Model Archive Summary for Suspended Sediment Concentration at Station 01472157; French Creek near Phoenixville, Pennsylvania [2020 Version]

This model archive summary describes the regression model developed to estimate continuous instantaneous (15-minute) suspended sediment concentrations starting in 2005. The model was updated in June 2020, using all available calibration data. The previous model (Sloto and Olsen, 2011) was in use from October 1, 2007 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: 01472157

Station name: French Creek near Phoenixville, Pennsylvania

Lat 40°09'05", long 75°36'06" referenced to North American Datum of 1927, Chester County, PA, Hydrologic Unit 02040203, on right bank 70 ft downstream from two-span county bridge on French Creek Road, 4.5 mi northwest of Phoenixville, and 7.3 mi upstream from mouth.

Date regression model was created: June 2020.

Period of data for model calibration: July 29, 2007 – August 13, 2018.

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) 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 right 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 111 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 Models

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

$SSC = 1.689\ Turb - 3.954$ (October 1, 2007 to September 30, 2019)

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 7 to 990 mg/L in the 111 analyzed samples.

Turbidity data were collected according to USGS Techniques and Methods 1-D3 (Wagner and Others, 2006). Turbidity concentrations ranged from 2.2 to 520 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*).

Model Summary

Final regression model for suspended sediment concentration (SSC) at site number 01472157:

$$\log_{10} SSC = .949 \log_{10} Turb + .332$$

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:

