Appendix 4. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07144780, North Fork Ninnescah River above Cheney Reservoir, Kansas, during January 1, 1999, through October 16, 2009

This model archive summary summarizes the suspended-sediment concentration (SSC) model developed to compute hourly or daily SSC during January 1, 1999, through October 16, 2009. This model supersedes all prior models used during this period. 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).

Site and Model Information

Site number: 07144780
Site name: North Fork Ninnescah River above Cheney Reservoir, Kansas
Location: Lat 37°51'45", long 98°00'49" referenced to North American Datum of 1927, in NE 1/4 SE 1/4 NE 1/4 sec.19, T.25 S., R.6 W., Reno County, Kans., Hydrologic Unit 11030014, on right bank at upstream side of county highway bridge, 10 miles south of Hutchinson, 18.1 miles upstream from Cheney Dam.

Equipment: A YSI 6600 Extended Deployment System water-quality monitor equipped with sensors for water temperature, specific conductance, pH, dissolved oxygen, and turbidity (a YSI Model 6026 turbidity sensor [November 9, 1998, to December 1, 2010] and a YSI Model 6136 turbidity sensor [October 17, 2009, to November 12, 2015; March 31, 2017, to June 7, 2017]) (YSI Incorporated, 2007, 2012). The YSI 6600 water-quality monitor was in operation during November 9, 1998, through November 12, 2015.

A Sutron Satlink 2 High Data Rate Collection Platform and a Design Analysis Water Log H350/355 nonsubmersible pressure transducer transfers real-time stage, precipitation, and water quality data via satellite. The primary reference gage is a Type-A wire-weight gage located on the downstream bridge guardrail. Check-bar elevation is 21.804 feet. The orifice is enclosed in 1 1/4-inch pipe, which runs from the gage house, under the bridge, and along an I-beam where it is attached to the concrete pier closest to the left edge of water.

Date model was developed: April 26, 2019

Model calibration data period: January 31, 1999, to August 25, 2010

Model Data

All data were collected using USGS protocols (U.S. Geological Survey, 2006; Wagner and others, 2006; Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010) 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, pH, dissolved oxygen, and turbidity. Seasonal components (sine and cosine variables) were also evaluated as explanatory variables.

The regression model is based on 69 concomitant values of discretely measured SSC samples, streamflow, and continuously measured turbidity during January 31, 1999, through August 25, 2010. Discrete samples were collected over a range of streamflows and turbidity conditions. No samples were less than laboratory detection limits. Summary statistics and the complete model-calibration data are provided below. Outliers were identified using studentized residuals (for values greater than 3 or less than −3). The sample collected on September 19, 2001, had large heterogeneity in the channel cross-sectional data during sample collection and was removed from the model calibration dataset. Outliers in previously published versions of this model (Christensen and others, 2006; Stone and others, 2013) were examined and retained in the dataset if there were no clear issues, explanations, or conditions that would cause a result to be invalid for model calibration.

Suspended Sediment

Discrete samples were collected from the downstream side of the bridge or instream within 50 feet of the bridge using equal-width-increment, multiple vertical, single vertical, or grab 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 3 to 10 samples per year with a Federal Interagency Sedimentation Project U.S. DH–95 or D–95 with a Teflon bottle, cap, and nozzle depth-integrating sampler; a DH–81 with a Teflon bottle, cap, and nozzle hand sampler; 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

Turbidity was measured using a YSI model 6026 sensor installed during November 9, 1998, through December 1, 2010. Concomitant turbidity values were time interpolated. If continuous data were not available (2 or more hours of specific conductance values bracketing the sample collection time were missing) because of fouling, changes in equipment, or unsuitable site conditions, then the field monitor turbidity value measured during sampling was substituted. If no concomitant continuous data were available, the sample was not included in the dataset.

The streamflow data were measured using a nonsubmersible pressure transducer during January 1, 1991, through December 31, 2019. The surrogate 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.

Model Development

Ordinary least squares regression analysis was done using R programming language (R Core Team, 2019) to relate discretely collected SSC to turbidity 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.

Turbidity and streamflow were selected as the best predictors 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 USGS site 07144780:

SSC-based model:

$$\log_{10}(SSC) = 0.471 \times \log_{10}(Q) + 0.626 \times \log_{10}(TBY6026) - 0.0555 ,$$

where,

SSC = suspended-sediment concentration, in milligrams per liter;
Q = streamflow in cubic feet per second; and
TBY6026 = turbidity, YSI model 6026, in formazin nephelometric units.

The use of turbidity as an explanatory variable is appropriate physically and statistically. Turbidity makes sense physically because suspended sediment is composed of particles that scatter light in water. The relation between turbidity and SSC can vary given varying concentrations of organic suspended particles that increase turbidity but are not included in the SSC analysis.

