved instantaneous turbidity and discharge data for NF-Toutle-SRS and Toutle-Tower used in this analysis are shown in figure 4.
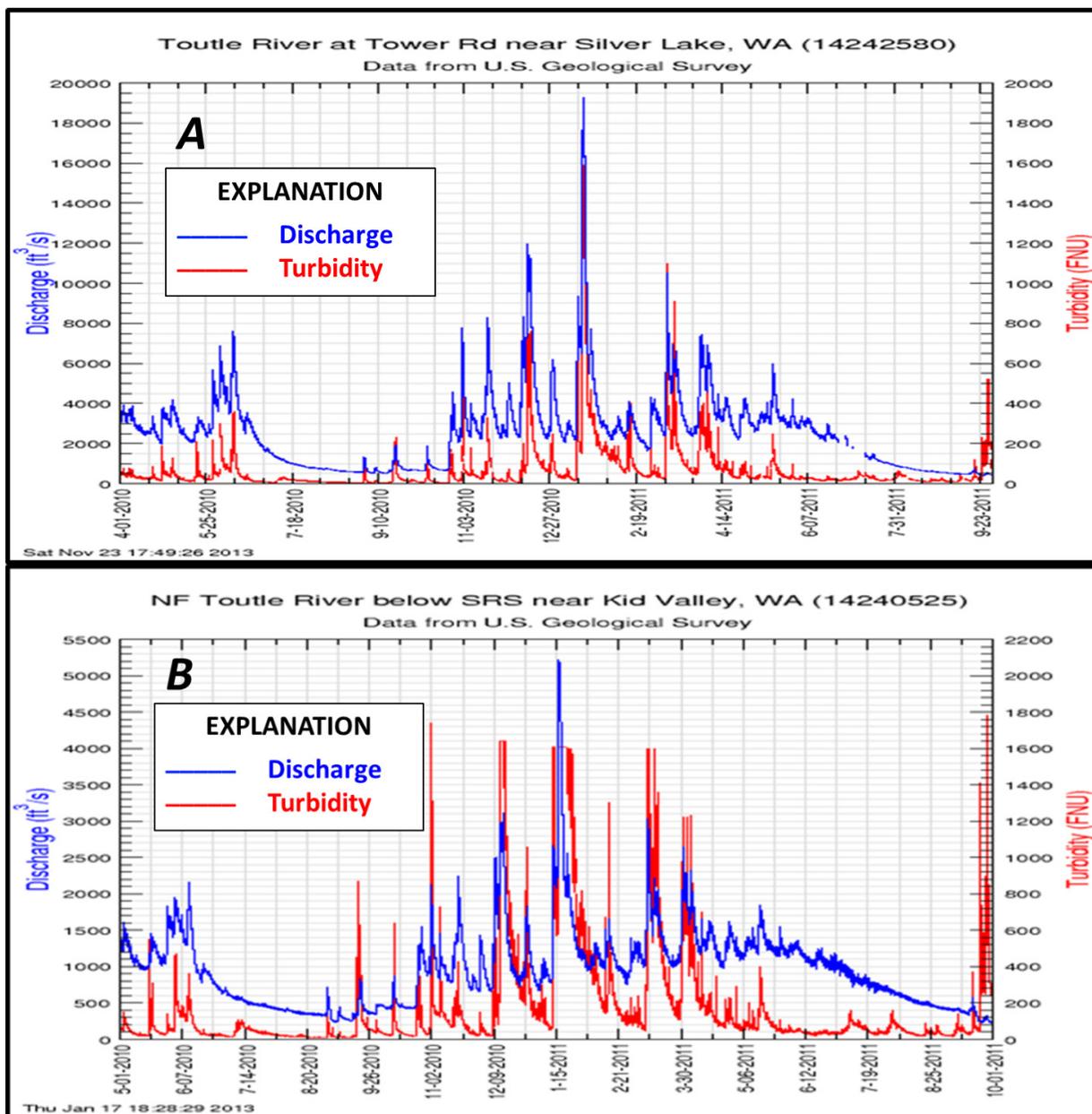
Figure 4. Graphs showing stream discharge and turbidity at two streamgages in Toutle River Basin, Washington, 2010–11. (*A*) April 1, 2010–September 30, 2011 and (*B*) May 1, 2010–September 30, 2011. (ft^3/s, cubic foot per second; FNU, Formazin Nephelometric Units.)

Turbidity Greater than Instrument Limits

All instream turbidity sensors have a maximum, instrument-specific reading. If turbidity surpasses that threshold, the sensor produces a false reading wherein the maximum value is reported repeatedly throughout the event. When graphed, this turbidity threshold displays as a horizontal line. After turbidity decreases to less than this threshold, the sensor again records valid measurements within the range of the probe.

The DTS-12 sensor™ threshold varies from sensor to sensor and ranges from about 1,600 to 2,100 FNU. Turbidity at NF Toutle-SRS reached the sensor threshold in December 2010, and in January and March 2011, with each threshold reading lasting from several hours to as long as 5 days. Turbidity at Toutle-Tower exceeded the sensor threshold on January 16, 2011, for 3 hours. Sediment samples collected during these turbidity sensor thresholds were not included in the regression models, as the true turbidity for the samples was unknown.

Existing alternative turbidity sensors suitable for instream monitoring that measure values greater than the DTS-12 sensor™ threshold potentially could provide a more consistent and complete turbidity record through peak events. Such a turbidity sensor was tested and routinely calibrated at NF Toutle-SRS, although the records for that sensor have not yet been approved. Data from this alternative sensor could be used to supplement periods when the DTS-12 sensor™ recorded threshold values and flat-lined. It then would be possible to run the regression models using these secondary values. The high-end turbidity values, if estimated or measured for missing or greater-than-threshold periods, also could be used to compute more complete and continuous model-generated SSC and SSQ values, which would be valuable given that these often are the periods of the greatest sediment transport. However, processing the high-end turbidity data would require further examination and review. Because the development of turbidity-surrogate regressions for this report was considered a proof of concept for the Toutle River gaging stations, processing high-end turbidity data was beyond the scope of this report; we, therefore, used only existing turbidity data that was approved and published. The potential utility of the high-turbidity data for refining the existing load estimates is considered in the section, "Discussion and Future Studies."

Selection of Turbidity and Sediment Concentration Data for Regression Analysis

Approved turbidity and SSC data were paired by matching the autosampler and EDI concentration to the closest-in-time turbidity value. If the EDI sample took more than 30 minutes to collect, the 15-minute turbidity values were averaged for the necessary time period in order to obtain a single value. Turbidity and sediment-sample data used for this report constitute roughly one-half of WY 2010 and the entire WY 2011 (April or May 2010–September 2011) for the Toutle-Tower and NF Toutle-SRS gaging stations, respectively. This provided a base dataset to begin construction of the regression models (appendix A). These relations can be evaluated from year to year, and can be compared with turbidity and SSC data collected in later years to determine any shift in turbidity-discharge to SSC relations and (or) transport regime.

To maintain consistency with the previously published sediment records for these periods (*http://wdr.water.usgs.gov/*), the identical sample concentrations (both EDIs and pumping samples) used in the sediment records were used in the regression analysis, except for turbidity and sample concentrations deleted from the analysis dataset when turbidity readings were at maximum threshold. The total number of EDI and pumping samples available for each gaging station collected from April or May 2010 through September 2011 are shown in table 1.

NF Toutle-SRS and Toutle-Tower discharge and turbidity with EDI samples collected from May 1 or April 1, 2010, through September 30, 2011, are shown in figure 5. Two EWI samples were collected, but neither sample was used because of contamination from the streambed.

Table 1. Number and type of sediment samples collected at North Fork Toutle River below Sediment Retention Structure near Kid Valley (NF Toutle-SRS) and Toutle River at Tower Road near Silver Lake (Toutle-Tower), Toutle River Basin, Washington, 2010–11.

