

# **Appendix 9. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07182250, Cottonwood River near Plymouth, 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: 07182250

Site name: Cottonwood River near Plymouth, Kansas

Location: Lat 38°23'51", long 96°21'21" referenced to North American Datum of 1927, in NE 1/4 NE 1/4 SE 1/4 sec.13, T.19 S., R.9 E., Chase County, Kans., hydrologic unit 11070203, on right bank at upstream side of county highway bridge, 0.8 mile (mi) downstream from Buckeye Creek, 1.5 mi southwest of Plymouth, and at mile 39.2.

Equipment: Design Analysis H350 pressure transducer/data logger, Design Analysis H355 bubbler, and Sutron Satlink High Data Rate transmitter housed in a 3x4x1.5-foot (ft) steel shelter attached to the right upstream wingwall of the county highway bridge. Primary data is from the DCP interfaced with the H350. Backup data is from the DCP interfaced with the FTS radar. Air is bubbled through black plastic tubing enclosed in metal conduit pipe attached to the underside of bridge to top of first pier from the right abutment, which is in 1 1/4-inch galvanized steel pipe anchored to the downstream side of the pier to stream where connection is made to a 1 1/4-inch open end orifice. The reference gage is a Type-A wire-weight gage attached to the downstream handrail of the bridge. Check-bar elevation is 42.050 ft.

Date model was developed: January 16, 2020

Model calibration data period: March 26, 2007, through March 13, 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 concurrent measurements of discretely collected SSC samples and continuously measured streamflow during March 26, 2007, through March 13, 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 400 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 2 to 15 samples per year with a Federal Interagency Sedimentation Project U.S. DH-48, DH-59, DH-75P, or a D-95 with a Teflon bottle, cap, and nozzle depth-integrating sampler 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 December 31, 2006, through March 15, 2019 (U.S. Geological Survey, 2018). The continuous streamflow data used 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 34,700; minimum 14.8; mean 866; median 237.

## 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).

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 07182250:  
SSC-based model:

$$\text{Log}_{10}(\text{SSC}) = 0.617 \times \text{Log}_{10}(Q) + 0.732$$

where

$\text{SSC}$  = 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 1.36 for this model and the formula for the retransformed model accounting for BCF is as follows:

$$\text{SSC} = 7.34 \times Q^{0.617}$$

## Previously Published Model

$$\text{Log}_{10}(\text{SSC}) = 0.742 \times \text{Log}_{10}(Q) + 0.21$$

Model author: Foster (2014)

Model data period: June 17, 2009, through September 27, 2012

## Model Statistics, Data, and Plots

### Model

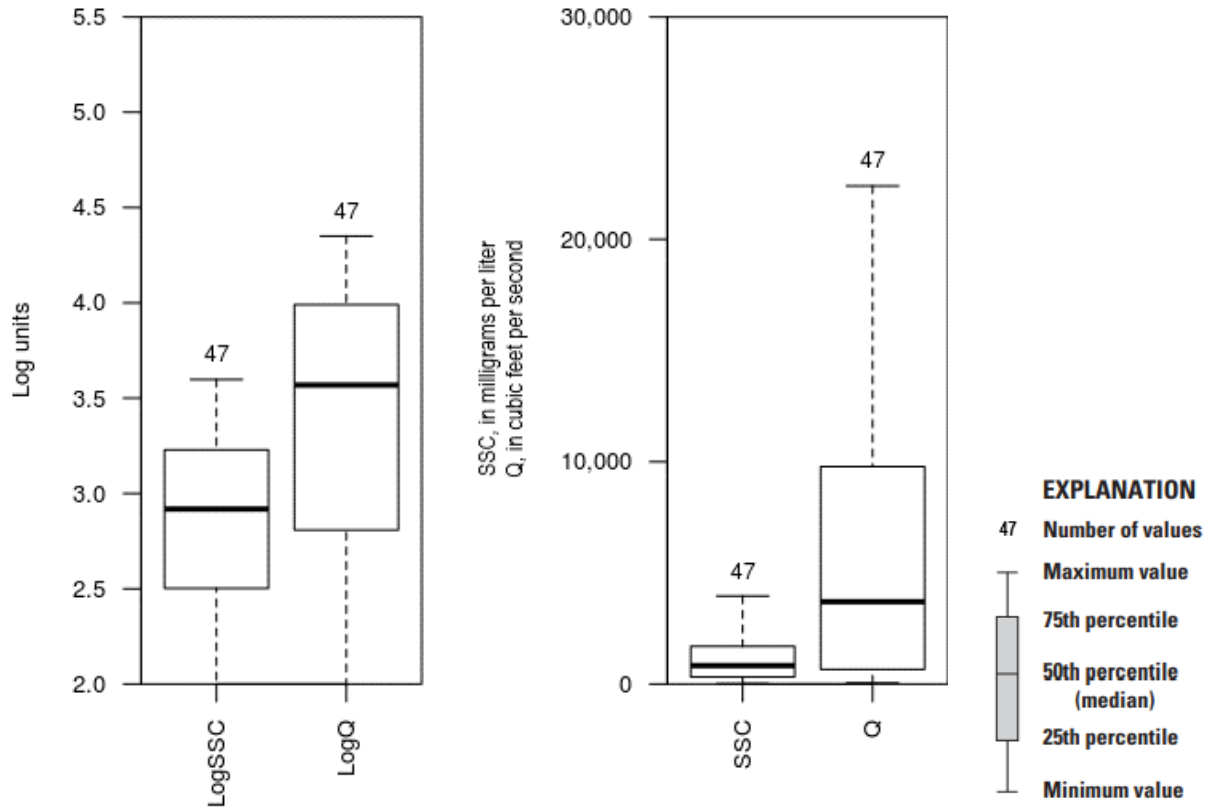
$$\text{Log}(\text{SSC}) = + 0.617 * \text{Log}(Q) + 0.732$$

