Appendix 4. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07182510, Neosho River at Burlington, Kansas, during January 1, 2010, through October 16, 2015

This model archive summary summarizes the suspended-sediment concentration (SSC) model developed to compute hourly or daily SSC during January 1, 2010, through October 16, 2015. 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, and the policy and guidance for approval of surrogate regression models for computation of time series SSCs and loads (Rasmussen and others, 2009; U.S. Geological Survey, 2016).

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: A YSI 6600 water-quality monitor equipped with sensors for water temperature, specific conductance, and turbidity (YSI model 6136 turbidity sensor). The YSI 6600 water-quality monitor was in operation during February 08, 2007, through October 16, 2015.

Date model was developed: January 16, 2020
Model calibration data period: March 26, 2007, through September 23, 2015

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 34 concurrent measurements of discretely collected SSC samples and continuously measured turbidity during March 26, 2007, through September 23, 2015. Discrete samples were collected over a range of streamflows and turbidity conditions. 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 Sediment 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 five-point grain size by the USGS Sediment Laboratory in Iowa City, Iowa.

Continuous Data

Continuously monitored turbidity was measured using a YSI 6600 model 6136 turbidity sensor installed during February 8, 2007, through October 16, 2015 (U.S. Geological Survey, 2018). Concomitant turbidity values were time interpolated. If continuous data were not available (2 or more hours of specific conductance values bracketing the sample collected 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 neither concomitant continuous data nor field monitor data were available, the sample was not included in the dataset. The range of continuous turbidity data of the YSI model 6136 sensor (in formazin nephelometric units) was as follows: maximum 760; minimum 3.90; mean 39.5; median 28.0.

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).

Turbidity was 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 site 07182510:

SSC-based model:

$$\log_{10}(SSC) = 0.904 \times \log_{10}(Turb6136) + 0.316$$

where

SSC = suspended-sediment concentration, in milligrams per liter, and

*Turb*6136 = turbidity, YSI model 6136, 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 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 1.03 for this model and the formula for the retransformed model accounting for BCF is as follows:

$$SSC = 2.13 \times Turb6136^{0.904}$$

Suspended-Sediment Concentration Record

The SSC record that is being used in this regression model is stored at the National Real-Time Water Quality (NRTWQ) website (https://nrtwq.usgs.gov/ks).

Previously Published Model

No previously published model.

