Appendix 12. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07182510, Neosho River at Burlington, 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: 07182510
Site name: Neosho River at Burlington, Kansas
Location: Lat 38°11'40", long 95°44'06" referenced to North American Datum of 1927, in NE 1/4 SW 1/4 sec.26, T.21 S., R.15 E., Coffey County, Kans., hydrologic unit 11070204, on right bank at upstream side of county highway bridge at Burlington, 0.3 mile upstream from Rock Creek, and at mile 338.4.

Equipment: Sutron Satlink data collection platform (DCP), Design Analysis H–350 pressure transducer. Primary gage is the pressure sensor, which is set to agree with the wire-weight readings. The connection to the stream is a standard open-end orifice buried under riprap directly streamward from the gage shelter. Reference gage is a Type-A wire-weight attached to the upstream side of the bridge rail about mid-stream. Check-bar elevation is 41.479 feet above gage datum.

Date model was developed: January 16, 2020
Model calibration data period: March 26, 2007, through March 15, 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 (sine and cosine variables) were also evaluated as explanatory variables.

The regression model is based on 52 concurrent measurements of discretely collected SSC samples and continuously measured streamflow during March 26, 2007, through March 15, 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 1,000 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 one to nine samples per year with a Federal Interagency Sedimentation Project U.S. DH–75P, DH–76 TM, DH–95, or D–95 with a Teflon bottle, cap, and nozzle depth-integrating sampler, a D–96 bag sampler, a weighted-bottle sampler, an open mouth bottle, 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 5-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 36,500; minimum 5.00; mean 1,710; median 336.

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 07182510:
SSC-based model:

$$\mathrm{Log}_{10}(SSC) = 0.333 \times \mathrm{Log}_{10}(Q) + 0.806$$

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 1.25 for this model and the formula for the retransformed model accounting for BCF is as follows:

$$SSC = 8.00 \times Q^{0.333}$$

Previously Published Model

No previously published model.

