

Model Archive Summary for Suspended Sediment Concentration at Station 014800870; East Branch Brandywine Creek below Downingtown, Pennsylvania [2020 Version]

This model archive summary describes the regression model developed to estimate continuous instantaneous (15-minute) suspended sediment concentrations starting in 2007. The model was updated in March 2020, using all available calibration data. The previous model (Sloto and Olsen, 2011) was in use from October 1, 2008 to September 30, 2019. The new model is used to compute suspended sediment concentrations beginning in October 2019.

Site and Model Information

U.S. Geological Survey (USGS) station number: 01480870

Station name: East Branch Brandywine Creek below Downingtown, Pennsylvania

Lat 39°58'07", long 75°40'25" referenced to North American Datum of 1927, Chester County, PA, Hydrologic Unit 02040205, on left bank at downstream side of Sugars Bridge (U.S. Highway 322), 2,000 ft upstream from Valley Creek, 1.5 mi north of Marshallton, and 3.3 mi southeast of Downingtown.

Date regression model was created: June 2020.

Period of data for model calibration: August 21, 2007 – July 12, 2019.

Model application date: October 2019 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) 6920V2 monitor equipped with sensors for temperature, specific conductance, pH, dissolved oxygen, and turbidity (6136 sensor) is deployed at this site. The monitor is housed in a 4-inch perforated plastic pipe placed in the stream about 5 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 76 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. Although several samples met these criteria; other regression

diagnostics show their influence on model performance to be minimal, and no samples were removed from the dataset.

Prior Model

This model supersedes the prior model that was originally published in Sloto and Olsen (2011).

$$\log_{10}SSC = 1.0115 \log_{10}Turb + .2596$$

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. Over time, these thresholds were changed to collect samples under different conditions, covering the range of measured 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. 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 21 to 1,390 mg/L in the 76 analyzed samples.

Turbidity data were collected according to USGS Techniques and Methods 1-D3 (Wagner and Others, 2006). Turbidity concentrations ranged from 10 to 790 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 on model performance.

Model Summary

Final regression model for suspended sediment concentration (SSC) at site number 01480870:

$$\log_{10}SSC = 0.955 \log_{10}Turb + 0.343$$

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 1.05. The retransformed model, using the BCF, is:

$$SSC = 2.313 Turb^{.955}$$

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.

Sloto, R.A., and Olson, L.E., Estimated suspended-sediment loads and yields in the French and Brandywine Creek Basins, Chester County, Pennsylvania, water years 2008–09: U.S. Geological Survey Scientific Investigations Report 2011–5109, 31 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

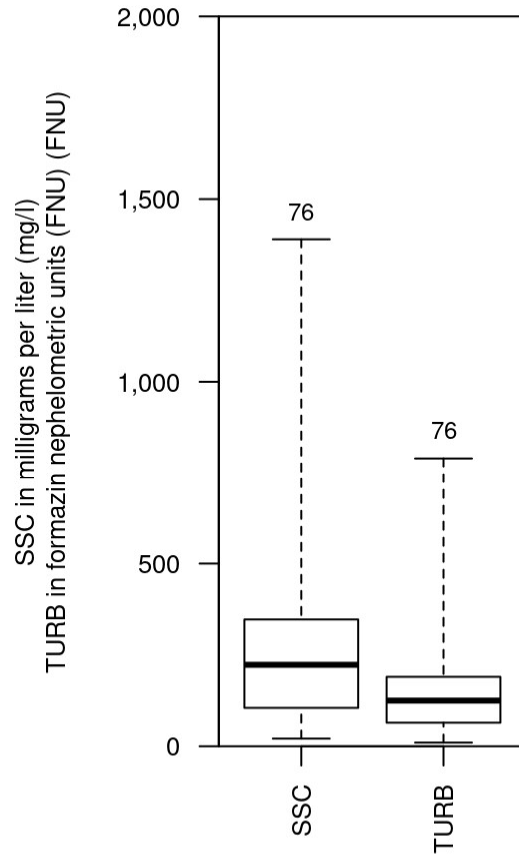
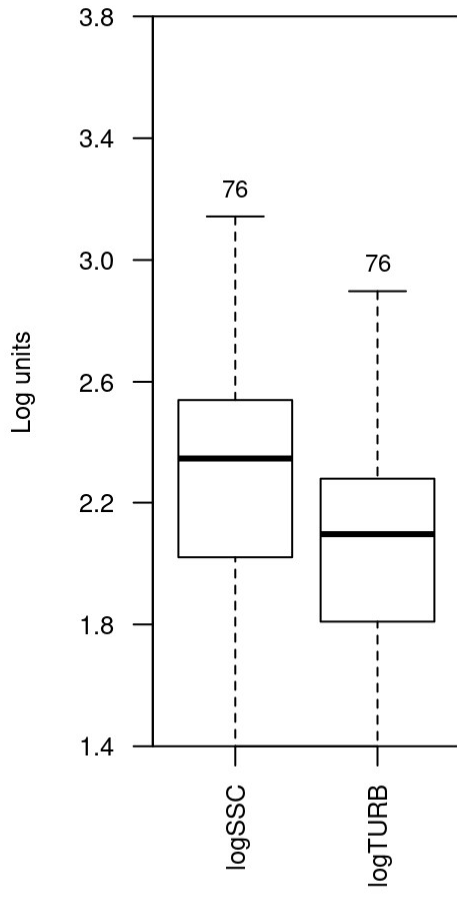
Model