Model Statistics, Data, and Plots

Model

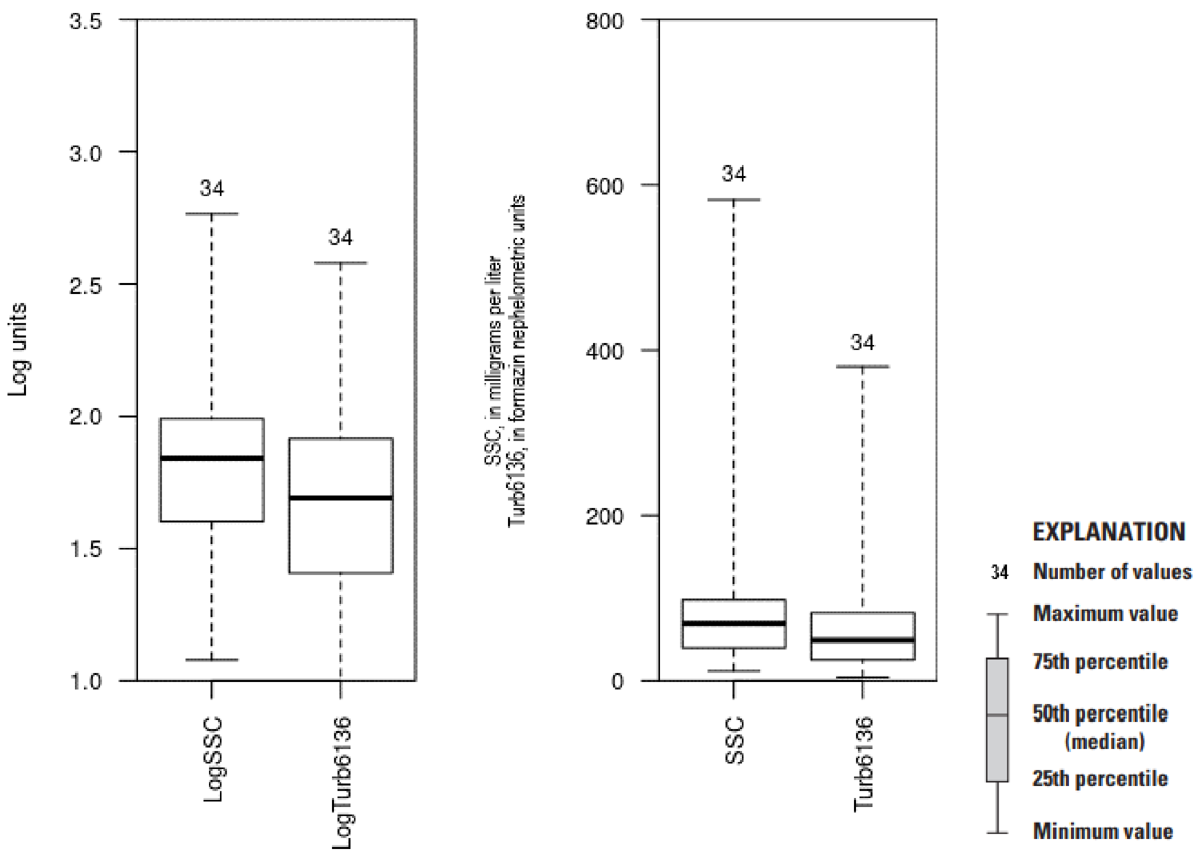
$$\text{Log}(SSC) = + 0.904 * \text{Log}(Turb6136) + 0.316$$

Variable Summary Statistics

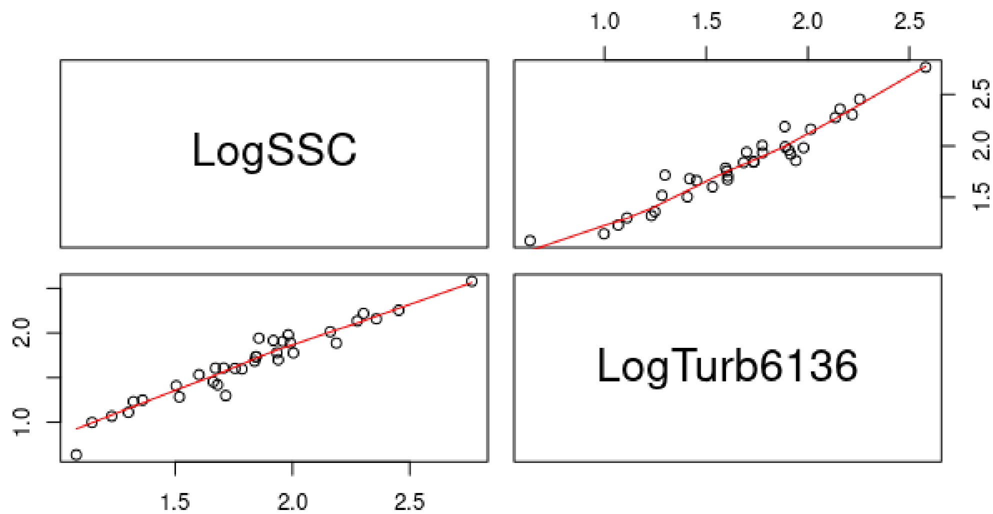
```
              LogSSC   SSC LogTurb6136 Turb6136
Minimum         1.08  12.0       0.633      4.3
1st Quartile    1.60  40.0       1.410     25.6
```

```
Median           1.84   69.5        1.690      49.2
Mean             1.81   96.5        1.660      68.3
3d Quartile      1.99   98.0        1.920      82.3
Maximum          2.76 582.0         2.580     380.0
```

