Appendix 5. Model Archive Summary for Suspended-Sediment Concentration at U.S. Geological Survey Site 07182250, Cottonwood River near Plymouth, Kansas, during April 22, 2015, through December 31, 2019

This model archive summary summarizes the suspended-sediment concentration (SSC) model developed to compute hourly or daily SSC during April 22, 2015, through December 31, 2019. 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: 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: A YSI EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, and turbidity. The YSI EXO2 water-quality monitor recorded readings every 15 minutes and transmitted hourly via satellite using a Sutron Satlink 2 high data rate collection platform. The EXO2 water-quality monitor began operation during April 22, 2015.

Date model was created: January 16, 2020
Model calibration data period: May 18, 2015, 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 (sine and cosine variables) were also evaluated as explanatory variables.

The regression model is based on 23 measurements of discretely collected SSC samples and continuously measured turbidity collected during May 18, 2015, through March 13, 2019. Samples were collected over a range of streamflow 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. The discrete sample on February 25, 2016, met the outlined outlier criteria however it was not excluded from the dataset as it was determined there was nothing anomalous about the sample collection or analysis. There were no other samples in this dataset deemed outliers. One sample was removed from the dataset (January 29, 2018) because the sample bottle was damaged during shipment to the U.S. Geological Survey Iowa Sediment Laboratory in Iowa City, Iowa.

Suspended-Sediment Sampling Details

Discrete samples were collected from the downstream side 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 eight 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, or a D–96 bag sampler, a DH–81 with a Teflon bottle, cap, and nozzle hand sampler. 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

Turbidity was measured using a YSI EXO turbidity sensor installed during April 22, 2015, through December 31, 2019 (U.S. Geological Survey, 2018). Concomitant turbidity values were time interpolated. If continuous data were not available (2 or more hours of turbidity 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 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 EXO2 sensor (in formazin nephelometric units) was as follows: maximum 1,640; minimum 0.00; mean 60.2; median 25.0.

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 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 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 SSC regression analysis at site 07182250:

SSC-based model:

$$\mathrm{Log}_{10}(SSC) = 1.04 \text{ x } \mathrm{Log}_{10}(Turb\mathrm{EXO}) + 0.371$$

where

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

$Turb$EXO = turbidity, YSI model EXO, in formazin nephelometric units.

The use of turbidity as an explanatory variable is appropriate physically and statistically. In a physical sense, particles comprised of suspended solids scatter light which affects turbidity. In a statistical sense using turbidity resulted in a model with a low standard error and high R^2 values. 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 1.06 for this model and the formula for the retransformed model accounting for BCF is as follows:

$$SSC = 2.49 \text{ x } Turb\mathrm{EXO}^{1.04}$$

Suspended-Sediment Concentration Record

The data are computed at 15-minute intervals and the computed 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