$$\log\text{SSC} = + 0.955 * \log\text{TURB} + 0.343$$

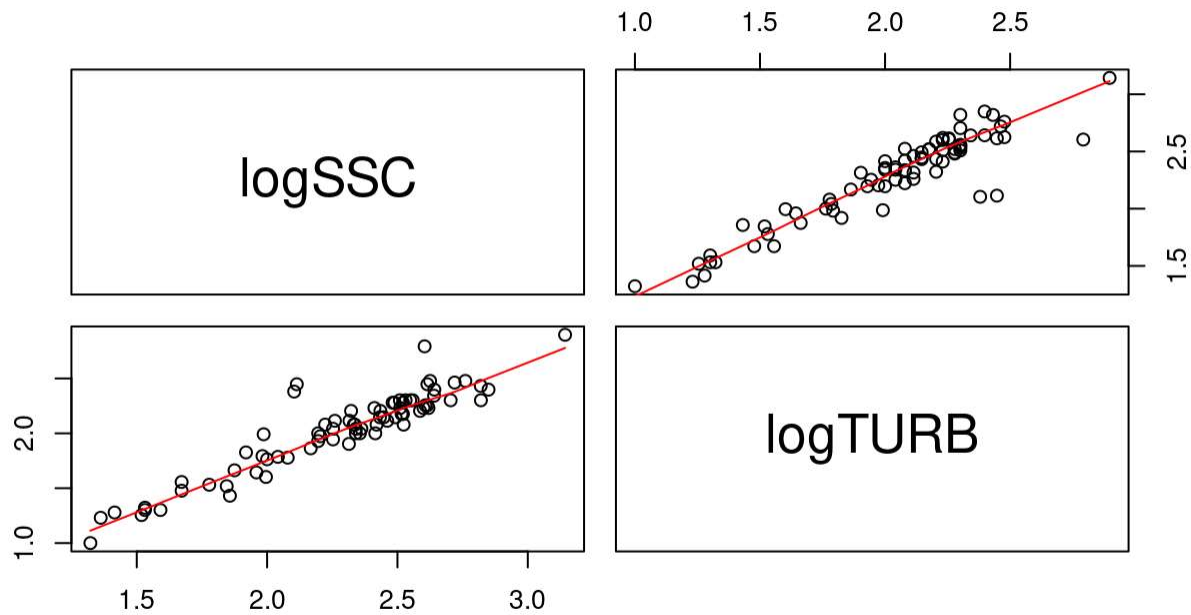
Variable Summary Statistics

	logSSC	SSC	logTURB	TURB
Minimum	1.32	21	1.00	10.0
1st Quartile	2.02	105	1.81	64.5
Median	2.35	223	2.10	125.0
Mean	2.28	262	2.03	146.0
3rd Quartile	2.54	347	2.28	190.0
Maximum	3.14	1390	2.90	790.0

Box Plots



Exploratory Plots



Basic Model Statistics

Number of Observations	76
Standard error (RMSE)	0.142
Average Model standard percentage error (MSPE)	33.2
Coefficient of determination (R^2)	0.868
Adjusted Coefficient of Determination (Adj. R^2)	0.866
Bias Correction Factor (BCF)	1.05