Model Statistics, Data, and Plots

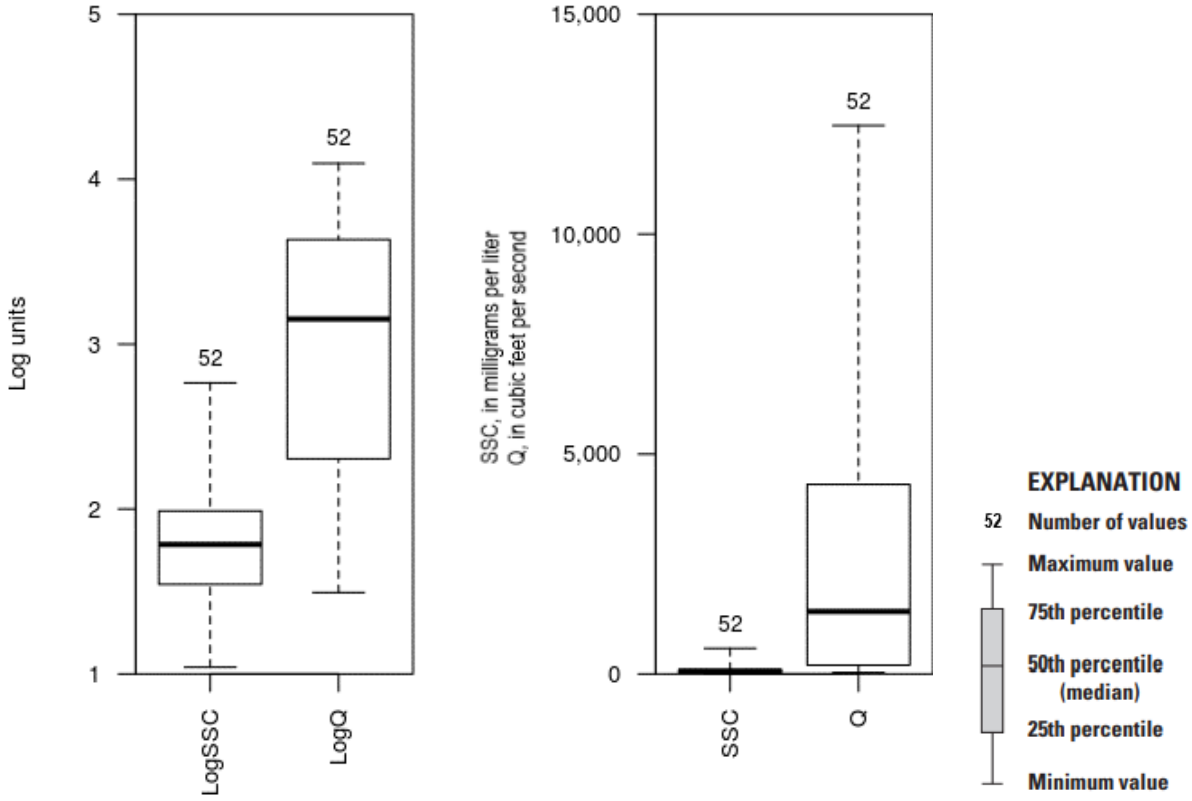
Model

Log(*SSC*) = + 0.333 * Log(*Q*) + 0.806

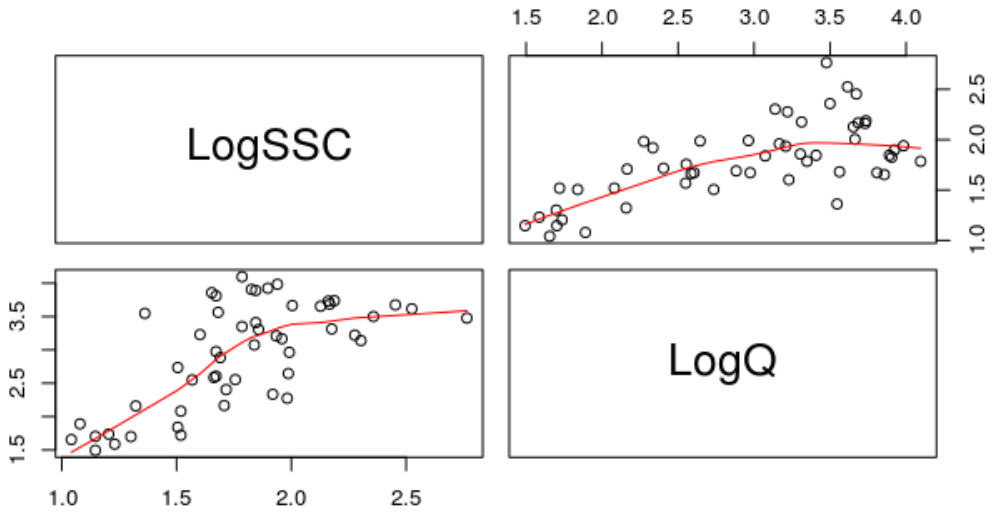
Variable Summary Statistics

	LogSSC	SSC	LogQ	Q
Minimum	1.04	11.0	1.49	31.2
1st Quartile	1.54	35.0	2.30	202.0
Median	1.79	61.0	3.15	1420.0
Mean	1.79	90.0	2.94	2530.0
3d Quartile	1.99	97.5	3.63	4310.0
Maximum	2.76	582.0	4.10	12500.0

Box Plots



Exploratory Plots



Basic Model Statistics

Number of Observations	52
Standard error (RMSE)	0.284
Average Model standard percentage error (MSPE)	70.2
Coefficient of determination (R^2)	0.457
Adjusted Coefficient of Determination (Adj. R^2)	0.446
Bias Correction Factor (BCF)	1.25

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.806	0.1560	5.17	4.16e-06
LogQ	0.333	0.0513	6.48	3.85e-08