### Variable Summary Statistics

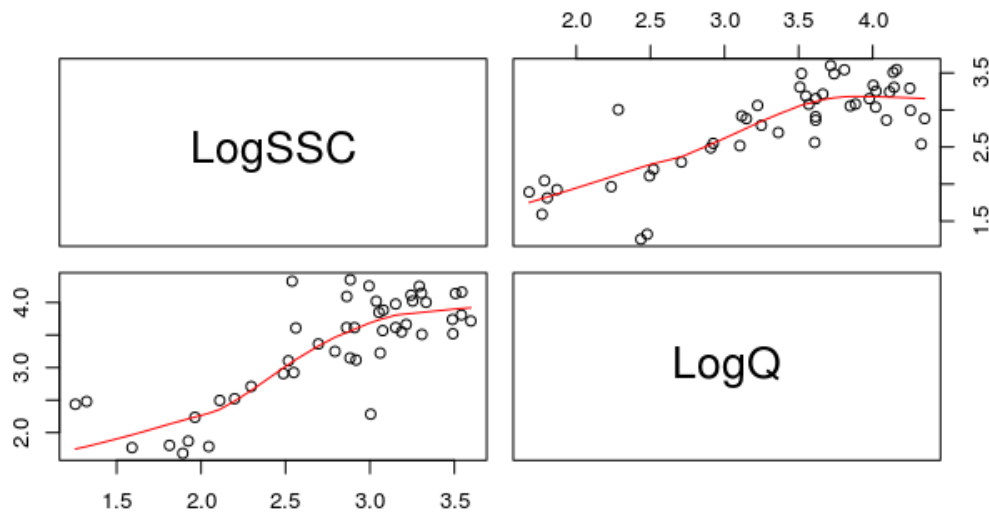
	LogSSC	SSC	LogQ	Q
Minimum	1.26	18	1.68	48
1st Quartile	2.49	307	2.71	512
Median	2.92	828	3.57	3700
Mean	2.78	1150	3.32	5700

3d Quartile	3.24	1750	4.00	10100
Maximum	3.60	3960	4.35	22400

### Box Plots



## Exploratory Plots



## Basic Model Statistics

Number of Observations	47
Standard error (RMSE)	0.369
Average Model standard percentage error (MSPE)	95.5
Coefficient of determination ( $R^2$ )	0.64
Adjusted Coefficient of Determination (Adj. $R^2$ )	0.632
Bias Correction Factor (BCF)	1.36

## Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t )
(Intercept)	0.732	0.236	3.11	3.28e-03
LogQ	0.617	0.069	8.94	1.54e-11

## Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.974
E.vars	-0.974	1.000

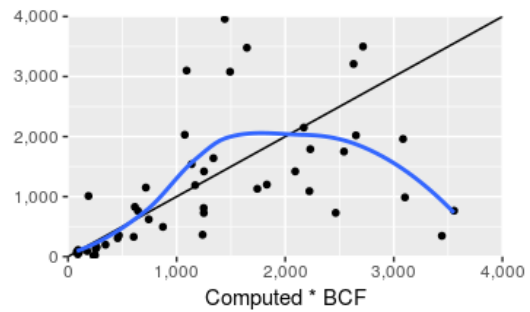
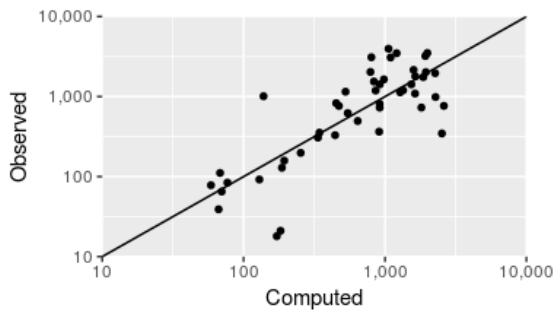
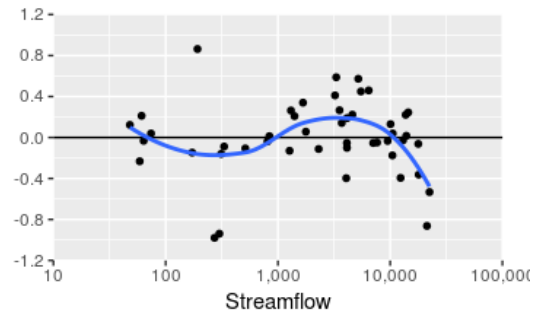
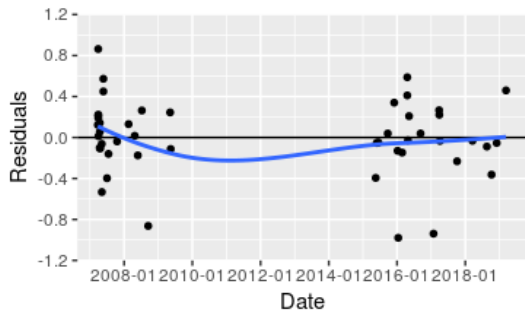
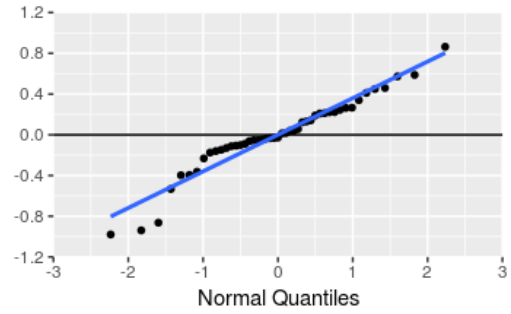
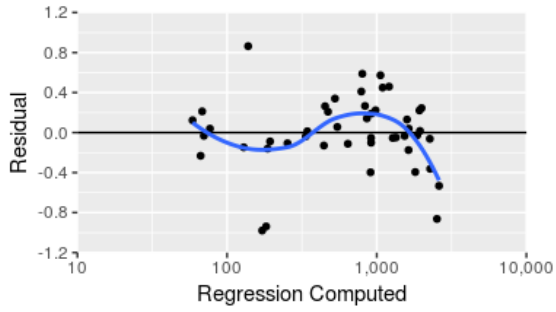
## Outlier Test Criteria

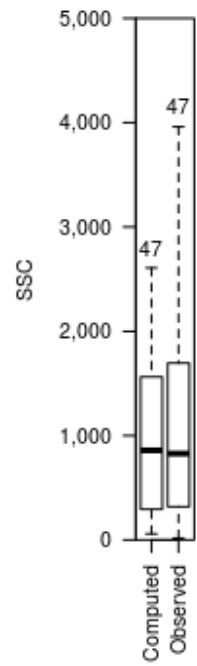
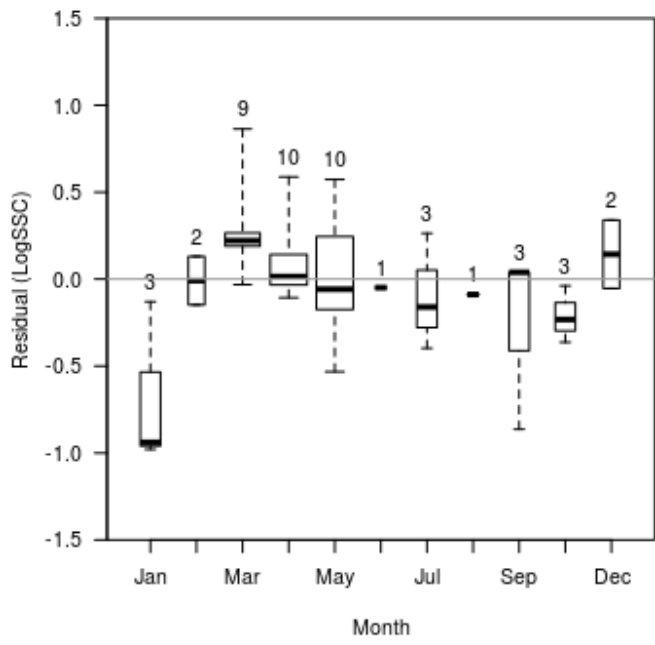
Leverage	Cook's D	DFFITS
0.128	0.194	0.413