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t)
(Intercept)	0.343	0.0893	3.84	2.59e-04
logTURB	0.955	0.0434	22.00	3.09e-34

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.983
E.vars	-0.983	1.000

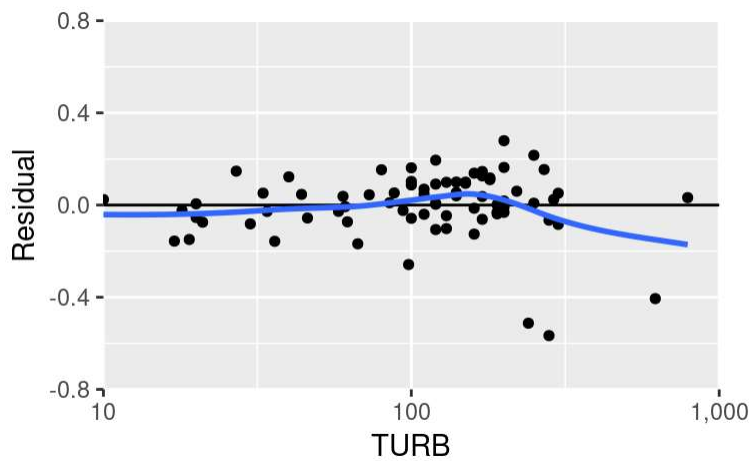
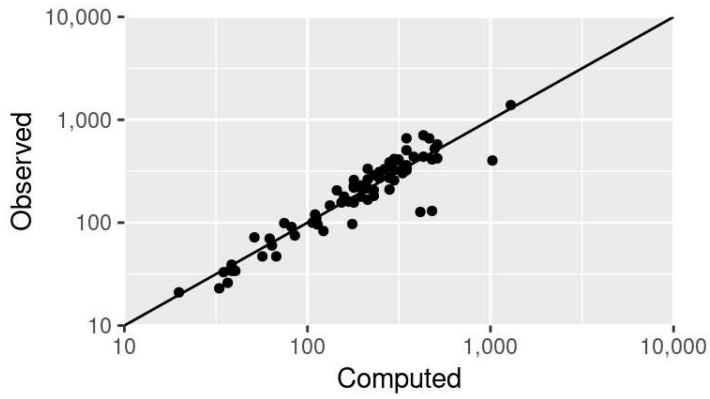
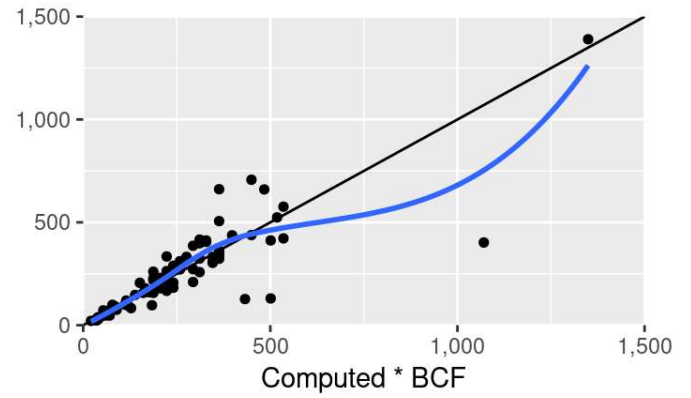
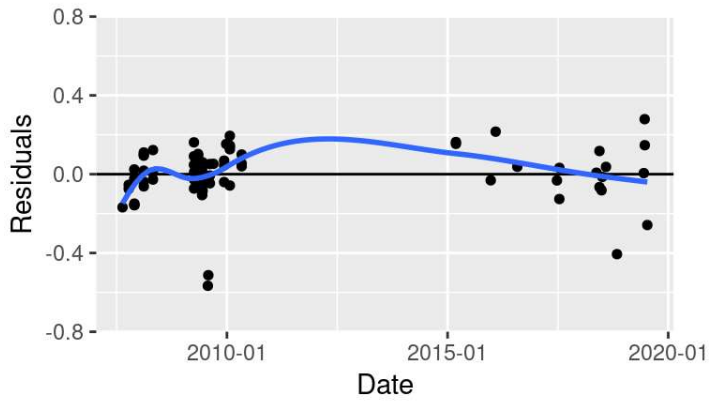
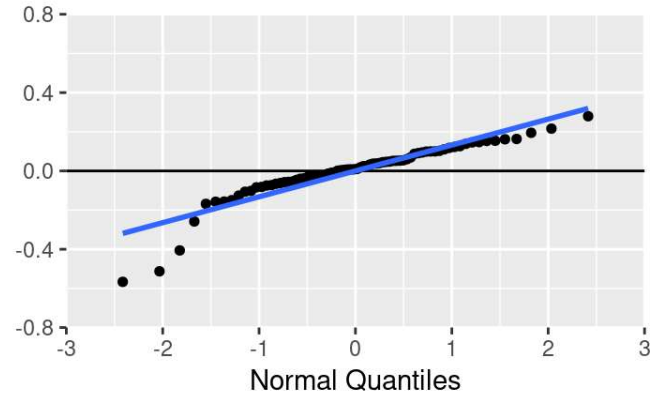
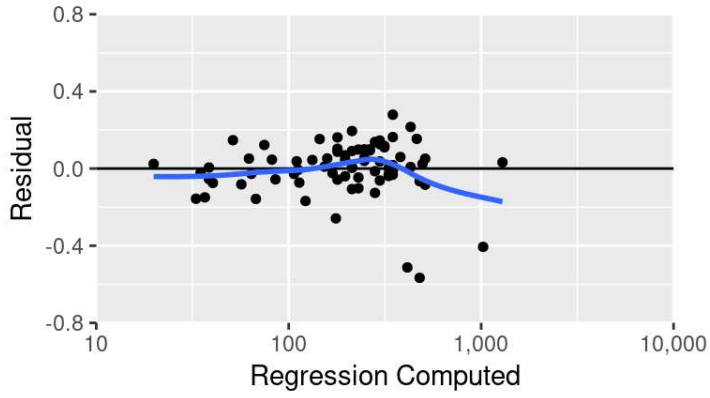
Outlier Test Criteria