$$SSC = 2.212\ Turb^{.949}$$

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

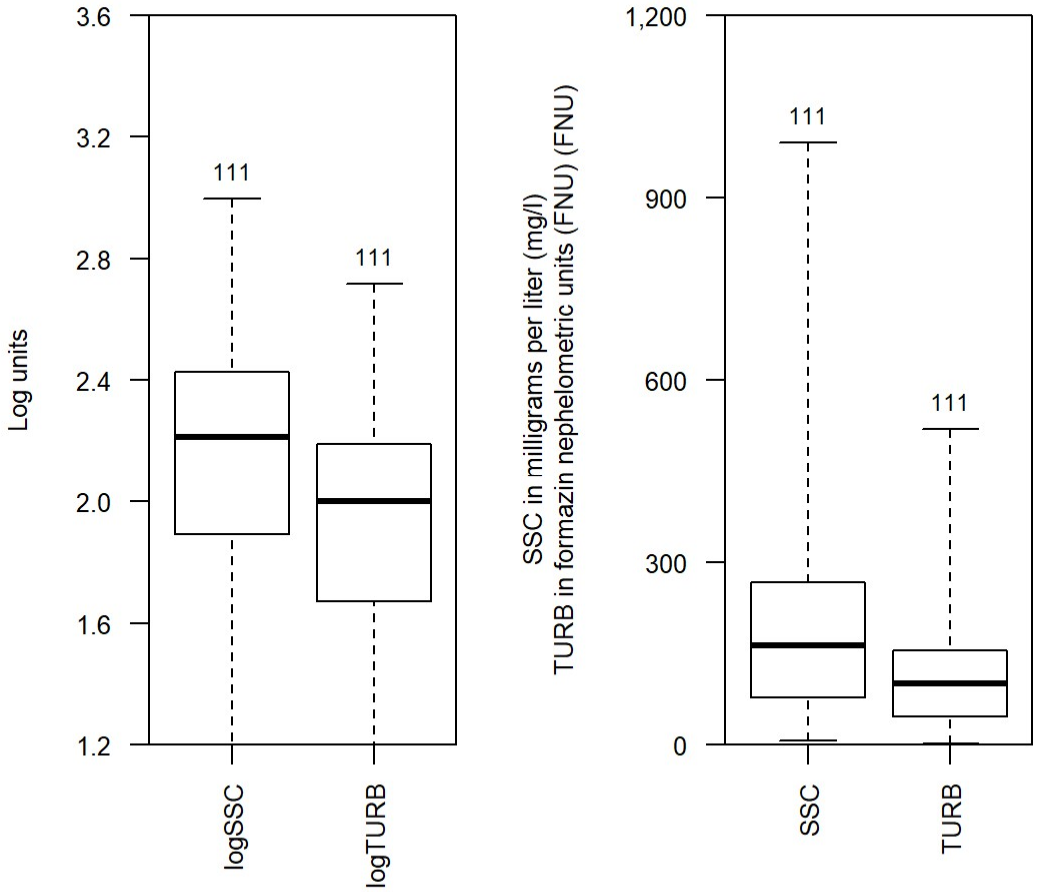
Model

logSSC = + 0.949 * logTURB + 0.332

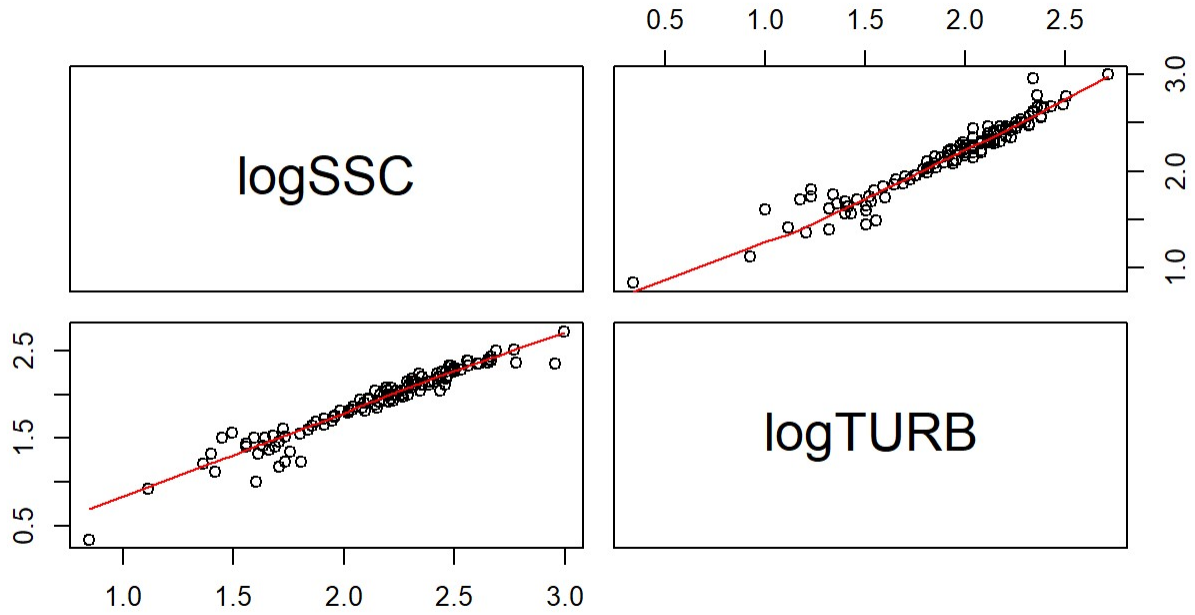
Variable Summary Statistics

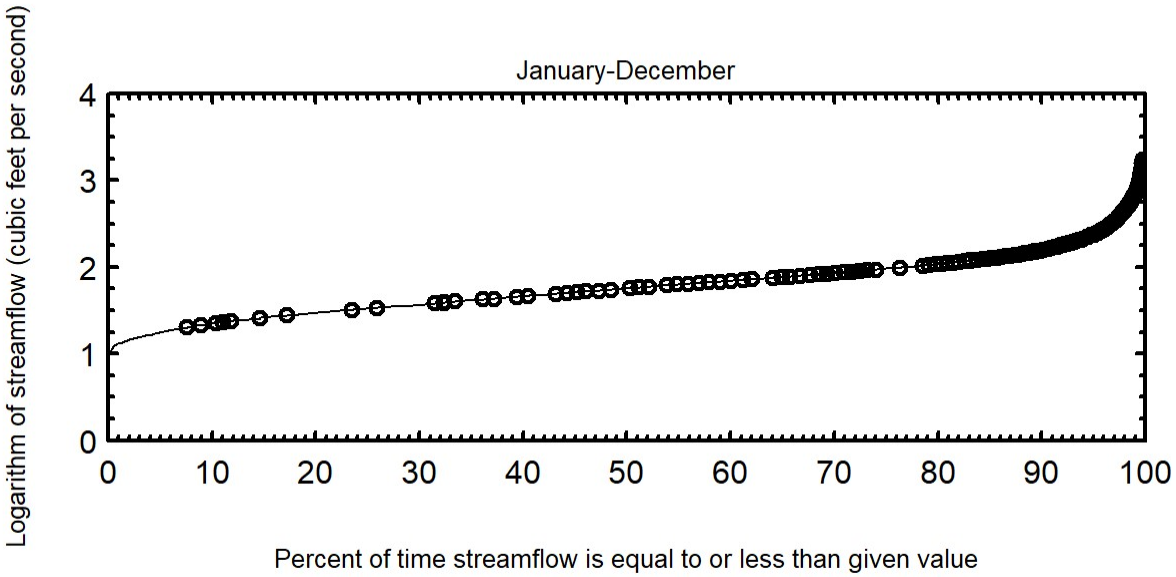
	logSSC	SSC	logTURB	TURB
Minimum	0.845	7	0.342	2.2
1st Quartile	1.880	75	1.650	45.0
Median	2.210	164	2.000	100.0
Mean	2.150	197	1.910	113.0
3rd Quartile	2.430	271	2.200	160.0
Maximum	3.000	990	2.720	520.0

