Model Archive Summary for Suspended Sediment Concentration at Station 01479820; Red Clay Creek near Kennett Square, Pennsylvania [2020 Version]

This model archive summary describes the regression model developed to estimate continuous instantaneous (15-minute) suspended sediment concentrations starting in 2015.

Site and Model Information

U.S. Geological Survey (USGS) station number: 01479820

Station name: Red Clay Creek near Kennett Square, Pennsylvania

Lat 39°49'00", long 75°41'31" referenced to North American Datum of 1927, Chester County, PA, Hydrologic Unit 02040205, on left bank along SR82 (Creek Road), and 3.0 mi south of the intersection of SR82 and U.S. Highway 1 at Kennett Square.

Date regression model was created: June 2020.

Period of data for model calibration: February 16, 2016 – August 15, 2019.

Model application date: December 2015 onward.

Database Reviewed by: Elizabeth A. Hittle, May 2020

Computed by: Matthew C. Gyves, June 2020.

Approved by: Joseph W. Duris, Water-Quality Specialist, June 2020.

Equipment

A Yellow Spring Instrument (YSI) 600OMS monitor equipped with sensors for temperature, specific conductance, and turbidity (6136 sensor) is deployed at this site. The monitor is housed in a 3-inch perforated plastic pipe placed in the stream about 2 ft from the left bank. Readings from the sensors are recorded every 15 minutes and transmitted hourly by way of GOES satellite.

Model Calibration Dataset

All data were collected using standard USGS protocols and are stored in the National Water Information System (NWIS) database. Linear regression models were developed using RStudio Version 1.2.1335. Primary packages included (dataRetrieval, v2.7.5; dplyr, v0.8.3; plotly v4.9.0, and the Model Archive Summary Shiny App (available at https://github.com/PatrickEslick/ModelArchiveSummary). Explanatory variables evaluated as inputs to linear regression were turbidity and streamflow.

The final regression model is based on 38 concurrent measurements of suspended sediment and turbidity concentrations. Suspended sediment concentrations were determined from analysis of discrete samples, and turbidity concentrations were determined from continuous record of 15-minute values, interpolated when necessary to correspond with collection time of the discrete sample for sediment analysis. Suspended sediment samples were collected over a range of hydrologic conditions throughout each year during sampling period. Studentized residuals for final model were inspected and considered for potential removal as outliers if residual values were greater than 3 or less than -3. None of the samples met these criteria, and no samples were removed from the dataset.

Data Collection

Discrete samples for suspended sediment analysis were collected using an ISCO automatic sampler which was set to sample based on stage and turbidity thresholds to provide data for a broad range of turbidity values.

Water was pulled by the sampler peristaltic pump through 0.5-inch tubing from an intake located near the turbidity probe and stored in 1-liter polypropylene bottles. The timing of sample collection was determined using thresholds for both stream stage and turbidity. Over time, these thresholds were changed to collect samples under different conditions, covering the range of measured values. Once collected, sample times were recorded in the field during bottle retrieval as reported by the ISCO. Samples were analyzed for suspended sediment concentration (SSC) in the laboratory at the USGS office in Exton, Pennsylvania and the USGS Kentucky Sediment Laboratory in Louisville, Kentucky using standard USGS methods. Suspended sediment concentrations ranged from 31 to 2,220 mg/L in the 38 analyzed samples.

Turbidity data were collected according to USGS Techniques and Methods 1-D3 (Wagner and Others, 2006). Turbidity concentrations ranged from 24 to 920 formazin nephelometric units (FNU) during collection of concurrent discrete suspended sediment samples used for model calibration.

Model Development

Regression analysis was done using R by examining turbidity (*Turb*) and streamflow (*Q*) as explanatory variables for SSC. A variety of linear regression models that predict *SSC* and $\log_{10}(SSC)$ were evaluated and computed using methods described by Rasmussen and others, 2011.

The model with $\log_{10}(Turb)$ as the explanatory variable was selected as the best predictor of $\log_{10}(SSC)$ on the basis of residual plots, maximizing adjusted coefficient of determination (adjusted R^2), and minimizing model residual standard error (or root mean square error, *RMSE*) and standard percentage error (*MSPE*). The addition of discharge as a variable showed minimal improvement of model performance.

Model Summary

Final regression model for suspended sediment concentration (SSC) at site number 01479820:

$$\log_{10} SSC = 1.14 \log_{10} Turb - .0658$$

where

SSC = suspended sediment concentration in milligrams per liter (mg/L) (parameter 80154);
Turb = turbidity in formazin nephelometric units (FNU) (parameter 63680);

Turb makes physical and statistical sense as explanatory variable for *SSC* because previous studies showed suspended sediment concentrations were directly related to turbidity concentrations. The transformed model may be retransformed to the original units so that SSC concentrations can be calculated directly. A potential bias that is introduced because of retransformation can be corrected using Duan's bias correction factor (BCF). For this model the BCF is 1.03. The retransformed model, using the BCF, is:

$$SSC = .885\, Turb^{1.14}$$

References

Rasmussen, P.P., Gray, J.R., Glysson, G.D., and Ziegler, A.C., 2009, revised 2011, 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.