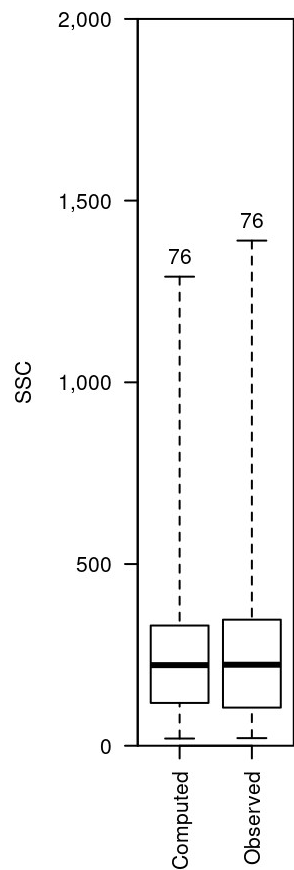
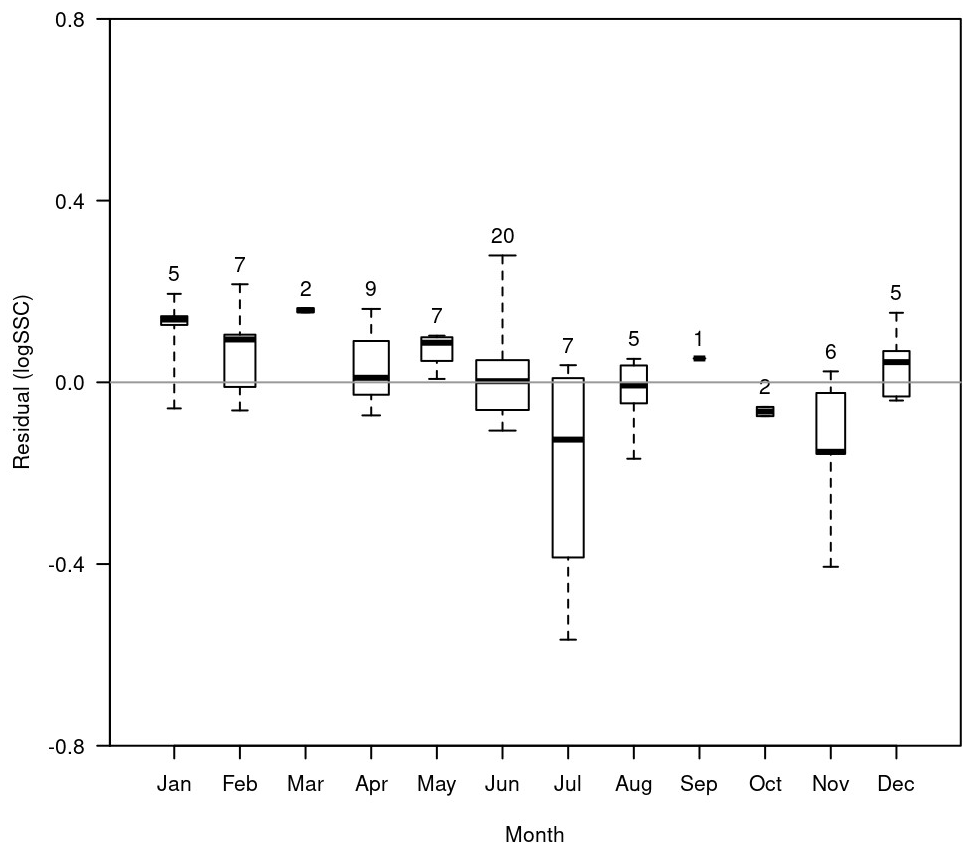
Leverage	Cook's D	DFBETS
0.0789	0.1943	0.3244