Box Plots



Exploratory Plots



Basic Model Statistics

```
Number of Observations                                  34
Standard error (RMSE)                               0.0995
Average Model standard percentage error (MSPE)        23.1
Coefficient of determination (R²)                    0.934
Adjusted Coefficient of Determination (Adj. R²)      0.932
Bias Correction Factor (BCF)                          1.03
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)        0.316         0.0722    4.37 1.21e-04
LogTurb6136        0.904         0.0423   21.40 1.70e-20
```

Correlation Matrix

```
          Intercept E.vars
Intercept     1.000 -0.972
E.vars       -0.972  1.000
```

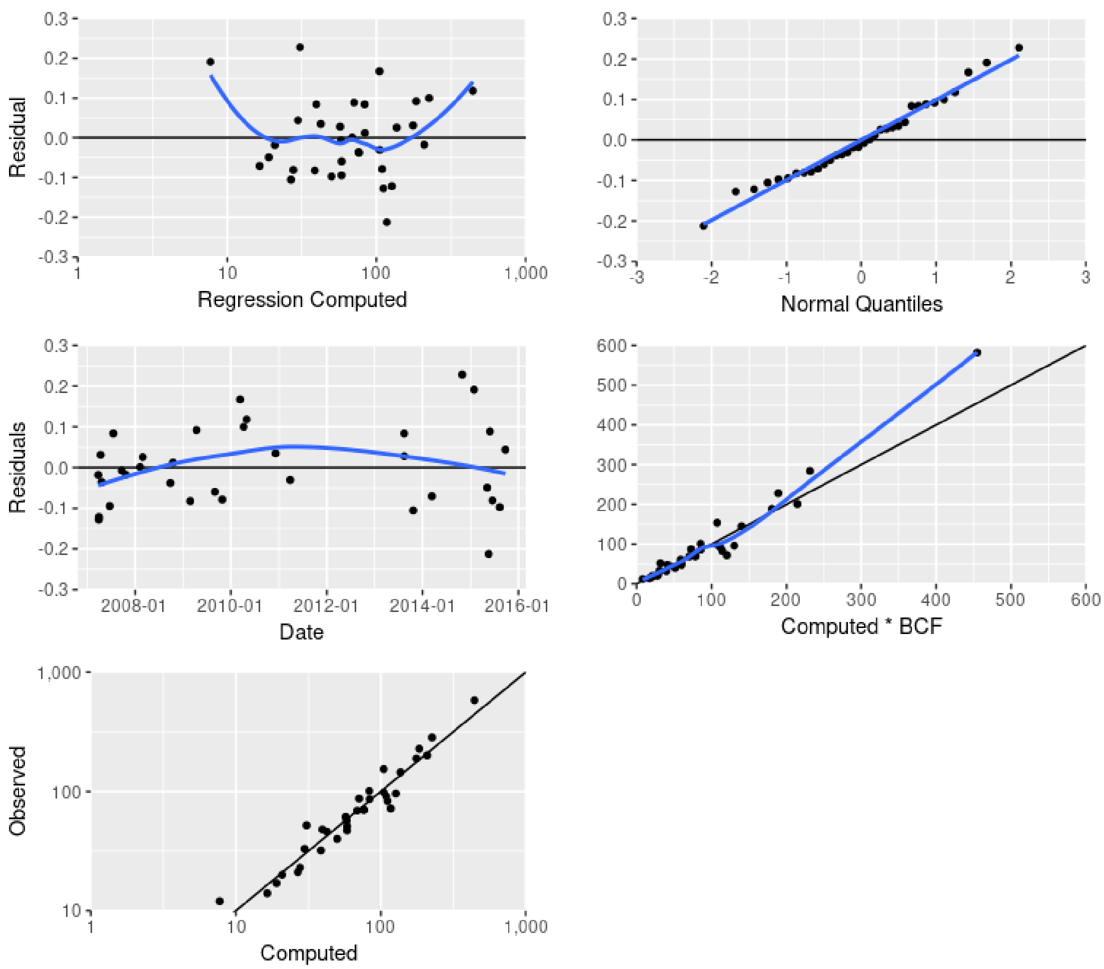
Outlier Test Criteria

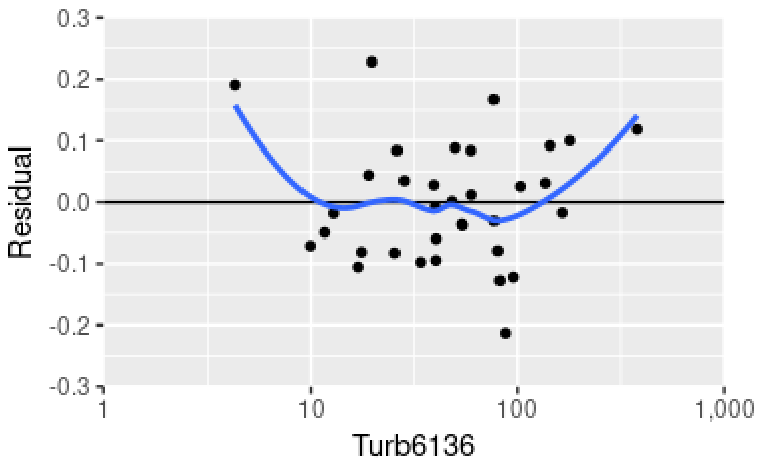
```
Leverage Cook's D   DFFITS
   0.176    0.194    0.485
```

