Appendix 10. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07179750, Neosho River at Burlingame Road near Emporia, Kansas, during January 1, 2010, through December 31, 2019

This model archive summary summarizes the suspended-sediment concentration (SSC) model developed to compute hourly or daily SSC during January 1, 2010, through December 31, 2019. This model is used concomitantly with other models during this period to calculate concentrations when other explanatory variables are not available for the purposes of load and concentration model calculations. The methods used follow U.S. Geological Survey (USGS) guidance as referenced in relevant Office of Surface Water/Office of Water Quality Technical Memoranda and USGS Techniques and Methods, book 3, chapter C4 (Rasmussen and others, 2009; U.S. Geological Survey, 2016), and other standard USGS methods (Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010).

Site and Model Information

Site number: 07179750
Site name: Neosho River at Burlingame Road near Emporia, Kansas
Location: Lat 38°25'43", long 96°09'29" referenced to North American Datum of 1927, in NE 1/4 SW 1/4 NE 1/4 sec.02, T.19 S., R.11 E., Lyon County, Kans., hydrologic unit 11070201, on left downstream side of bridge at Burlingame Road, 1.5 miles north of Interstate Highway 35, and at mile 391.8.

Equipment: Sutron Satlink data collection platform (DCP), Design Analysis H–350XL pressure transducer, and DA H–355 gas purge system. Orifice/bubble line extends from shelter and runs under bridge to enter stream on left downstream side of bridge. A wire-weight gage is located on the downstream side bridge handrail. The reference gage is the wire-weight gage attached to the downstream handrail on the bridge. Check-bar elevation is 38.621 feet.

Date model was developed: January 16, 2020
Model calibration data period: August 12, 2009, through March 14, 2019

Model Data

All data were collected using USGS protocols (Wagner and others, 2006; Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010; U.S. Geological Survey, variously dated) and are stored in the National Water Information System (NWIS) database

(https://doi.org/10.5066/F7P55KJN; U.S. Geological Survey, 2020). Explanatory variables were evaluated individually and in combination. Potential explanatory variables included streamflow, water temperature, specific conductance, and turbidity. Seasonal components were also evaluated as explanatory variables.

The regression model is based on 47 concomitant values of discretely measured SSC samples and continuously measured streamflow during August 12, 2009, through March 14, 2019. Discrete samples were collected over a range of streamflows. No samples had concentrations below laboratory detection limits. Identification of potential outliers included any values that exceeded the Cook's D test (Cook, 1977) and any point for which the studentized residual was greater than 3 or less than −3. None of the samples in this dataset were deemed outliers or removed from the model calibration dataset.

Suspended-Sediment Sampling Details

Discrete samples were collected from the downstream side of the bridge or instream within 350 feet of the bridge using equal-width-increment, multiple vertical, single vertical, or grab-dip methods following U.S. Geological Survey (2006) and Rasmussen and others (2014). Discrete samples were collected on a semifixed to event-based schedule ranging from four to nine samples per year with a Federal Interagency Sedimentation Project U.S. D–77, DH–95, or D–95 with a Teflon bottle, cap, and nozzle depth-integrating sampler, a D–96 bag sampler, a DH–81 with a Teflon bottle, cap, and nozzle hand sampler, DH–48, or a grab sample with a Teflon bottle depending on sample location. Samples were analyzed for SSC, loss on ignition, and occasionally five-point grain size by the USGS Sediment Laboratory in Iowa City, Iowa.

Continuous Data

Streamflow was measured using a nonsubmersible pressure transducer during August 12, 2009, through March 14, 2019 (U.S. Geological Survey, 2018). The continuous streamflow data used for model calibration were time interpolated values from the continuous time series. If the continuous data were not available, the sample was not included in the dataset. The range of continuous streamflow data (in cubic feet per second) was as follows: maximum 17,400; minimum 0.445; mean 335; median 33.5.

Model Development

Ordinary least squares regression analysis was done using R programming language (R Core Team, 2019) to relate discretely collected SSC to streamflow and other continuously measured data. The distribution of residuals was examined for normality and plots of residuals (the difference between the measured and model calculated values) compared to model calculated SSC were examined for homoscedasticity (departures from zero did not change substantially over the range of model calculated values). Previously published explanatory

variables were also strongly considered for continuity however, the best explanatory variable(s) was ultimately selected.