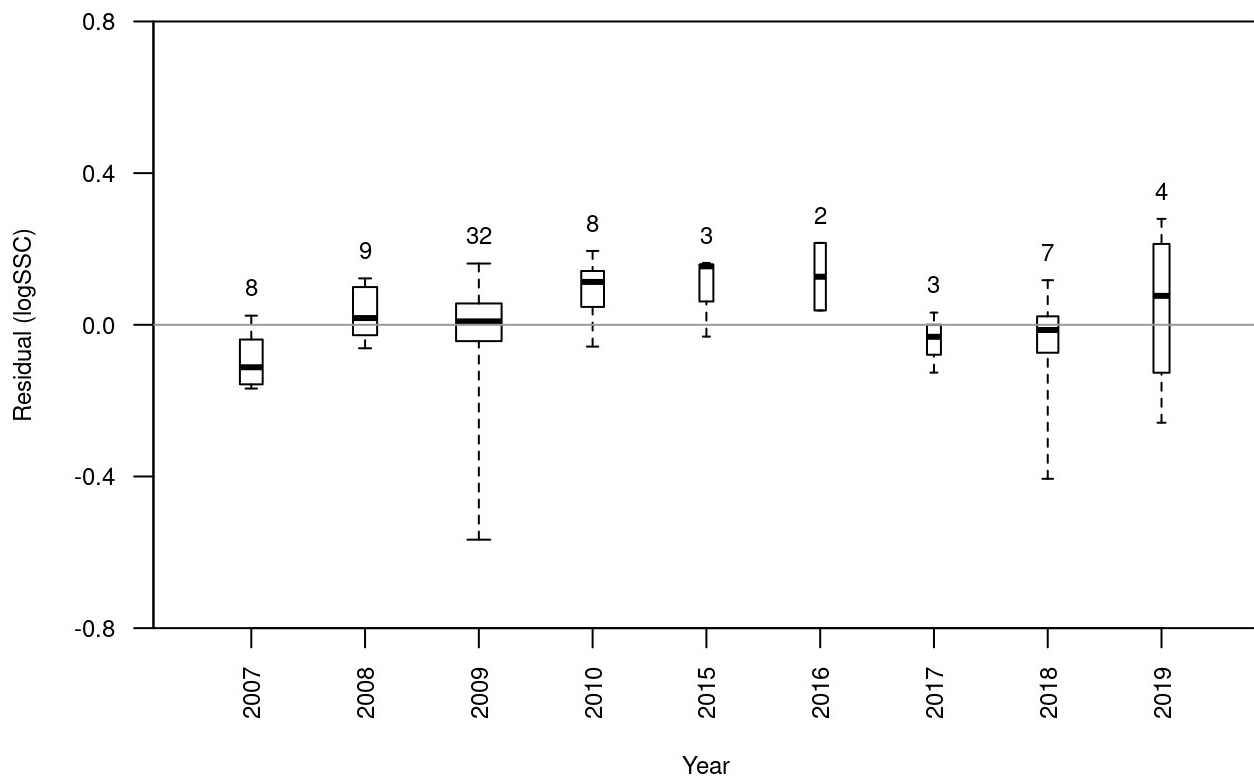
Flagged Observations

	logSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFBETS
2007-11-26 13:16	1.32	1.30	0.0241	0.181	0.180	0.1120	0.00205	0.0637
2009-07-27 00:26	2.11	2.68	-0.5670	-4.060	-4.580	0.0298	0.25300	-0.8020
2009-07-31 19:43	2.10	2.62	-0.5130	-3.670	-4.030	0.0249	0.17200	-0.6440
2017-07-14 16:00	3.14	3.11	0.0322	0.238	0.236	0.0844	0.00260	0.0717
2018-11-03 03:00	2.60	3.01	-0.4060	-2.970	-3.140	0.0682	0.32300	-0.8510