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.968
E.vars	-0.968	1.000

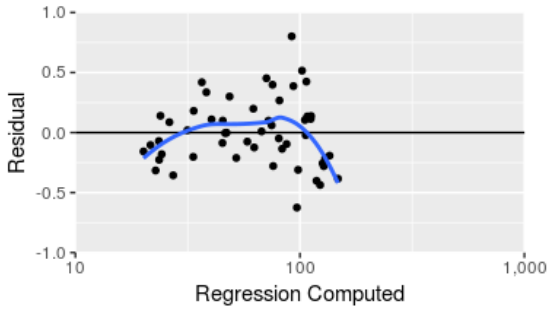
Outlier Test Criteria

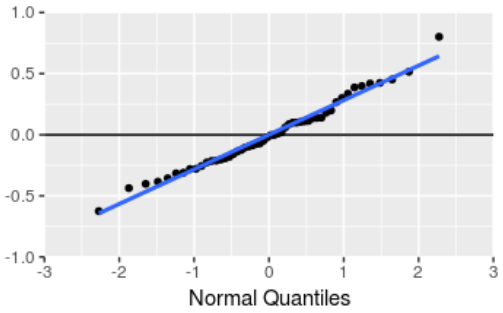
Leverage	Cook's D	DFFITS
0.115	0.194	0.392

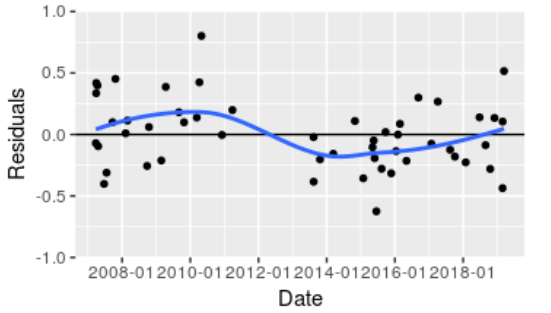
Flagged Observations

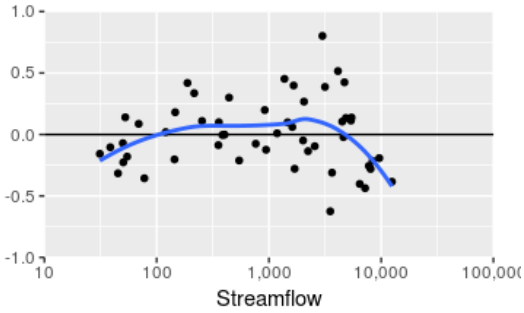
	LogSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
4/30/2010 11:15	2.76	1.96	0.802	2.86	3.10	0.0286	0.1210	0.532
6/17/2015 13:20	1.36	1.99	-0.625	-2.23	-2.33	0.0312	0.0804	-0.418

