Model Archive Summary for Suspended-Sediment Concentration at Station 04015330; Knife River near Two Harbors, Minnesota

This model archive summary (MAS) documents the suspended-sediment concentration (SSC) model developed to compute 15-minute SSC from turbidity readings. This is the first model developed for the site.

SITE AND MODEL INFORMATION

Site number: 04015330
Site name: Knife River near Two Harbors, Minnesota (MN), Lake County
Location: Latitude N 46°56'49", Longitude W 91°47'32", referenced to NAD27
Hydrologic Unit: 04010102
Drainage area: 83.6 square miles

Date rating model was created: 1/3/2017
Model calibration data period: 5/24/2012–7/12/2016
Model application date: 4/1/2017
Computed by: Joel Groten (jgroten@usgs.gov), Hydrologist, Minnesota Water Science Center
Reviewed by: Chris Ellison (cellison@usgs.gov), Supervisory Hydrologist, Wyoming-Montana Water Science Center
Reviewed by: Jeffrey Ziegeweid (jrziege@usgs.gov), Hydrologist, Minnesota Water Science Center
Approved by: James Stark (stark@usgs.gov), Center Director, Minnesota Water Science Center

Additional Information

Type	Scheme	Key
doi	http://dx.doi.org	doi:10.5066/F71N7ZCN
IPDS	https://ipds.usgs.gov	IP-085866

PHYSICAL SAMPLING EQUIPMENT AND SAMPLING DETAILS

Samples were collected 7–10 times per year during the open-water season (April through September) throughout the range of continuously observed hydrologic conditions. No samples were collected during the winter months because sediment transport is low (Tornes, 1986) in Minnesota because streamflow is generally contained under ice and receives little sediment input from the surrounding landscape. Water samples were collected using isokinetic samplers and equal-width-increment (EWI) and depth-integrating techniques, following procedures by Edwards and Glysson (1999). Most samples were collected using a D-74 rigid bottle sampler suspended from the downstream side of the Church Road Bridge, or upstream side using a modified cableway system, during nonwadeable flows, and a DH-48 hand-held sampler during wadeable flows at a cross-section 1,000 feet upstream of the bridge, close to where the USGS gage is located. The total stream width was divided into 10 equal-width increments, and individual depth-integrated samples were collected at the centroid of each increment. Following collection, samples were transported to the USGS Sediment Laboratory in Iowa City, Iowa, where they were composited into a single sample and analyzed for SSC and particle-size fractions less than 0.0625 millimeters (fines). Analyses results were stored in the National Water Information System (NWIS) database (U.S. Geological Survey, 2017) and can be found at http://mn.water.usgs.gov.

Collecting a range of turbidity samples at Knife River is difficult because of the seasonal conditions and flashy nature of the river. During spring snowmelt, when turbidity readings are high, large volumes of ice flowing in the stream prevents installation of the turbidity sensor. Once the sensor is installed, rain events often produce rapid turbidity responses that peak prior to streamflow and quickly return to pre-event conditions. The rapid turbidity events, in combination with the prolonged travel time to the site (approximately 100 miles) makes it difficult to collect samples during elevated turbidity events. Also, hydrologic technicians often scheduled their sampling days based on bankfull streamflow, which was a requirement from

another study that incorporated this site into its sampling plan. These factors limited the number of samples that are suitable for calibrating the model.

SURROGATE EQUIPMENT AND SETUP DETAILS

A Forest Technology Systems DTS-12 digital turbidity sensor is installed at the site. The turbidity sensor was mounted on the streamward face of a large dolomite rock at the right bank, inside a 3-inch, schedule 80 polyvinyl (PVC) tube that protects the probe and makes it easier to regularly pull it out for calibration and re-deployment. The turbidity sensor water-level stage may be adjusted to keep the bottom of the probe a few feet under water. The battery and datalogger are located in an enclosure near the turbidity sensor.