Statistical Plots

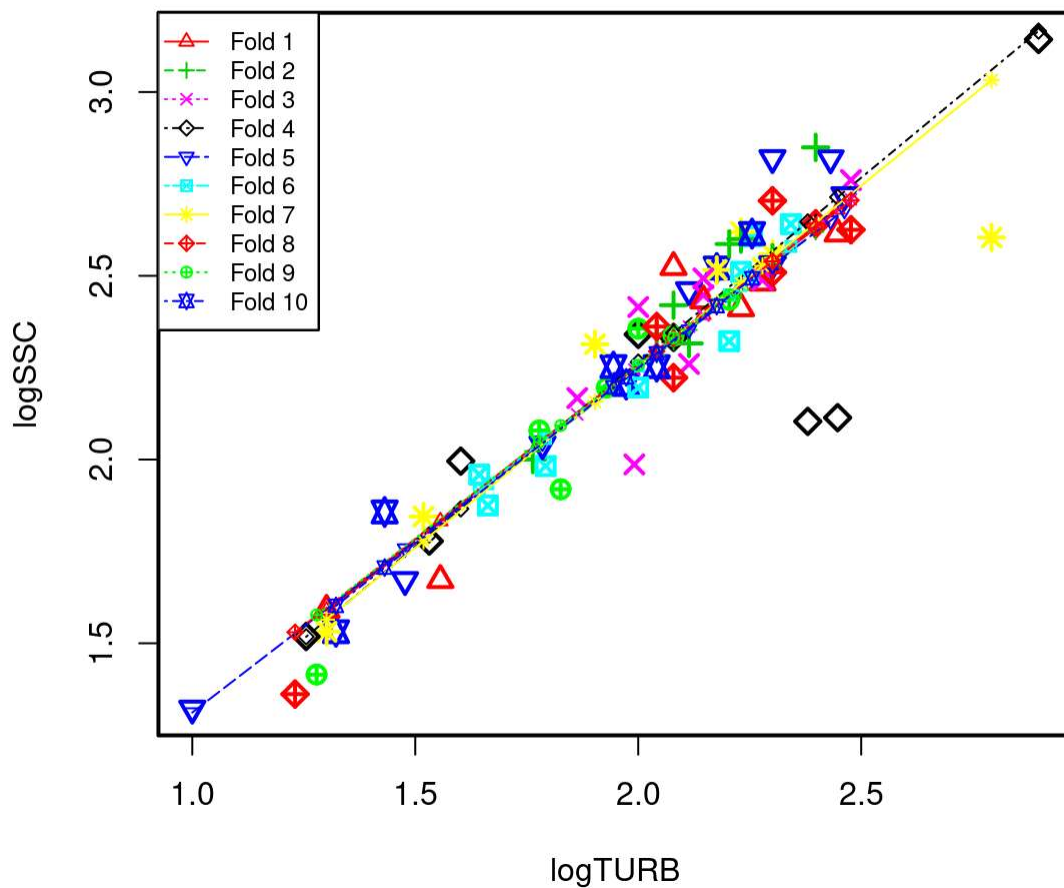






Cross Validation

Cross-validation



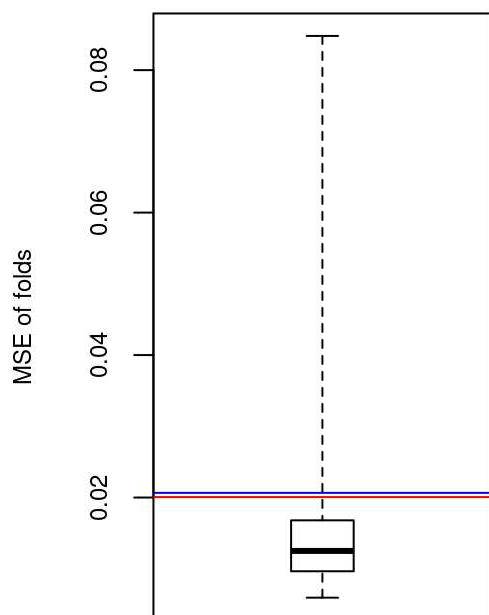
Minimum MSE of folds: 0.00594

Mean MSE of folds: 0.02070

Median MSE of folds: 0.01250

Maximum MSE of folds: 0.08480

(Mean MSE of folds) / (Model MSE): 1.03000



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
0										
1	2007-08-21	1.92	1.83	83	67	2.09	128	-0.168	-1.55	--
2	2007-10-11	1.53	1.32	34	21	1.61	42.2	-0.0744	-0.921	--
3	2007-10-11	1.53	1.3	34	20	1.59	40.3	-0.0542	-0.652	--
4	2007-11-26	1.32	1	21	10	1.3	20.8	0.0241	0.0823	--
5	2007-11-26	1.52	1.26	33	18	1.54	36.4	-0.0234	-0.318	--
6	2007-11-27	1.36	1.23	23	17	1.52	34.5	-0.157	-1.36	--
7	2007-11-27	1.41	1.28	26	19	1.56	38.3	-0.149	-1.28	--
8	2007-11-27	1.67	1.56	47	36	1.83	70.6	-0.157	-1.45	--
9	2008-02-13	2.52	2.18	332	150	2.42	276	0.0996	0.779	--
10	2008-02-13	2.52	2.18	328	150	2.42	276	0.0943	0.693	--
11	2008-02-13	2.61	2.26	405	180	2.5	328	0.11	0.921	--

12	2008-02-13	2.56	2.3	362	200	2.54	363	0.0178	0.0494	--
13	2008-02-13	2.48	2.28	303	190	2.52	346	-0.0382	-0.534	--
14	2008-02-13	2.41	2.23	258	170	2.47	311	-0.0619	-0.779	--
15	2008-04-29	2.55	2.3	354	200	2.54	363	0.00809	-0.0164	--
16	2008-04-29	2	1.76	100	58	2.03	111	-0.0274	-0.388	--
17	2008-04-29	2	1.6	99	40	1.87	78.1	0.122	1.03	--
18	2009-04-03	2.41	2	260	100	2.25	187	0.162	1.55	--
19	2009-04-03	2.42	2.08	263	120	2.33	223	0.091	0.652	--
20	2009-04-03	2.34	2.04	219	110	2.29	205	0.0475	0.353	--
21	2009-04-04	2.2	1.93	157	85	2.19	160	0.00997	0.0164	--