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 1–D3, 51 p. + 8 attachments; accessed April 10, 2006, at http://pubs.water.usgs.gov/tm1d3

Model Statistics, Data, and Plots

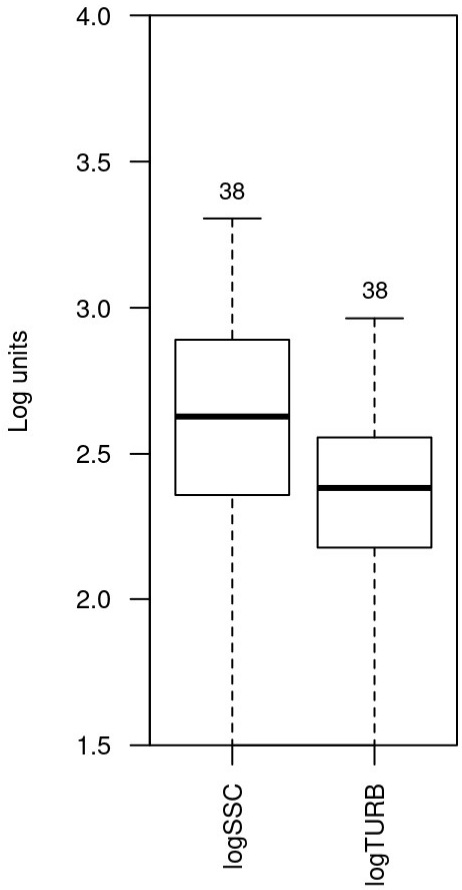
Model

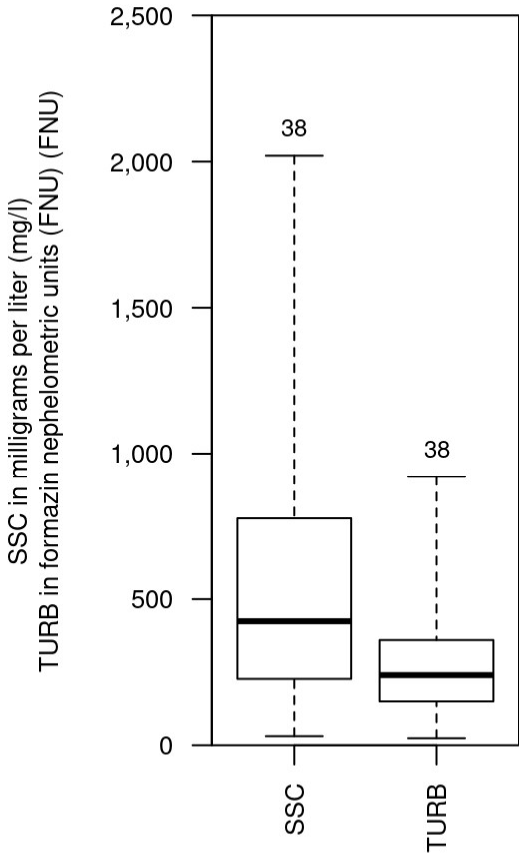
logSSC = + 1.14 * logTURB - 0.0658

Variable Summary Statistics

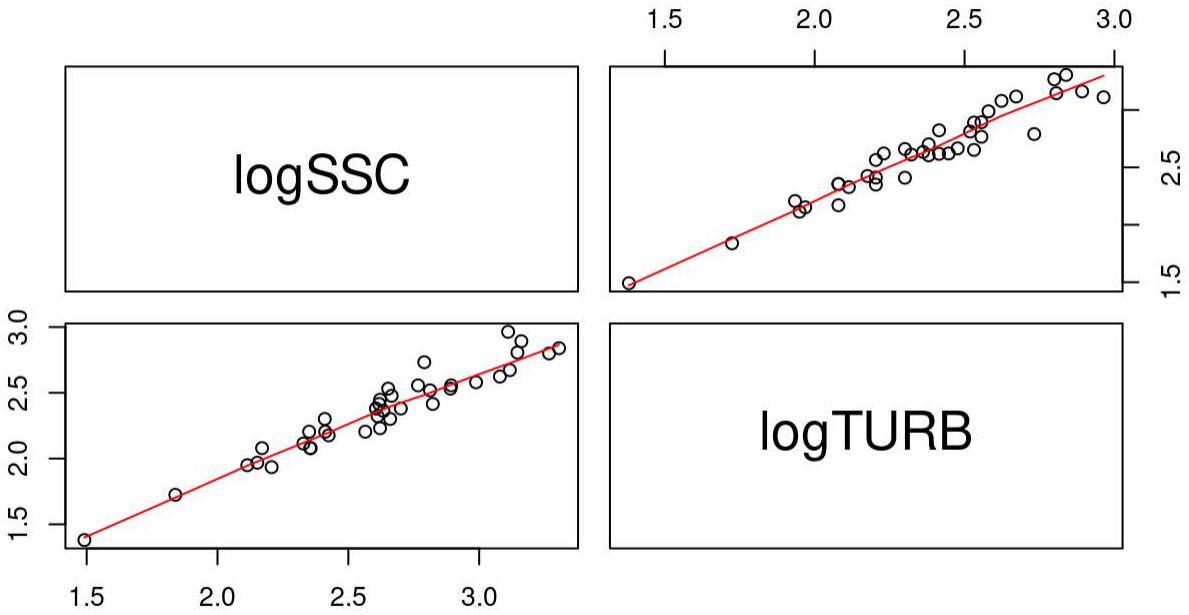
	logSSC	SSC	logTURB	TURB
Minimum	1.49	31	1.38	24
1st Quartile	2.36	227	2.18	150
Median	2.63	425	2.38	240
Mean	2.62	596	2.36	296
3rd Quartile	2.89	777	2.56	360
Maximum	3.31	2020	2.96	920