Turbidity Sensor Manufactured characteristics and configuration	
Make	Forest Technology Systems (FTS)
Model	DTS-12 Digital Turbidity Sensor
Serial number	039762 (installed on 05/18/2012) 044284 (installed on 06/28/2012)
Sensor Type	Optical nephelometer (sidescatter)
Range	0 to 1,600 Formazin Nephelometric Unit (FNU)
Accuracy (@ 25°C)	± 2% of reading + 0.2 FNU (0–399 FNU) ± 4% of reading (400–1,600 FNU)
Resolution	0.01 FNU
Operating Temperature	+32˚F to 104˚F
Measurement interval	900 sec
Measurement returned	Mean from 100 instantaneous samples at a sample rate of 20 Hz

MODEL-CALIBRATION DATASET

All data are stored in the National Water Information System (NWIS) database (U.S. Geological Survey, 2017). The complete water-quality record can be found at http://mn.water.usgs.gov. The regression model dataset considered 33 concurrent measurements of suspended-sediment concentration (SSC) and turbidity collected from 5/24/2012 through 7/12/2016 during the period while the turbidity sensor was deployed. All available SSC samples had a reasonable percentage (mean of 86 percent) of suspended-sediment smaller than 0.0625 millimeters and outliers were not identified in the dataset.

Initially, all of the 33 samples were used to calibrate a model. This analysis indicated there was no relation between turbidity (TURB) less than 16 Formazin Nephelometric Unit (FNU) and SSC. Seventeen values less than 16 FNU were removed from the dataset, and the streamflow (Q) ranged from 6 cubic feet per second (cfs) to 53 cfs when these samples were collected. SSC samples collected in 2013 and 2014 were below the turbidity threshold. Values that were collected from the same streamflow event also were removed from the dataset (4 values) to reduce autocorrelation. This resulted in 12 SSC and turbidity pairs.

Also, because a sampling event usually takes longer than the 15-minute interval between turbidity measurements, the turbidity values used for model development were an average of three turbidity values: the turbidity value closest to when the SSC was collected (start time), 15-minutes after, and 30-minutes after.

MODEL DEVELOPMENT

An ordinary linear least squares regression analysis was carried out in the R open source environment (R Development Core Team, 2011), using different combinations of untransformed (Appendix 1) and $\log_{10}$-transformed data. SSC concentrations were evaluated for normal distribution of the data and potential outliers. The outlier test criteria flagged potential outliers and they were investigated. None of the SSC samples were deemed outliers. The distribution of residuals were examined for normality, and plots of residuals (the difference between the measured and predicted values) as compared to predicted SSC were examined for homoscedasticity (meaning that their departures from zero did not change substantially

over the range of predicted values). This comparison indicated that the data were not normally distributed and led to the conclusion that the most appropriate and reliable model would be one that estimated $\log_{10}$(SSC). An ordinary linear least squares regression analysis with $\log_{10}$-transformed data was performed. Log_{10}(TURB) and Log10(Q) were selected as the best predictors of $\log_{10}$(SSC) based on residual plots, relatively high adjusted coefficient of determination (adjusted R^2), and relatively low model standard percentage error (MSPE). Values of the aforementioned statistics and metrics were computed and are included below, along with all relevant sample data and more in-depth statistical information. The final model was chosen because at the highest measured SSC values, the predicted SSCs were closer to the measured than just using $\log_{10}$(TURB) (Appendix 2).

MODEL SUMMARY

Summary of final regression analysis for suspended-sediment concentration at site number 04015330. In the figures and tables below, the "log" in the R outputs refers to "$\log_{10}$". The model does not contain the lower ranges of streamflow because of non-significance. The model encompasses the range of streamflow from 57 to 2,860 cfs.
Suspended-sediment concentration-based model:
$\log_{10}(SSC) = [(1.18)\log_{10}(TURB)] + [(0.107)\log_{10}(Q)] - 0.892$,

where
SSC = suspended-sediment concentration, in milligrams per liter (mg/L),
TURB = Turbidity, in formazin nephelometric units (FNU) ; and,
Q = Streamflow, in cubic feet per second (cfs).

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. Streamflow makes sense physically because higher volumes of water tend to dilute these particles. The relation between turbidity and SSC can vary given varying concentrations of organic suspended particles that increase turbidity, but analysis of organic suspended particles was not included in the SSC analysis. The log-transformed model may be retransformed to the original units so that SSC can be calculated directly. The retransformation introduces a bias in the calculated constituent. This bias may be corrected using Duan's Bias Correction Factor (BCF) (Duan, 1983). For this model, the calculated BCF is 1.04. The retransformed model, accounting for BCF is:

$$SSC = 0.128 \times TURB^{1.18} \times Q^{0.107} \text{ x } 1.04.$$

Model Statistics, Data, and Plots (In the figures and tables below, the "log" in the R outputs refers to "$\log_{10}$")