22	2009-04-04	1.98	1.79	96	62	2.06	119	-0.0728	-0.872	--
23	2009-04-04	1.78	1.53	60	34	1.81	66.8	-0.0276	-0.424	--
24	2009-05-07	2.46	2.11	289	130	2.36	241	0.0987	0.736	--
25	2009-05-07	2.34	2	219	100	2.25	187	0.0871	0.612	--
26	2009-05-07	2.36	2	227	100	2.25	187	0.103	0.872	--
27	2009-06-06	1.88	1.66	75	46	1.93	89.2	-0.0561	-0.693	--
28	2009-06-06	1.96	1.64	91	44	1.91	85.5	0.0463	0.318	--
29	2009-06-09	2.76	2.48	576	300	2.71	535	0.0513	0.388	--
30	2009-06-09	2.72	2.46	524	290	2.7	518	0.0243	0.115	--
31	2009-06-09	2.64	2.4	438	250	2.63	450	0.00799	-0.0494	--
32	2009-06-09	2.52	2.28	332	190	2.52	346	0.0015	-0.182	--
33	2009-06-10	2.53	2.3	340	200	2.54	363	-0.00944	-0.249	--
34	2009-06-10	2.26	2.11	182	130	2.36	241	-0.102	-1.08	--
35	2009-06-10	2.22	2.08	167	120	2.33	223	-0.106	-1.15	--
36	2009-06-14	2.64	2.34	437	220	2.58	398	0.06	0.534	--
37	2009-06-14	2.63	2.48	422	300	2.71	535	-0.0838	-1.03	--
38	2009-06-14	2.33	2.08	215	120	2.33	223	0.00345	-0.149	--
39	2009-06-14	2.2	1.97	160	94	2.23	177	-0.0236	-0.353	--
40	2009-07-27	2.11	2.45	130	280	2.68	501	-0.567	-2.41	--
41	2009-07-31	2.1	2.38	127	240	2.62	432	-0.513	-2.03	--
42	2009-08-13	2.32	2.11	207	130	2.36	241	-0.0462	-0.612	--
43	2009-08-13	2.04	1.79	110	61	2.05	117	-0.0069	-0.215	--
44	2009-08-13	1.85	1.52	70	33	1.79	65	0.0517	0.424	--
45	2009-09-11	2.25	1.94	179	88	2.2	166	0.0525	0.46	--
46	2009-12-09	2.36	2.04	230	110	2.29	205	0.0688	0.573	--
47	2009-12-09	2.25	2.04	179	110	2.29	205	-0.04	-0.573	--
48	2009-12-09	2.17	1.86	147	73	2.12	139	0.0445	0.283	--
49	2009-12-26	2.31	1.9	206	80	2.16	151	0.153	1.36	--