When the SSC and turbidity model could not be applied the streamflow, model was selected instead as a good predictor of logarithm base 10 ($\log_{10}$) (SSC) based on residual plots, relatively high coefficient of determination (R^2), and relatively low model standard percentage error (MSPE).

Model Summary

Summary of final SSC regression analysis at site 07179750:
SSC-based model:

$$\text{Log}_{10}(SSC) = 0.707 \times \text{Log}_{10}(Q) + 0.533$$

where

SSC = suspended-sediment concentration, in milligrams per liter, and
Q = streamflow, in cubic feet per second.

The log-transformed model may be retransformed to the original units to calculate SSC directly. A bias is introduced in the calculated constituent during retransformation and may be corrected using the Duan's bias correction factor (BCF; Duan, 1983). The calculated BCF is 1.32 for this model and the formula for the retransformed model accounting for BCF is as follows:

$$SSC = 4.50 \times Q^{0.707}$$

Previously Published Model

$$\text{Log}_{10}(SSC) = 0.744 \times \text{Log}_{10}(Q) + 0.50$$

Model author: Foster (2014)
Model data period: June 17, 2009, through September 27, 2012
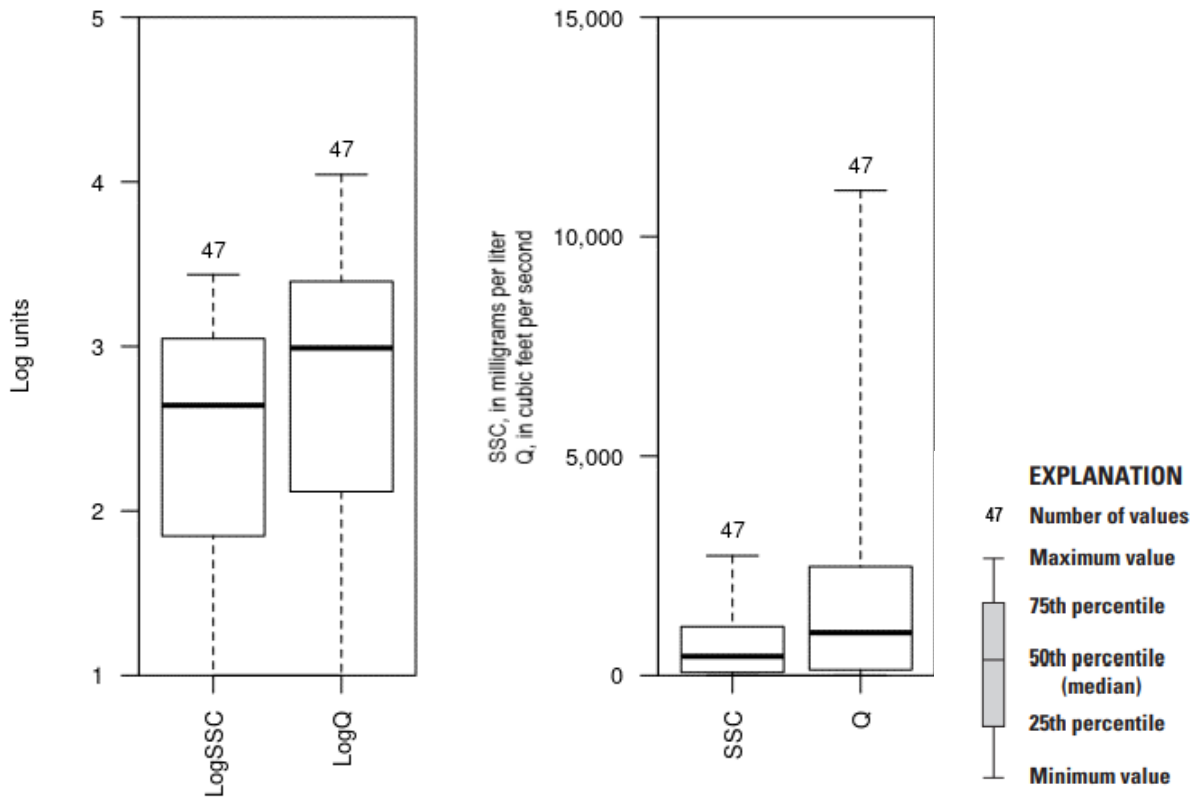
Model Statistics, Data, and Plots