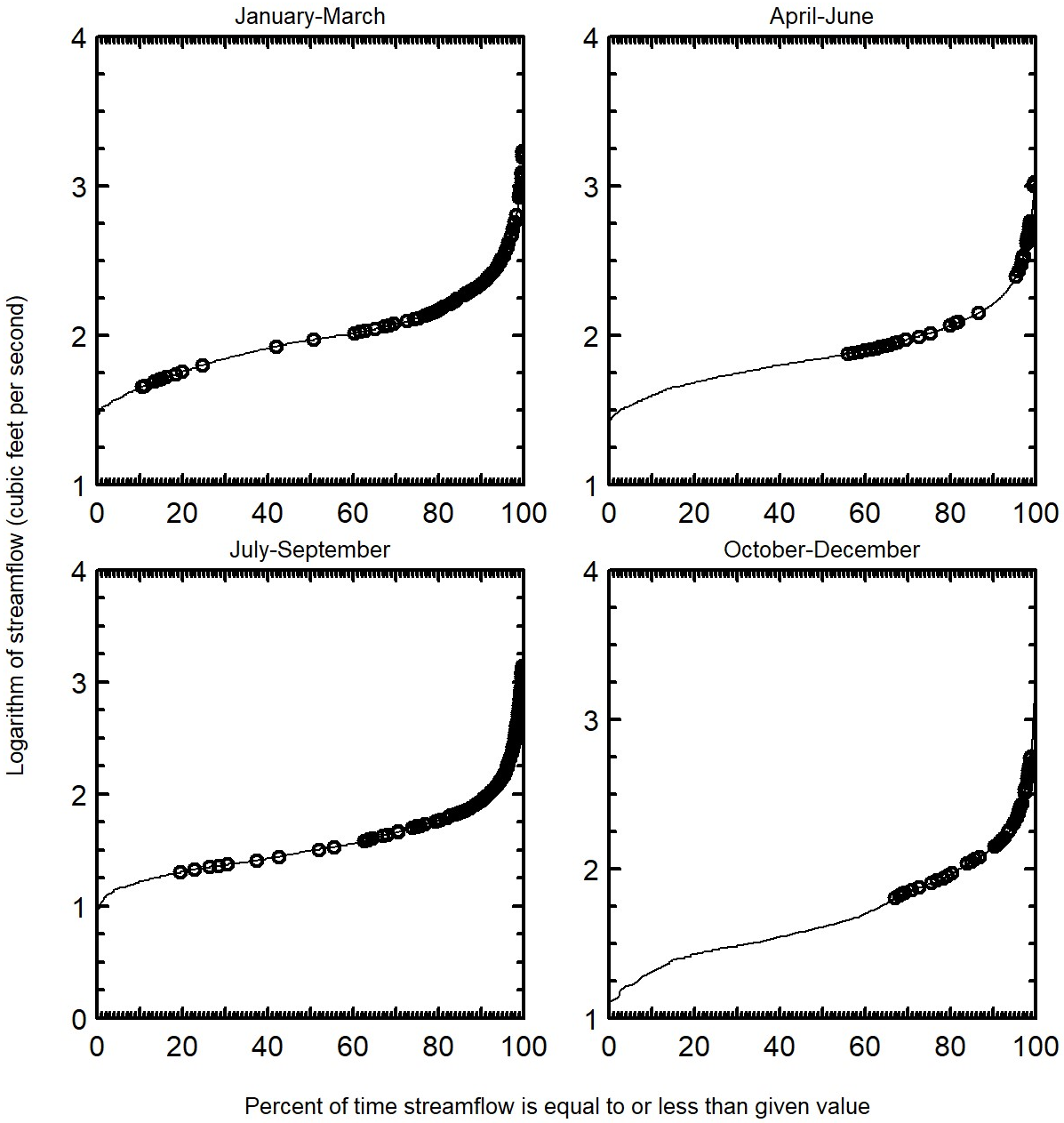
Box Plots


Log units
3.6
3.2
2.8
2.4
2.0
1.6
1.2
111
111
logSSC
logTURB
SSC in milligrams per liter (mg/l)
TURB in formazin nephelometric units (FNU) (FNU)
1,200
900
600
300
0
111
111
SSC
TURB

Exploratory Plots







Basic Model Statistics

Number of Observations	111
Standard error (RMSE)	0.105
Average Model standard percentage error (MSPE)	24.4
Coefficient of determination (R^2)	0.929
Adjusted Coefficient of Determination (Adj. R^2)	0.928
Bias Correction Factor (BCF)	1.03

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.332	0.0492	6.74	7.88e-10
logTURB	0.949	0.0252	37.60	2.74e-64

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.979
E.vars	-0.979	1.000

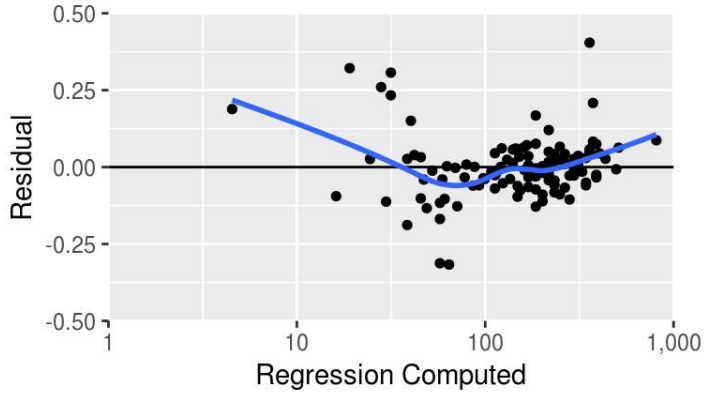
Outlier Test Criteria

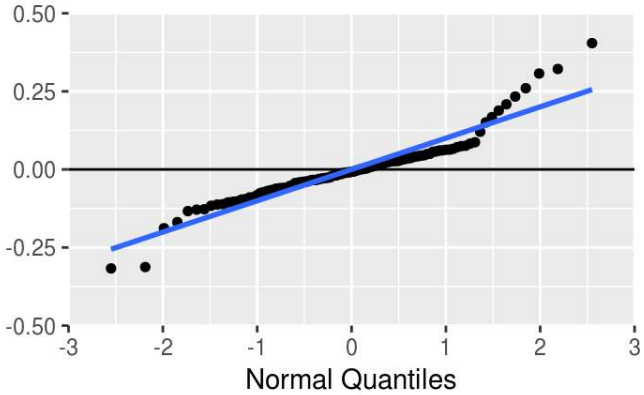
Leverage	Cook's D	DFFITS
0.0541	0.1944	0.2685