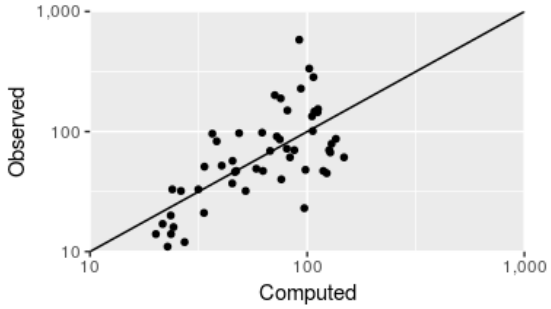
Statistical Plots

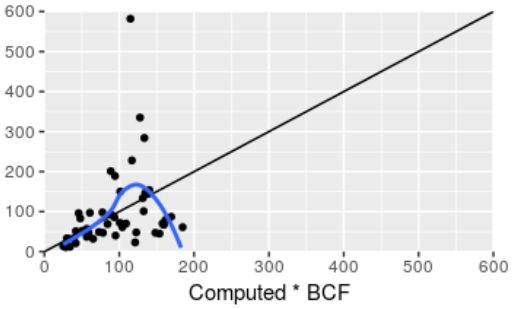


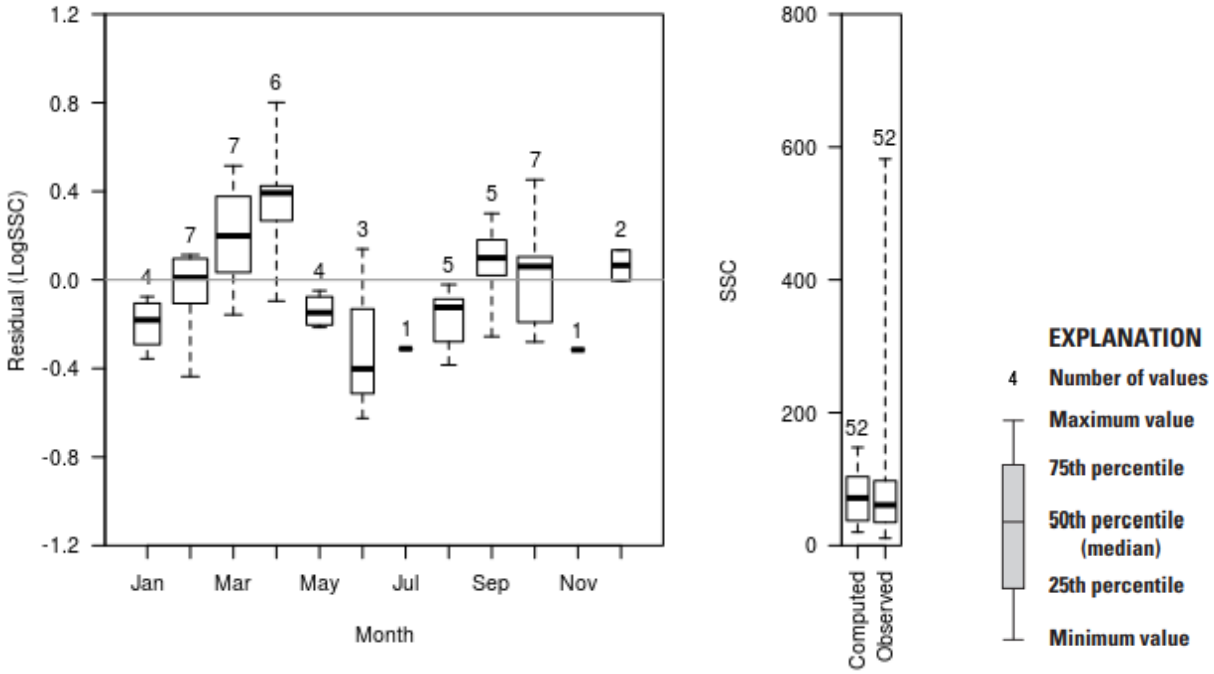


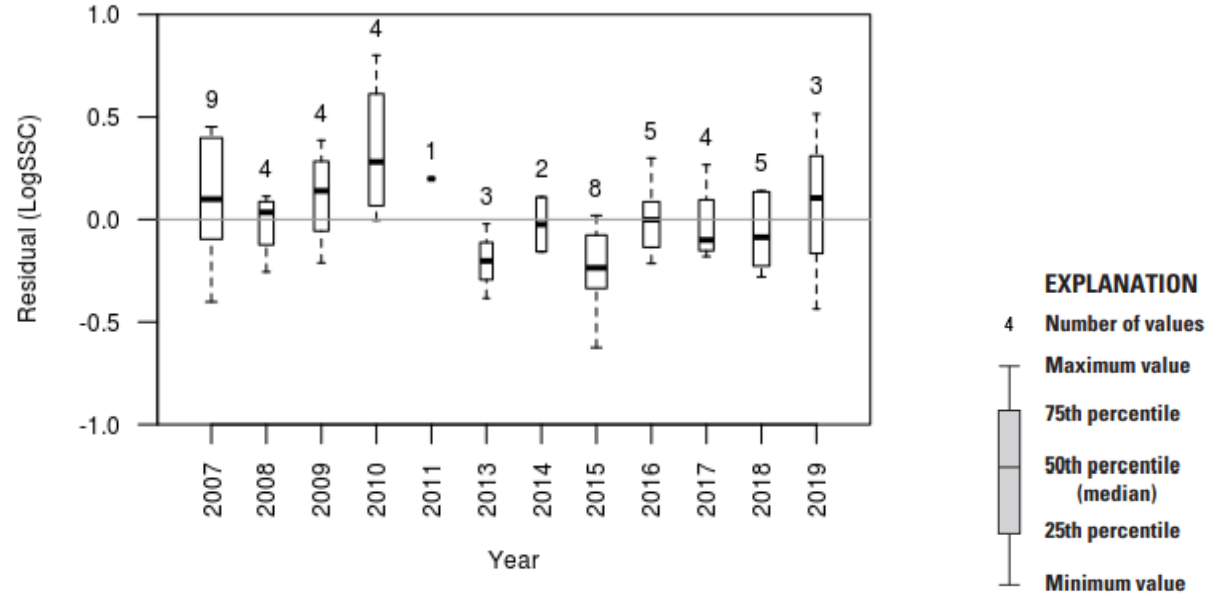




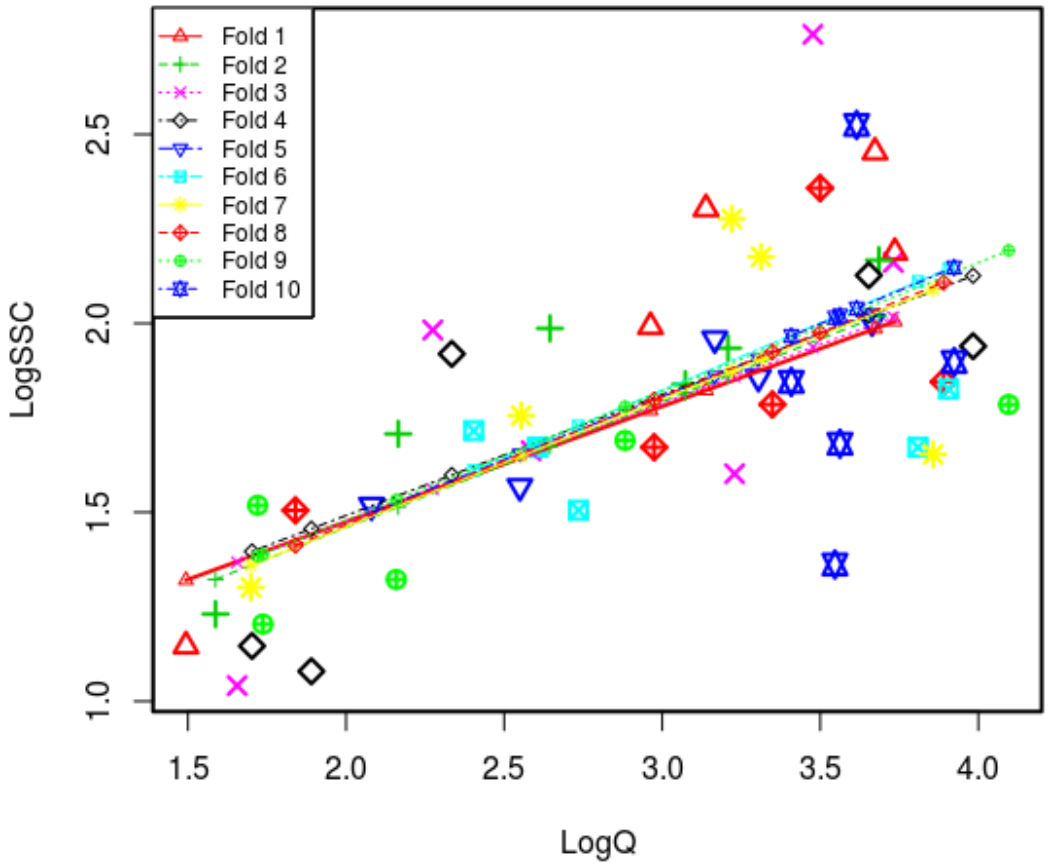




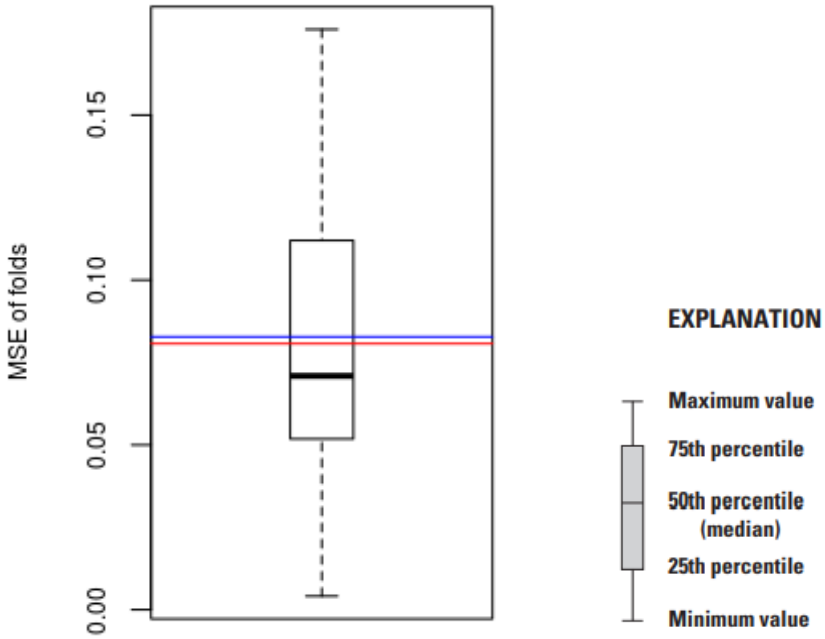

Residual (LogSSC)
1.2
0.8
0.4
0.0
-0.4
-0.8
-1.2
4
7
7
6
4
3
1
5
5
7
1
2
Jan
Mar
May
Jul
Sep
Nov
Month
SSC
800
600
400
200
0
52
52
Computed
Observed
EXPLANATION
4 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value


Residual (LogSSC)
1.0
0.5
0.0
-0.5
-1.0
9
4
4
4
1
3
2
8
5
4
5
3
2007
2008
2009
2010
2011
2013
2014
2015
2016
2017
2018
2019
Year
EXPLANATION
4 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Cross Validation




```