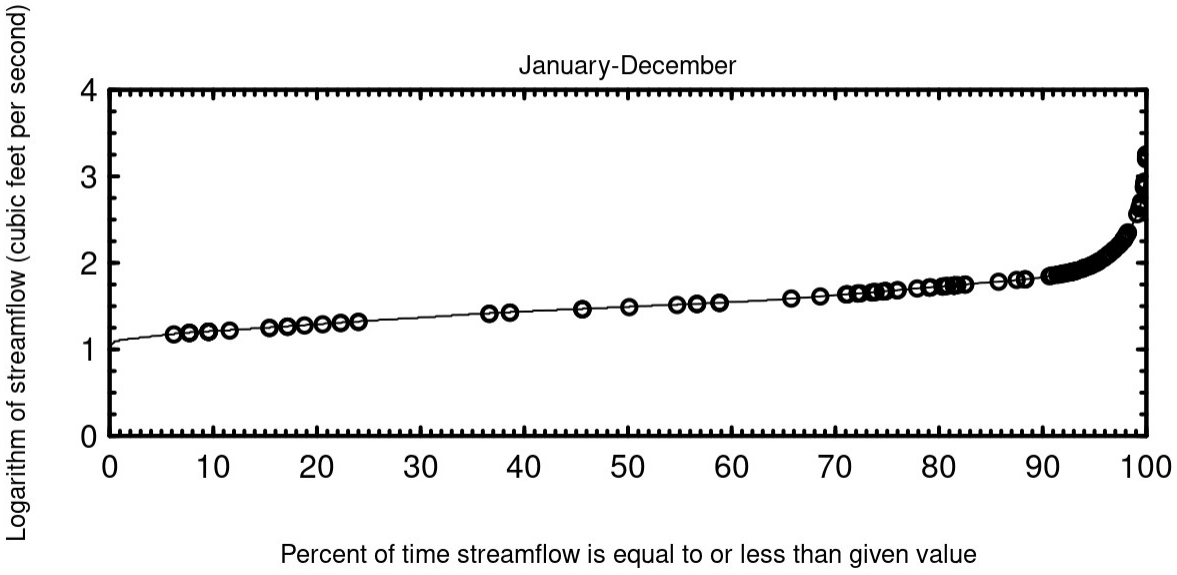
Box Plots

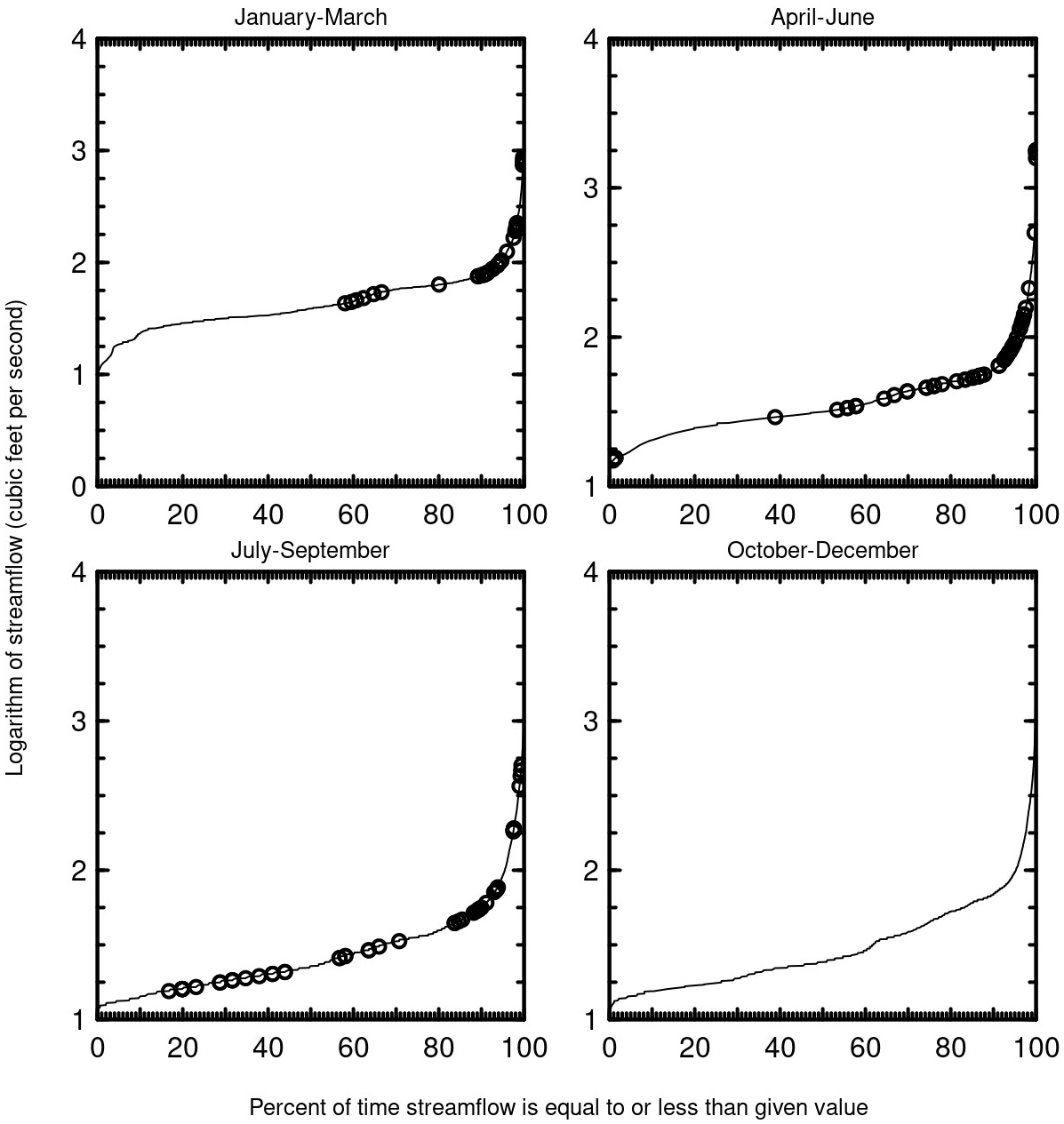

4.0
3.5
3.0
2.5
2.0
1.5
Log units
38
38
logSSC
logTURB


2,500
2,000
1,500
1,000
500
0
SSC in milligrams per liter (mg/l)
TURB in formazin nephelometric units (FNU) (FNU)
38
38
SSC
TURB

Exploratory Plots







Basic Model Statistics

Number of Observations	38
Standard error (RMSE)	0.106
Average Model standard percentage error (MSPE)	24.8
Coefficient of determination (R^2)	0.929
Adjusted Coefficient of Determination (Adj. R^2)	0.927
Bias Correction Factor (BCF)	1.03

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.0658	0.1250	-0.525	6.03e-01
logTURB	1.1400	0.0525	21.700	2.97e-22

Correlation Matrix

	Intercept	E.vars
Intercept	1.00	-0.99
E.vars	-0.99	1.00

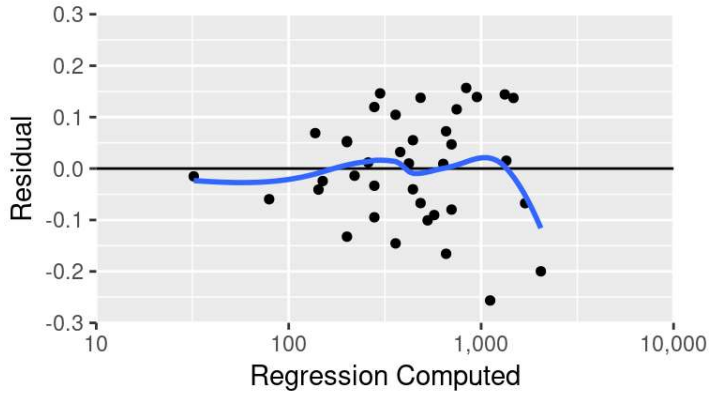
Outlier Test Criteria

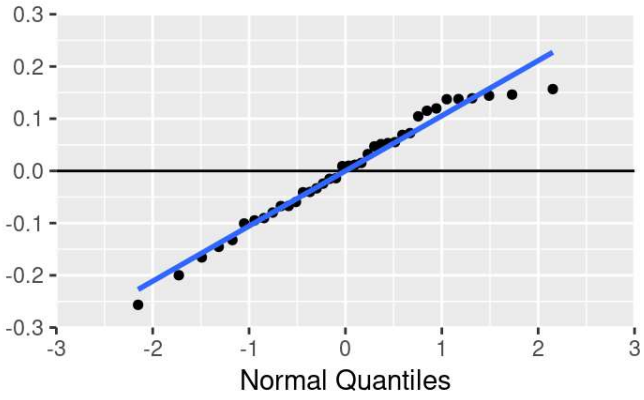
Leverage	Cook's D	DFFITS
0.158	0.194	0.459