Model

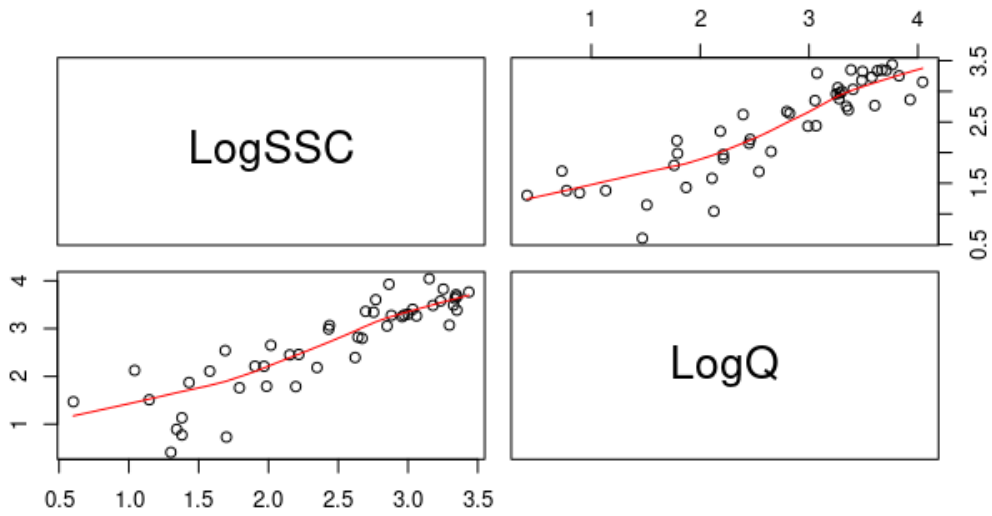$$\text{Log}(SSC) = +\ 0.707 * \text{Log}(Q) + 0.533$$

Variable Summary Statistics

	LogSSC	SSC	LogQ	Q
Minimum	0.602	4	0.41	2.57
1st Quartile	1.790	62	2.11	128.00
Median	2.640	438	2.99	978.00
Mean	2.430	730	2.68	1860.00
3d Quartile	3.060	1150	3.40	2530.00
Maximum	3.440	2730	4.04	11100.00

Box Plots

Exploratory Plots

Basic Model Statistics

```
Number of Observations                                 47
Standard error (RMSE)                               0.364
Average Model standard percentage error (MSPE)   93.8
Coefficient of determination (R²)                   0.778
Adjusted Coefficient of Determination (Adj. R²) 0.773
Bias Correction Factor (BCF)                        1.32
```

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.533	0.1600	3.34	1.71e-03
LogQ	0.707	0.0563	12.60	2.58e-16

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.943
E.vars	-0.943	1.000

Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.128	0.194	0.413

Flagged Observations

	LogSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
12/7/2010 12:30	0.602	1.570	-0.970	-2.75	-2.98	0.0563	0.225	-0.727
9/24/2015 10:50	1.300	0.823	0.478	1.42	1.44	0.1450	0.171	0.591
1/26/2017 14:20	1.040	2.040	-0.996	-2.78	-3.02	0.0286	0.114	-0.518
6/28/2018 10:00	1.700	1.050	0.650	1.90	1.96	0.1120	0.227	0.695