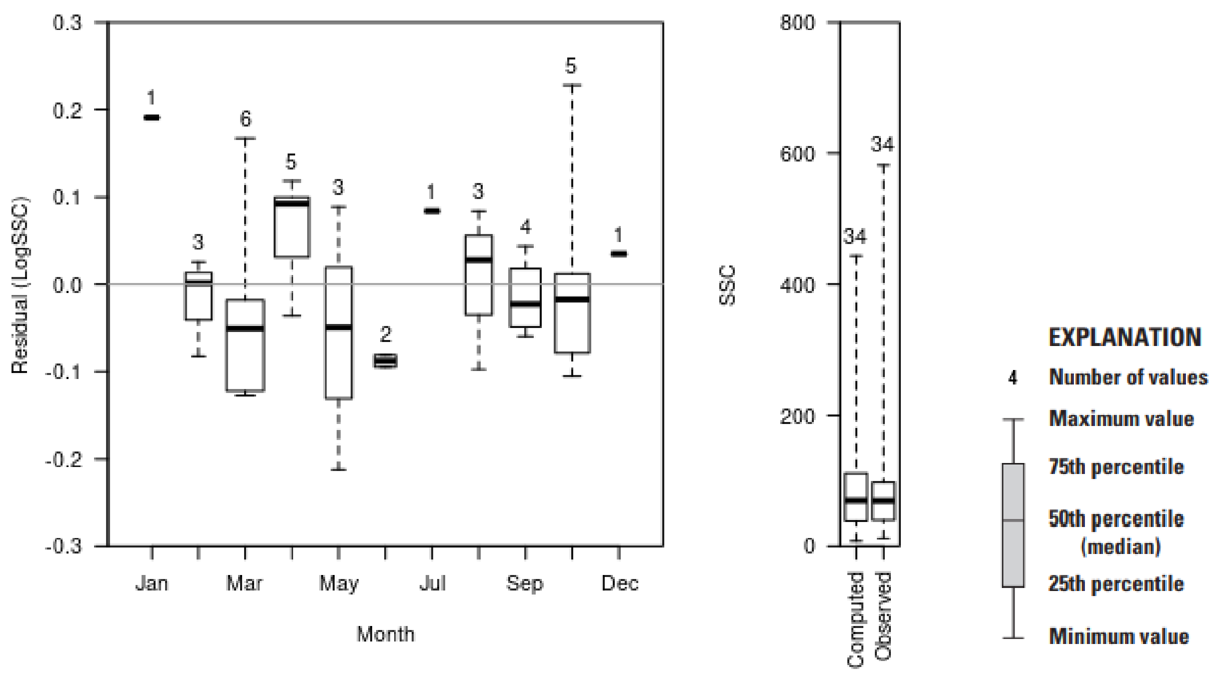
Flagged Observations

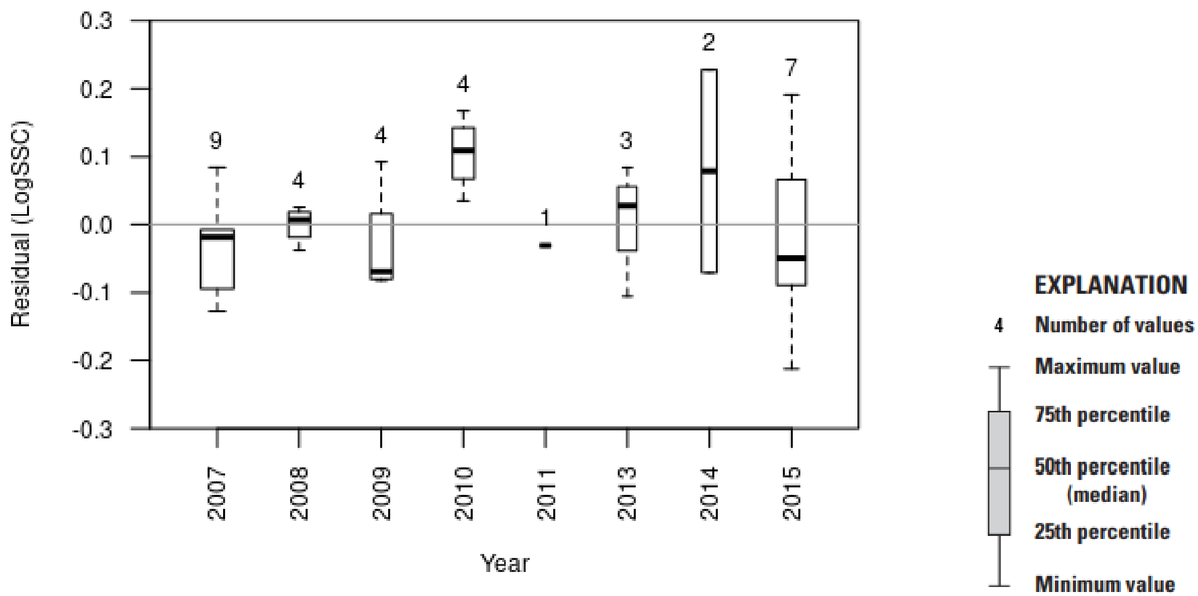
	LogSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
4/30/2010 11:15	2.76	2.650	0.118	1.31	1.33	0.1830	0.193	0.629
10/30/2014 12:10	1.72	1.490	0.228	2.35	2.55	0.0529	0.155	0.602
1/28/2015 13:00	1.08	0.888	0.191	2.17	2.31	0.2190	0.661	1.230
5/19/2015 16:50	1.86	2.070	-0.212	-2.18	-2.33	0.0439	0.109	-0.499