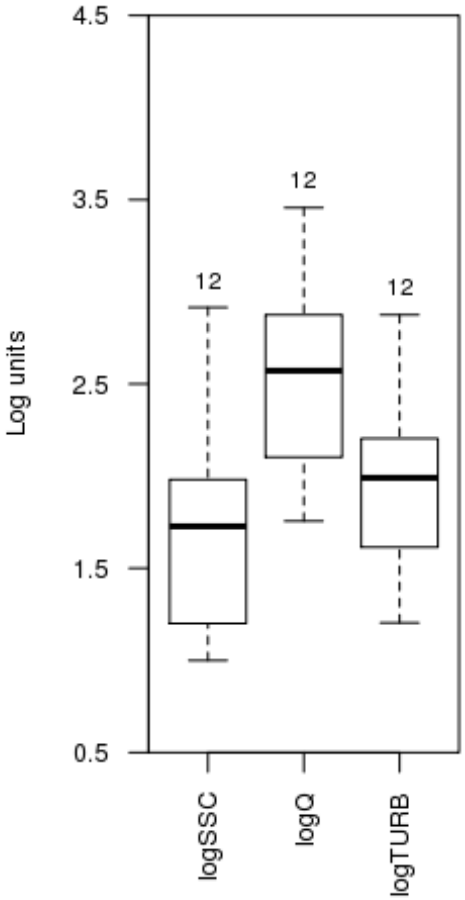
Model

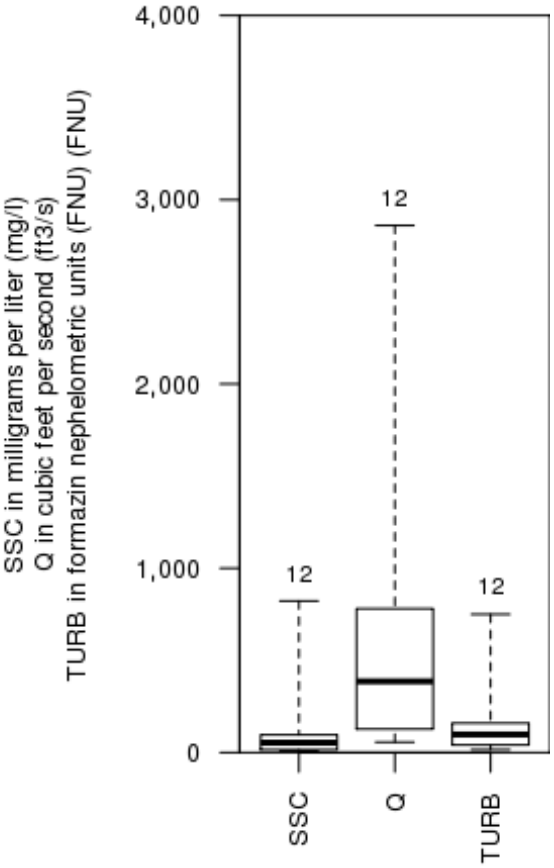
$\text{Log}_{10}\text{SSC} = [(1.18)\log_{10}(TURB)] + [(0.107)\log_{10}(Q)] - 0.892$