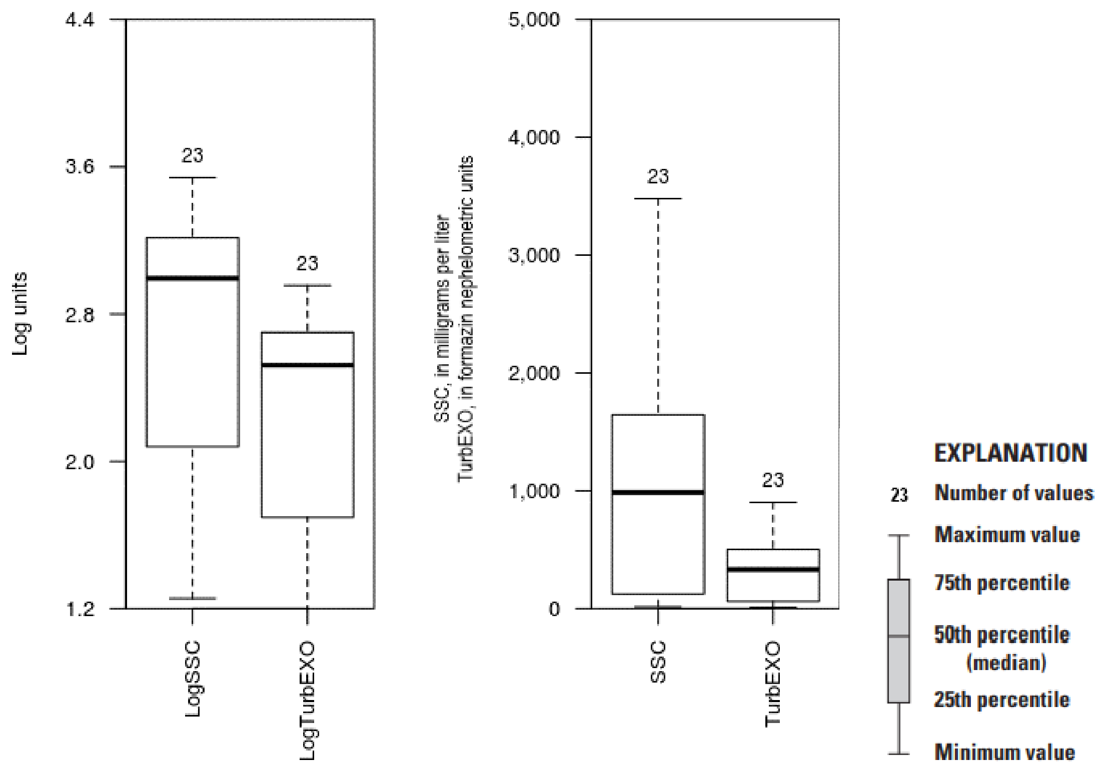
Model

$$\mathrm{Log}(SSC) = + 1.04 * \mathrm{Log}(Turb\mathrm{EXO}) + 0.371$$

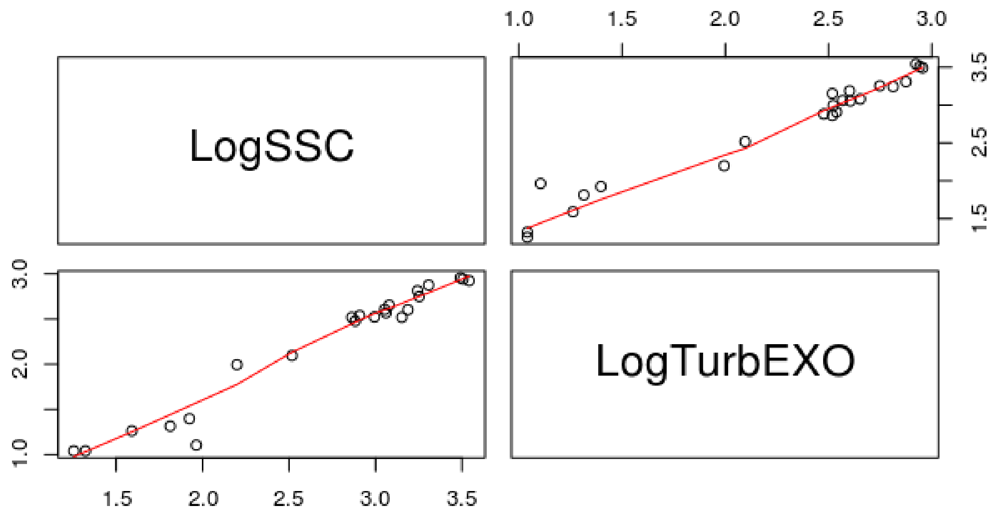
Variable Summary Statistics

	LogSSC	SSC	LogTurbEXO	TurbEXO
Minimum	1.26	18	1.04	11
1st Quartile	1.96	92	1.40	25
Median	2.99	987	2.52	333
Mean	2.70	1130	2.24	355
3d Quartile	3.24	1750	2.75	560
Maximum	3.54	3480	2.96	903

Box Plots



Exploratory Plots



Basic Model Statistics

Number of Observations	23
Standard error (RMSE)	0.145
Average Model standard percentage error (MSPE)	34.1
Coefficient of determination (R^2)	0.961
Adjusted Coefficient of Determination (Adj. R^2)	0.959
Bias Correction Factor (BCF)	1.06

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.371	0.1070	3.48	2.22e-03
LogTurbEXO	1.040	0.0456	22.80	2.70e-16