Statistical Plots



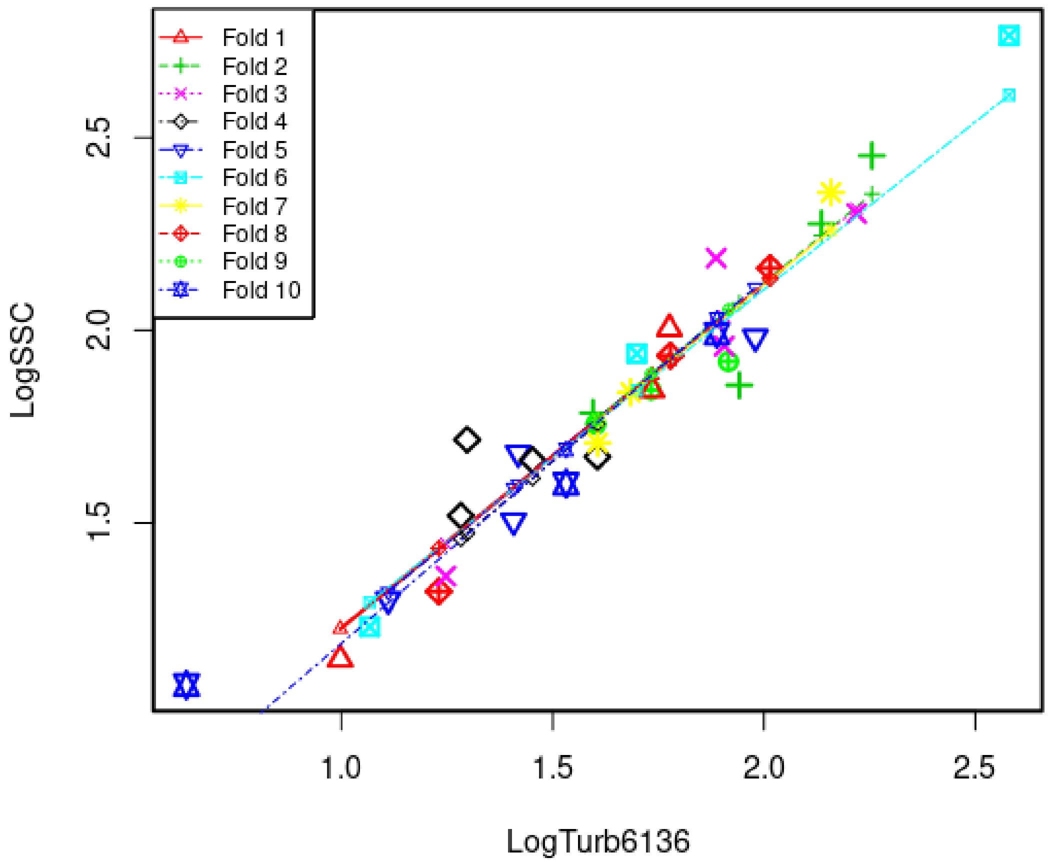

Residual
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
1
10
100
1,000
Turb6136


Residual (LogSSC)
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
1
3
6
5
3
2
1
3
4
5
1
Jan
Mar
May
Jul
Sep
Dec
Month
SSC
800
600
400
200
0
34
34
Computed
Observed
EXPLANATION
4 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