Gaging station and sample dates	Equal-Discharge-Increment samples collected	Pumping samples collected
NF Toutle-SRS, May 2010–September 2011	48	605
Toutle-Tower, April 2010–September 2011	9	696
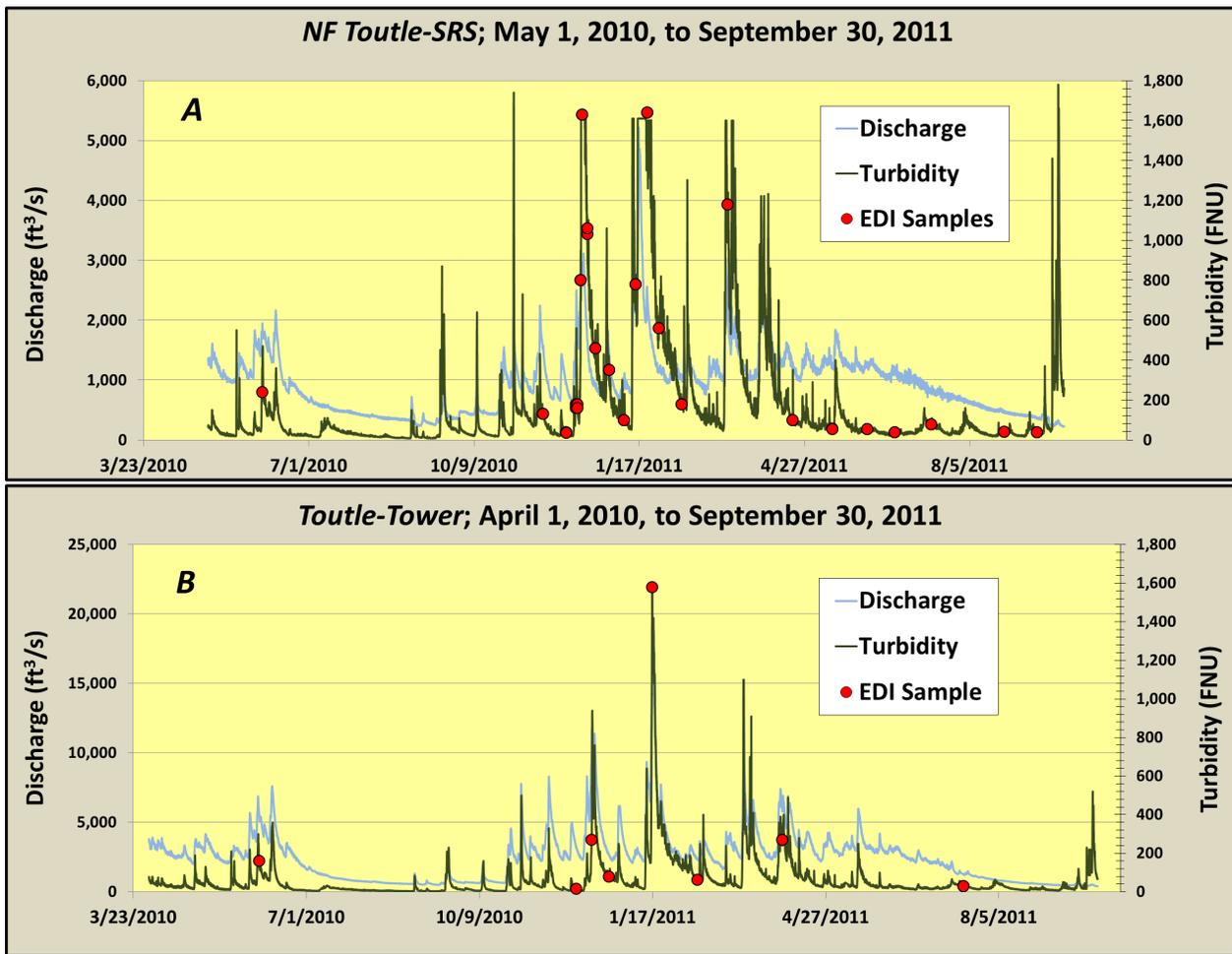
Figure 5. Graphs showing stream discharge and turbidity with time of equal-discharge-increment samples collected overlaid on turbidity, at (*A*) North Fork Toutle River below Sediment Retention Structure near Kid Valley (NF Toutle-SRS), May 1, 2010–September 30, 2011, and (*B*) Toutle River at Tower Road near Silver Lake (Toutle-Tower), April 1, 2010–September 30, 2011, Toutle River Basin, Washington. Discharge is measured in cubic feet per second (ft^3/s) and turbidity is measured in Formazin Nephelometric Units (FNU). Not all points are visible because of overlap of set A and B sample points collected close in time to each other.

9

These two sample sets indicate a strong reliance on autosamples, as is the normal routine in working a sample-based sediment-discharge record. As mentioned in the "Suspended-Sediment Sampling; Point Samples" section, autosamples, by nature of their position and orientation along the side of a channel cross section and as single-point samples, may not typically represent a concentration equal to the manual depth-integrated, cross-sectional samples. Therefore, a regression-based approach ideally would rely more on EDI/EWI samples than on pumping samples because of the differences in uncorrected concentrations between the two sample types.

Discharge Data

Water discharge values used for this analysis were computed from a stage-discharge rating developed from current-meter measurements and a 15-minute, time-series stage record, using established USGS techniques (Buchanan and Somers, 1976; Rantz and others, 1983). Streamflow measurements typically, but not always, accompanied cross-section EDI and EWI samples. Current-meter instruments were used exclusively for discharge data in this report. According to Sauer and Meyer (1992), the standard errors associated with individual discharge measurements can range from 2 to 20 percent, although most standard errors range from 3 to 6 percent. Discharge data for the Toutle River sites are available at *http://wdr.water.usgs.gov/*.

Statistical Methods

Regression Models Applied

We used OLS linear regression (Helsel and Hirsch, 2002) with turbidity and discharge as explanatory variables and the EDI/EWI and auto-sampled SSC data as the response variable. Regression model development for SSC is covered extensively in Rasmussen and others (2009), including various correlation and data transformation measures and use of available explanatory variables. We selected the best candidate model on the basis of supportive diagnostic statistics, the fit of the explanatory and response variables, and hydrographer knowledge of sediment dynamics and data collection at the individual sites.

Following visual and statistical analysis of the SSC, turbidity, and discharge datasets, as well as examination of the residuals from preliminary OLS models, we log-transformed the datasets of both streamgages to improve distributional normality. We also tried natural log, square, and cube root transformations. The log transformation worked best overall by compressing tailings and outliers, as well as addressing possible heteroscedasticity, thereby improving the fit of the regression (Helsel and Hirsch, 2002). We also tested using a univariate model with turbidity as the sole explanatory variable. Finally, the addition of discharge statistically improved the sum of squares error (SSE) and coefficient of determination (R^2) and, therefore, was used in a multiple linear regression (MLR). However, the log transformation and MLR did not alleviate time-related auto-correlation, as indicated by low Durban-Watson statistics (Helsel and Hirsch, 2002). Although this transformation improved overall model fit by decreasing the SSE and normalizing the residuals, autocorrelation was still a concern.

One method to address autocorrelation and to increase accuracy in the regression model was the inclusion of time lags of turbidity and discharge as additional explanatory variables. The final MLR used a single lag of turbidity as a third variable. The inclusion and importance of lagged turbidity is explained in the section, "Accounting for Autocorrelation."

Statistical Diagnostics and Analysis of Variance

      Analysis of Variance (ANOVA) statistics generated for each gaging station regression are shown in tables 2 and 3. The structure of the ANOVA is written from left to right, with each column broadening the understanding and role that each "Source" statistic contributes to the development and significance of the final regression model. A base understanding of the terminology and structure of the statistics is necessary to better interpret the results.

      The Sequential Sum of Squares (Seq. SS) consists of the decomposition of the sum of the squared difference between the individual observed value of the log of SSC and the mean of log of SSC into the "*Regression*" part and the "*Error*" part. The *Regression* part is the sum of the squared difference between the predicted value and the overall mean of log of SSC, whereas the *Error* part is the difference between the observed value and the predicted value. Because there are multiple values for each day, the *Error* is further decomposed into *Lack-of-Fit* (sum of square of difference between local average and fitted) and *Pure Error* (sum of square of difference between observed and local average). These SS values then are corrected for bias by their respective degrees of freedom (df) with the unbiased estimation value under Sequential Mean Square (Seq. MS). The Seq. MS functions as the value for the estimations of variance for the distributions of the *Regression* and the *Error*.

Table 2. Final Analysis of Variance (ANOVA) for logSSC compared to LogT, logT-lag, logQ, for North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington. [See text for explanation of statistical terms].

Source	Seq. SS	df	Seq. MS	F-statistic	P > F
Regression	145.62527	3	48.5419	929.71	0.000
Error	**33.8854**	649	0.05221		
Lack-of-Fit	33.798	624	0.0541635	15.48589	0.000
Pure Error	0.08744	25	0.0034976		
Total	179.5112	652			

$$\sqrt{MS_E} = 0.228499 \; R^2 = \mathbf{81.1\%} \; R^2_{adj} = 81.0\% \; PRESS = 34.4759 \; DW = 0.168913 \; Cp = 4$$

where

Seq. SS	*is Sequential Sum of Squares,*
Seq. MS	*is Sequential Mean Squares,*
df	*is degrees of freedom,*
Regression SS	*is Sum of Squares from Regression; Regression SS/Regression df*
Regression MS	*is Mean Squares from Regression (MS$_E$)*
Error SS	*is Sum of Squares Error (SSE); Pure Error SS +Lack-of-Fit SS*
Error MS	*is Mean Squares Error (MS$_E$); SSE/Error df or Error SS/Error df*
Pure Error SS	*is True Error*
Lack-of-Fit SS	*is Error from poor estimation*
Total SS	*is Total Sum of Squares; Regression SS + SSE*

The *SSE* and R^2 values are important statistical and comparative diagnostics referred to in the "Accounting for Autocorrelation" section, hence appear bolded to emphasis.