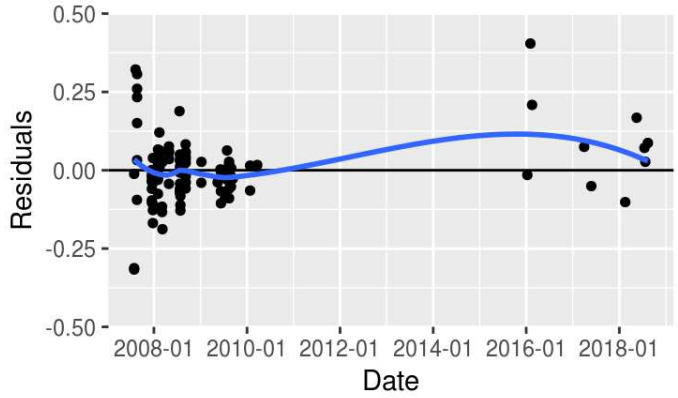
Flagged Observations

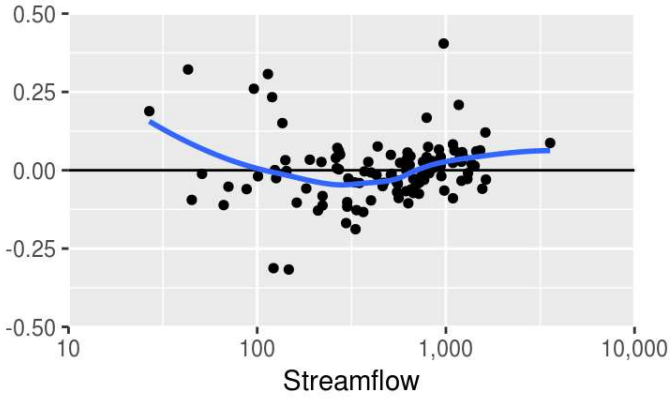
	logSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2007-07-29 20:30	1.490	1.810	-0.3170	-3.040	-3.160	0.0163	0.0766	-0.407
2007-07-30 01:30	1.450	1.760	-0.3130	-3.000	-3.120	0.0185	0.0852	-0.429
2007-08-10 04:45	1.600	1.280	0.3220	3.150	3.290	0.0569	0.3000	0.808
2007-08-21 00:00	1.110	1.210	-0.0946	-0.932	-0.931	0.0652	0.0303	-0.246
2007-08-22 13:00	1.730	1.500	0.2330	2.260	2.310	0.0357	0.0948	0.444
2007-08-22 14:00	1.810	1.500	0.3070	2.980	3.090	0.0357	0.1640	0.595
2007-08-22 16:00	1.710	1.450	0.2600	2.530	2.590	0.0402	0.1340	0.530
2008-03-08 01:00	1.400	1.590	-0.1880	-1.820	-1.840	0.0290	0.0494	-0.318
2008-07-19 22:00	0.845	0.657	0.1890	1.950	1.970	0.1510	0.3370	0.832
2016-02-03 18:15	2.960	2.550	0.4040	3.890	4.170	0.0197	0.1520	0.591
2016-02-16 16:45	2.780	2.570	0.2090	2.010	2.040	0.0206	0.0424	0.296