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.959
E.vars	-0.959	1.000

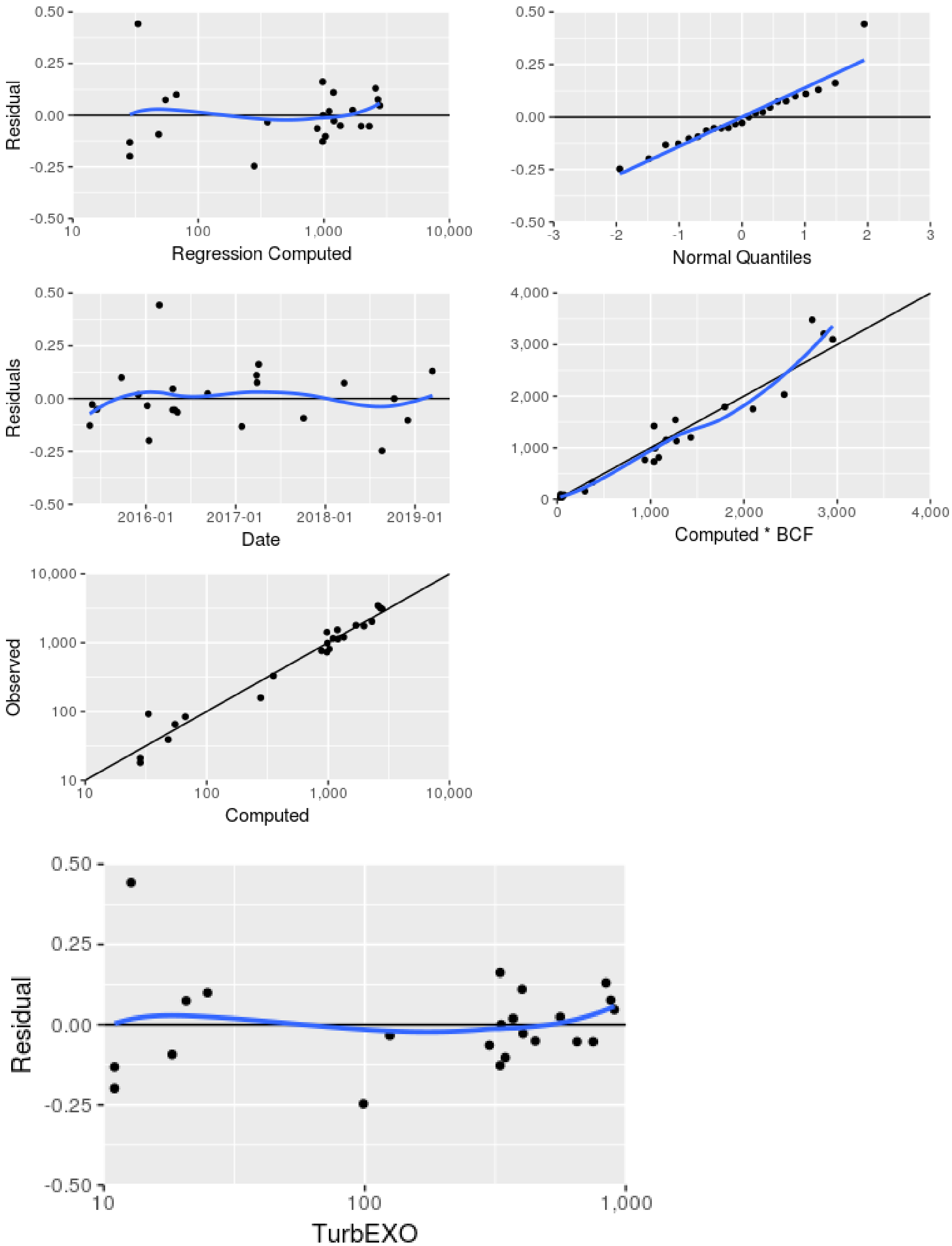
Outlier Test Criteria

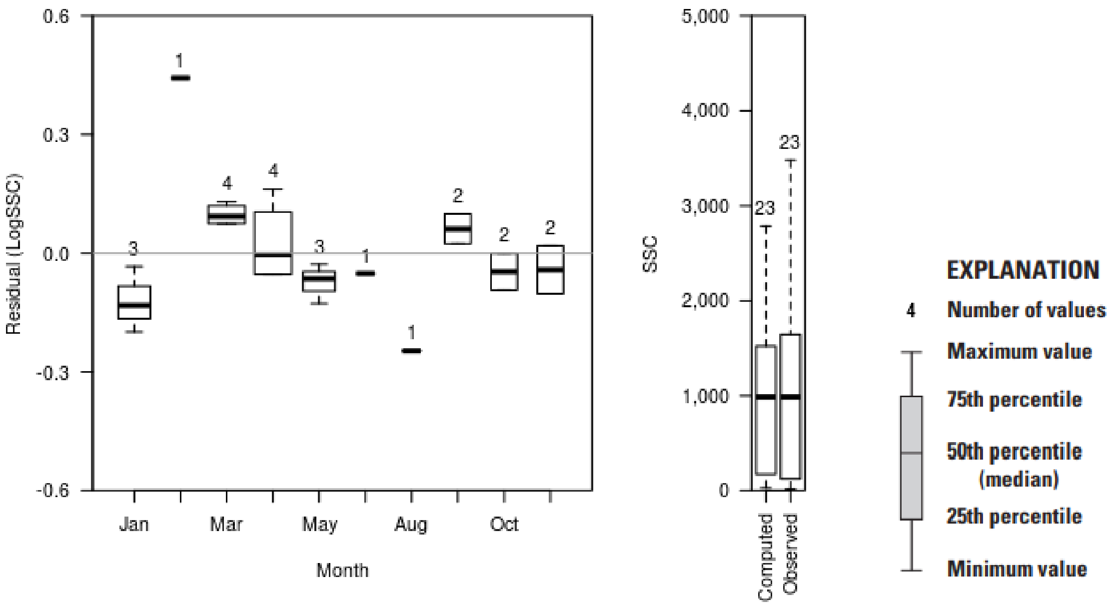
Leverage	Cook's D	DFFITS
0.261	0.193	0.590