Table 3. Final Analysis of Variance (ANOVA) for logSSC compared to LogT, logT-lag, logQ, for Toutle River at Tower Road near Silver Lake, Washington.

Source	Seq. SS	df	Seq. MS	F-statistic	P > F
Regression	428.1358	3	142.7112	2,968.57	0.000
Residual Error	**33.70**	701	0.04807		
Lack-of-Fit	33.55	693	0.0484127	2.5803	0.074
Pure Error	0.1501	8	0.0187625		
Total	461.8359	704			

$$\sqrt{MS_E} = 0.219259 \ R^2 = 92.7\% \ R^2_{adj} = 92.7\% \ PRESS = 34.3041 \ DW = 0.686354 \ Cp = 4$$

In testing the significance or statistical fit of the regression equation for the NF Toutle-SRS and Toutle-Tower gaging stations, the ANOVA F-statistic from tables 2 and 3 indicates a significant relation between log of turbidity and log of SSC with a 1-percent probability of a type I error or the probability of incorrectly rejecting a true null hypothesis. The significance is determined by comparison to a critical F* value on the F-distribution with 3, and 649 or 701 df for the NF Toutle-SRS and Toutle-Tower sites, respectively, as determined by the numerator ($MS_{regression}$ = *Regression**SS**/Regression**df***) and the denominator ($MS_E = ErrorSS/Errordf$). This F-statistic is formed through the ratio of two probability distributions: the explained regression to the unexplained errors. The resulting ratio is an indicator of the overall fit of the regression model without involving units of measure or implying multiplicative effects.

The F-statistics for both regressions indicated that a significant proportion of the variation in log (SSC) was explained by the relation with log T (Turbidity) and log Q (Discharge) relative to the unexplained variation in log (SSC). Because the variance estimator Seq. ME (or MS_E from the ANOVA table) is expressed as *SSE* divided by df of the error, focusing on minimizing *SSE* was important for minimizing the estimate of the variance and standard deviation ($\sqrt{MS_E}$) of the model. The ANOVA tables 2 and 3 also included a "*Lack-of-Fit*" statistic that for both regressions was significant, indicating a poor overall fit. The discrepancy between the F and Lack-of-Fit statistics indicates a high variation within the data, including the possibility of autocorrelation of the errors observed through the distribution of the residuals, as reflected in the low Durban-Watson scores.

One of the best methods for determining the quality of a regression is the *PRESS* or prediction sum of squares. In general terms, the *PRESS* is a cross-validation calculation that provides a regression-fit summary that measures how well the model will perform in predicting new data. *PRESS* values were included and evaluated so that the best candidate model would have the lowest *PRESS*, and, thus, the best structure.

The Variance Inflation Factors (VIFs) in tables 4 and 5 also help determine the quality of a regression; VIFs measure the extent to which multicollinearity was present between the explanatory variables. Multicollinearity occurs when two or more variables are linear combinations of the other variables. A VIF greater than 5 is cause for concern, whereas a VIF greater than 10 is a major sign of colinearity, indicating that the predictors are highly correlated. Also provided in tables 2 and 3 are Mallow's Cp statistics, which are designed to minimize bias and standard error by keeping the number of coefficients low and in balance. Too few model variables cause bias, whereas too many predictors result in an imprecise model. Mallow's Cp is used so that the precision and bias of the full MLR is compared to the best subsets of predictors. The desired Mallow's Cp is a value that is close to the number of beta or explanatory variable coefficients plus the constant or y-intercept. This provides a model that is relatively precise and unbiased in estimating the correct regression coefficients, as well as predicting future responses or SSCs. Overall, the ANOVA results in tables 2 and 3 indicate that the final regressions between log SSC and the log transformed turbidity and discharge data are significant, with these variables explaining much of the variation in SSC. However, the strength of these relations is lessened because of the presence of significant autocorrelation.

Autocorrelation

The large number of daily and sub-daily pumping samples and paired EDI A and B sets, collected close in time to each other and available for this analysis, opened the dataset to potential problems associated with autocorrelation, or the serial correlation of a variable such as turbidity and (or) suspended-sediment concentration with itself over successive time intervals. When a variable indicates autocorrelation, one observation is related to another observation such that both observations will change together to some extent. In this case, the individual values of SSC, turbidity, or discharge are essentially similar to their previous value in the time series, such as during a storm event, and, therefore, do not represent random or independent occurrences. This presents a problem because statistically sound OLS regression models are assumed to have independent and normally distributed errors. When the errors, as observed through the residuals, show autocorrelation, the OLS method tends to underestimate the standard errors and coefficients of the model, thereby producing erroneously narrow confidence and prediction interval bands. These patterns typically can be identified through graphical analysis. For instance, if several samples are collected during a particular event, such as on the rising or falling limb of a hydrograph, the residuals may appear grouped together for that event in a non-random pattern.

Initial attempts to minimize the effects of autocorrelation led to averaging EDI-paired A and B sample sets, as well as randomly subsampling the autosamples. These smaller datasets then were tested by applying different regressions on the reduced number of autosamples and EDIs as suggested by Helsel and Hirsch (2002). Although these attempts reduced the potential for autocorrelation, the resulting graphical and statistical analysis showed minimal reduction. Therefore, additional methods were used to develop a model using all data (EDIs and autosamples), while also reducing the autocorrelation and *SSE*.

Evaluating Autocorrelation

The Durbin-Watson statistic (DW) (tables 2 and 3) essentially is the measure of the Sum of Error generated from the difference between a residual at index i and index i-1 taken over all residuals. A DW statistic between 0 and 1.6 generally indicates a positive auto-correlation for large sample sizes, especially when DW is less than 1. Because the DWs for both regressions were close to 0, there is a strong indication that positive autocorrelation was present. Because there was reason to be concerned about the variability of the residuals, a closer analysis of residual graphs for normality was warranted.