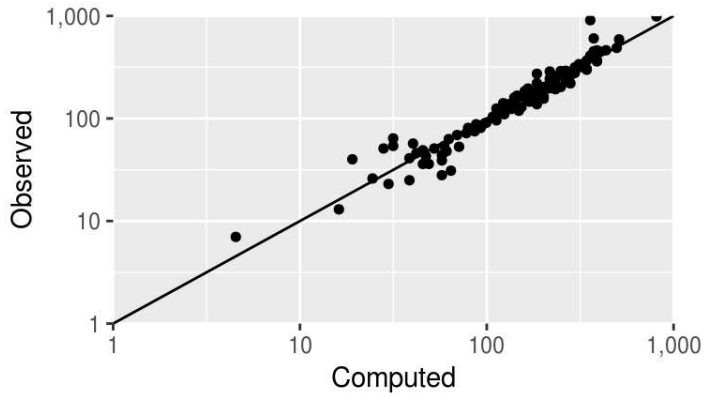
Statistical Plots

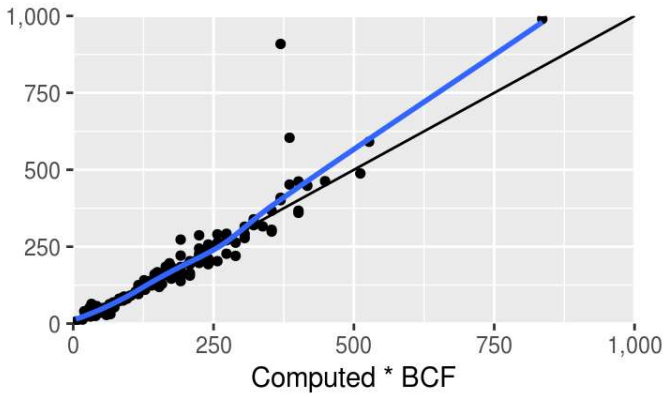


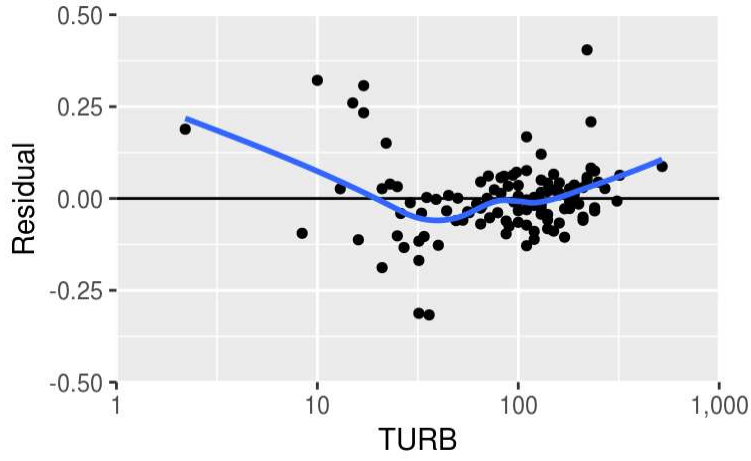


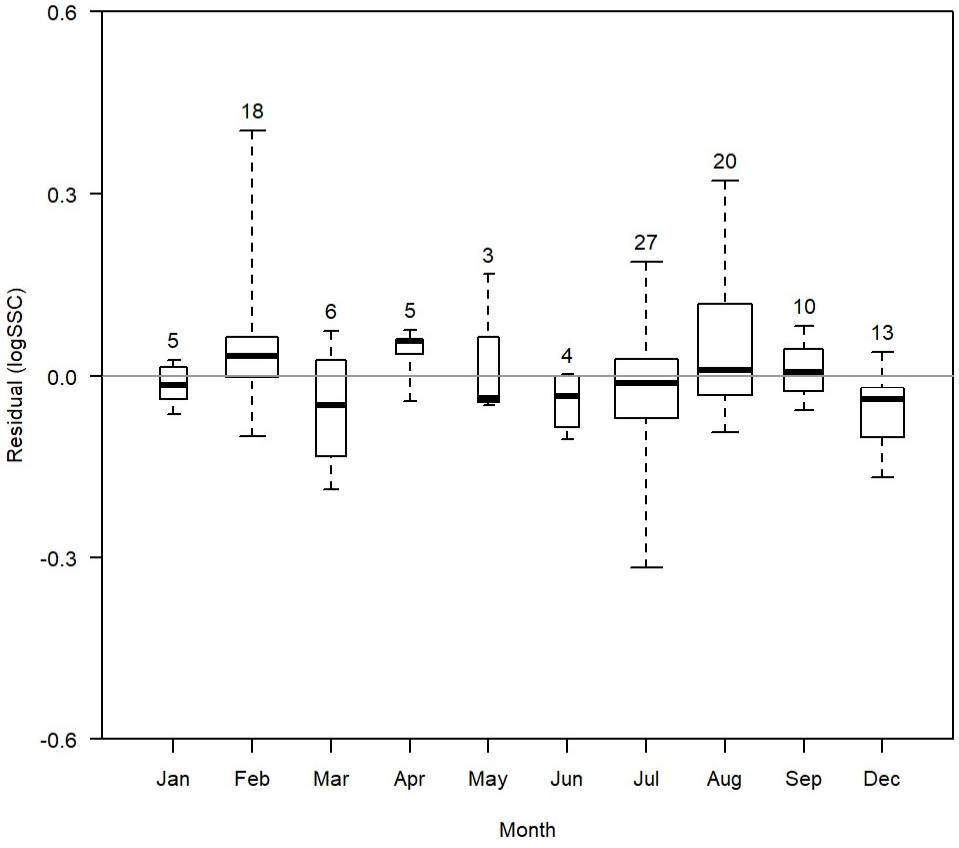


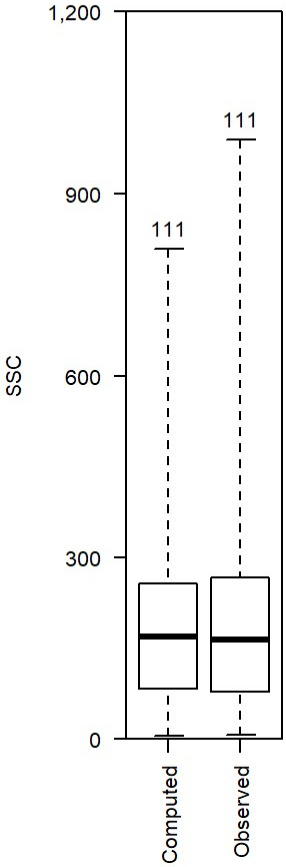


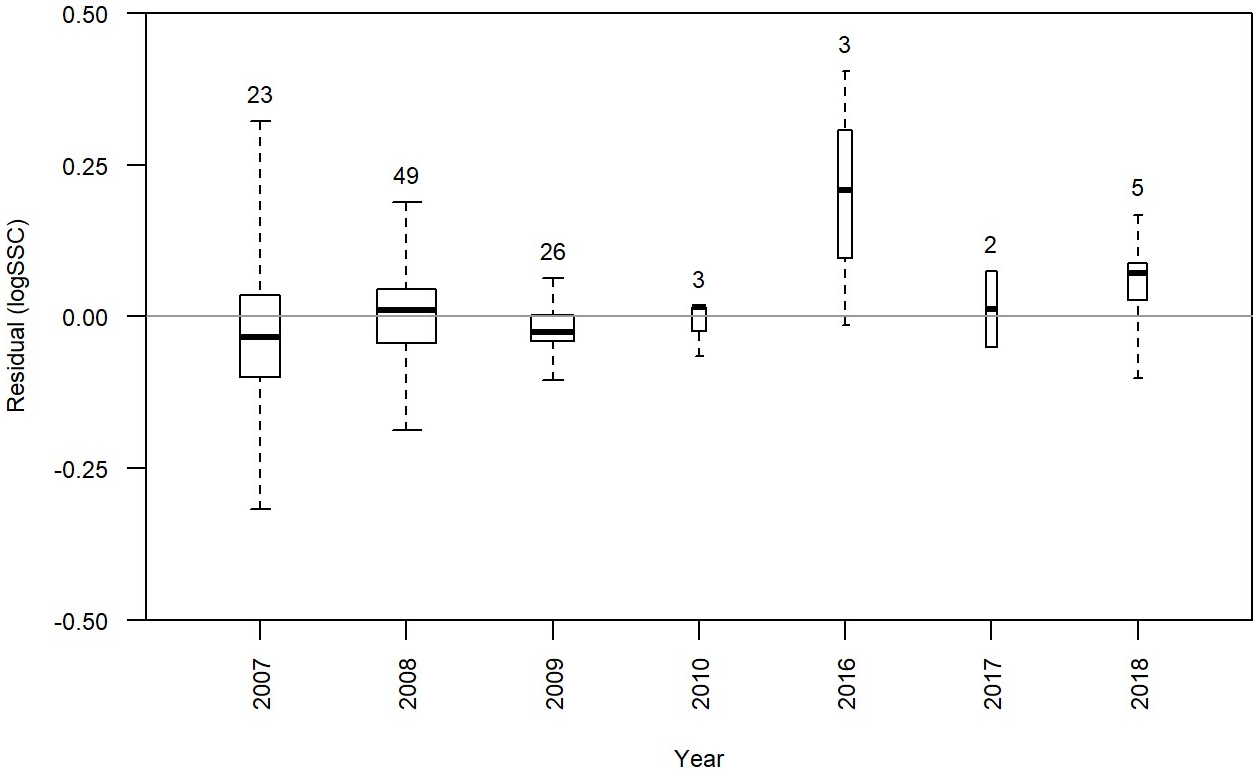




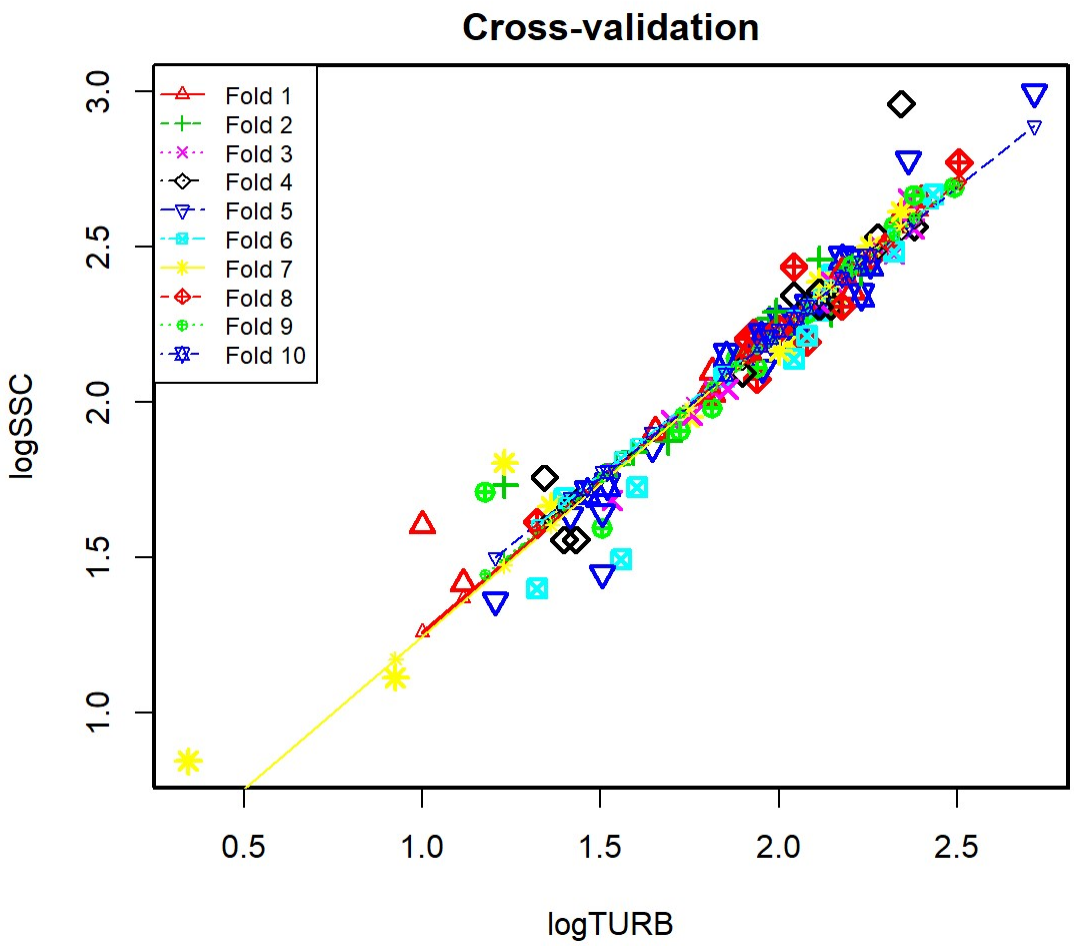



0.6