Flagged Observations

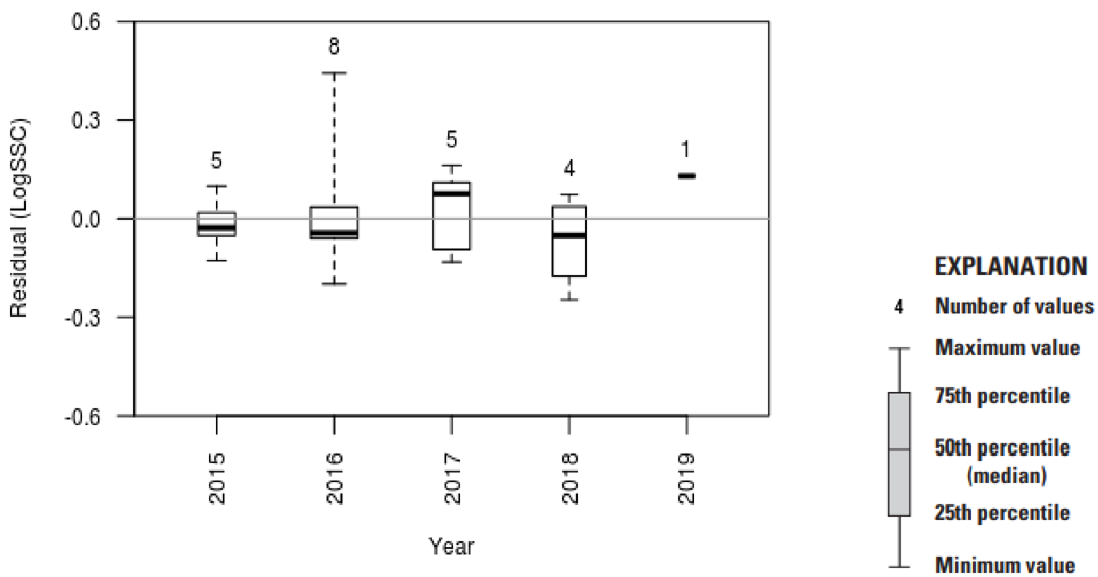
	LogSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
1/14/2016 14:50	1.26	1.45	-0.199	-1.52	-1.57	0.185	0.261	-0.747
2/25/2016 12:00	1.96	1.52	0.443	3.35	4.78	0.170	1.150	2.170