The normal probability graphs for the NF Toutle-SRS and Toutle-Tower gaging stations (figs. 6A and 7A) showed no strong deviation from a normal distribution of residuals. However, a comparison of the histogram (figs. 6B and 7B) and the "fitted values" against their residuals (figs. 6C and 7C) showed some abnormal grouping and tailing. Collectively, these three graphs show no reason for concern; for each station, the graphs of residuals against "observation order" (figs. 6D and 7D) showed that the residuals were related to each other across time, substantiating the DW statistic.
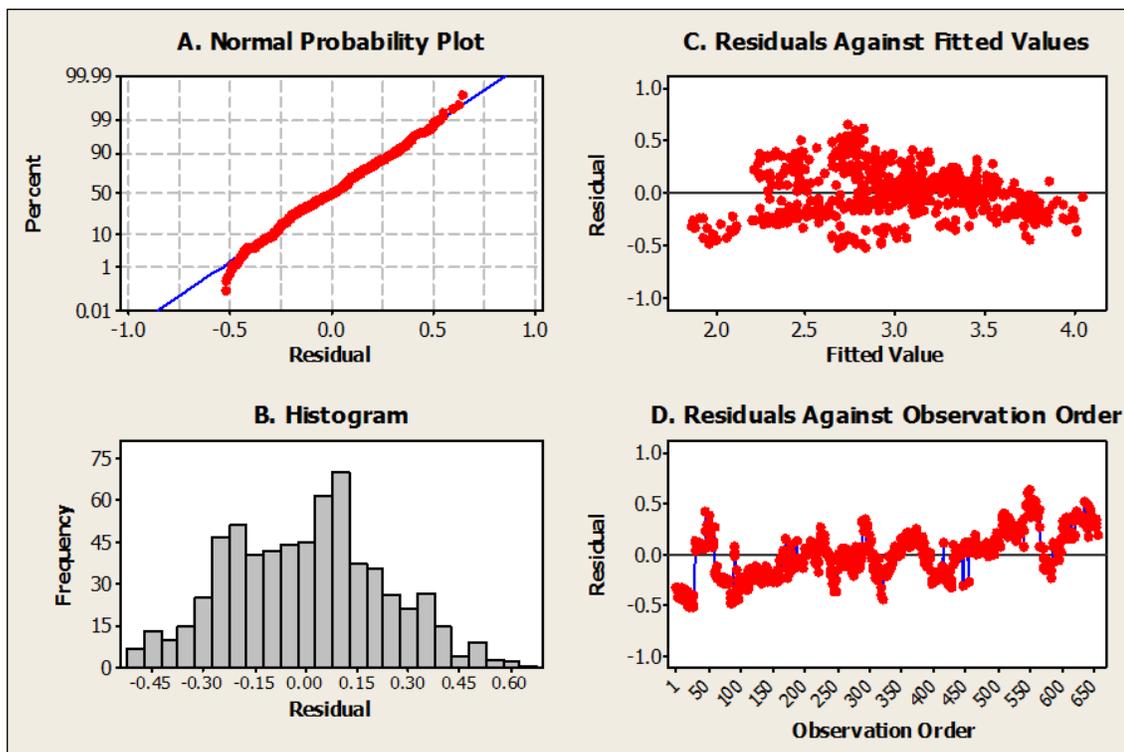
Figure 6. Graphs showing (*A*) normal probability distribution of residuals; (*B*) frequency distribution of residuals; (*C*) comparison of residuals with fitted values; and (*D*) comparison of residuals with observation order, for a multivariate regression of log(SSC) against log(T), log(Tlag), and log(Q), for North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington. Figure made from Minitab® software as 4-in-1 plots (*www.minitab.com*).
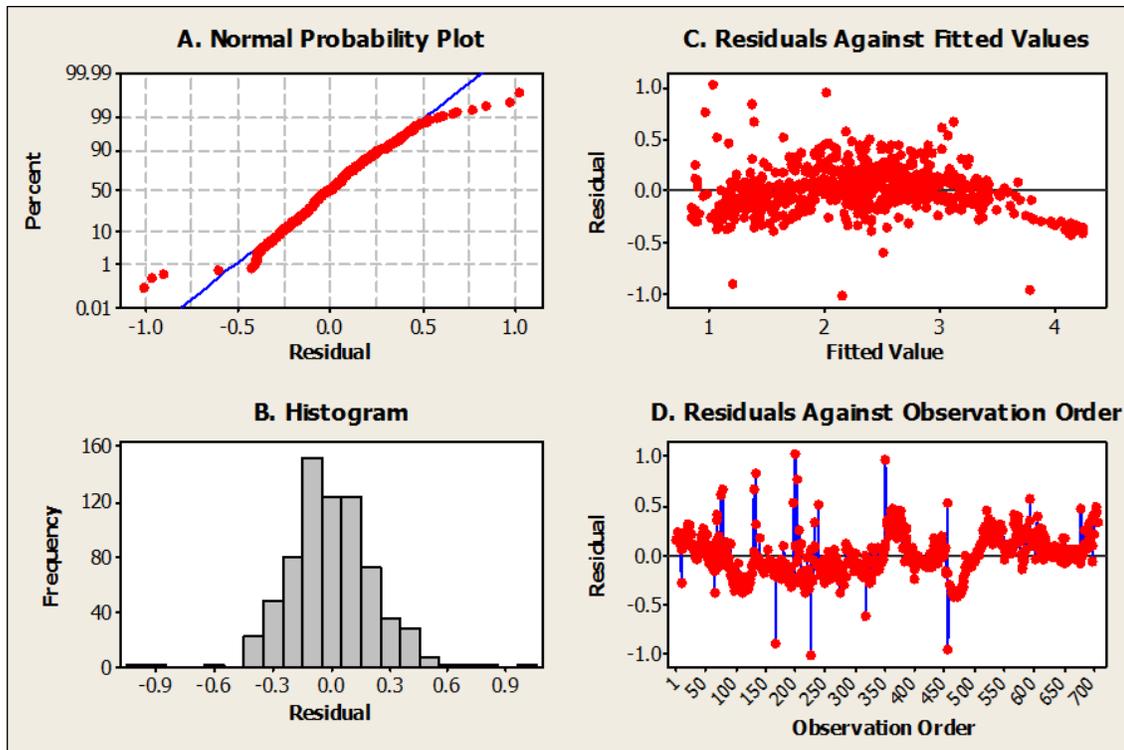
Figure 7. Graphs showing (*A*) normal probability distribution of residuals; (*B*) frequency distribution of residuals; (*C*) comparison of residuals with fitted values; and (*D*) comparison of residuals with observation order, for a multivariate regression of log(SSC) against log(T), log(Tlag), and log(Q), for Toutle River at Tower Road near Silver Lake, Washington. Figure made from Minitab® software as 4-in-1 plots (*www.minitab.com*).

Accounting for Autocorrelation

There are various options to account for time-related autocorrelation, including Auto-Regressive Moving Average (ARMA) modeling (Box and Jenkins, 1976); state-space modeling (SSM) using a Kalman filter (Harvey, 1989); and variable lagging, among others. Because this particular application was for real-time estimation and not for future forecasting, more extensive autocorrelation modeling techniques such as ARMA and SSM were not used. Additionally, the collection time difference between paired observations reduced the necessity for more extensive modeling as described in "Autocorrelations Die Out." Thus, regressions were run adding lag values of discharge and turbidity to account for some of the autocorrelation. The inclusion of lags increased R^2, lowered the standard error (*SSE*), and improved the DW statistic. The final R^2 and *SSE* are listed with tables 2 and 3.

Autocorrelations Die Out

Although this analysis indicated that autocorrelation was present in the datasets, use of more extensive time-series modeling was impractical for real-time application given that the correlation of logSSC with the most recent observed value of SSC died out after about 30 days. That is, the daily statistical dependence or strength of the relation between the 96 values of 15-minute logSSC variables decreases to near zero in about 1 month, such that the change in one 15-minute SSC will correspond to a change in another 15-minute SSC for only about 30 days. Because it normally takes more than 30 days for a sample concentration to become available from the laboratory and accessible for analysis, and because this dataset contained breaks in pump and manual sample collection that were longer than 30 days, this model used lagged values instead of a time-series component to increase accuracy in the regression model. Given these conditions, the regression developed using 2010–11 data worked adequately because the SSC correlations went to zero in such a relatively short time. In other words, the 30-day die out and the availability of new SSC sample data will almost never overlap, making the value of time-series models negligible in real-time estimation of SSC. If SSC were to be predicted into the future, a time-series model would be necessary.

Lagging Turbidity and Discharge

Regressions using lagged values of turbidity and discharge were tested for significance and improvement over the non-lagged MLR. A lag is a past value of the variable; a turbidity lag of 1 would use the previous 15-minute value, a turbidity lag of 2 would use the previous 30-minute value, and so on. In our case, we evaluated using 1 lag of turbidity and 1 lag of water discharge by adding these values as third and fourth explanatory variables. On the basis of the regression diagnostics and ANOVA, we decided to use a single turbidity lag of 1, without using lags of discharge.

Robustness Checks

The term "robustness" here refers to statistics with good performance with the data, such that the coefficients are resistant to errors in the results and not unduly influenced by outliers. Robustness checks look for consistency of coefficient estimates by subsampling the original dataset and then estimating the model with out-of-sample data, along with other means of testing the validity of regression results.

The consistency of the OLS regression coefficient estimates for each streamgage was checked using the following methodology: Data for each streamgage first was condensed to a single matched pair per day. Days with only one value were automatically included. On days with multiple observations, one observation per day was randomly selected. For NF Toutle-SRS, the 653 observations came from 341 days, and for Toutle-Tower, the 705 observations came from 355 days.

This condensed dataset of 341 and 355 observations, respectively, was further subsampled. Each observation was given a random number and then sorted by that number from highest to lowest. The top 90 percent of the data with the highest random number were selected for use in generating the five potential OLS regression models. The five sets of explanatory variables included (1) logT; (2) logT and logT-lag; (3) LogT and logQ; (4) LogT, logT-lag, logQ; and (5) logT, logT-lag, logQ, logQ-lag. The remaining 10 percent of the data were used as out-of-sample or sequestered data and input to the 90-percent regression equation. That is, the turbidity, lagged turbidity, and discharge values from the 10-percent group were input to the 90-percent subsampled regression equation. The estimated SSC results and associated *SSE* were compared between the 90- and 10-percent datasets. For NF Toutle-SRS, 307 observations were used for the 90-percent regression and 34 observations were used for comparison. For Toutle-Tower, 320 observations were used for the 90-percent regression and 35 observations were used for the 10-percent comparison (appendix B).

OLS regressions were run on the 90-percent subsampled data for each of the five models for each gaging station. In using the 90-percent subsampled data for each model, two means of comparison were used. First, the coefficient estimates and standard errors were compared to their full data counterparts for consistency. Second, SSC was estimated and *SSE* was calculated using the 10-percent sequestered data. The model with the smallest *SSE* and most consistent estimates was considered the best model. If the 90-percent OLS estimates were grossly different and (or) had different signs from the full dataset OLS, then this model would not be the best to use.

Across models, although there was some deviation in the magnitude of the lagged turbidity estimate, the positive or negative sign remained the same and estimates for turbidity and discharge fluctuated within reason. Testing the subsample models on the 10-percent sequestered data showed that the ideal model using turbidity, lagged turbidity, and discharge (number 4 in the explanatory variable list) as explanatory variables had the lowest out-of-sample *SSE*. These results support the use of the final model and coefficient estimates.

Final Regression Models

Regression Model Coefficients

The log-log regression model analysis of SSC (response variable) with turbidity, turbidity-lag, and discharge (predictor variables) for the NF Toutle-SRS and Toutle-Tower gaging stations provided the output shown in tables 4 and 5. The coefficients are used to generate the regression equations listed as equations 1 and 2. The ANOVA statistics in tables 2 and 3 apply to these equations.

Table 4. Regression coefficients for North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington.