Statistical Plots
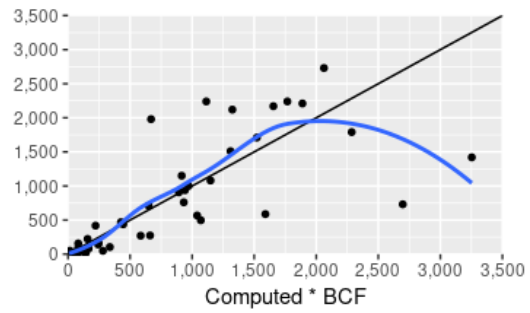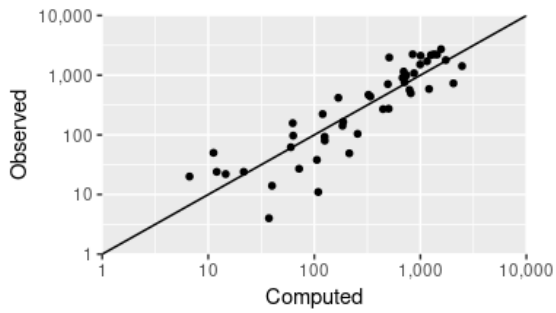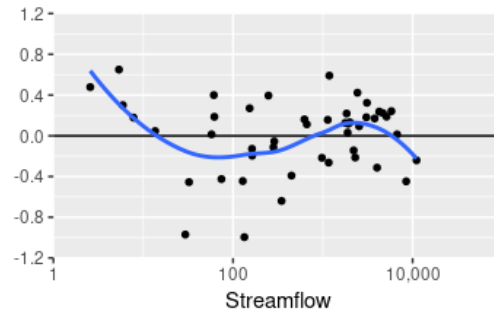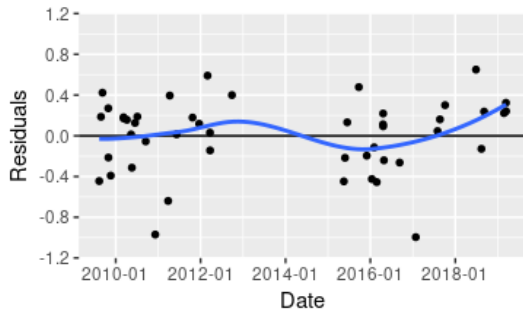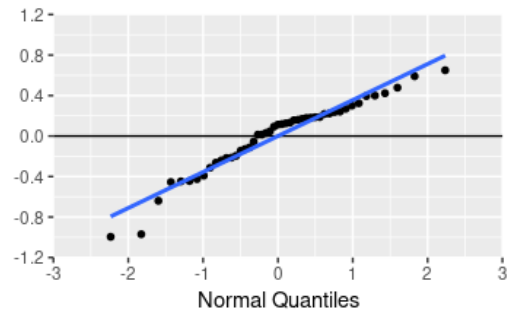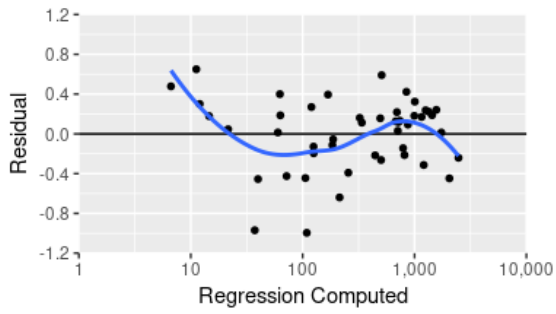