The log-transformed model may be retransformed to original units so that SSC can be calculated directly. The retransformation introduces a bias in the calculated constituent. This bias may be corrected using Duan's bias correction factor (BCF; Duan, 1983). For this model, the calculated BCF is 1.12. The retransformed model, accounting for BCF, is as follows:

$$SSC = (Q^{0.471} \times TBY6026^{0.626} \times 10^{-0.0555}) \times 1.12$$

Previous Models

Version	Model Equation	Reference
1.0	$\log_{10}(SSC) = 1.10 \times \log_{10}(TBY6026) + 0.0037$	Christensen and others (2006)
1.1	$\log_{10}(SSC) = 0.540 \times \log_{10}(Q) + 0.559 \times \log_{10}(TBY6026) - 0.11$	Stone and others (2013)

Model Statistics, Data, and Plots

Definitions for terms used in this output can be found at the end of this document.

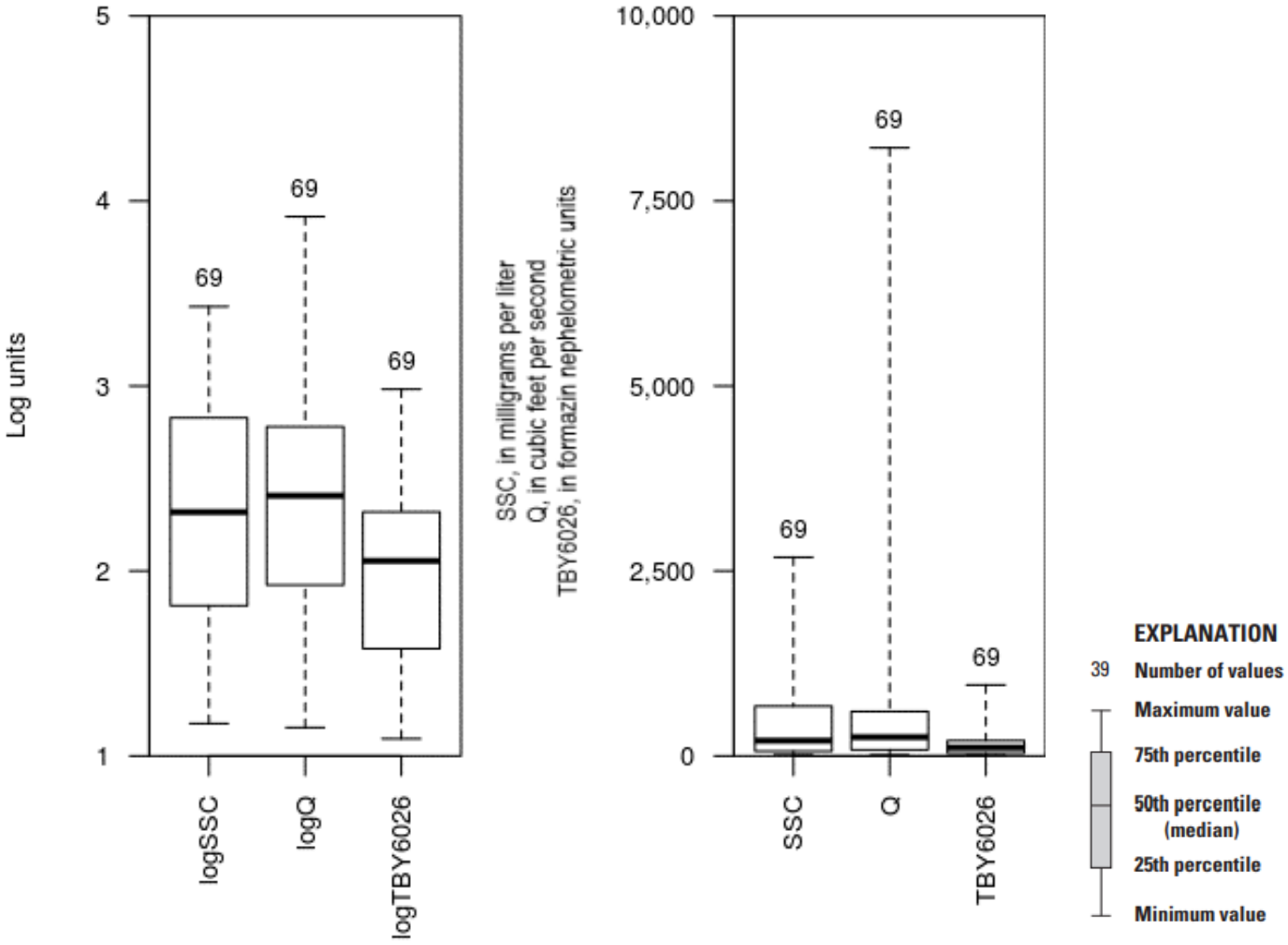
Model

logSSC = + 0.471 * log$Q$ + 0.626 * log$TBY6026$ – 0.0555

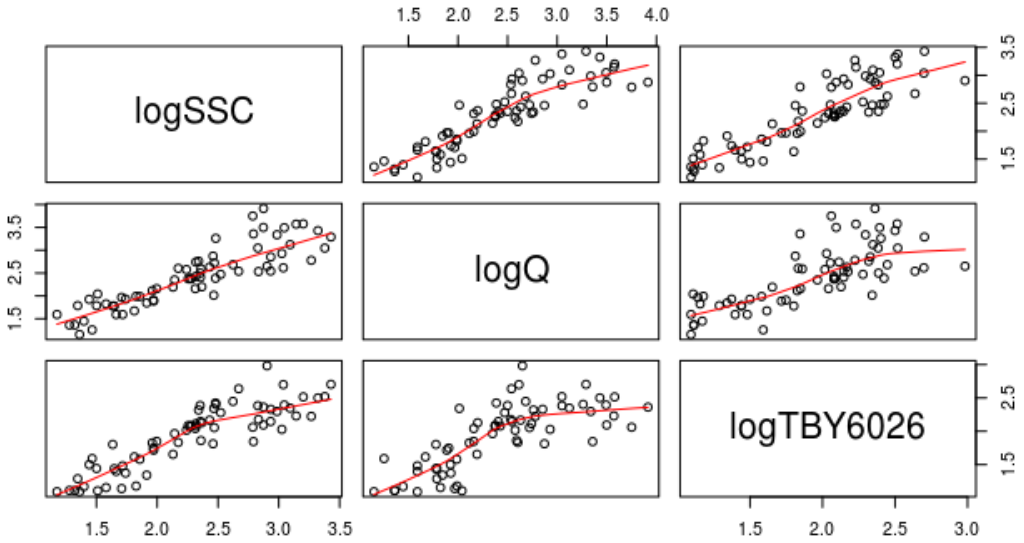
Variable Summary Statistics

	logSSC	SSC	logQ	logTBY6026	Q	TBY6026
Minimum	1.18	15.0	1.15	1.09	14.2	12.4
1st Quartile	1.81	64.9	1.92	1.58	84.0	38.0
Median	2.32	208.0	2.41	2.05	255.0	113.0
Mean	2.29	452.0	2.42	1.93	799.0	143.0