Parameter	Coefficient	SE	t-statistic	P-value	VIF
logT	0.1854	0.2882	0.64	0.52	322.304
logT lag	0.3545	0.2897	1.22	0.221	321.817
logQ	0.89518	0.04497	19.91	0	1.601
Constant	-0.8054	0.1135	-7.10	0	

$$\log_t SSC = -0.8054 + 0.1854 \log_t T + 0.3545 \log_t Tlag + 0.8952 \log_t Q \quad (1)$$

where

T	is turbidity,
Q	is discharge,
$Tlag$	is lag turbidity value for the previous 15-minute period, and
t	is the 15-minute interval time.

Table 5. Regression coefficients for Toutle River at Tower Road near Silver Lake, Washington..

Parameter	Coefficient	SE	t-statistic	P-value	VIF
logT	0.5676	0.1456	3.90	0	115.711
logT lag	0.1612	0.1449	1.11	0.266	112.942
logQ	0.9101	0.03587	25.37	0	3.149
Constant	-1.99049	0.09096	-21.88	0	

$$\log_t SSC = -1.9905 + 0.5676 \log_t T + 0.1612 \log_t Tlag + 0.9101 \log_t Q \quad (2)$$

Selecting the Predictor Variables for Model

Using the coefficients from the $\log_t SSC$ to $\log_t T$, $\log_t Tlag$, and $\log_t Q$ regression model, the unlogged or untransformed final equations became:

NF Toutle-SRS:
Equation (1) is converted to power form as equation 3,
$$SSC_t = 0.156531 * T_t^{0.1854} * Tlag_t^{0.3545} * Q_t^{0.8952} \quad (3)$$

Toutle-Tower:
Equation (2) is converted to power form as equation 4,
$$SSC_t = 0.010221 * T_t^{0.5676} * Tlag_t^{0.1612} * Q_t^{0.9101} \quad (4)$$

Applying the Bias Correction Factor

Because regressions were conducted on log-transformed variables, a bias was introduced that distorts the estimated SSC when the log values are converted back to their original linear form. Duan's smearing bias correction factor (BCF) was computed using the average of the unlogged residuals, as a best estimate of this introduced bias (Helsel and Hirsch, 2002; Rasmussen and others, 2009; Uhrich and Bragg, 2003). The BCF result for each station is computed as:

Bias Correction Factor (BCF): $\frac{\sum_i^n 10^{r_i}}{N} = 1.1491573 \; and \; 1.14909,$

for **NF Toutle-SRS** and **Toutle-Tower**, respectively, and where r = logged residual values.

The right side of the regressions (equations 3 and 4) then are multiplied by the BCF to obtain the final equation:

NF Toutle-SRS:

$$SSC_t = 0.179879 * T_t^{0.1854} * Tlag_t^{0.3545} * Q_t^{0.8952} \text{ (5)}$$

Toutle-Tower:

$$SSC_t = 0.011745 * T_t^{0.5676} * Tlag_t^{0.1612} * Q_t^{0.9101} \text{ (6)}$$

Equations 5 and 6 are considered the general regression analysis (GRA) in this report and can be used normally to estimate SSC, with no further derivations.

The BCF accounts only for model error with no corrections for sample error, or error arising when estimating regression coefficients from a more finite dataset. That is, if one wanted to calculate the daily mean turbidity and averaged just three 15-minute values for that day, the sample error would be higher than if the mean turbidity was averaged using all ninety-six 15-minute values available for that day. Hence, larger sample sets, such as those used in this analysis, will tend to have a lower or negligible sample error. Although the BCF for model error increases SSC, the sample error correction has the inverse effect. Smaller sample sets without a sample error correction tend to overestimate the SSC. Because sample error was negligible in this analysis, no correction was applied.

Applying the Regression Models to Future Data

As new turbidity and discharge data are collected, they can be added to the original 15-minute turbidity and discharge datasets or kept separate as their own unique dataset. This distinction depends on the new GRA assembled from the additional SSC samples, which are paired with a turbidity and discharge value at the specific time of each sample. Analysis of covariance or ANCOVA can be used to test the significance of the original regression against future data added to the dataset. This would help determine if a change in the turbidity-SSC relation warrants developing a model for the new dataset (Helsel and Hirsch, 2002, p. 316). Rasmussen and others (2009) suggest that each water year be worked separately, and that the data from that water year then be compared to the data from the previous water year. If there is no significant difference in the slope and y-intercept between water years, the data could be joined together to refine the model and to generate a single multi-water year GRA. As a potential benefit, the refined model may have a lower *SSE* and reduced prediction interval. If the difference in regression models is significant, then a new GRA equation must be developed, using the methods described in this section, for the additional water year and (or) period of record. The new GRA equation then would be used until the analysis is reiterated using data from subsequent water years.

Final Regression Model Graphs

Graphs of the logSSC (measured) against logSSC (estimated) from final GRA equations for both gaging stations are shown in figures 8 and 9. The OLS lines in figures 8 and 9 represent the GRA relation defined by equations 5 and 6. The 95-percent prediction and confidence intervals are shown in figures 8 and 9, as provided by the statistical package used (*www.minitab.com*). A prediction interval is always wider than a confidence interval because it must account for both the uncertainty of the population mean and data scatter, also described as the model and sampling uncertainty. The distinction is that prediction intervals provide information on the distribution of values and not the uncertainty in determining the population mean, whereas confidence intervals provide information on how well the population mean was determined. The key point here is that confidence intervals provide information on the true population parameter, whereas prediction intervals represent ranges of values within which there is a 95-percent certainty (in this case) that the true population (SSC) occurs.
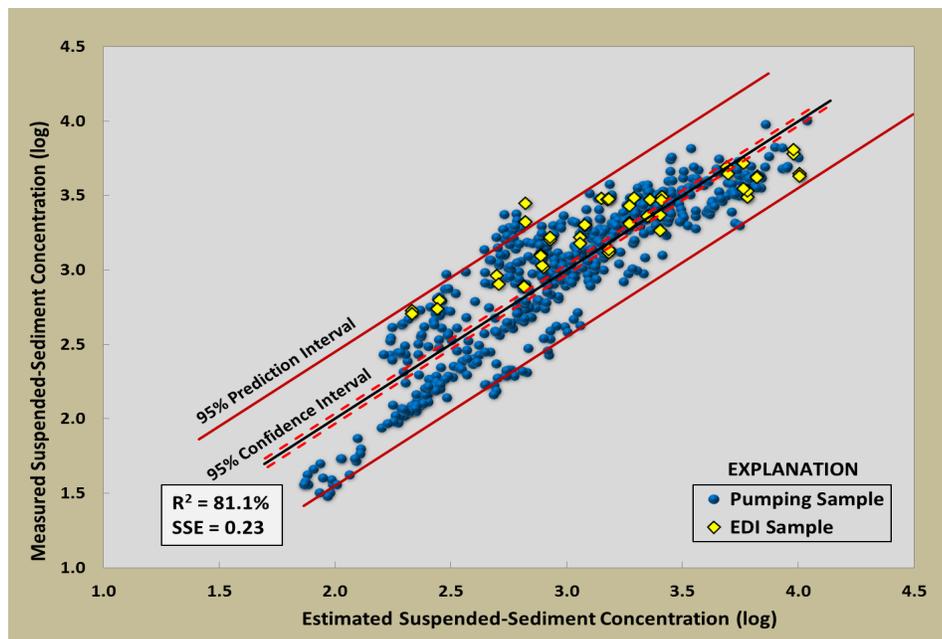
Figure 8. Final multiple linear regression model showing the general regression analysis line (equation 5) superimposed over measured and estimated suspended-sediment concentrations for pump and equal discharge increment samples, for North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington, water years 2010–11. Graph also shows 95-percent prediction and confidence intervals.
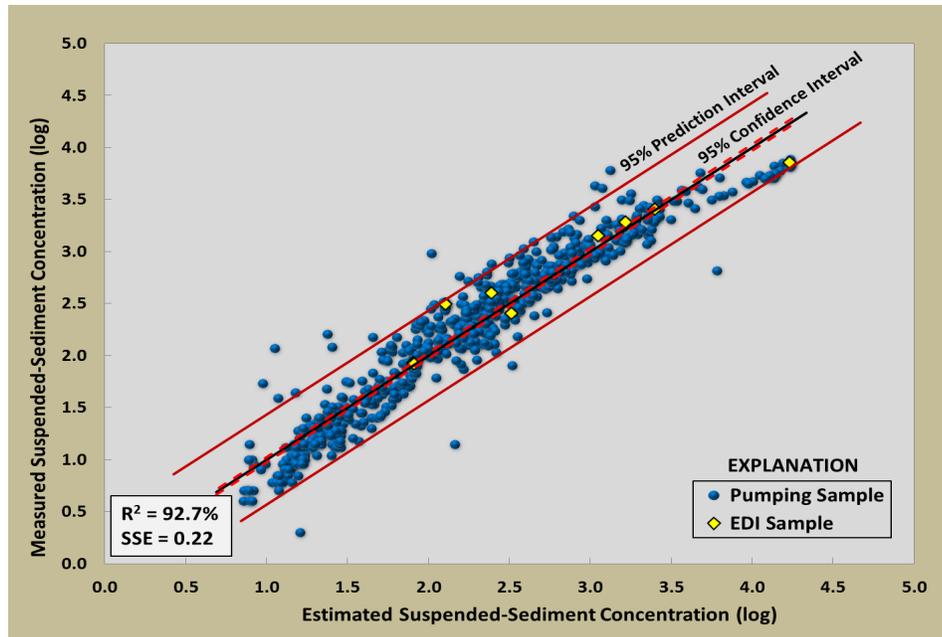
Figure 9. Final multiple linear regression model showing the general regression analysis line (equation 6) superimposed over measured and estimated suspended-sediment concentrations for pump and equal discharge increment samples, for Toutle River at Tower Road near Silver Lake, Washington, water years 2010–11. Graph also shows 95-percent prediction and confidence intervals.