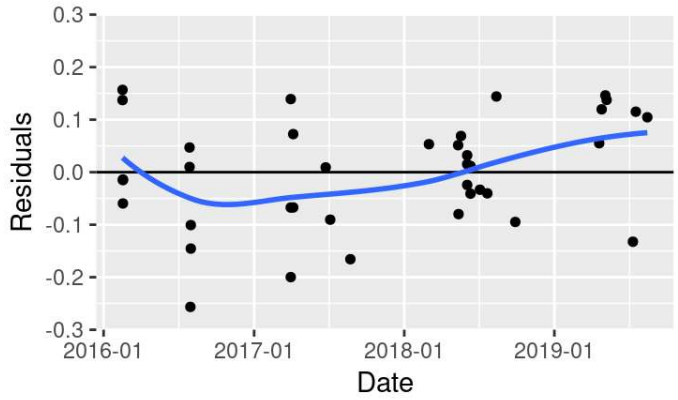
Flagged Observations

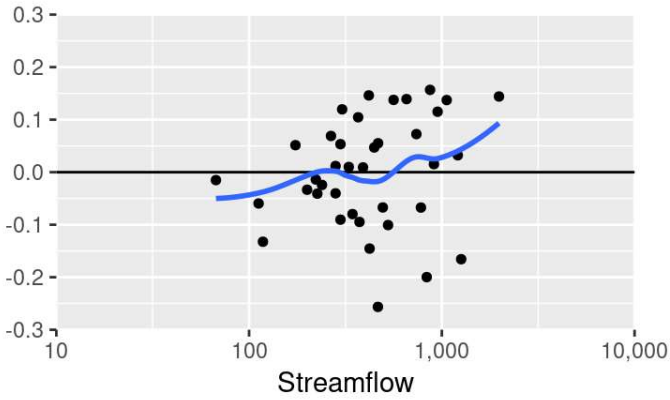
	logSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2016-02-17 12:00	1.49	1.51	-0.0151	-0.165	-0.163	0.2610	0.00479	-0.0965
2016-07-30 18:45	2.79	3.05	-0.2560	-2.480	-2.690	0.0598	0.19600	-0.6790
2017-03-31 16:30	3.11	3.31	-0.2000	-1.990	-2.080	0.1150	0.25700	-0.7500