## Flagged Observations

	LogSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
3/30/2007 18:00	3.00	2.14	0.864	2.41	2.56	0.0592	0.183	0.642
9/15/2008 16:00	2.54	3.40	-0.863	-2.41	-2.55	0.0566	0.174	-0.625
1/14/2016 14:50	1.26	2.23	-0.979	-2.72	-2.95	0.0489	0.191	-0.668
1/26/2017 15:20	1.32	2.26	-0.939	-2.61	-2.80	0.0463	0.165	-0.616

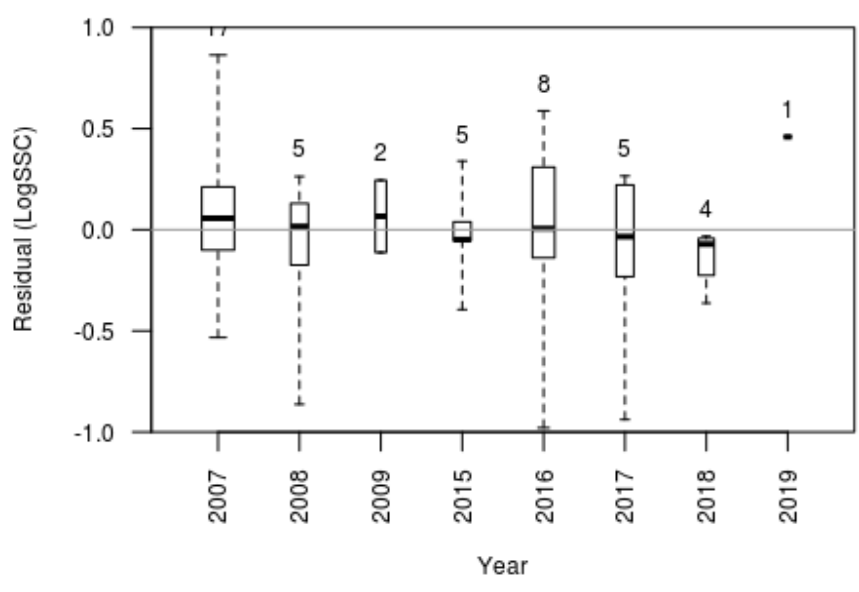
## Statistical Plots





**EXPLANATION**

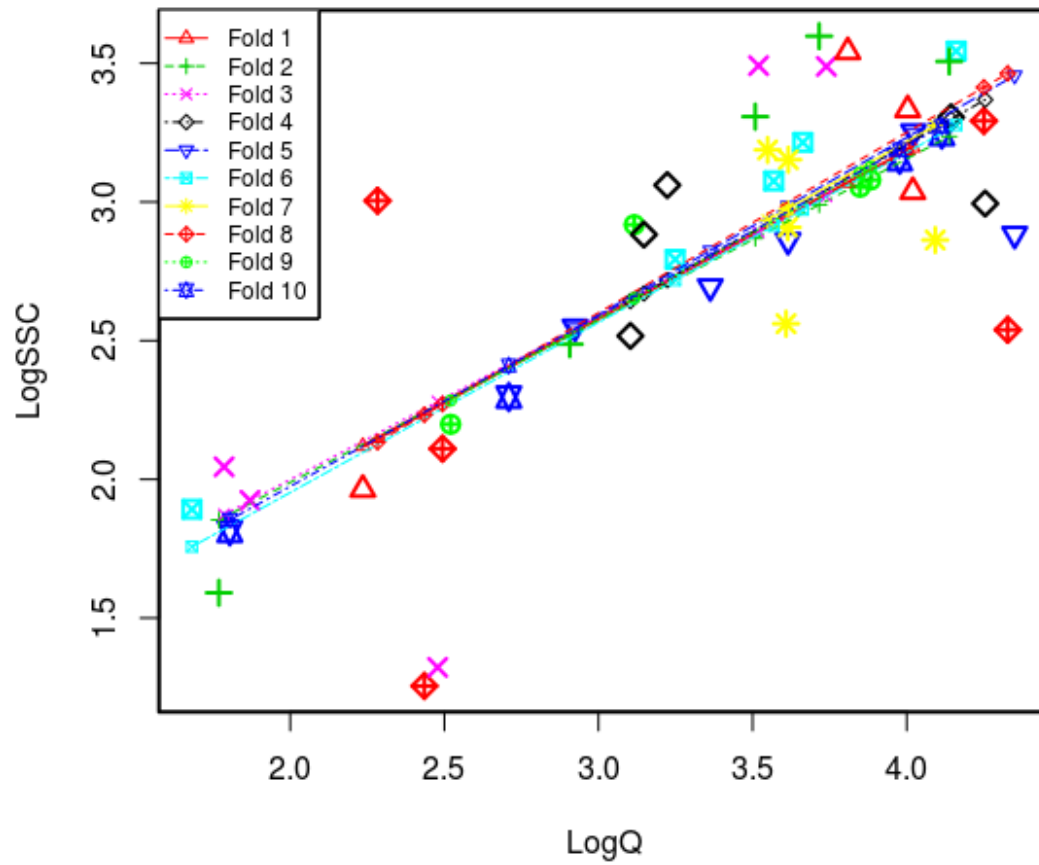
- 2 Number of values