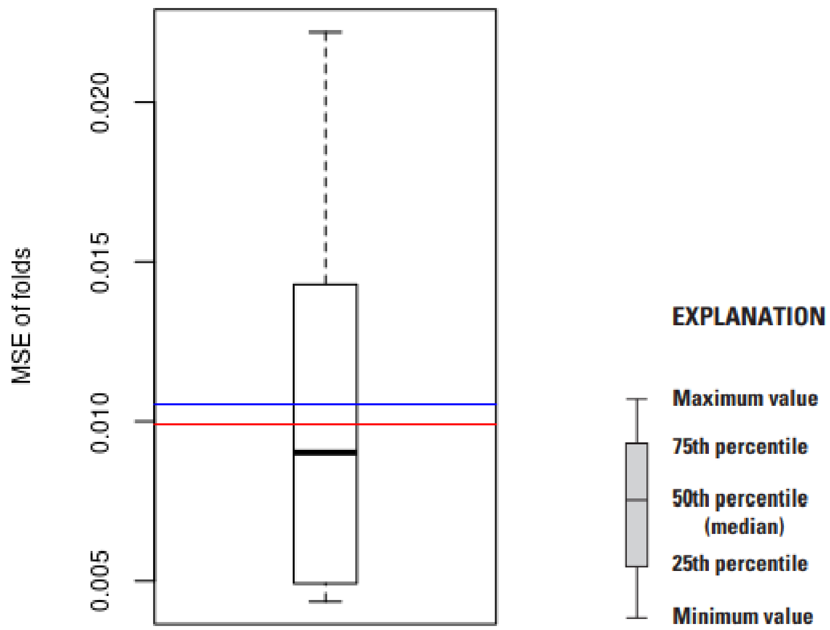


0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
Residual (LogSSC)
9
4
4
4
1
3
2
7
2007
2008
2009
2010
2011
2013
2014
2015
Year
EXPLANATION
4
Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Cross Validation




```
Minimum mean squared error (MSE) of folds:  0.00436
                        Mean MSE of folds:  0.01050
                      Median MSE of folds:  0.00902
                     Maximum MSE of folds:  0.02220
        (Mean MSE of folds) / (Model MSE):  1.06000
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Dataset

	Date	LogSSC	LogTurb6136	SSC	Turb6136	Computed LogSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
0										
1	2007-03-26	1.3	1.11	20	12.9	1.32	21.4	-0.0182	-0.11	--
2	2007-03-30	1.92	1.92	83	82.3	2.05	114	-0.128	-1.68	--
3	2007-03-31	1.98	1.98	96	95.3	2.1	130	-0.122	-1.43	--
4	2007-04-12	2.28	2.14	189	137	2.25	180	0.031	0.415	--
5	2007-04-20	1.85	1.73	70	54	1.88	78	-0.036	-0.259	--
6	2007-06-22	1.67	1.61	47	40.3	1.77	59.9	-0.0945	-0.979	--
7	2007-07-19	1.68	1.42	48	26.2	1.6	40.6	0.0839	0.765	--
8	2007-09-21	1.76	1.6	57	40	1.76	59.5	-0.00745	0.0367	--
9	2007-10-22	2.3	2.22	201	166	2.32	215	-0.0175	-0.0367	--
10	2008-02-08	1.84	1.68	69	48.3	1.84	70.6	0.00127	0.11	--
11	2008-02-28	2.16	2.01	145	103	2.14	140	0.0256	0.259	--
12	2008-09-25	1.85	1.73	70	54.3	1.88	78.4	-0.0382	-0.336	--