Variable Summary Statistics

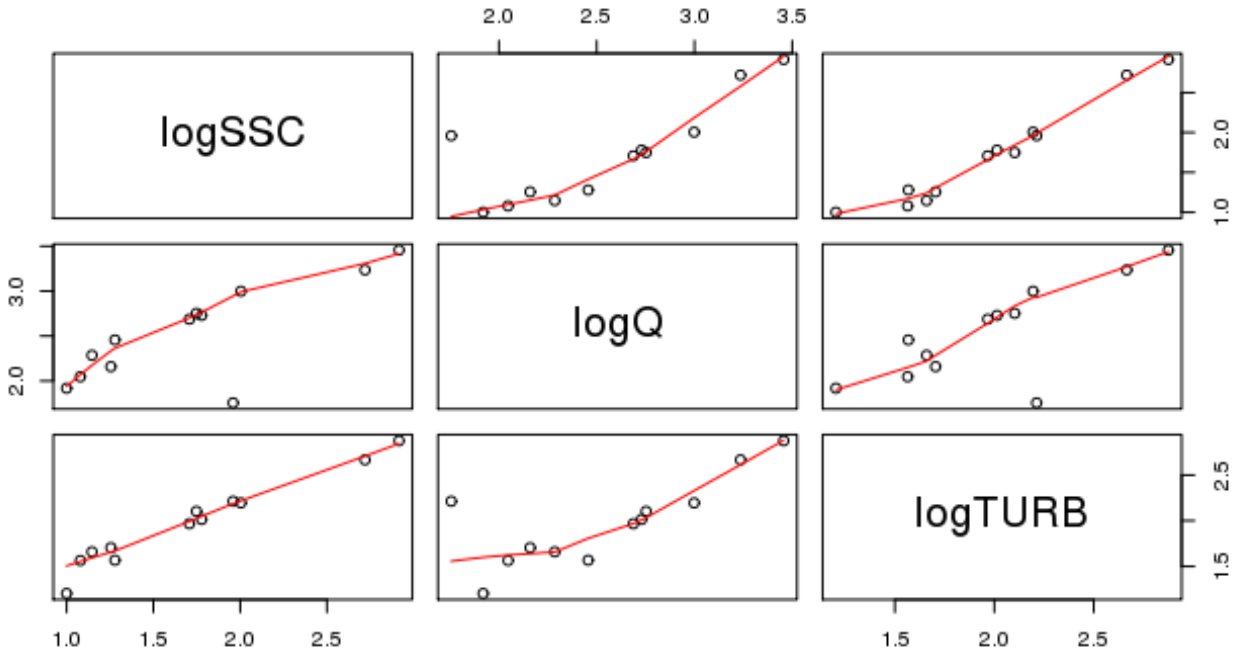
	Log_{10}SSC	SSC	$\log_{10}$Q	$\log_{10}$TURB	Q	TURB
Minimum	1.00	10.0	1.76	1.20	57	16.0
1st Quartile	1.20	16.0	2.10	1.61	128	41.3
Median	1.73	53.5	2.57	1.99	386	98.0
Mean	1.72	148.0	2.54	1.98	670	170.0
3rd Quartile	1.98	96.0	2.88	2.20	781	160.0
Maximum	2.92	823.0	3.46	2.88	2860	750.0

Box Plots


Log units
4.5
3.5
2.5
1.5
0.5
12
12
12
logSSC
logQ
logTURB


SSC in milligrams per liter (mg/l)
Q in cubic feet per second (ft3/s)
TURB in formazin nephelometric units (FNU) (FNU)
4,000
3,000
2,000
1,000
0
12
12
12
SSC
Q
TURB

Exploratory Plots



Basic Model Statistics

```
Number of Observations                                12
Standard error (RMSE)                              0.132
Average Model standard percentage error (MSPE)      30.9
Coefficient of determination (R²)                  0.963
Adjusted Coefficient of Determination (Adj. R²) 0.955
Bias Correction Factor (BCF)                        1.04

Variance Inflation Factors (VIF)
   logQ logTURB
   2.47    2.47
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)       -0.892          0.196  -4.540 1.40e-03
logQ               0.107          0.118   0.904 3.90e-01
logTURB            1.180          0.131   9.020 8.35e-06
```

Correlation Matrix

```
Intercept   logQ logTURB
Intercept     1.000 -0.508  -0.142
logQ         -0.508  1.000  -0.771
logTURB      -0.142 -0.771   1.000
```

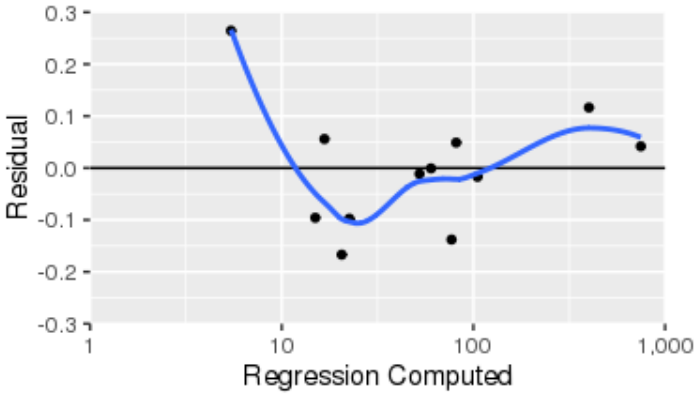
Outlier Test Criteria

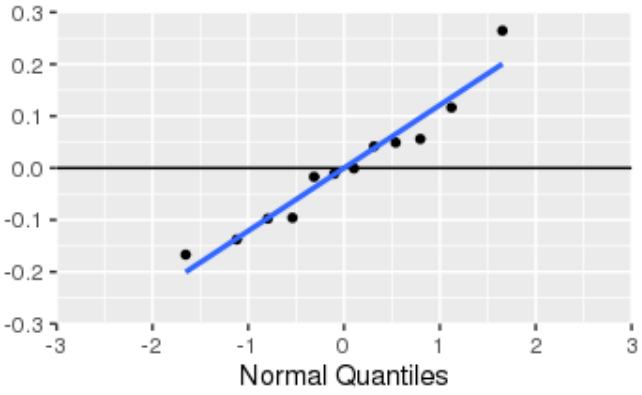
```
Leverage Cook's D   DFFITS
   0.500    0.191    0.816
```