Discussion and Future Studies

The use of surrogates for high-density measurements in real time offers many opportunities for improved understanding of hydrologic processes, along with well-characterized and reduced uncertainty, and ultimately better informed decision-making. In this study, we used turbidity as a surrogate for SSC in the sand-dominated Toutle River Basin; as a proof-of-concept approach to evaluate the feasibility of improving estimates of sediment loading and transport in the drainage basin; and possibly to reduce costs, compared to historical, manual techniques. The results of the study indicate that the potential for such improvements is high, with relatively robust regressions developed at both the NF-Toutle-SRS and Toutle-Tower sites. Although beyond the scope of this report, use of these regressions, together with discharge data from the two gaging stations, could be used to calculate 15-minute and daily concentrations and loads from the WYs 2010–11 dataset. The calculations also could be extended through WY 2012, with each computation being a relatively straightforward exercise. Future refinement and other uses of these regression techniques, beyond calculation of concentration and load, could provide additional information for understanding changes over time in sediment sources, transport, and deposition. Additionally, there remain some limitations and criteria to the regressions obtained in this report and to the overall use of surrogate technologies, which must be considered when using these results for decision-making.

Appropriate Uses of Turbidity-SSC Surrogate Regressions

Development of turbidity-SSC regressions are not conventionally universal across all water systems and riverine environments. The models developed herein can be used solely for the Toutle River Basin and cannot be transferred to other drainage basins. In addition, some waterways do not lend themselves to this type of analysis because of variability in the sediment-water matrix, as well as unacceptable monitoring conditions. From a monitoring standpoint, the turbidity-SSC surrogate regressions generally assume a consistent amount of light scattering by particles in transport over the range of the regression data. However, sediment grain-size distributions usually change during events or by season, based on the energy of the stream and the sediment sources, which can add uncertainty to the regression-based estimates of concentration or load. Based on past sediment events, it may be advantageous to subdivide the data by seasonal time frames or by increasing or decreasing streamflow and (or) turbidity components. This might produce a suite of regression models that could be used in conjunction with each other, each invoked by assessing in real time the sequential changes in turbidity and (or) discharge to determine which model to use, and thereby improve the estimates of sediment concentration.

By using these refined models, potential future work could compare results from the single regression model developed in this report to such a combined seasonal or event model approach. The combined approach likely would provide a tighter fit with a near-zero covariance between the residuals. The Toutle River Basin is a complex fluvial system that, upon further analysis, might lend itself to this type of event-based sediment-transport regime. Seasonal or event models may better estimate sediment-transport events that are unrelated to streamflow, which show up as tailings outside the confidence interval of the single regression line. Examples include volcanic- or glacial-influenced events from the Mount St. Helens crater, as well as landslides or localized streambank sloughing. These types of studies would provide insight into how the Mount St. Helens sediment-source terrain and depositional areas evolve over time, along with insight into management of sedimentation in the lower Toutle River Basin.

It also would be informative to test the comparability, cross-sectional representation, and cost effectiveness of a sampling regime that emphasizes more pump samples, as used for this study, compared to one composed of a greater number of manual EDIs. Such an evaluation could bolster the cost-effectiveness and usability of the data-collection program by assuring samples would be collected at the appropriate time and frequency. Additionally, the autosample sediment-size mixture of coarse and fine sediment can differ greatly from the EDIs and EWIs because of various pipe-hose lengths and configurations, hydraulic head required to pump and disperse the sample, and variable stream velocity and bed movement near the autosampler intake. Many of the pumping samples collected and used in this study were targeted to capture the full range of suspended sediment during peak discharge and turbidity events, when manual samples could not be collected. These samples provided valuable confirmation at critical sediment flux periods that otherwise would not have been possible. Regardless of this, if the surrogate-regression based approach is used, the number of pump samples collected and analyzed in the future could be decreased without significantly increasing manual EDI samples.

As a reasonable next step in processing these data, future work could include SSC as an online near-real-time parameter, using the regressions shown in this report with the ongoing continuous turbidity values. SSC could be added to the parameters of turbidity, stage, and discharge for each station, and also could be used as a comparison to the previous sample-based, sediment-record results.

As mentioned in the section, "Turbidity Greater than Instrument Limits," when the turbidity sensors are at their flat-line threshold, the data are not used in turbidity-discharge to SSC model development, as well as in any continuous SSC estimation. This is the most critical limitation to this study, as sediment transport is highest during these episodes and, therefore, is most vital in quantifying the sediment flux. Future work could include estimates through these peak periods or use alternative high-end turbidity sensors to provide a complete record of model-estimated SSC and SSQ. However, the high-end turbidity sensors would need to have their own instrument- and site-specific regressions generated. This prerequisite is owing to differences in light scattering and detection between a high-end sensor and the DTS-12 sensor™ used in this study (Rasmussen and others, 2009). One such high-end sensor initially was deployed at the NF Toutle-SRS site in 2011; therefore, a dataset with paired SSC sample results already is available, and can be used as the starting point from which to begin this work.

Finally, no inferences were drawn with respect to the sediment-size data. All manual samples and many of the autosamples include size-fraction data; however, none of these data were taken into account for this study. Regression models could be constructed for the individual sand/silt fraction, such that concentration and load for coarse- or fine-grain sizes could be determined separately. Additional work could use the size-fraction data to suggest source areas and to develop a synopsis of how specific areas have eroded and evolved over time, as well as to estimate the volumes of different size classes transported downstream past the NF Toutle-SRS gaging station to the main-stem Toutle and Cowlitz Rivers.

Updating Existing Regressions

The regression models in this report use data only from April 2010 to September 2011, as the time frame and scope for this work coincided with WYs 2010–11 approved and published turbidity and SSC records. The regression models and equations can easily be applied to or updated to include later water years. Inclusion of additional manual and pumping samples, the data for which already are available for WYs 2012–13, would better define the turbidity-discharge to SSC relation and improve the regression development and structure. By periodically evaluating the latest, finalized turbidity and discharge data, by water year, major changes in the sediment-transport system could be documented.

Trends and Use of State-Space Models

Sediment flux in the Toutle River Basin at both gaging stations responds to regional hydrology, but also responds to localized events and patterns. Specific erosional events from the Mount St. Helens crater and debris avalanche, and areas directly upstream of the SRS have all caused spikes and anomalies that are outside the typical sediment-transport pattern. These types of events can produce a hysteresis or differential pattern between sediment concentration and turbidity or discharge over varying parts of the event hydrograph. These patterns could reveal source or process information that, with closer evaluation, could be used to more effectively understand and manage sediment transport throughout the Toutle River Basin. The debris-avalanche deposit and braided channels formed through the entire valley, upstream and downstream of the SRS, also have implications for other environmental factors, such as fish survival and migration, along with the health and restoration of other aquatic species and habitats. Additional explanatory variables that weigh supplementary factors (such as seasonality, specific events, antecedent conditions, water temperature, and other water-quality parameters) could be incorporated in the model to help understand these wide-ranging ecological conditions.

In working within the 30-day autocorrelation die-off period, if SSC sample results, including laboratory analysis and database entry, could be routinely performed on a more real-time, continuous basis, such that SSC values were provided in less than 30 days, then autocorrelation modeling with a time-series component would be relevant to the results and should be applied. Realistically, however, most processing of sediment samples takes more than 30 days to generate an SSC value. One benefit of more real-time SSC data would be improved event-based estimation. Additionally, understanding and correcting for time-series properties of SSC would be most useful when interpolating between missing values of observed SSC. However, to apply these types of time-series corrections would require all SSC samples to be in an even time-step (Jones, 1986); although adjustments could be made using SSMs to alleviate this concern. Other time-series components, such as an ARMA model, also might be required, along with smoothing techniques to estimate intermediate values of SSC, such as using a Kalman filter in a SSM.