13	2008-10-15	1.93	1.78	86	60	1.92	85.8	0.0121	0.184	--
14	2009-02-24	1.51	1.41	32	25.6	1.59	39.7	-0.0824	-0.867	--
15	2009-04-13	2.36	2.16	228	144	2.27	189	0.092	0.979	--
16	2009-09-01	1.71	1.61	51	40.4	1.77	60	-0.0597	-0.496	--
17	2009-10-26	1.96	1.91	91	80.5	2.04	112	-0.0787	-0.67	--
18	2010-03-11	2.19	1.89	154	77	2.02	107	0.167	1.43	--
19	2010-04-08	2.45	2.26	284	180	2.35	231	0.0998	1.11	--
20	2010-04-30	2.76	2.58	582	380	2.65	455	0.118	1.25	--
21	2010-12-06	1.66	1.45	46	28.3	1.63	43.6	0.0347	0.496	--
22	2011-03-28	1.99	1.89	98	77.3	2.02	108	-0.0308	-0.184	--
23	2013-08-13	2	1.78	101	59.7	1.92	85.4	0.0839	0.67	--

24	2013-08-14	1.79	1.6	61	39.4	1.76	58.7	0.0279	0.336	--
25	2013-10-21	1.32	1.23	21	17	1.43	27.5	-0.105	-1.25	--
26	2014-03-12	1.15	0.997	14	9.94	1.22	16.9	-0.0709	-0.581	--
27	2014-10-30	1.72	1.3	52	19.8	1.49	31.6	0.228	2.11	--
28	2015-01-28	1.08	0.633	12	4.3	0.888	7.93	0.191	1.68	--
29	2015-05-07	1.23	1.07	17	11.7	1.28	19.5	-0.0494	-0.415	--
30	2015-05-19	1.86	1.94	72	87.3	2.07	120	-0.212	-2.11	--
31	2015-06-17	1.36	1.25	23	17.7	1.44	28.4	-0.0809	-0.765	--
32	2015-05-29	1.94	1.7	87	50	1.85	72.8	0.0886	0.867	--
33	2015-08-12	1.6	1.53	40	34	1.7	51.4	-0.0975	-1.11	--
34	2015-09-23	1.52	1.28	33	19.2	1.47	30.6	0.0439	0.581	--

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)

Turb6136: Turbidity, in formazin nephelometric units (63680)

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.]