Minimum mean squared error (MSE) of folds:  0.00412
                          Mean MSE of folds:  0.08280
                        Median MSE of folds:  0.07090
                       Maximum MSE of folds:  0.17600
          (Mean MSE of folds) / (Model MSE):  1.03000
```



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
1	2007-03-26	1.3	1.7	20	50	1.37	29.4	-0.0707	-0.12	--
2	2007-03-30	1.92	2.33	83	216	1.58	47.9	0.336	1.06	--
3	2007-03-31	1.98	2.27	96	188	1.56	45.7	0.419	1.35	--
4	2007-04-12	2.28	3.22	189	1660	1.88	94.4	0.398	1.24	--
5	2007-04-20	1.85	3.41	70	2560	1.94	109	-0.0955	-0.267	--
6	2007-06-22	1.67	3.81	47	6440	2.07	148	-0.402	-1.65	--
7	2007-07-19	1.68	3.56	48	3650	1.99	123	-0.311	-1.14	--
8	2007-09-21	1.76	2.55	57	358	1.66	56.6	0.0996	0.368	--
9	2007-10-22	2.3	3.14	201	1380	1.85	88.7	0.452	1.65	--
10	2008-02-08	1.84	3.07	69	1180	1.83	84.3	0.00977	0.12	--
11	2008-02-28	2.16	3.73	145	5380	2.05	140	0.113	0.528	--
12	2008-09-25	1.85	3.89	70	7760	2.1	158	-0.256	-0.901	--
13	2008-10-15	1.93	3.21	86	1620	1.87	93.5	0.0605	0.218	--
14	2009-02-24	1.51	2.73	32	543	1.72	65.1	-0.211	-0.702	--
15	2009-04-13	2.36	3.5	228	3160	1.97	117	0.387	1.14	--
16	2009-09-01	1.71	2.16	51	146	1.53	42	0.181	0.765	--
17	2009-10-26	1.96	3.17	91	1470	1.86	90.6	0.099	0.317	--
18	2010-03-11	2.19	3.74	154	5440	2.05	140	0.138	0.642	--
19	2010-04-08	2.45	3.67	284	4710	2.03	134	0.425	1.48	--
20	2010-04-30	2.76	3.48	582	3000	1.96	115	0.802	2.27	--