3d Quartile	2.83	673.0	2.78	2.32	602.0	208.0
Maximum	3.43	2,690.0	3.91	2.98	8,220.0	960.0

Box Plots



Exploratory Plots



Red line shows the locally weighted scatterplot smoothing (LOWESS).

Basic Model Statistics

For a detailed definition and explanation of the terms used below, refer to Helsel and Hirsch (2002).

```
Number of Observations                              69
Standard error (RMSE)                            0.206
Average Model standard percentage error (MSPE)    49.3
Coefficient of determination (R²)                0.886
Adjusted Coefficient of Determination (Adj. R²) 0.883
Bias Correction Factor (BCF)                      1.12

Variance Inflation Factors (VIF)
      logQ logTBY6026
      2.26       2.26
```

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.0555	0.1070	-0.517	6.07e-01
logQ	0.4710	0.0573	8.220	1.09e-11
logTBY6026	0.6260	0.0788	7.940	3.41e-11

Correlation Matrix

	Intercept	logQ	logTBY6026
Intercept	1.000	-0.235	-0.453
logQ	-0.235	1.000	-0.746
logTBY6026	-0.453	-0.746	1.000

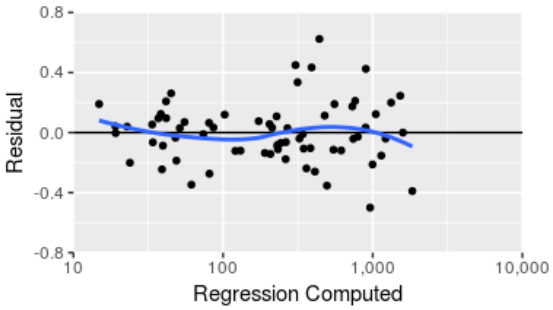
Outlier Test Criteria

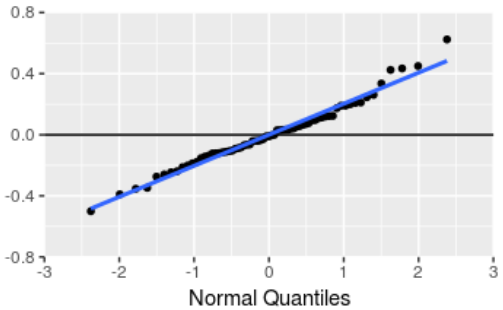
Leverage	Cook's D	DFFITS
0.130	0.264	0.417