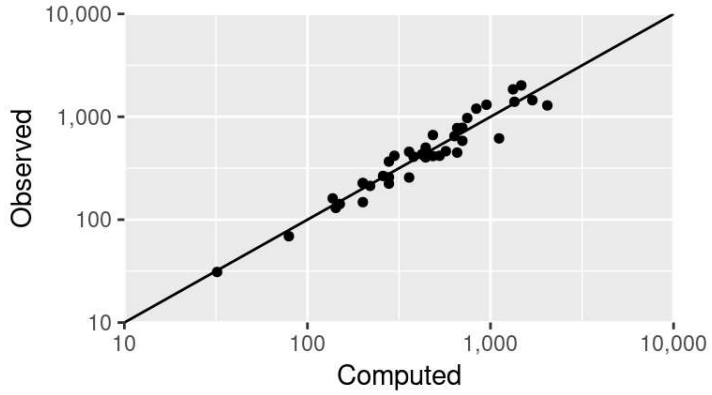
Statistical Plots

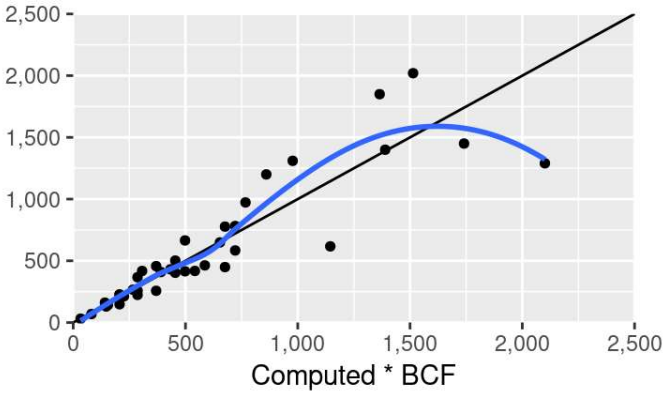


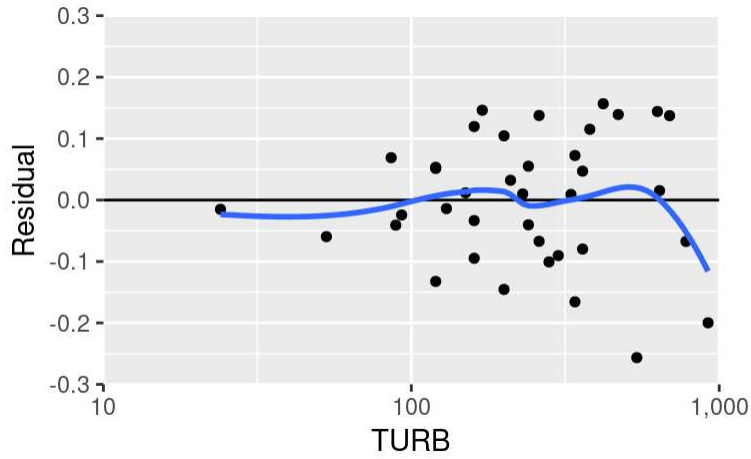