0.3
0.0
-0.3
-0.6
Residual (logSSC)
5
18
6
5
3
4
27
20
10
13
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Dec
Month

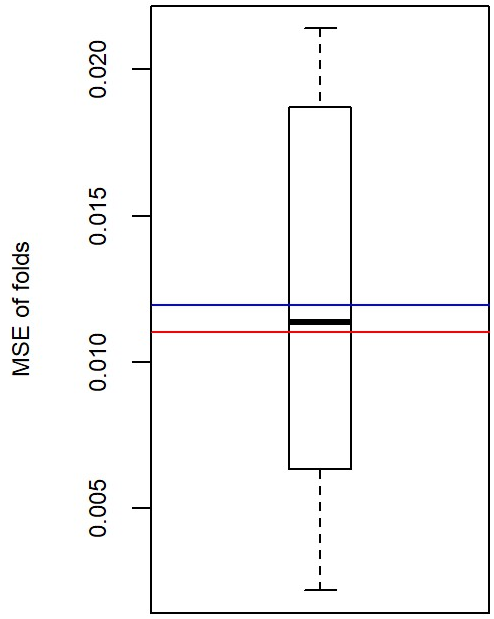

1,200
900
600
300
0
SSC
111
111
Computed
Observed


0.50
0.25
0.00
-0.25
-0.50
Residual (logSSC)
23
49
26
3
3
2
5
2007
2008
2009
2010
2016
2017
2018
Year