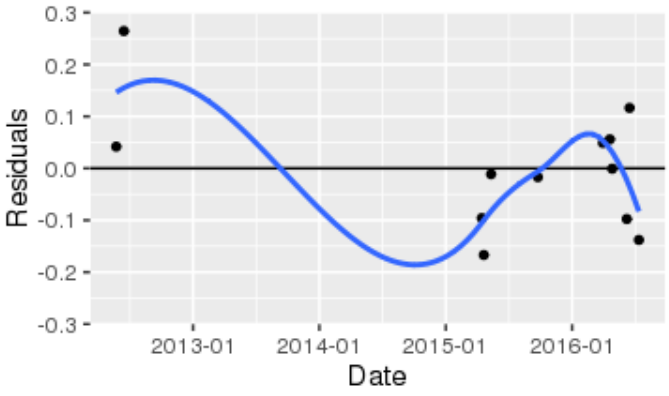
Flagged Observations

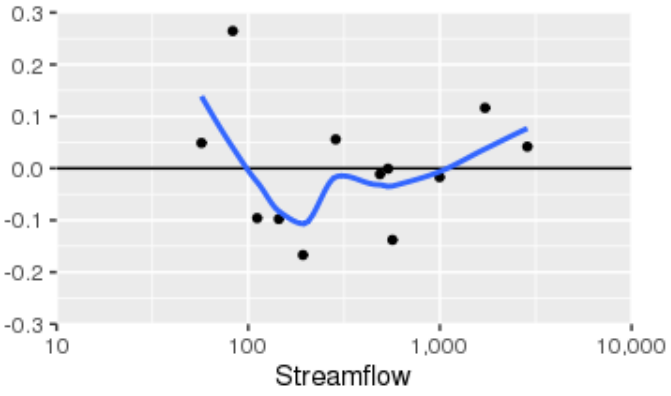
```
                 logSSC Estimate Residual Standard Residual Studentized Residual Leverage Cook's D DFFITS
6/15/2012 13:15   1.00    0.735   0.2650              2.44                 3.93    0.323    0.945   2.72
3/30/2016 14:45   1.96    1.910   0.0489              1.07                 1.08    0.880    2.810   2.93
```