Statistical Plots



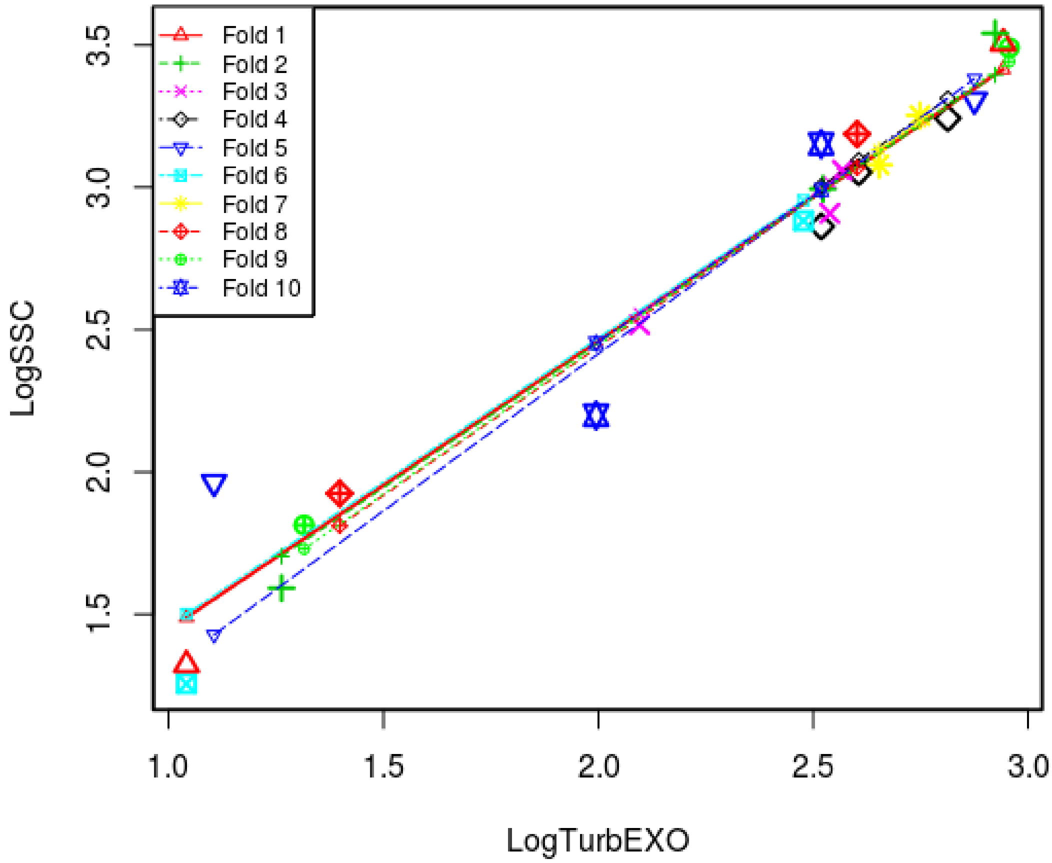


Residual (LogSSC)
0.6
0.3
0.0
-0.3
-0.6
1
4
4
3
3
1
2
1
2
2
Jan
Mar
May
Aug
Oct
Month
SSC
5,000
4,000
3,000
2,000
1,000
0
23
23
Computed
Observed
EXPLANATION
4 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