50	2010-01-25	2.52	2.08	334	120	2.33	223	0.195	1.82	--
51	2010-01-25	2.62	2.23	416	170	2.47	311	0.146	1.21	--
52	2010-01-25	2.6	2.23	398	170	2.47	311	0.126	1.08	--
53	2010-01-25	2.59	2.2	386	160	2.45	293	0.138	1.15	--
54	2010-01-25	2.2	2	157	100	2.25	187	-0.0575	-0.736	--
55	2010-05-03	2.49	2.15	311	140	2.39	258	0.0998	0.825	--
56	2010-05-03	2.45	2.15	280	140	2.39	258	0.0542	0.497	--
57	2010-05-03	2.43	2.15	271	140	2.39	258	0.04	0.249	--
58	2015-03-11	2.7	2.3	506	200	2.54	363	0.163	1.67	--
59	2015-03-11	2.82	2.43	660	270	2.67	484	0.154	1.45	--
60	2015-12-24	2.49	2.28	308	190	2.52	346	-0.0311	-0.46	--
61	2016-02-03	2.85	2.4	707	250	2.63	450	0.216	2.03	--
62	2016-07-31	2.08	1.78	120	60	2.04	115	0.0378	0.215	--
63	2017-06-24	2.51	2.3	323	200	2.54	363	-0.0317	-0.497	--
64	2017-07-14	3.14	2.9	1390	790	3.11	1350	0.0322	0.149	--
65	2017-07-14	2.32	2.2	210	160	2.45	293	-0.126	-1.21	--
66	2018-05-17	2.34	2.08	217	120	2.33	223	0.00747	-0.0823	--
67	2018-06-11	2.61	2.26	412	180	2.5	328	0.118	0.973	--
68	2018-06-11	2.61	2.45	412	280	2.68	501	-0.0656	-0.825	--
69	2018-06-28	1.67	1.48	47	30	1.75	59.3	-0.0818	-0.973	--
70	2018-07-04	2.43	2.2	272	160	2.45	293	-0.0138	-0.283	--
71	2018-08-04	2.51	2.23	324	170	2.47	311	0.0371	0.182	--
72	2018-11-03	2.6	2.79	402	620	3.01	1070	-0.406	-1.82	--
73	2019-06-13	1.59	1.3	39	20	1.59	40.3	0.00541	-0.115	--
74	2019-06-20	2.82	2.3	661	200	2.54	363	0.279	2.41	--
75	2019-06-21	1.86	1.43	72	27	1.71	53.6	0.147	1.28	--
76	2019-07-12	1.99	1.99	97	98	2.24	184	-0.258	-1.67	--

Definitions

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

TURB: Turbidity in FNU (63680)