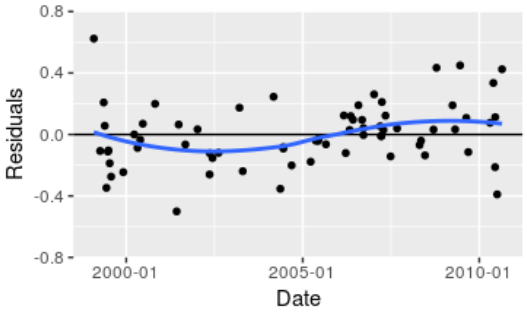
Flagged Observations

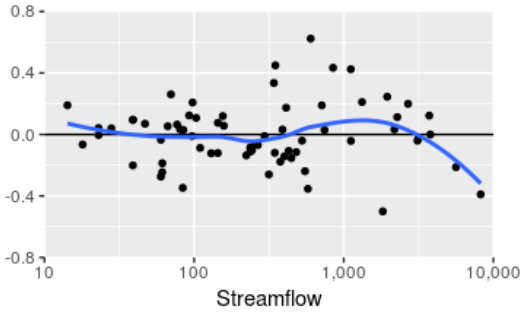
	logSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
1/31/1999 14:25	3.27	2.64	0.624	3.060	3.270	0.0202	0.0643	0.470
6/6/2001 11:35	2.48	2.98	-0.500	-2.480	-2.580	0.0385	0.0817	-0.516
6/12/2002 11:10	2.90	3.06	-0.153	-0.803	-0.801	0.1420	0.0355	-0.326
6/14/2010 11:30	2.79	3.00	-0.212	-1.100	-1.100	0.1260	0.0582	-0.419
7/6/2010 10:30	2.87	3.26	-0.390	-2.010	-2.050	0.1120	0.1680	-0.728
8/25/2010 11:00	3.38	2.95	0.425	2.100	2.160	0.0367	0.0560	0.421