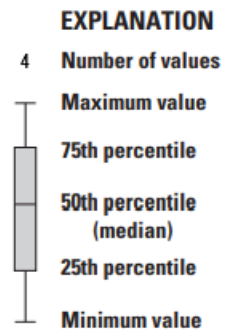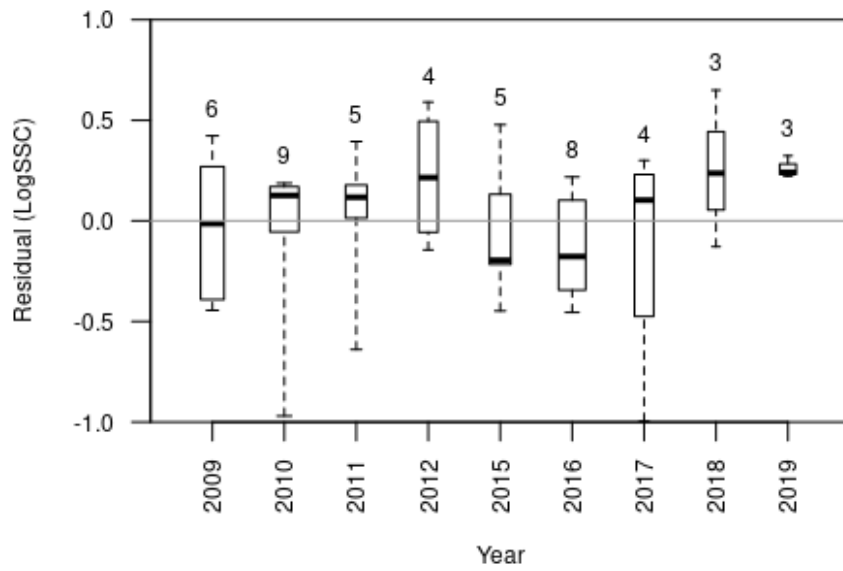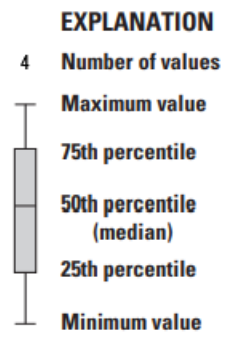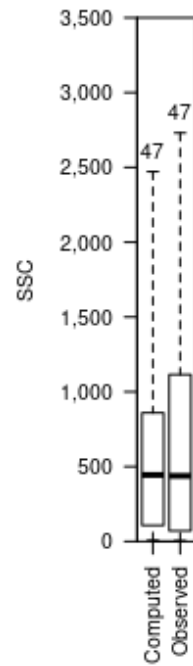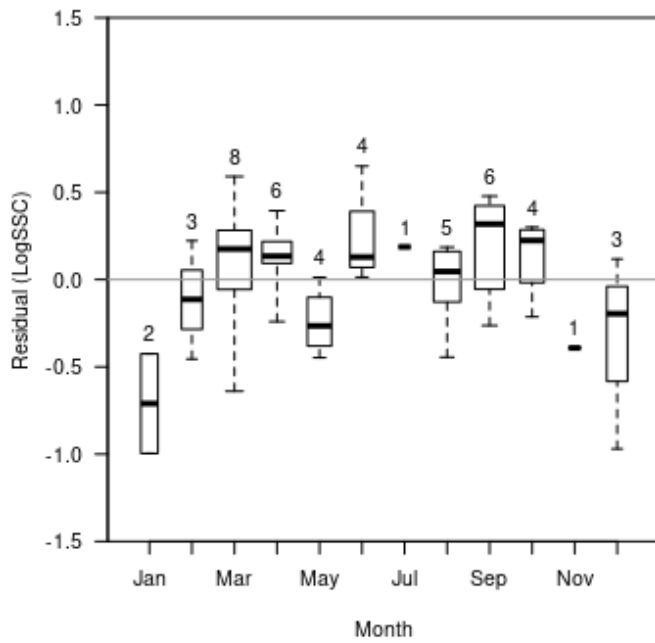
Cross Validation
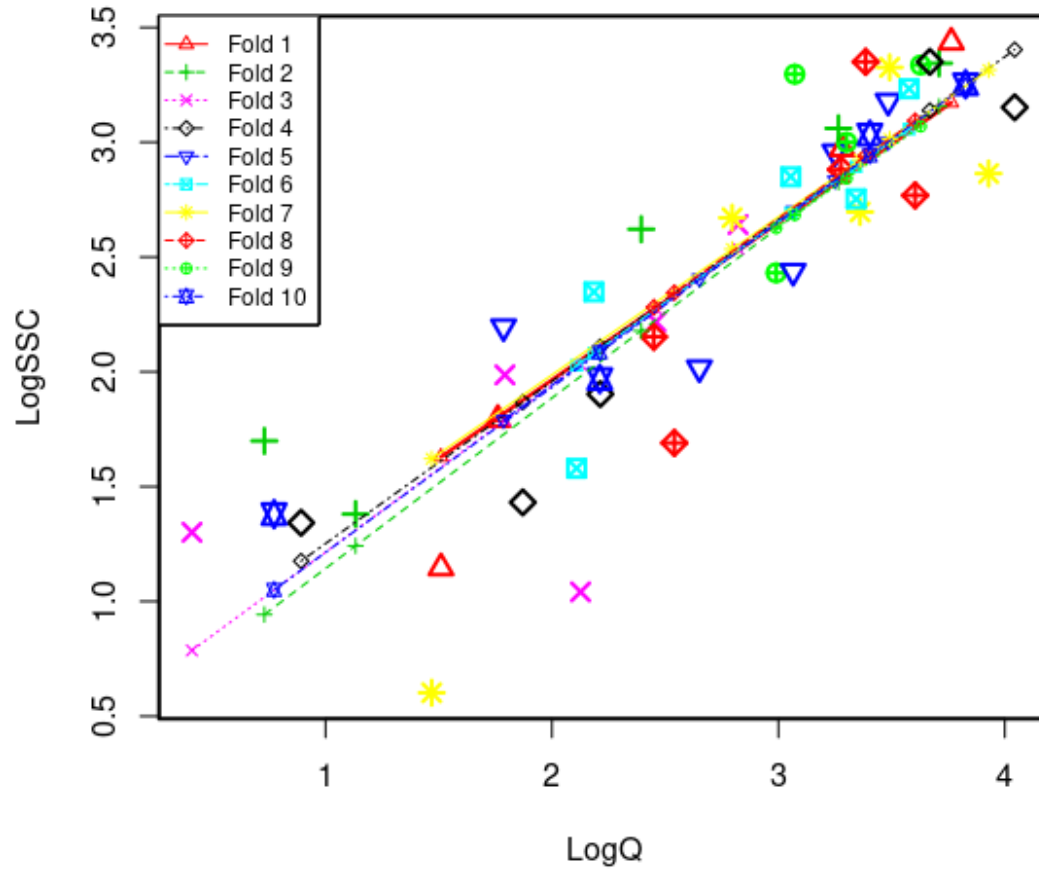

```
Minimum mean squared error (MSE) of folds:  0.0327
                      Mean MSE of folds:  0.1330
                    Median MSE of folds:  0.1070
                   Maximum MSE of folds:  0.2820
     (Mean MSE of folds) / (Model MSE):  1.0100
```