Such sophisticated techniques as ARMA processing and SSMs, if employed, would better simulate the trends in observed SSC by incorporating seasonality and rise/fall hysteresis variables. One possible parameter to better define rise/fall dynamics in fluvial constituent studies is use of the square of streamflow (Cohn and others, 1992). Particularly powerful are SSMs that use dynamic optimization techniques to define the best "path" through a deterministic or stochastic dataset. One such data-fusion procedure is a Kalman filter, which works by smoothing linear data and then estimating missing SSC in a feed-forward and feed-back manner by minimizing the mean square error of the estimated SSC (Maybeck, 1979). For instance, noisy, erratic data could be smoothed and estimates made in past, present, and future states. The Kalman filter works much like GCLAS by melding the observed sampled SSCs with estimated SSCs and interpolating missing values, although the two methods have their distinctions. GCLAS by its design is a human-based, more time-dependent interactive process, whereas SSM with a Kalman filter can be entirely automated. The distinct advantages of the SSM method are its reproducibility and reduced processing time, as well as the ability to estimate error metrics of the interpolated values. Thus, the more frequently samples are collected, the less the error estimate. After turbidity and discharge records are available, a sediment discharge record could be generated automatically with a defined uncertainty.

High-End Turbidity Sensor

As mentioned in the "Turbidity Greater Than Instrument Limits" section, a high-end turbidity sensor capable of monitoring suspended-sediment at levels at least one order of magnitude higher than the current turbidity sensor is in operation, on a trial basis, at NF Toutle-SRS. Future work could include this high-end sensor as part of the normal turbidity calibration and records-processing work, which could be published as a second turbidity parameter. Separate regressions for the high-end sensor also would need to be developed. See figure 10 for comparison of instream DTS-12 sensor[™] and high-end sensor readings.
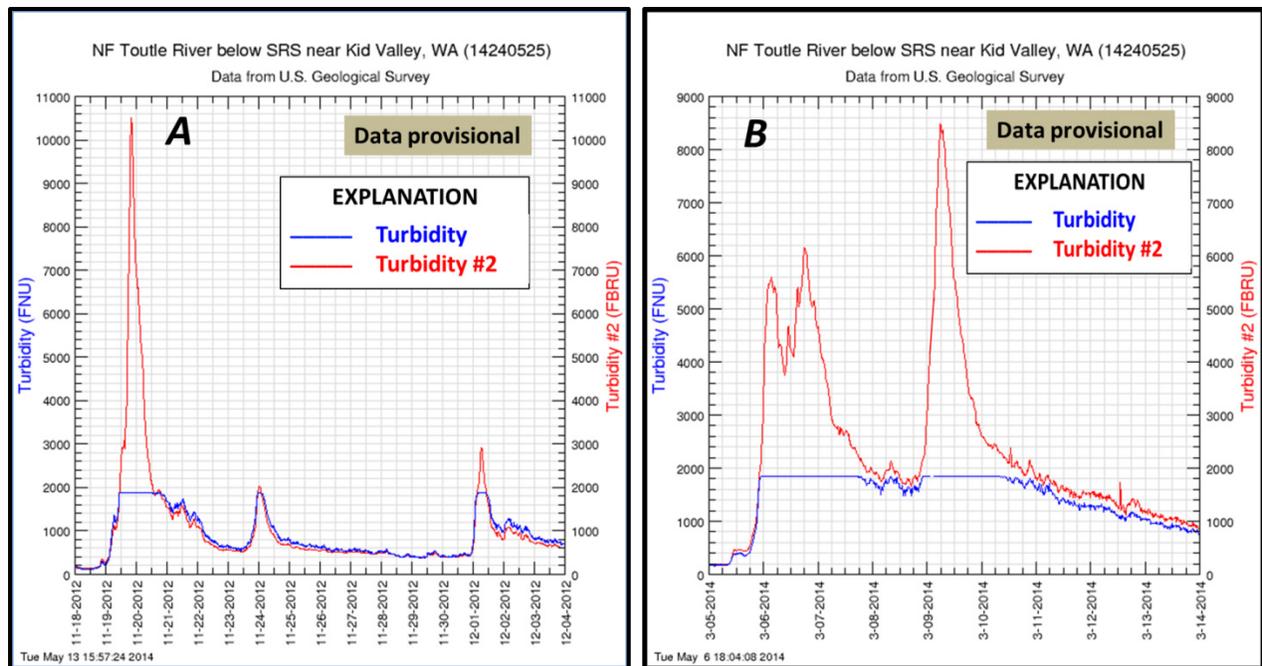
Figure 10. Graphs showing turbidity at North Fork Toutle River below Sediment Retention Structure near Kid Valley, Toutle River Basin, Washington, 2012 and 2014. Graph in blue shows how *NF Toutle-SRS* turbidity, for (*A*) November 18–December 4, 2012, and (*B*) March 5–March 14, 2014, reached the sensor maximum at near 2000 FNU. Graph in red shows, for the same time period and scale, how a high-end sensor (Turbidity #2) recorded turbidity (in formazin backscatter ratio units, FBRU) beyond the threshold level in blue.

Expected Effects of Raising SRS-Spillway

The SRS spillway was raised in elevation by 7 ft in September–October 2012, (back cover photograph; U.S. Army Corps of Engineers, 2012). The effects of this higher spillway on sediment transport and downstream channel morphology are not yet quantified. Additional analysis could integrate future turbidity and streamflow data into the established regression model, and also serve as a contrast to previous turbidity-discharge to SSC relations. For example, data directly preceding and following the spillway construction, such as data for WYs 2012 and 2013, could be compared. Any change in this relation would help to define and quantify new trends in sediment transport affected and (or) caused by this spillway raise. Similarly, future longer-term modifications to the SRS could be evaluated for any changes to the turbidity-SSC, turbidity-streamflow, and turbidity plus streamflow-SSC relation. Finally, the spillway raise may have affected the sediment-size fraction transported downstream past the SRS; one possible effect would be that relatively more coarse sediment is retained upstream of the SRS, with relatively more fine material transported downstream. Existing data on both size fractions and the nature of the turbidity-SSC regression could shed light on the degree to which these changes have occurred. The suspended-sediment loads could be computed with the percentage of certain size classes quantified by volume and compared year-to-year.

25

Conclusions

Despite the more than 30 years since the eruption of Mount St. Helens, sediment management in the Toutle River remains a daunting task. To help improve estimates of sediment transport and to reduce costs, the feasibility of instream turbidity measurement as a surrogate for suspended-sediment concentration (SSC) was tested in the Toutle River Basin. The results indicate that turbidity can be used reliably to augment the existing SSC sample collection, and possibly to improve the final estimates, as well as to reduce future data collection costs. The Toutle River at Tower Road (Toutle-Tower) near Silver Lake, Washington, and North Fork Toutle River below Sediment Retention Structure (NF Toutle-SRS) near Kid Valley, Washington, gaging stations each had sensors installed and data collected for the periods April 1 and May 1, 2010, through September 30, 2011, respectively. Multiple linear regression models using ordinary least square methods were generated and equations were provided for both gaging stations that use the instream turbidity and discharge data to enable prediction of real-time SSC. The equations to calculate SSC were corrected for bias using a smearing estimator.

The turbidity-SSC regressions were relatively successful, and could be improved in the future by employing sensors that have a higher maximum range. The use of pump samplers also could be optimized by finding a balance between cost savings from their unattended sampling capabilities, and the uncertainty they introduce. Uncertainty from pump samplers occurred because the sample represented a single point rather than a cross section, and the large number (as used herein) contributed to autocorrelation. Scheduling manual equal-discharge-increment sampling for times providing the most desirable and broadest range in streamflow and turbidity levels also could help to streamline the data-collection program.

In addition to the regression statistics, other tests and improvement measures were applied, such as the Durbin-Watson statistic to test for serial correlation and the use of lagged turbidity and discharge variables. The regression with the best supportive diagnostic statistics and best fit of the explanatory and response variables, along with minimal serial correlation, was selected as the final model and equation. The final regression equation used logged values of turbidity, discharge, and a single 15-minute lag of turbidity as explanatory variables in estimating SSC.

The dataset used in this study was confined to roughly 1.5 years of turbidity and SSC; however, additional years of data were made available after this data analysis was underway. Future water years could readily be added to better define and fine-tune these correlations. Sediment-size data were not used in this analysis, which prevented any inferences regarding sediment transport of various size fractions. Future models could be constructed for separate fine- or coarse-grain sediment transport.

Despite these limitations, the proof of concept described in the initial study objectives has shown that, even in a high sand-transport environment, rugged in-stream turbidity instrumentation, robust measuring technology, and appropriate statistical modeling methods may produce a more efficient and less costly alternative to conventional sample-based, sediment-record methods currently (2014) in use. More sophisticated statistical analysis would be useful for this dataset and future Toutle-Tower and NF Toutle-SRS datasets, as this would broaden the understanding of turbidity-discharge to SSC correlations by incorporating seasonality, trends, and rise/fall hysteresis terms. Future use of an Auto-Regressive Moving Average component and State-Space Models using a Kalman filter also would automate sediment-discharge computations, deliver reproducibility, and provide an error measurement of the load estimate.

Acknowledgments

References Cited

Anderson, C.W., 2005, Turbidity (ver. 2.1): U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. A6, section 6.7, 64 p., accessed May 6, 2014, at http://water.usgs.gov/owq/FieldManual/Chapter6/6.7_contents.html.