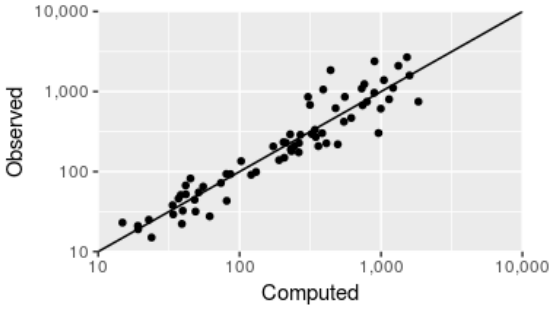
Statistical Plots

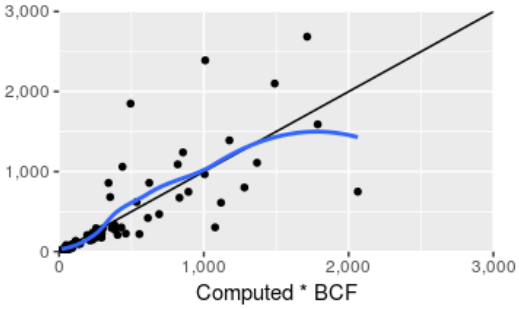


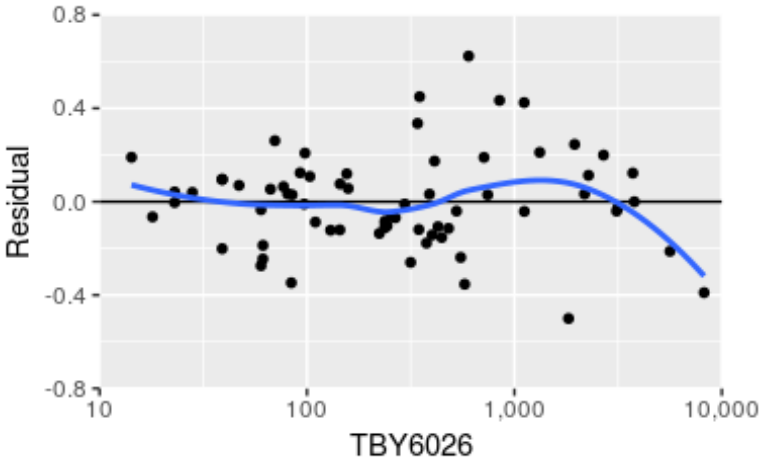


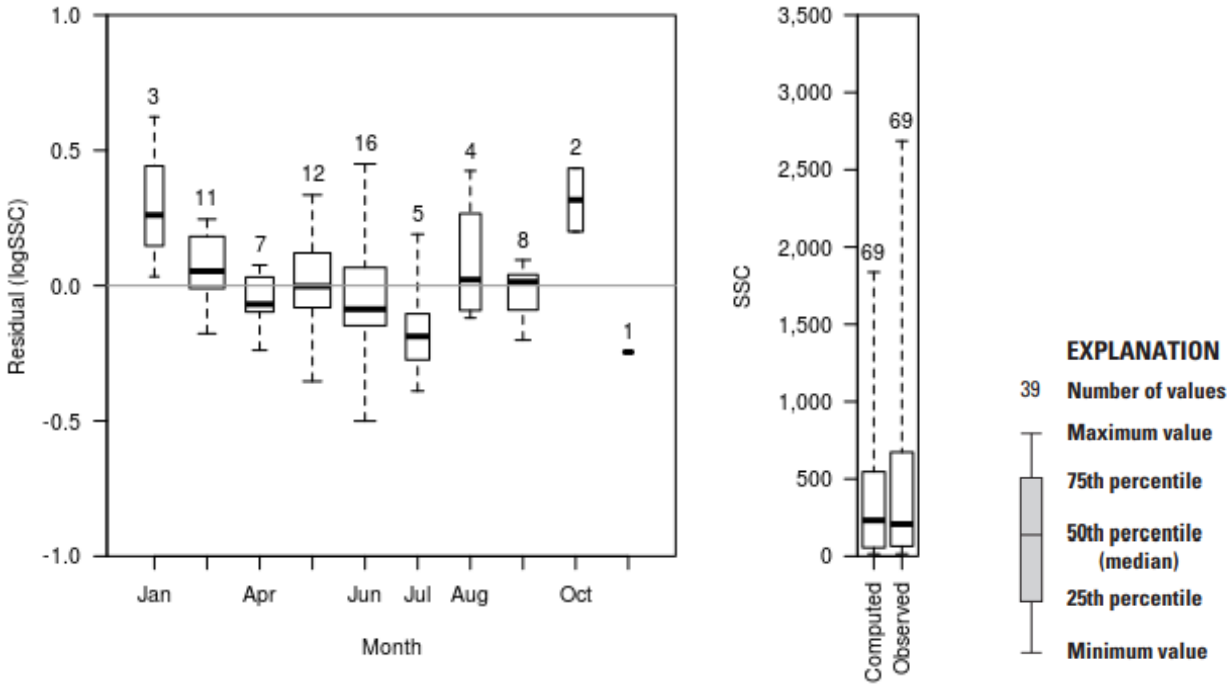


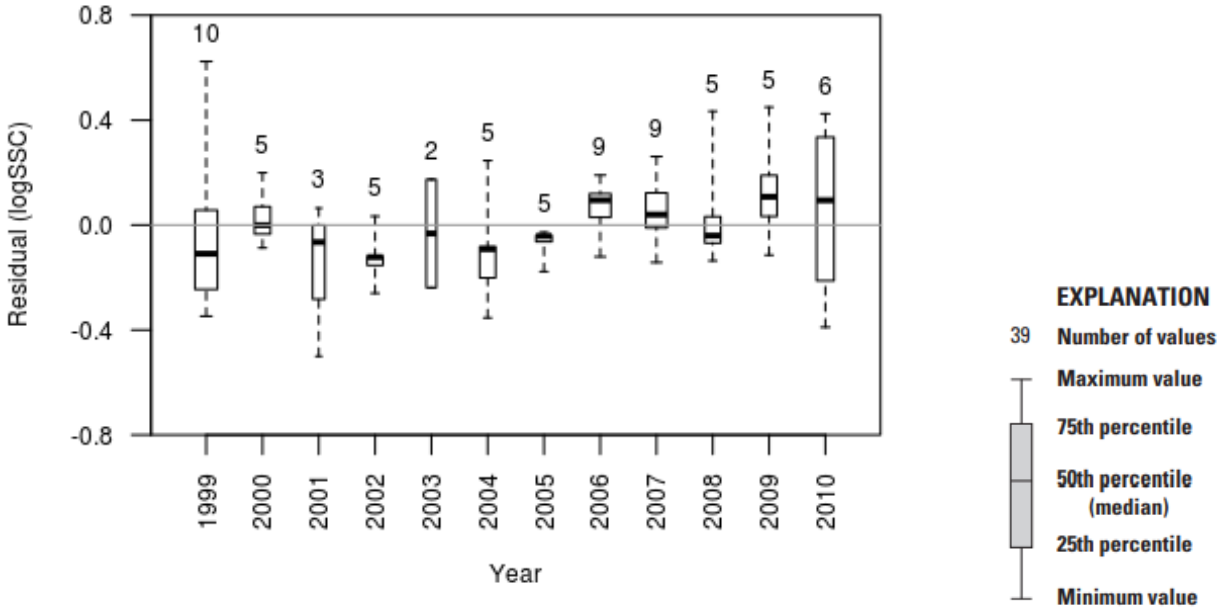




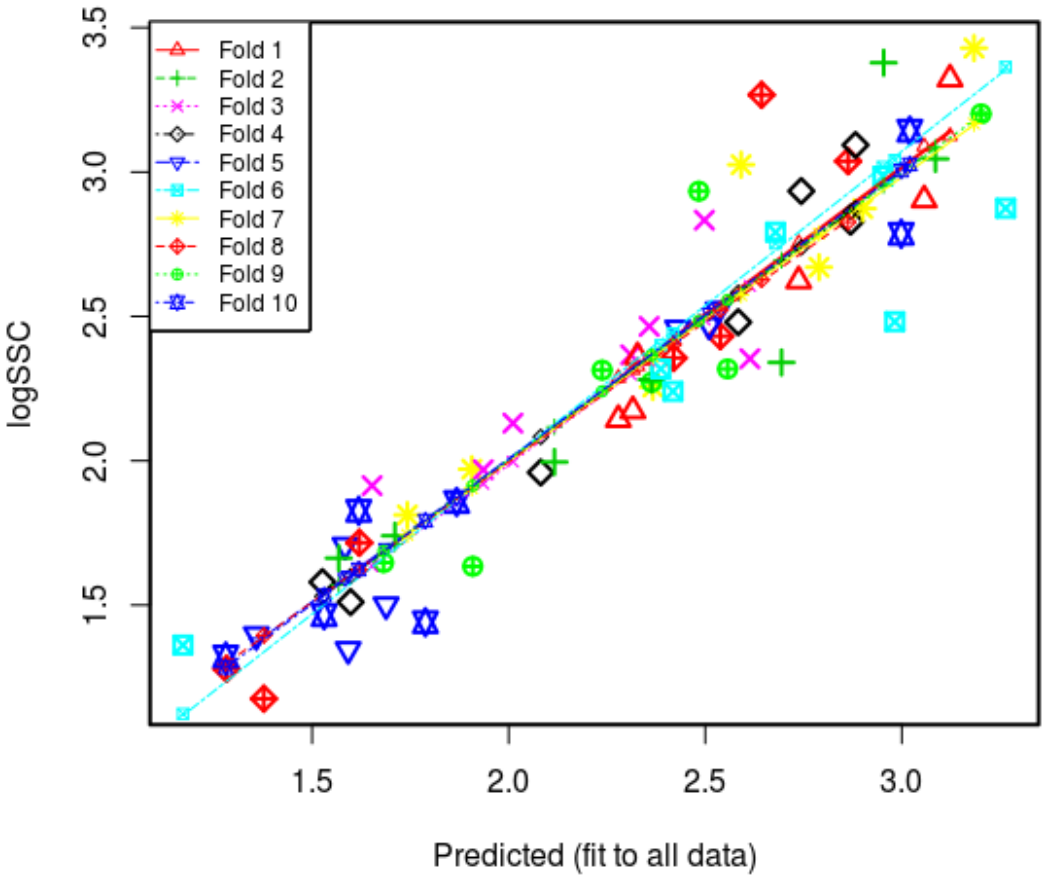





1.0
0.5
0.0
-0.5
-1.0
Residual (logSSC)
3
11
7
12
16
5
4
8
2
1
Jan
Apr
Jun
Jul
Aug
Oct
Month
3,500
3,000
2,500
2,000
1,500
1,000
500
0
SSC
69
69
Computed
Observed
EXPLANATION
39 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value


0.8
0.4
0.0
-0.4
-0.8
Residual (logSSC)
10
5
3
5
2
5
5
9
9
5
5
6
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
Year
EXPLANATION
39 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Cross Validation