MSE of folds

EXPLANATION

Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Dataset

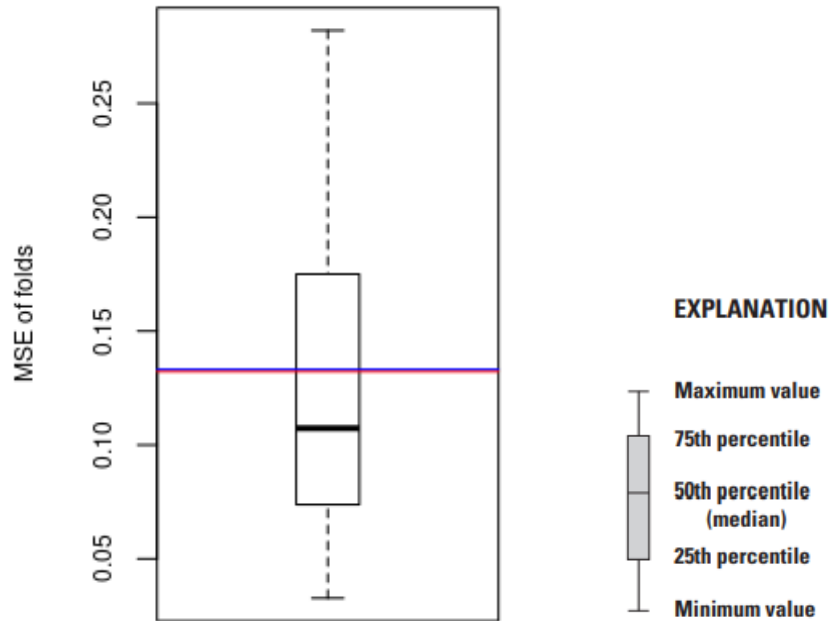|	Date	LogSSC	LogQ	SSC	Q	Computed LogSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
0										
1	2009-08-12	1.58	2.11	38	128	2.02	139	-0.445	-1.18	--
2	2009-08-27	1.99	1.79	97	62	1.8	83.2	0.186	0.498	--
3	2009-09-09	3.35	3.39	2240	2430	2.93	1110	0.423	1.43	--
4	2009-10-29	2.35	2.18	223	153	2.08	158	0.27	0.907	--
5	2009-10-30	2.7	3.36	497	2290	2.91	1070	-0.214	-0.622	--
6	2009-11-19	2.02	2.65	104	448	2.41	337	-0.391	-0.99	--
7	2010-03-09	3.18	3.48	1510	3050	3	1310	0.181	0.439	--
8	2010-03-11	3.23	3.58	1710	3770	3.06	1520	0.17	0.324	--
9	2010-04-07	2.85	3.05	708	1130	2.69	650	0.156	0.214	--
10	2010-05-13	3.25	3.83	1790	6710	3.24	2290	0.0132	-0.269	--
11	2010-05-20	2.77	3.6	587	4020	3.08	1590	-0.313	-0.907	--
12	2010-06-16	2.96	3.25	907	1780	2.83	893	0.126	0.106	--
13	2010-07-06	3.34	3.71	2210	5120	3.16	1890	0.188	0.559	--
14	2010-09-16	2.22	2.46	165	288	2.27	246	-0.055	-0.324	--
15	2010-12-07	0.602	1.47	4	29.4	1.57	49.1	-0.97	-1.83	--
16	2011-03-28	1.69	2.54	49	347	2.33	281	-0.64	-1.6	--
17	2011-04-11	2.62	2.39	418	248	2.23	222	0.395	1.18	--
18	2011-06-09	1.79	1.76	62	57.7	1.78	79	0.0139	-0.214	--
19	2011-10-24	1.34	0.892	22	7.8	1.16	19.2	0.178	0.381	--

20	2011-12-20	2.97	3.28	938	1920	2.85	942	0.117	0.0531	--
21	2012-03-01	3.3	3.07	1980	1180	2.71	669	0.59	1.83	--
22	2012-03-22	2.88	3.28	760	1890	2.85	933	0.0304	-0.16	--
23	2012-03-23	2.75	3.34	567	2200	2.9	1040	-0.144	-0.498	--
24	2012-09-27	2.2	1.79	157	61.1	1.8	82.4	0.399	1.3	--
25	2015-05-18	2.86	3.93	731	8480	3.31	2700	-0.448	-1.3	--
26	2015-06-16	3	3.3	1000	2000	2.87	971	0.132	0.16	--
27	2015-05-28	2.43	2.99	270	978	2.65	585	-0.217	-0.688	--
28	2015-09-24	1.3	0.41	20	2.57	0.823	8.76	0.478	1.6	--
29	2015-12-01	1.9	2.21	80	163	2.1	165	-0.195	-0.559	--
30	2016-01-14	1.43	1.87	27	74.3	1.86	94.6	-0.425	-1.08	--
31	2016-02-03	2.15	2.45	142	282	2.27	243	-0.114	-0.381	--
32	2016-02-25	1.15	1.51	14	32.4	1.6	52.5	-0.455	-1.43	--
33	2016-04-19	2.64	2.82	438	663	2.53	445	0.113	0	--
34	2016-04-20	3.06	3.27	1150	1840	2.84	916	0.218	0.622	--
35	2016-04-20	3.03	3.4	1080	2530	2.94	1150	0.0928	-0.0531	--
36	2016-04-27	3.15	4.04	1420	11100	3.39	3250	-0.241	-0.757	--
37	2016-09-09	2.44	3.07	274	1160	2.7	661	-0.263	-0.83	--
38	2017-01-26	1.04	2.13	11	134	2.04	143	-0.996	-2.23	--
39	2017-08-22	2.67	2.8	469	625	2.51	426	0.161	0.269	--
40	2017-08-01	1.38	1.13	24	13.6	1.33	28.4	0.046	-0.106	--
41	2017-10-05	1.38	0.773	24	5.93	1.08	15.8	0.301	0.99	--
42	2018-06-28	1.7	0.73	50	5.37	1.05	14.7	0.65	2.23	--
43	2018-08-14	1.97	2.21	93	163	2.1	165	-0.129	-0.439	--
44	2018-09-06	3.34	3.63	2170	4250	3.1	1650	0.237	0.757	--
45	2019-02-24	3.35	3.67	2240	4660	3.13	1770	0.222	0.688	--
46	2019-03-13	3.33	3.49	2120	3100	3	1320	0.324	1.08	--
47	2019-03-14	3.44	3.76	2730	5800	3.19	2060	0.241	0.83	--