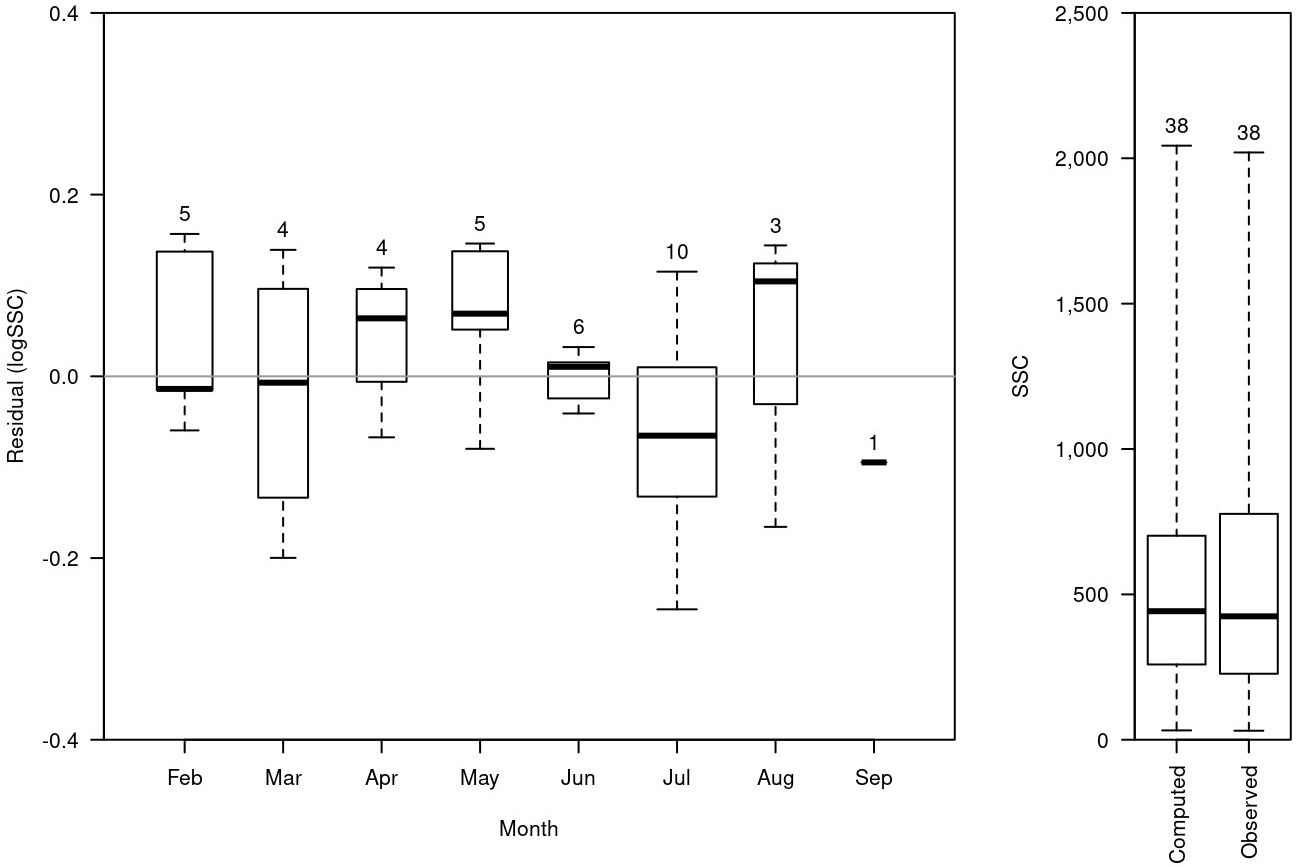


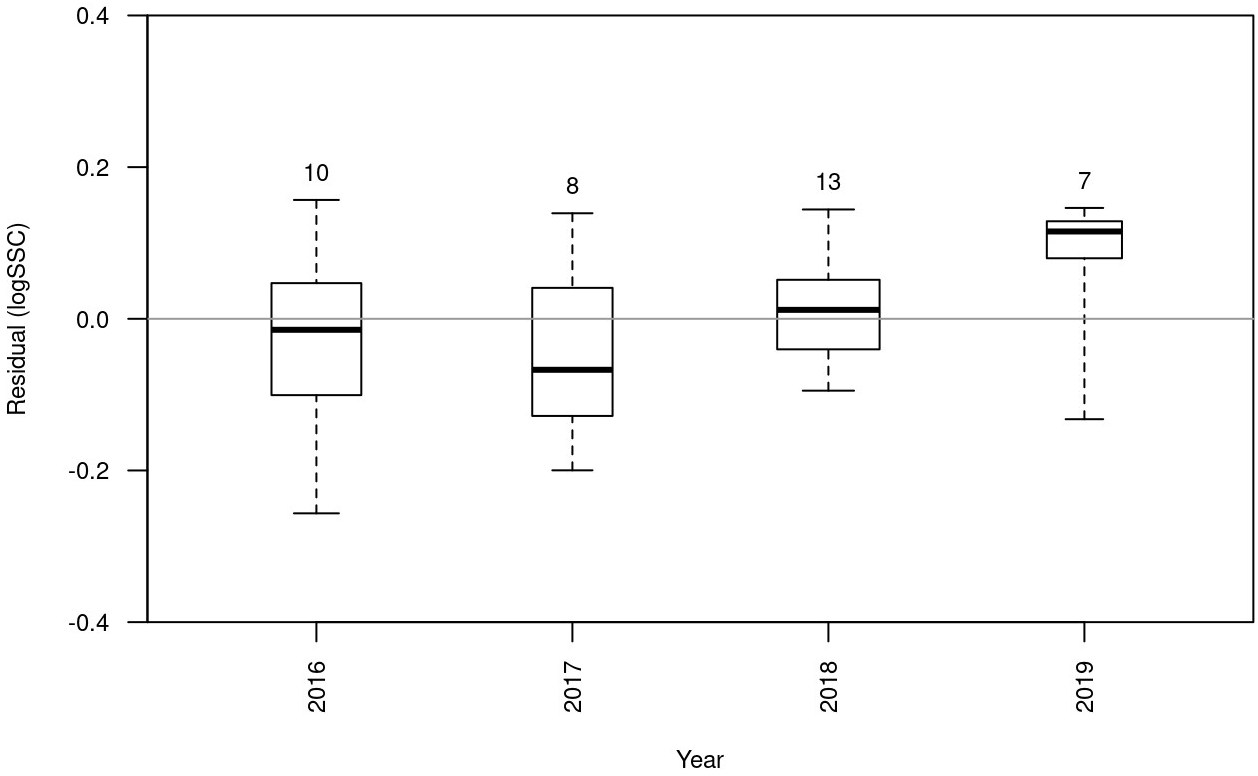




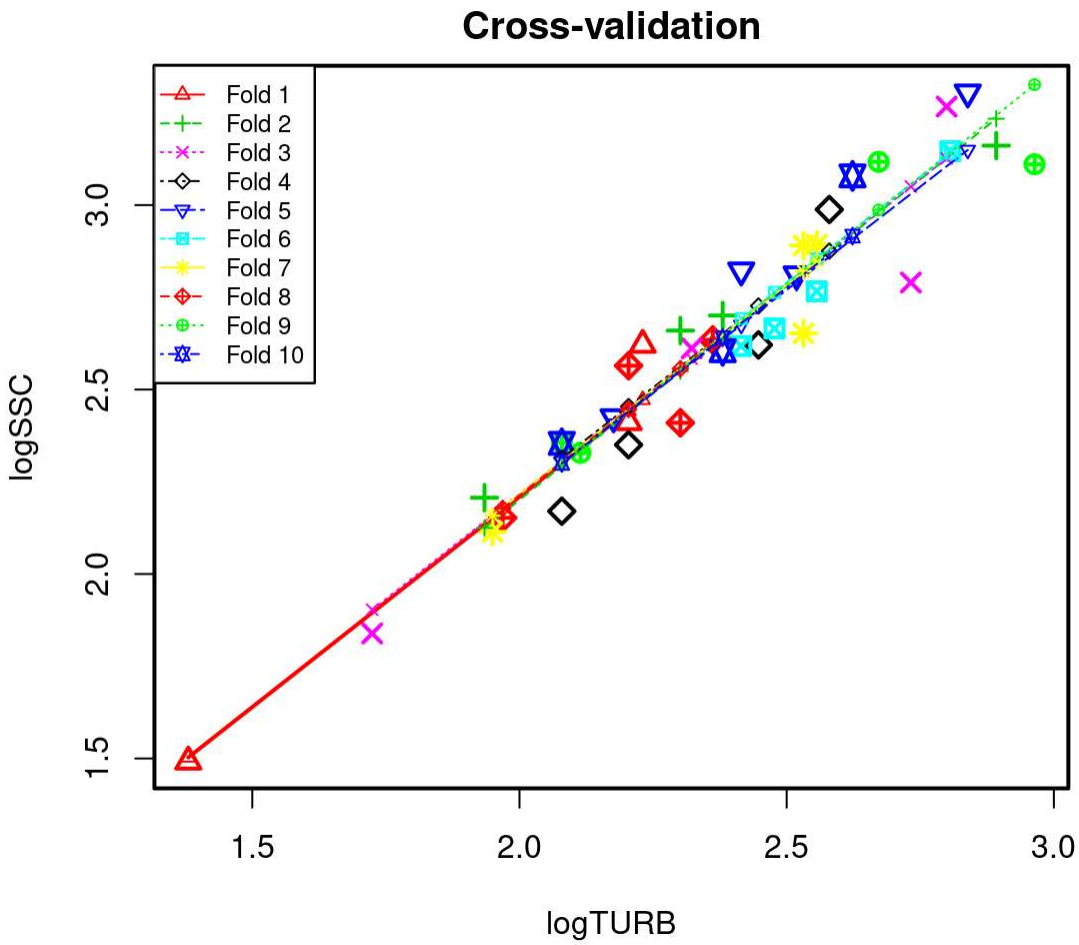




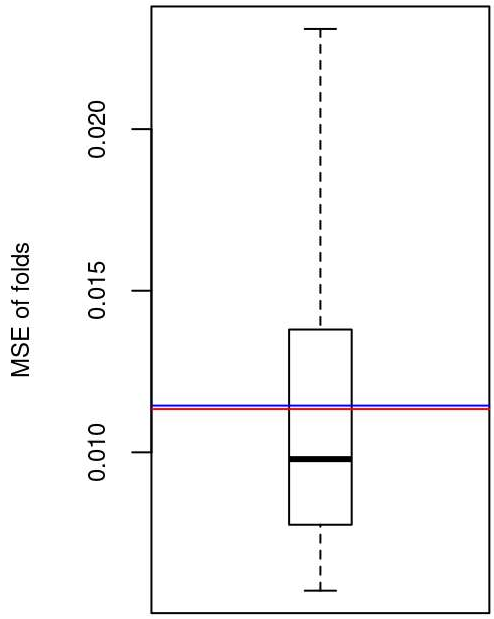

0.4
0.2
0.0
-0.2
-0.4
Residual (logSSC)
5
4
4
5
6
10
3
1
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Month
2,500
2,000
1,500
1,000
500
0
SSC
38
38
Computed
Observed