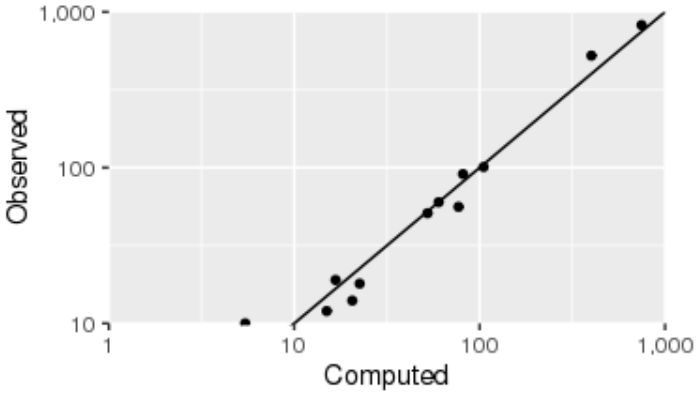
Statistical Plots

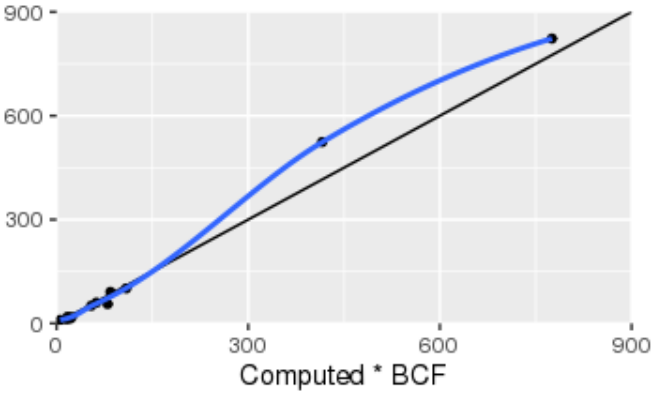


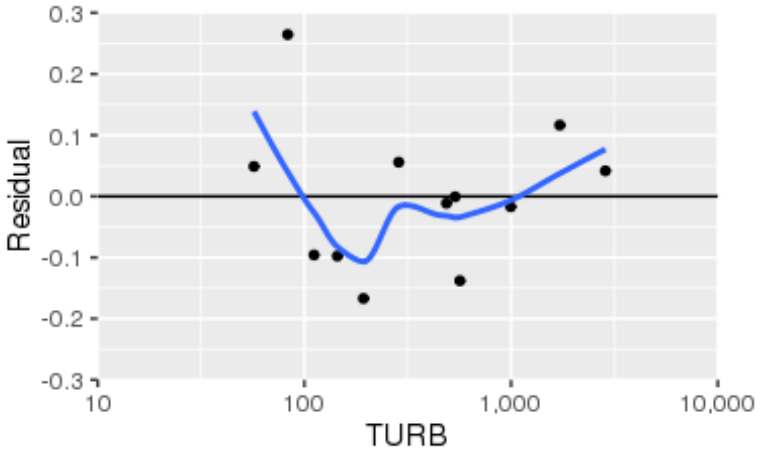


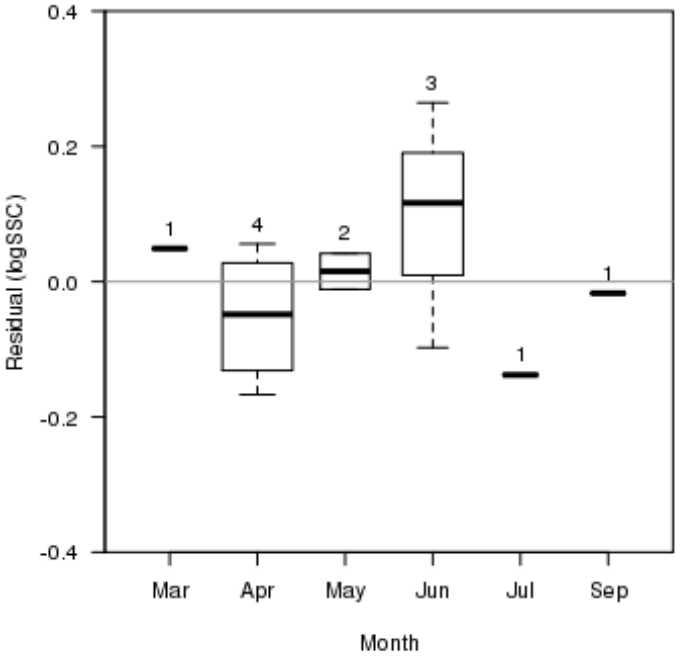