- Maximum value
- 75th percentile
- 50th percentile (median)
- 25th percentile
- Minimum value



**EXPLANATION**

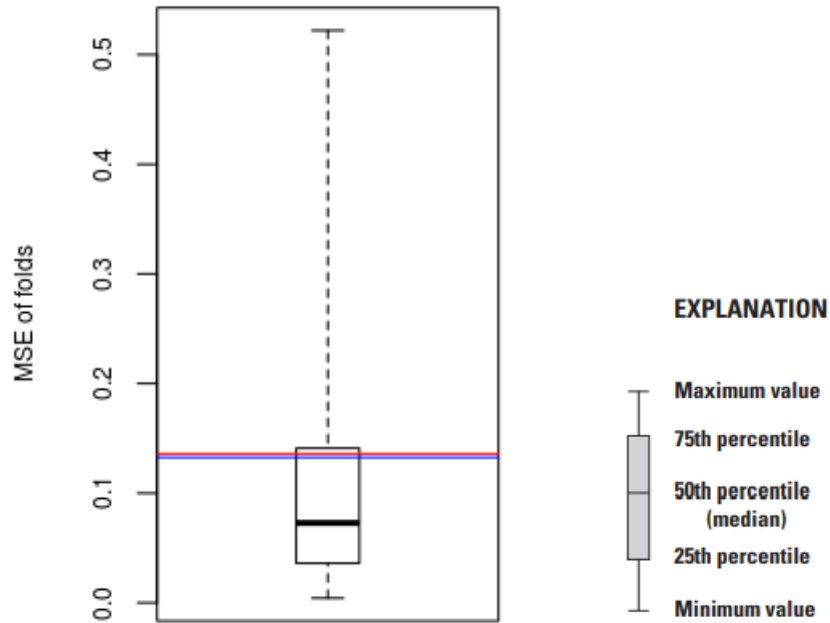
- 2 Number of values
- Maximum value
- 75th percentile
- 50th percentile (median)
- 25th percentile
- Minimum value

## Cross Validation



Minimum mean squared error (MSE) of folds: 0.00419  
Mean MSE of folds: 0.13300  
Median MSE of folds: 0.07280  
Maximum MSE of folds: 0.52200  
(Mean MSE of folds) / (Model MSE): 0.97500