0.4
0.2
0.0
-0.2
-0.4
Residual (logSSC)
10
8
13
7
2016
2017
2018
2019
Year

Cross Validation




```
                Minimum MSE of folds:  0.00572
                   Mean MSE of folds:  0.01140
                 Median MSE of folds:  0.00979
                Maximum MSE of folds:  0.02310
(Mean MSE of folds) / (Model MSE):  1.01000
```



Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Data Set

	Date	logSSC	logTURB	SSC	TURB	Computed logSSC	Computed SSC	Residual	Normal Quantiles	Censored Values
1	2016-02-16	3.31	2.84	2020	690	3.17	1510	0.137	1.05	--
2	2016-02-16	3.08	2.62	1200	420	2.92	860	0.157	2.15	--
3	2016-02-17	2.33	2.11	213	130	2.34	226	-0.0139	-0.0986	--
4	2016-02-17	1.84	1.72	69	53	1.9	81.4	-0.0595	-0.514	--
5	2016-02-17	1.49	1.38	31	24	1.51	33	-0.0151	-0.165	--
6	2016-07-28	2.89	2.56	782	360	2.85	721	0.047	0.3	--
7	2016-07-28	2.63	2.36	431	230	2.62	433	0.00994	0.0328	--
8	2016-07-30	2.79	2.73	617	540	3.05	1140	-0.256	-2.15	--
9	2016-07-31	2.62	2.45	418	280	2.72	542	-0.101	-1.05	--
10	2016-07-31	2.41	2.3	257	200	2.56	369	-0.145	-1.32	--
11	2017-03-31	3.12	2.67	1310	470	2.98	977	0.139	1.32	--

12	2017-03-31	3.16	2.89	1450	780	3.23	1740	-0.0673	-0.67	--
13	2017-03-31	3.11	2.96	1290	920	3.31	2100	-0.2	-1.73	--
14	2017-04-06	2.89	2.53	777	340	2.82	676	0.0725	0.67	--
15	2017-04-06	2.62	2.41	415	260	2.69	498	-0.0671	-0.59	--
16	2017-06-24	2.81	2.52	649	330	2.8	653	0.00911	-0.0328	--
17	2017-07-05	2.67	2.48	463	300	2.76	586	-0.0904	-0.845	--
18	2017-08-23	2.65	2.53	449	340	2.82	676	-0.166	-1.49	--
19	2018-03-02	2.36	2.08	227	120	2.3	206	0.0533	0.44	--
20	2018-05-12	2.35	2.08	226	120	2.3	206	0.0514	0.369	--
21	2018-05-13	2.77	2.56	584	360	2.85	721	-0.0798	-0.755	--
22	2018-05-19	2.21	1.93	161	86	2.14	141	0.069	0.59	--
23	2018-06-03	3.15	2.81	1400	640	3.13	1390	0.0153	0.165	--
24	2018-06-03	2.61	2.32	409	210	2.58	390	0.0322	0.232	--
25	2018-06-03	2.15	1.97	142	93	2.18	154	-0.0243	-0.232	--
26	2018-06-11	2.42	2.18	266	150	2.41	266	0.0118	0.0986	--
27	2018-06-11	2.11	1.95	130	89	2.15	147	-0.0409	-0.44	--
28	2018-07-04	2.41	2.2	258	160	2.45	286	-0.0334	-0.3	--
29	2018-07-22	2.61	2.38	403	240	2.65	455	-0.0403	-0.369	--
30	2018-08-13	3.27	2.8	1850	630	3.12	1360	0.144	1.49	--
31	2018-09-28	2.35	2.2	224	160	2.45	286	-0.0948	-0.943	--
32	2019-04-20	2.7	2.38	502	240	2.65	455	0.0551	0.514	--
33	2019-04-26	2.56	2.2	367	160	2.45	286	0.12	0.943	--
34	2019-05-05	2.62	2.23	418	170	2.47	307	0.146	1.73	--
35	2019-05-08	2.82	2.41	665	260	2.69	498	0.138	1.17	--
36	2019-07-11	2.17	2.08	148	120	2.3	206	-0.132	-1.17	--
37	2019-07-18	2.99	2.58	973	380	2.87	767	0.115	0.845	--
38	2019-08-15	2.66	2.3	457	200	2.56	369	0.105	0.755	--

Definitions

SSC: Suspended sediment concentration (SSC) in mg/l (80154)

TURB: Turbidity in FNU (63680)

App Version 1.0