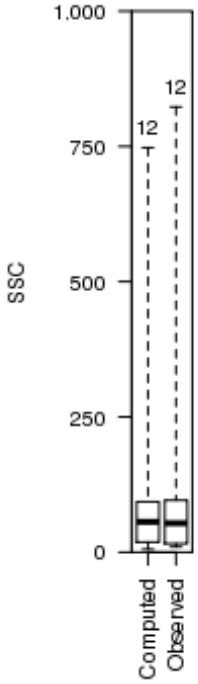


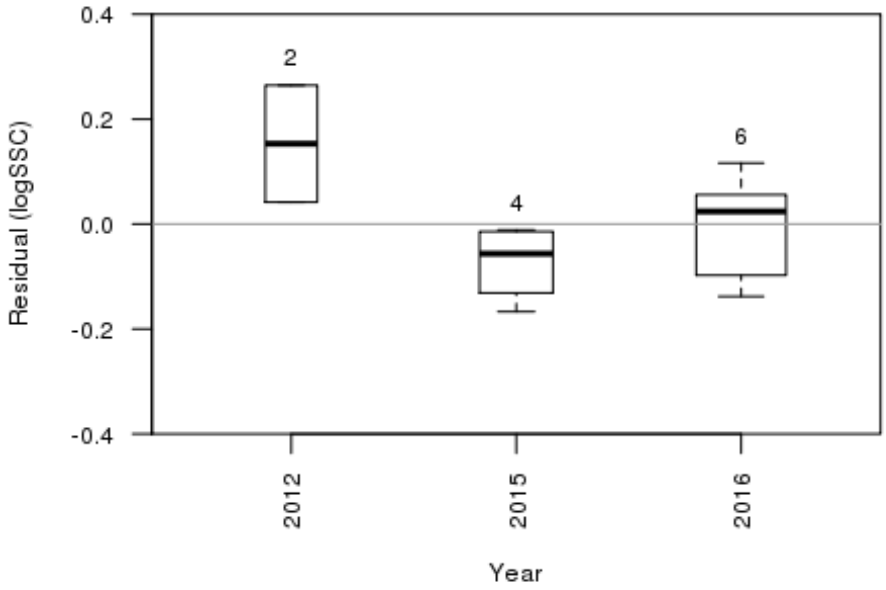




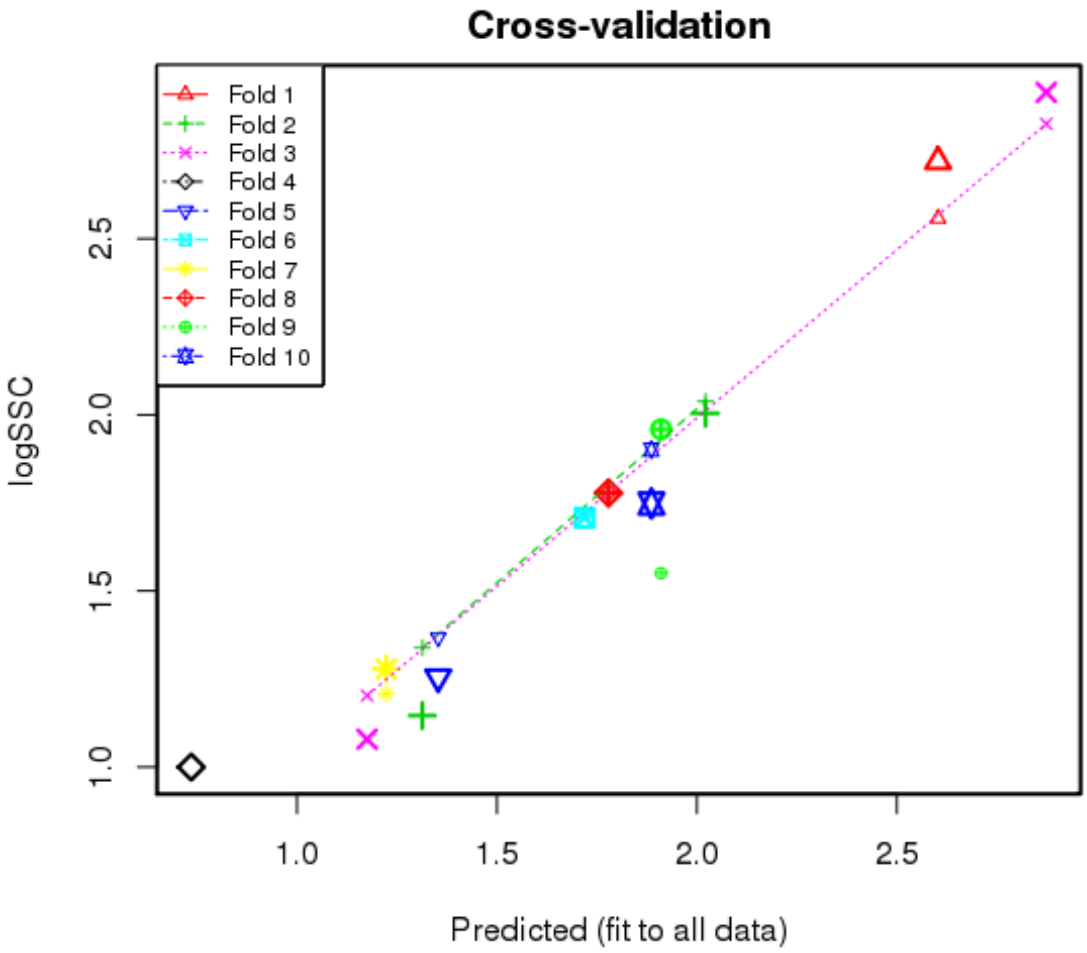

Residual
0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
10
100
1,000
10,000
TURB