Definitions

Adj R^2: Adjusted coefficient of determination

BCF: Bias correction factor

DFFITS: Studentized difference in fits

Log: logarithm base 10

MSE: Mean squared error

MSPE: Model standard percentage error

R^2: Coefficient of determination

RMSE: Root mean square error

```
SSC: Suspended-sediment concentration, in milligrams per liter (80154)

Q: Streamflow, mean daily, in cubic feet per second (00060)
```

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

References Cited

Cook, R.D., 1977, Detection of influential observations in linear regression: Technometrics, v. 19, no. 1, p. 15–18. [Also available at https://doi.org/10.2307/1268249.]

Duan, N., 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610. [Also available at https://doi.org/10.1080/01621459.1983.10478017.]

Foster, G.M., 2014, Relations between continuous real-time turbidity data and discrete suspended-sediment concentration samples in the Neosho and Cottonwood Rivers, east-central Kansas, 2009–2012: U.S. Geological Survey Open-File Report 2014–1171, 20 p., accessed September 2019 at https://doi.org/10.3133/ofr20141171.

R Core Team, 2019, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed September 2019 at https://www.Rproject.org/.

Rasmussen, T.J., Bennett, T.J., Stone, M.L., Foster, G.M., Graham, J.L., and Putnam, J.E., 2014, Quality-assurance and data-management plan for water-quality activities in the Kansas Water Science Center, 2014: U.S. Geological Survey Open-File Report 2014–1233, 41 p., accessed September 2019 at https://doi.org/10.3133/ofr20141233.

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. C4, 54 p. [Also available at https://doi.org/10.3133/tm3c4.]

Sauer, V.B., and Turnipseed, D.P., 2010, Stage measurement at gaging stations: U.S. Geological Survey Techniques and Methods, book 3, chap. A7, 45 p., accessed September 2019 at https://doi.org/10.3133/tm3A7.

Turnipseed, D.P., and Sauer, V.B., 2010, Discharge measurements at gaging stations: U.S. Geological Survey Techniques and Methods, book 3, chap. A8, 87 p., accessed September 2019 at https://doi.org/10.3133/tm3A8.

U.S. Geological Survey, 2006, Collection of water samples (ver. 2.0): U.S. Geological Survey Techniques of Water Resources Investigations, book 9, chap. A4 [variously paged]. [Also available at https://pubs.water.usgs.gov/ twri9A4/.]

U.S. Geological Survey, 2016, Policy and guidance for approval of surrogate regression models for computation of time series suspended-sediment concentrations and loads: U.S. Geological Survey, Office of Water Quality Technical Memorandum 2016.10, 40 p., accessed November 20, 2020, at https://water.usgs.gov/admin/memo/SW/sw.2016.07+wq.2016.10.pdf.

U.S. Geological Survey, 2018, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed January 2017 at https://doi.org/10.5066/F7P55KJN.

U.S. Geological Survey, 2020, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed January 2020, at https://doi.org/10.5066/F7P55KJN.

U.S. Geological Survey, variously dated, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1–A9 [variously paged]. [Also available at https://water.usgs.gov/owq/FieldManual/.]

Wagner, R.J., Boulger, R.W., Jr., Oblinger, C.J., and Smith, B.A., 2006, Guidelines and standard procedures 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. [Also available at https://doi.org/10.3133/tm1D3.]