Red line - Model MSE

Blue line - Mean MSE of folds

### Model-Calibration Dataset

$\theta$	Date	LogSSC	LogQ	SSC	Q	Computed LogSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
1	2007-03-26	1.89	1.68	78	48	1.77	80.1	0.123	0.324	--
2	2007-03-30	2.05	1.79	111	61	1.83	92.9	0.212	0.622	--
3	2007-03-30	3	2.28	1010	192	2.14	189	0.864	2.23	--
4	2007-03-31	3.21	3.66	1640	4600	2.99	1340	0.223	0.757	--
5	2007-03-31	3.15	3.62	1420	4130	2.96	1250	0.189	0.498	--
6	2007-04-02	2.55	2.92	354	839	2.54	468	0.013	0.0531	--
7	2007-04-15	2.86	3.61	727	4120	2.96	1250	-0.101	-0.498	--
8	2007-04-16	3.08	3.57	1190	3700	2.93	1170	0.142	0.439	--
9	2007-04-17	2.79	3.25	621	1770	2.74	743	0.0566	0.269	--
10	2007-04-20	2.3	2.71	198	512	2.4	345	-0.107	-0.559	--
11	2007-05-07	3.29	4.25	1960	17800	3.35	3080	-0.0623	-0.381	--
12	2007-05-08	2.88	4.35	765	22400	3.42	3550	-0.533	-1.43	--
13	2007-05-24	3.6	3.72	3960	5200	3.02	1440	0.573	1.6	--
14	2007-05-24	3.49	3.74	3080	5490	3.04	1490	0.449	1.3	--
15	2007-07-02	2.56	3.61	364	4060	2.96	1240	-0.397	-1.3	--
16	2007-07-18	2.11	2.49	129	312	2.27	254	-0.16	-0.83	--
17	2007-10-18	2.49	2.91	307	807	2.53	457	-0.0384	-0.16	--
18	2008-02-18	3.33	4	2150	10100	3.2	2170	0.13	0.381	--
19	2008-04-24	3.31	4.14	2020	13900	3.29	2650	0.0165	0.106	--

20	2008-05-27	3.04	4.02	1090	10500	3.21	2220	-0.175	-0.907	--
21	2008-07-09	2.92	3.12	828	1310	2.65	615	0.263	0.907	--
22	2008-09-15	2.54	4.33	346	21300	3.4	3440	-0.863	-1.6	--
23	2009-05-08	3.54	4.16	3500	14500	3.3	2720	0.245	0.83	--
24	2009-05-14	2.69	3.36	495	2310	2.81	874	-0.112	-0.622	--
25	2015-05-18	2.86	4.09	729	12400	3.26	2460	-0.394	-1.18	--
26	2015-06-16	3.08	3.88	1200	7640	3.13	1830	-0.0489	-0.214	--
27	2015-05-28	3.05	3.85	1130	7060	3.11	1740	-0.0536	-0.324	--
28	2015-09-24	1.92	1.87	84	74	1.89	105	0.0389	0.16	--
29	2015-12-02	3.06	3.22	1150	1670	2.72	717	0.34	1.08	--
30	2016-01-07	2.52	3.1	329	1270	2.65	605	-0.13	-0.688	--
31	2016-01-14	1.26	2.44	18	272	2.23	234	-0.979	-2.23	--
32	2016-02-25	1.96	2.24	92	172	2.11	176	-0.147	-0.757	--
33	2016-04-20	3.49	3.52	3100	3310	2.9	1090	0.588	1.83	--
34	2016-04-20	3.31	3.51	2030	3230	2.9	1080	0.411	1.18	--
35	2016-04-27	3.24	4.11	1750	13000	3.27	2540	-0.0274	0	--
36	2016-05-09	2.88	3.15	763	1400	2.67	644	0.209	0.559	--
37	2016-09-09	3.25	4.02	1790	10500	3.21	2230	0.039	0.214	--
38	2017-01-26	1.32	2.48	21	301	2.26	249	-0.939	-1.83	--
39	2017-03-27	3.19	3.55	1540	3540	2.92	1140	0.266	0.99	--
40	2017-03-29	3.51	4.14	3210	13700	3.29	2630	0.221	0.688	--
41	2017-04-05	3.15	3.98	1420	9490	3.19	2090	-0.0336	-0.106	--
42	2017-10-05	1.59	1.77	39	58.7	1.82	90.7	-0.232	-0.99	--
43	2018-03-19	1.81	1.8	65	63.7	1.84	95.4	-0.032	-0.0531	--
44	2018-08-20	2.2	2.52	158	332	2.29	264	-0.0885	-0.439	--
45	2018-10-09	2.99	4.25	987	18000	3.36	3100	-0.363	-1.08	--
46	2018-12-03	2.91	3.61	810	4110	2.96	1250	-0.0535	-0.269	--
47	2019-03-13	3.54	3.81	3480	6440	3.08	1650	0.459	1.43	--

## Definitions

Adj R<sup>2</sup>: 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<sup>2</sup>: 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>.]