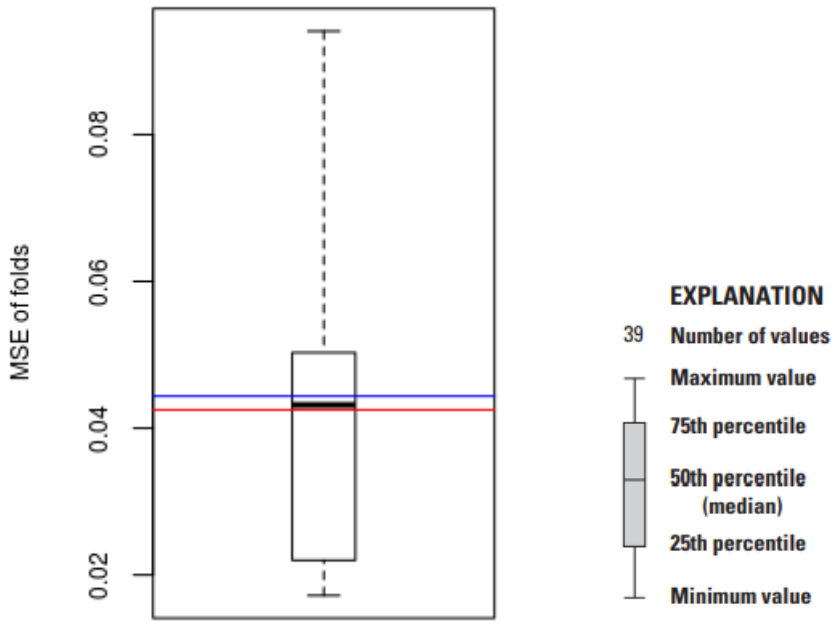

```
Fold – equal partition of the data (10 percent of the data)

Large symbols – observed value of a data point removed in a fold

Small symbols – recomputed value of a data point removed in a fold

Recomputed regression lines – adjusted regression line with one fold removed


              Minimum MSE of folds:  0.0172
                 Mean MSE of folds:  0.0444
               Median MSE of folds:  0.0432
              Maximum MSE of folds:  0.0941
 (Mean MSE of folds) / (Model MSE):  1.0400
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Data Set

	Date	logSSC	logQ	logTBY6026	SSC	Q	TBY6026	Computed logSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
1	1999-01-31	3.27	2.78	2.22	1850	602	167	2.64	494	0.624	2.38	--
2	1999-04-06	2.43	2.63	2.17	270	428	147	2.54	388	-0.107	-0.531	--
3	1999-05-13	1.83	1.99	1.18	67.1	97.7	15.1	1.62	46.7	0.208	1.15	--
4	1999-05-24	2.37	2.2	2.13	233	158	134	2.31	229	0.0567	0.41	--
5	1999-06-10	1.44	1.92	1.5	27.6	84	31.6	1.79	69	-0.347	-1.63	--
6	1999-06-25	2.26	2.38	2.09	180	237	122	2.37	261	-0.111	-0.573	--
7	1999-07-02	2.48	2.39	2.42	302	243	266	2.58	431	-0.104	-0.49	--
8	1999-07-14	1.5	1.79	1.44	31.7	61.3	27.7	1.69	54.8	-0.187	-1.02	--
9	1999-07-29	1.63	1.78	1.8	43.1	60	63.4	1.91	91	-0.274	-1.5	--
10	1999-12-02	1.35	1.79	1.29	22.2	61.2	19.4	1.59	43.9	-0.245	-1.31	--
11	2000-03-24	3.2	3.58	2.51	1590	3790	327	3.2	1790	-0.000215	0.0725	--
12	2000-04-27	1.51	2.04	1.11	32.4	110	12.8	1.6	44.5	-0.0874	-0.41	--
13	2000-05-25	1.65	1.78	1.44	44.4	60	27.5	1.68	54	-0.0343	-0.109	--
14	2000-06-21	1.81	1.67	1.62	64.9	47	41.3	1.74	62	0.07	0.49	--
15	2000-10-26	3.32	3.43	2.5	2100	2690	316	3.12	1490	0.2	1.09	--
16	2001-06-06	2.48	3.26	2.4	303	1820	253	2.98	1080	-0.5	-2.38	--
17	2001-06-27	1.97	1.89	1.72	93.5	77	52.1	1.91	90.5	0.0645	0.449	--
18	2001-09-04	1.47	1.26	1.59	29.2	18	39	1.53	38.1	-0.0651	-0.294	--
19	2002-01-08	1.97	1.91	1.75	93.1	80.7	56	1.94	96.8	0.0336	0.256	--
20	2002-05-13	2.35	2.5	2.39	226	316	243	2.61	461	-0.26	-1.4	--
21	2002-05-15	1.96	2.11	1.83	91.1	130	67	2.08	136	-0.122	-0.754	--
22	2002-06-12	2.9	2.65	2.98	801	446	960	3.06	1280	-0.153	-0.907	--
23	2002-08-14	2.67	2.54	2.64	468	346	434	2.79	691	-0.119	-0.661	--
24	2003-03-18	3.04	2.61	2.7	1090	412	500	2.86	820	0.174	0.907	--
25	2003-04-21	2.32	2.74	2.11	208	551	130	2.56	405	-0.239	-1.23	--
26	2004-03-05	3.43	3.29	2.7	2690	1950	504	3.18	1710	0.246	1.31	--
27	2004-05-14	2.34	2.76	2.32	219	576	208	2.69	555	-0.354	-1.78	--
28	2004-06-14	2.27	2.38	2.08	187	238	120	2.36	259	-0.0917	-0.449	--
29	2004-06-14	2.28	2.38	2.08	191	238	120	2.36	259	-0.0825	-0.371	--
30	2004-09-08	1.18	1.59	1.09	15	39	12.4	1.38	26.8	-0.201	-1.09	--
31	2005-03-24	2.24	2.58	2.02	174	377	104	2.42	294	-0.177	-0.964	--

32	2005-05-16	2.47	2.72	2.05	295	526	113	2.51	363	-0.0403	-0.182	--
33	2005-06-10	2.83	3.05	2.38	673	1120	242	2.87	832	-0.0417	-0.219	--
34	2005-06-13	2.87	3.5	2.09	747	3150	124	2.9	894	-0.0274	-0.0725	--
35	2005-08-29	2.36	2.41	2.15	227	255	140	2.42	295	-0.0636	-0.256	--
36	2006-03-02	1.71	1.97	1.14	51	92.7	13.8	1.58	43.1	0.124	0.854	--
37	2006-03-22	2	2.16	1.85	99	144	70.5	2.12	147	-0.121	-0.707	--
38	2006-05-01	1.74	1.93	1.37	55	84.8	23.6	1.71	57.7	0.0296	0.145	--
39	2006-05-12	2.13	2.19	1.65	135	155	45.1	2.01	115	0.12	0.754	--
40	2006-06-05	1.72	1.59	1.48	52	39	30.3	1.62	46.8	0.096	0.616	--
41	2006-07-31	1.36	1.15	1.09	23	14.2	12.4	1.17	16.7	0.19	1.02	--
42	2006-09-07	1.66	1.59	1.4	46	39	25	1.57	41.5	0.095	0.573	--
43	2006-09-21	1.28	1.36	1.11	19	23	13	1.28	21.5	-0.00343	0.0362	--
44	2006-09-21	1.32	1.36	1.11	21	23	12.9	1.28	21.4	0.0418	0.332	--
45	2007-01-09	1.91	1.85	1.34	82	70	22	1.65	50.5	0.261	1.4	--
46	2007-03-14	1.58	1.82	1.16	38	66.7	14.3	1.53	37.7	0.0536	0.371	--
47	2007-03-22	1.86	1.99	1.58	72	97	38	1.87	82.8	-0.0104	-0.0362	--
48	2007-03-26	2.52	2.47	2.28	333	296	190	2.53	383	-0.0103	0	--
49	2007-03-31	3.09	3.12	2.35	1240	1320	222	2.88	856	0.211	1.23	--