Cross Validation




```
                Minimum MSE of folds:  0.0022
                   Mean MSE of folds:  0.0120
                 Median MSE of folds:  0.0113
                Maximum MSE of folds:  0.0214
(Mean MSE of folds) / (Model MSE):  1.0800
```



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	2007-07-29	1.49	1.56	31	36	1.81	66.3	-0.317	-2.55	--
2	2007-07-30	1.45	1.51	28	32	1.76	59.3	-0.313	-2.19	--
3	2007-07-30	1.71	1.46	51	29	1.72	54	-0.0116	-0.0903	--
4	2007-08-10	1.6	1	40	10	1.28	19.7	0.322	2.19	--
5	2007-08-21	1.11	0.924	13	8.4	1.21	16.7	-0.0946	-1.12	--
6	2007-08-22	1.69	1.4	49	25	1.66	46.9	0.0322	0.544	--
7	2007-08-22	1.76	1.34	57	22	1.61	41.6	0.151	1.42	--
8	2007-08-22	1.73	1.23	54	17	1.5	32.5	0.233	1.74	--
9	2007-08-22	1.81	1.23	64	17	1.5	32.5	0.307	1.99	--
10	2007-08-22	1.71	1.18	51	15	1.45	28.9	0.26	1.85	--
11	2007-12-16	1.68	1.53	48	34	1.78	62.8	-0.103	-1.26	--

12	2007-12-16	1.86	1.64	72	44	1.89	80.2	-0.0336	-0.393	--
13	2007-12-16	2.08	1.94	119	87	2.17	153	-0.0963	-1.16	--
14	2007-12-16	2.02	1.79	104	62	2.03	111	-0.0152	-0.136	--
15	2007-12-16	1.95	1.75	90	56	1.99	101	-0.0361	-0.442	--
16	2007-12-16	1.94	1.7	88	50	1.94	90.6	0.000884	0.113	--
17	2007-12-16	1.91	1.72	81	53	1.97	95.7	-0.0591	-0.709	--
18	2007-12-16	2.04	1.82	109	66	2.06	118	-0.0206	-0.204	--
19	2007-12-24	1.66	1.36	46	23	1.62	43.4	0.0391	0.652	--
20	2007-12-24	1.63	1.41	43	26	1.67	48.7	-0.0407	-0.544	--
21	2007-12-24	1.72	1.6	53	40	1.85	73.3	-0.127	-1.56	--
22	2007-12-24	1.59	1.51	39	32	1.76	59.3	-0.169	-1.85	--
23	2007-12-29	1.36	1.2	23	16	1.47	30.7	-0.112	-1.42	--
24	2008-02-01	1.84	1.59	69	39	1.84	71.6	-0.00237	0.0677	--
25	2008-02-01	2.03	1.81	106	65	2.05	116	-0.0264	-0.251	--
26	2008-02-01	2.14	1.88	138	76	2.12	135	0.0237	0.393	--
27	2008-02-01	2.41	2.15	257	140	2.37	241	0.0421	0.68	--
28	2008-02-02	2.46	2.18	290	150	2.4	257	0.0661	1.08	--
29	2008-02-02	2.47	2.2	292	160	2.42	273	0.0425	0.709	--
30	2008-02-02	2.43	2.2	271	160	2.42	273	0.0101	0.227	--
31	2008-02-02	2.31	2.08	203	120	2.3	208	0.00316	0.158	--
32	2008-02-02	2.24	2.04	173	110	2.27	191	-0.0304	-0.345	--
33	2008-02-02	2.11	1.95	129	90	2.19	158	-0.0752	-0.965	--
34	2008-02-13	2.46	2.11	287	130	2.34	224	0.121	1.36	--
35	2008-02-14	2.27	1.97	185	94	2.2	165	0.0635	1.04	--
36	2008-02-14	2.22	1.93	167	85	2.16	150	0.0605	0.93	--
37	2008-02-14	2.1	1.81	125	65	2.05	116	0.0452	0.768	--
38	2008-02-14	1.91	1.65	81	45	1.9	82	0.0083	0.204	--
39	2008-03-05	1.64	1.51	44	32	1.76	59.3	-0.116	-1.49	--
40	2008-03-05	1.56	1.43	36	27	1.69	50.5	-0.133	-1.74	--
41	2008-03-08	1.41	1.11	26	13	1.39	25.2	0.0264	0.418	--
42	2008-03-08	1.4	1.32	25	21	1.59	39.8	-0.188	-1.99	--
43	2008-04-28	2.15	1.85	141	71	2.09	126	0.0611	0.965	--
44	2008-04-28	2.34	2.04	221	110	2.27	191	0.0759	1.21	--
45	2008-04-28	2.32	2.15	211	140	2.37	241	-0.0436	-0.597	--
46	2008-04-29	2.26	2	184	100	2.23	175	0.0356	0.597	--
47	2008-04-29	2.2	1.91	160	82	2.15	145	0.0567	0.862	--
48	2008-07-20	0.845	0.342	7	2.2	0.657	4.68	0.189	1.56	--
49	2008-07-24	1.88	1.69	75	49	1.94	88.9	-0.0602	-0.768	--