Bent, G.C., Gray, J.R., Smith, K.P., and Glysson, G.D., 2000, A synopsis of technical issues for monitoring sediment in highway and urban runoff: U.S. Geological Survey Open-File Report 00-497, accessed May 6, 2014, at *http://pubs.usgs.gov/of/2000/ofr00-497/*.

Box, G.E.P., and Jenkins, G.M., 1976, Time series analysis—forecasting and control (revised edition): San Francisco, California, Holden-Day, 575 p.

Buchanan, T.J., and Somers, W.P., 1976, Discharge measurements at gaging stations: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. A8, 65 p., accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri3a8/*.

Cohn, T., 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. 2,353–2,364.

Dinehart, R.L., 1998, Sediment transport at gaging stations near Mount St. Helens, Washington, 1980–90—Data collection and analysis: U.S. Geological Survey Professional Paper 1573, 111 p., accessed May 6, 2014, at http://pubs.usgs.gov/pp/1573/.

Edwards, T.K., and Glysson, G.D., 1999, Field methods for measurement of fluvial sediment: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. C2, 89 p., accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri3-c2/*.

Gibson, S., Nygaard, C., and Sclafani, P., 2010, Mobile bed modeling of the Cowlitz River using HEC-RAS—Assessing flooding risk and impact due to system sediment: U.S. Army Corps of Engineers, Portland District, 2nd Joint Federal Interagency Conference, Las Vegas, Nevada, June 27–July 1, 2010, accessed May 6, 2014, at *http://acwi.gov/sos/pubs/2ndJFIC/Contents/6C_Gibson_Nygaard_Modeling_of_Cowlitz_paper.pdf*.

Glysson, G.D., 2008, Guidelines and standards procedures for the use of automatic samplers for the collection of surface water-quality and sediment data: National Water Quality Monitoring Council, 6th National Monitoring Conference, Session C4, Atlantic City, New Jersey, May 18–22, 2008, accessed May 6, 2014, at *http://acwi.gov/monitoring/conference/2008/papers/C4-B.pdf*.

Guy, H.P., 1977, Laboratory theory and methods for sediment analysis: U.S. Geological Survey Techniques of Water-Resources Investigations, book 5, chap. C1, accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri5c1/html/pdf.html*.

Guy, H.P., 1978, Fluvial sediment concepts: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. C1, p. accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri3-c1/*.

Harvey, A.C., 1989, Forecasting, structural time series models and the kalman filter: New York, Cambridge University Press, p. 554.

Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. C3, p. 221–263, accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri4a3/*.

Jones, R.H., 1986, Time series regression with unequally spaced data: Journal of Applied Probability, v. 23A, Special Volume—Essays in Times Series and Allied Processes, p. 89–98.

Koltun, G.F., Eberle, Michael, Gray, J.R., and Glysson, G.D, 2006, User's manual for the Graphical Constituent Loading Analysis System (GCLAS): U.S. Geological Survey Techniques and Methods, book 4, chap. C1, 51 p., accessed May 6, 2014, at *http://pubs.usgs.gov/tm/2006/tm4C1/pdf/tm4C1.pdf*.

Major, J.J., and Spicer, K.R., 2003, Clearing Toutle River sediment issue—An op-ed opinion: The Columbian, August 9, accessed May 6, 2014, at *http://vulcan.wr.usgs.gov/Projects/JJMajor/Publications/TheColumbian/clearing_toutle_sediment_issue_aug9_03.html*.

Maybeck, P.S., 1979, Stochastic models, estimation and control—Volume 1, Chapter 1: New York, Academic Press, p. 3–16.

Porterfield, G., 1977, Computation of fluvial-sediment discharge: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. C3; accessed May 6, 2014, at *http://pubs.usgs.gov/twri/twri3-c3/html/pdf.html*.

Rantz, S.E., and others, 1983, Measurement and computation of streamflow—Volume 1, Measurement of stage and discharge—Volume 2, Computation of discharge: U.S. Geological Survey Water Supply Paper 2175, 631 p., accessed May 6, 2014, at *http://pubs.usgs.gov/wsp/wsp2175*.

Rasmussen, P.P., Gray, J.R., Glysson, G.D., and Ziegler, A.C., 2009, Guidelines and procedures for computing time-series suspended-sediment concentrations and loads from in-stream turbidity-sensor and streamflow data: U.S. Geological Survey Techniques and Methods, book 3, chap. 4, 53 p., accessed May 6, 2014, at *http://pubs.usgs.gov/tm/tm3c4/*.

Sauer, V.B., and Meyer, R.W., 1992, Determination of error in individual discharge measurements: U.S. Geological Survey Open-File Report 92–144, 21 p., accessed May 6, 2014, at *http://pubs.usgs.gov/of/1992/ofr92-144/*.

Topping, D.J., Rubin, D.M., Wright, S.A., and Melis, T.S., 2011, Field evaluation of the error arising from inadequate time averaging in the standard use of depth-integrating suspended-sediment samplers: U.S. Geological Survey Professional Paper, 95 p., accessed May 6, 2014, at http://pubs.usgs.gov/pp/1774/.

Uhrich, M.A., and Bragg, H.M., 2003, Monitoring instream turbidity to estimate continuous suspended-sediment loads and yields and clay-water volumes in the upper North Santiam River Basin, Oregon, 1998–2000: U.S. Geological Survey Water-Resources Investigations Report 03-4098, accessed May 6, 2014, at, *http://pubs.usgs.gov/wri/WRI03-4098/*.

U.S. Army Corps of Engineers, 2012, Final environmental assessment—Sediment Retention Structure (SRS) spillway raise project, Mount St. Helens sediment management for flood risk reduction: U.S. Army Corps of Engineers, Portland District, accessed May 6, 2014, at *http://www.nwp.usace.army.mil/Portals/24/docs/locations/msh/Final_EA_SRS_Spillway_Raise.pdf*.

U.S. Geological Survey, 2010, Water-resources data for the United States, water year 2010: U.S. Geological Survey Water-Data Report WDR-US-2010, accessed May 6, 2014, at *http://wdr.water.usgs.gov/wy2010/pdfs/14240525.2010.pdf*, and *http://wdr.water.usgs.gov/wy2010/pdfs/14242580.2010.pdf*.

U.S. Geological Survey, 2011, Water-resources data for the United States, water year 2011: U.S. Geological Survey Water-Data Report WDR-US-2011, accessed May 6, 2014, at *http://wdr.water.usgs.gov/wy2011/pdfs/14240525.2011.pdf*, and *http://wdr.water.usgs.gov/wy2011/pdfs/14242580.2011.pdf*.

Wagner, R.J., Boulger, R.W., Jr., Oblinger, C.J., and Smith, B.A., 2006, Guidelines and standard procedure for continuous water-quality monitors—Station operation, record computation, and data reporting: U.S. Geological Survey Techniques and Methods, book 1, chap. D3, 51 p., plus 8 attachments, accessed May 6, 2014, at *http://pubs.water.usgs.gov/tm1d3*.

Appendix A. Suspended-Sediment Sample, Discharge, and Turbidity Data

Suspended-sediment sample data are presented for the discharge and turbidity monitoring gaging stations North Fork Toutle River below Sediment Retention Structure near Kid Valley, Washington (NF Toutle-SRS; U.S. Geological Survey [USGS] gaging station No. 14240525) and Toutle River at Tower Road near Silver Lake, Washington (Toutle-Tower; USGS gaging station No. 14242580). The data were used for regression model calibration and consist of the total suspended-sediment concentration and the corresponding discharge and turbidity recorded at the monitoring station during sample collection. When available, sample concentrations with fine suspended-sediment data are presented for particle sizes less than 0.0625 millimeters. Separate bar plot worksheets also are included that depict the number of samples collected at each gaging station. Each bar plot is differentiated by sample type as Equal-Discharge-Increment or pump sample, and by nine different ranges in turbidity that span the entire sensor measurement scale.

[Appendix A is a Microsoft$^©$ Excel file and can be downloaded at *http://pubs.usgs.gov/ofr/2014/1204/.*]

Appendix B. Robustness Check Data

Data used in checking for robustness is presented as two separate worksheets for each station in an Excel spreadsheet. Included are the sampled suspended-sediment concentration (SSC) data, and adjoining discharge (Q), turbidity (T) and lagged turbidity (T-lag) values for the 90- and 10-percent subsampled and sequestered out-of-sample data, respectively. Samples were randomly selected for the 90-percent group. One sample was randomly selected for days with more than one sample. The remaining samples were used for the 10-percent sequestered data. For NF Toutle-SRS, 653 samples were reduced to 307 and 34 samples for the 90- and 10-percent groups, respectively. For Toutle-Tower, 705 samples were reduced to 320 and 35 samples for the 90- and 10-percent groups, respectively.

[Appendix B is a Microsoft$^©$ Excel file and can be downloaded at *http://pubs.usgs.gov/ofr/2014/1204/.*]