50	2007-04-16	2.46	2.87	1.81	288	744	65	2.43	302	0.0296	0.109	--
51	2007-05-07	3.14	3.57	2.23	1390	3720	170	3.02	1180	0.123	0.803	--
52	2007-06-29	2.17	2.6	1.83	149	401	68	2.32	232	-0.143	-0.854	--
53	2007-09-04	1.4	1.45	1.17	25	28	14.8	1.36	25.6	0.0397	0.294	--
54	2008-04-24	2.32	2.43	2.08	208	267	120	2.39	274	-0.0687	-0.332	--
55	2008-05-09	3.05	3.49	2.39	1110	3110	247	3.09	1370	-0.0402	-0.145	--
56	2008-06-19	2.14	2.35	1.96	139	223	92.3	2.28	213	-0.136	-0.803	--
57	2008-09-15	2.36	2.59	1.86	229	390	72.6	2.33	239	0.0321	0.182	--
58	2008-10-16	3.03	2.93	2.03	1060	848	107	2.59	438	0.434	1.78	--
59	2009-03-31	2.93	2.85	2.33	859	713	213	2.74	623	0.19	0.964	--
60	2009-04-27	2.99	3.34	2.3	967	2180	198	2.95	1010	0.0332	0.219	--
61	2009-06-17	2.93	2.54	2.15	858	349	140	2.48	342	0.45	1.99	--
62	2009-08-20	2.47	2.01	2.34	292	103	220	2.36	256	0.108	0.661	--
63	2009-09-10	2.62	2.68	2.45	420	481	280	2.74	614	-0.114	-0.616	--
64	2010-04-23	2.31	2.16	2.04	206	144	110	2.24	194	0.0762	0.531	--
65	2010-05-27	2.83	2.53	2.18	681	341	150	2.5	353	0.335	1.5	--
66	2010-06-14	2.79	3.75	2.06	611	5630	115	3	1120	-0.212	-1.15	--
67	2010-06-16	2.79	3.36	1.85	619	2280	70	2.68	536	0.113	0.707	--
68	2010-07-06	2.87	3.91	2.36	749	8220	230	3.26	2060	-0.39	-1.99	--
69	2010-08-25	3.38	3.05	2.52	2390	1110	330	2.95	1010	0.425	1.63	--

Definitions

SSC: suspended sediment concentration, in milligrams per liter (80154)
TBY6026: turbidity, YSI model 6026, in formazin nephelometric units (63680)
Q: streamflow, instantaneous, in cubic feet per second (00061)

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

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., and Ziegler, A.C., 2006, Water quality and relation to taste-and-odor compounds in the North Fork Ninnescah River and Cheney Reservoir, south-central Kansas, 1997–2003: U.S. Geological Survey Scientific Investigations Report 2006–5095, 43 p. [Also available at https://doi.org/10.3133/sir20065095.]

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.]

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. A3, 522 p. [Also available at https://doi.org/10.3133/tm4A3.]

R Core Team, 2019, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed August 2019 at https://www.R-project.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 April 2020 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, 52 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 April 2020 at https://doi.org/10.3133/tm3A7.

Stone, M.L., Graham, J.L., and Gatotho, J.W., 2013, Continuous real-time water-quality monitoring and regression analysis to compute constituent concentrations and loads in the North Fork Ninnescah River upstream from Cheney Reservoir, south-central Kansas, 1999–2012: U.S. Geological Survey Scientific Investigations Report 2013–5071, 44 p., accessed July 2020 at https://doi.org/10.3133/sir20135071.

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 July 2020 at https://doi.org/10.3133/tm3A8.

U.S. Geological Survey, 2006, Collection of water samples (ver. 2.0, September 2006): U.S. Geological Survey Techniques of Water Resources Investigations, book 9, chap. A4 [variously paged]. [Also available at https://doi.org/10.3133/twri09A4.]

U.S. Geological Survey, 2020, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed April 20, 2020, at https://doi.org/10.5066/F7P55KJN.

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. plus 8 attachments. [Also available at https://doi.org/10.3133/tm1D3.]

YSI Incorporated, 2007, YSI 6136 turbidity sensor: YSI Incorporated, 2 p., accessed November 2019 at https://www.ysi.com/File%20Library/Documents/Specification%20Sheets/E56-6136-Turbidity-Sensor.pdf.

YSI Incorporated, 2012, 6-series multiparameter water quality sondes—User manual, revision J: YSI Incorporated, 379 p., accessed November 2019 at https://www.ysi.com/File%20Library/Documents/Manuals/069300-YSI-6-Series-Manual-RevJ.pdf.