21	2010-12-06	1.66	2.59	46	385	1.67	58	-0.00401	0.024	--
22	2011-03-28	1.99	2.96	98	918	1.79	77.5	0.199	0.831	--
23	2013-08-13	2	3.66	101	4610	2.03	133	-0.0213	-0.024	--
24	2013-08-14	1.79	4.1	61	12500	2.17	185	-0.384	-1.48	--
25	2013-10-21	1.32	2.16	21	144	1.52	41.8	-0.203	-0.642	--
26	2014-03-12	1.15	1.49	14	31.2	1.3	25.1	-0.157	-0.473	--
27	2014-10-30	1.72	2.4	52	254	1.61	50.5	0.11	0.473	--
28	2015-01-28	1.08	1.89	12	77.7	1.44	34.1	-0.356	-1.35	--
29	2015-05-07	1.23	1.59	17	38.6	1.33	27	-0.104	-0.317	--
30	2015-05-19	1.86	3.3	72	2010	1.91	101	-0.0487	-0.0721	--
31	2015-06-17	1.36	3.55	23	3520	1.99	121	-0.625	-2.27	--
32	2015-05-29	1.94	3.98	87	9620	2.13	169	-0.192	-0.584	--
33	2015-08-12	1.6	3.23	40	1690	1.88	95	-0.279	-0.975	--
34	2015-09-23	1.52	2.08	33	120	1.5	39.4	0.0198	0.169	--
35	2015-11-23	1.04	1.66	11	45.2	1.36	28.4	-0.316	-1.24	--
36	2016-01-14	1.79	3.35	61	2230	1.92	104	-0.135	-0.42	--
37	2016-02-03	1.67	2.6	47	403	1.67	58.9	-0.00116	0.0721	--
38	2016-02-24	1.51	1.84	32	69.1	1.42	32.8	0.0866	0.267	--
39	2016-05-04	1.9	3.92	79	8380	2.11	162	-0.214	-0.765	--
40	2016-09-09	1.99	2.64	97	441	1.69	60.7	0.3	0.975	--
41	2017-01-26	1.69	2.88	49	763	1.77	72.9	-0.0754	-0.169	--
42	2017-04-05	2.18	3.31	150	2050	1.91	101	0.267	0.901	--
43	2017-08-15	1.67	2.97	47	943	1.8	78.2	-0.124	-0.368	--
44	2017-10-05	1.2	1.74	16	54.6	1.38	30.3	-0.18	-0.528	--
45	2018-01-29	1.15	1.7	14	50.4	1.37	29.5	-0.227	-0.831	--
46	2018-06-26	1.52	1.72	33	52.5	1.38	29.9	0.14	0.702	--
47	2018-08-29	1.57	2.55	37	355	1.65	56.5	-0.0867	-0.218	--
48	2018-10-18	1.83	3.91	67	8050	2.11	160	-0.28	-1.06	--
49	2018-12-03	2.17	3.69	147	4850	2.03	135	0.134	0.584	--
50	2019-02-27	1.65	3.86	45	7190	2.09	154	-0.437	-1.87	--
51	2019-02-28	2.13	3.65	134	4500	2.02	132	0.105	0.42	--
52	2019-03-15	2.53	3.61	335	4120	2.01	128	0.516	1.87	--

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.]