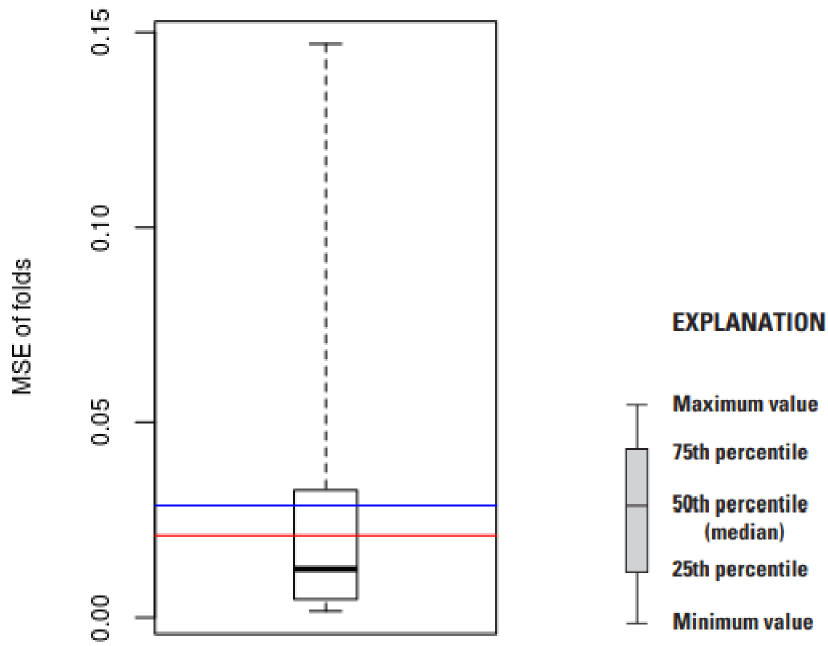


Residual (LogSSC)
0.6
0.3
0.0
-0.3
-0.6
5
8
5
4
1
2015
2016
2017
2018
2019
Year
EXPLANATION
4 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Cross Validation




```
Minimum mean squared error (MSE) of folds:  0.00164
                        Mean MSE of folds:  0.02870
                      Median MSE of folds:  0.01240
                     Maximum MSE of folds:  0.14700
         (Mean MSE of folds) / (Model MSE):  1.36000
```



Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Dataset

0	Date	LogSSC	LogTurbEXO	SSC	TurbEXO	Computed LogSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
1	2015-05-18	2.86	2.52	729	330	2.99	1040	-0.127	-1.01	--
2	2015-06-16	3.08	2.65	1200	450	3.13	1430	-0.0511	-0.218	--
3	2015-05-28	3.05	2.61	1130	403	3.08	1280	-0.0277	0	--
4	2015-09-24	1.92	1.4	84	25	1.82	70.8	0.0994	0.848	--
5	2015-12-02	3.06	2.57	1150	370	3.04	1170	0.0188	0.218	--
6	2016-01-07	2.52	2.1	329	125	2.55	377	-0.0335	-0.108	--
7	2016-01-14	1.26	1.04	18	11	1.45	30.2	-0.199	-1.48	--
8	2016-02-25	1.96	1.11	92	12.8	1.52	35.2	0.443	1.95	--
9	2016-04-20	3.49	2.96	3100	903	3.44	2950	0.0464	0.447	--
10	2016-04-20	3.31	2.88	2030	750	3.36	2430	-0.0535	-0.447	--
11	2016-04-27	3.24	2.81	1750	650	3.3	2100	-0.0533	-0.33	--
12	2016-05-09	2.88	2.48	763	300	2.95	938	-0.0646	-0.57	--
13	2016-09-09	3.25	2.75	1790	560	3.23	1800	0.0238	0.33	--
14	2017-01-26	1.32	1.04	21	11	1.45	30.2	-0.132	-1.22	--
15	2017-03-27	3.19	2.6	1540	400	3.08	1270	0.11	1.01	--
16	2017-03-29	3.51	2.94	3210	875	3.43	2860	0.076	0.702	--
17	2017-04-05	3.15	2.52	1420	330	2.99	1040	0.162	1.48	--
18	2017-10-05	1.59	1.26	39	18.3	1.68	51.2	-0.0931	-0.702	--
19	2018-03-19	1.81	1.32	65	20.7	1.74	58.1	0.074	0.57	--
20	2018-08-20	2.2	1.99	158	98.8	2.45	295	-0.247	-1.95	--
21	2018-10-09	2.99	2.52	987	333	2.99	1050	-6.05e-05	0.108	--
22	2018-12-03	2.91	2.54	810	346	3.01	1090	-0.103	-0.848	--
23	2019-03-13	3.54	2.92	3480	838	3.41	2730	0.13	1.22	--

Definitions

```
Adj R²: 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

R²: Coefficient of determination

RMSE: Root mean square error

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

TurbEXO: Turbidity, in formazin nephelometric units (63680)
```

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.]

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.]