50	2008-07-24	2.08	1.85	121	70	2.08	125	0.000542	0.0903	--	
51	2008-07-24	2.21	1.95	164	89	2.18	157	0.0337	0.57	--	
52	2008-07-24	2.39	2.11	244	130	2.34	224	0.0501	0.83	--	
53	2008-07-24	2.6	2.34	401	220	2.55	369	0.049	0.799	--	
54	2008-07-24	2.56	2.32	367	210	2.53	353	0.0297	0.518	--	
55	2008-07-24	2.51	2.28	320	190	2.49	321	0.0115	0.251	--	
56	2008-07-24	2.46	2.26	291	180	2.47	305	-0.00752	0	--	
57	2008-07-24	2.29	2.11	197	130	2.34	224	-0.0429	-0.57	--	
58	2008-07-24	2.2	2.04	157	110	2.27	191	-0.0726	-0.93	--	
59	2008-07-24	2.11	1.94	129	87	2.17	153	-0.0612	-0.799	--	
60	2008-07-24	1.98	1.81	96	65	2.05	116	-0.0694	-0.896	--	
61	2008-07-27	2.19	2.08	156	120	2.3	208	-0.111	-1.36	--	
62	2008-07-27	2.29	2.15	193	140	2.37	241	-0.0823	-1	--	
63	2008-07-27	2.14	2.04	138	110	2.27	191	-0.129	-1.64	--	
64	2008-09-06	2.31	2.15	204	140	2.37	241	-0.0582	-0.68	--	
65	2008-09-06	2.65	2.4	448	250	2.61	417	0.0445	0.738	--	
66	2008-09-06	2.66	2.36	452	230	2.57	385	0.0827	1.26	--	
67	2008-09-07	2.61	2.34	409	220	2.55	369	0.0576	0.896	--	
68	2008-09-07	2.53	2.28	339	190	2.49	321	0.0365	0.624	--	
69	2008-09-07	2.42	2.18	263	150	2.4	257	0.0237	0.369	--	
70	2008-09-07	2.29	2.08	193	120	2.3	208	-0.0188	-0.158	--	
71	2008-09-07	2.2	1.98	158	95	2.21	167	-0.00942	-0.0451	--	
72	2008-09-07	2.09	1.9	124	79	2.13	140	-0.0387	-0.493	--	
73	2009-01-07	1.73	1.52	54	33	1.77	61.1	-0.04	-0.518	--	
74	2009-01-08	1.61	1.32	41	21	1.59	39.8	0.0266	0.442	--	
75	2009-05-15	1.96	1.76	91	57	2	103	-0.0385	-0.467	--	
76	2009-06-06	1.8	1.54	63	35	1.8	64.6	0.00271	0.136	--	
77	2009-06-09	2.26	2.04	184	110	2.27	191	-0.00367	0.0451	--	
78	2009-06-10	2.36	2.2	227	160	2.42	273	-0.0669	-0.862	--	
79	2009-06-10	2.34	2.23	220	170	2.45	289	-0.105	-1.31	--	
80	2009-07-12	2.56	2.38	367	240	2.59	401	-0.0253	-0.227	--	
81	2009-07-12	2.69	2.49	488	310	2.7	511	-0.00699	0.0225	--	
82	2009-07-12	2.31	2.18	203	150	2.4	257	-0.0888	-1.04	--	
83	2009-07-12	2.33	2.11	212	130	2.34	224	-0.011	-0.0677	--	
84	2009-07-28	2.77	2.51	591	320	2.71	527	0.0631	1	--	
85	2009-07-28	2.31	2.11	203	130	2.34	224	-0.0298	-0.321	--	
86	2009-08-13	2.2	2	157	100	2.23	175	-0.0333	-0.369	--	
87	2009-08-13	2.67	2.43	463	270	2.64	449	0.0271	0.493	--	

88	2009-08-13	2.56	2.38	360	240	2.59	401	-0.0336	-0.418	--
89	2009-08-13	2.42	2.23	263	170	2.45	289	-0.0279	-0.297	--
90	2009-08-13	2.39	2.15	244	140	2.37	241	0.0195	0.345	--
91	2009-08-13	2.22	2	166	100	2.23	175	-0.0091	-0.0225	--
92	2009-08-13	2.16	1.91	145	82	2.15	145	0.0139	0.274	--
93	2009-08-13	2.48	2.32	299	210	2.53	353	-0.0593	-0.738	--
94	2009-08-13	2.21	2.08	164	120	2.3	208	-0.0895	-1.08	--
95	2009-08-28	2.32	2.11	208	130	2.34	224	-0.0193	-0.181	--
96	2009-08-28	2.04	1.86	110	72	2.09	128	-0.0525	-0.652	--
97	2009-08-30	2.24	2	172	100	2.23	175	0.00632	0.181	--
98	2009-09-11	2.44	2.26	278	180	2.47	305	-0.0274	-0.274	--
99	2010-01-25	2.44	2.2	274	160	2.42	273	0.0149	0.297	--
100	2010-01-25	2.16	2	146	100	2.23	175	-0.0649	-0.83	--
101	2010-03-22	2.35	2.11	226	130	2.34	224	0.0168	0.321	--
102	2016-01-10	2.5	2.3	316	200	2.51	337	-0.0151	-0.113	--
103	2016-02-03	2.96	2.34	909	220	2.55	369	0.404	2.55	--
104	2016-02-16	2.78	2.36	604	230	2.57	385	0.209	1.64	--
105	2017-03-31	2.66	2.38	462	240	2.59	401	0.0747	1.16	--
106	2017-05-26	2.48	2.32	305	210	2.53	353	-0.0506	-0.624	--
107	2018-02-16	1.56	1.4	36	25	1.66	46.9	-0.102	-1.21	--
108	2018-05-17	2.44	2.04	273	110	2.27	191	0.168	1.49	--
109	2018-07-17	2.29	1.99	196	98	2.22	172	0.0714	1.12	--
110	2018-07-24	2.5	2.26	315	180	2.47	305	0.0269	0.467	--
111	2018-08-13	3	2.72	990	520	2.91	836	0.0871	1.31	--

Definitions

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

TURB: Turbidity in FNU (63680)