Residual (logSSC)
0.4
0.2
0.0
-0.2
-0.4
1
4
2
3
1
1
Mar
Apr
May
Jun
Jul
Sep
Month


SSC
1,000
750
500
250
0
12
12
Computed
Observed

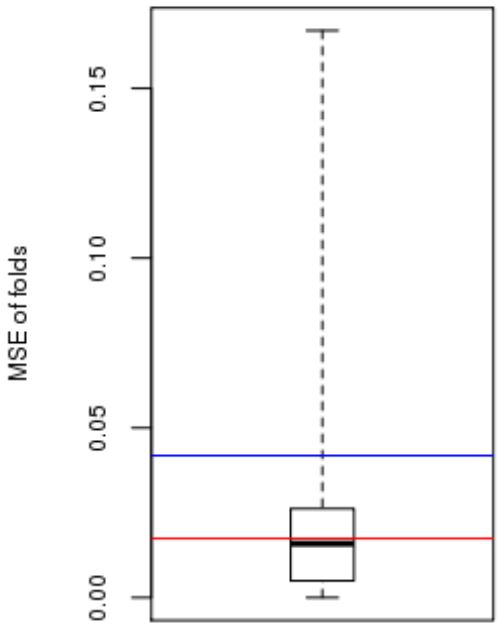


Cross Validation




```
Minimum MSE of folds:  6.53e-07
   Mean MSE of folds:  4.18e-02
```

```
              Median MSE of folds:  1.59e-02
             Maximum MSE of folds:  1.67e-01
(Mean MSE of folds) / (Model MSE):  2.40e+00
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Data Set

	Date	logSSC	logQ	logTURB	SSC	Q	TURB	Computed logSSC	Computed SSC	Residual	Normal Quantiles
1	5/24/2012	2.92	3.46	2.88	823	2860	750	2.87	775	0.0418	0.313
2	6/15/2012	1	1.92	1.2	10	83	16	0.735	5.64	0.265	1.65
3	4/15/2015	1.08	2.05	1.56	12	111	36.7	1.18	15.5	-0.0959	-0.539
4	4/21/2015	1.15	2.29	1.66	14	193	45.7	1.31	21.3	-0.167	-1.65
5	5/12/2015	1.71	2.69	1.97	51	487	92.7	1.72	54.3	-0.0113	-0.103
6	9/24/2015	2	3	2.2	101	996	157	2.02	109	-0.017	-0.313
7	3/30/2016	1.96	1.76	2.21	91	57	163	1.91	84.3	0.0489	0.539
8	4/19/2016	1.28	2.46	1.57	19	286	37	1.22	17.3	0.0558	0.796
9	4/26/2016	1.78	2.73	2.01	60	536	103	1.78	62.3	-0.00072	0.103
10	6/7/2016	1.26	2.16	1.71	18	144	50.7	1.35	23.4	-0.0976	-0.796
11	6/15/2016	2.72	3.24	2.67	524	1720	463	2.6	416	0.116	1.12
12	7/12/2016	1.75	2.75	2.1	56	566	127	1.89	79.8	-0.138	-1.12

Definitions

```
SSC: Suspended sediment concentration (SSC) in mg/l (80154)
TURB: Turbidity in FNU (63680)
Q: Stream flow, mean. daily in ft3/s (00060)
```

App Version 1.0

References

Duan, Naihua, 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610. [Also available at http://dx.doi.org/10.2307/2288126.]

Edwards, T.K., and Glysson, G.D., 1999, Field methods for measurement of fluvial sediment: U.S. Geological Survey Techniques of Water-Resources Investigations, book 3, chap. C2, 89 p. [Also available at http://pubs.usgs.gov/twri/twri3-c2/.]

R Development Core Team, 2011, R installation and administration, Version 2.14.1, 2011-12-22: 73 p. [Also available at http://streaming.stat.iastate.edu/CRAN/doc/manuals/Radmin.pdf.]

Tornes, L.H., 1986, Suspended sediment in Minnesota streams: U.S. Geological Survey Water-Resources Investigations Report 85–4312, 33 p. [Also available at http://pubs.er.usgs.gov/publication/wri854312.]

U.S. Geological Survey, 2017, National Water Information System (NWISWeb)—USGS surface-water data for Minnesota: U.S. Geological Survey database, accessed January, 2017, at http://waterdata.usgs.gov/mn/nwis/sw/.[Also available at http://dx.doi.org/10.5066/F7P55KJN.]

Appendix 1

Model Statistics, Data, and Plots

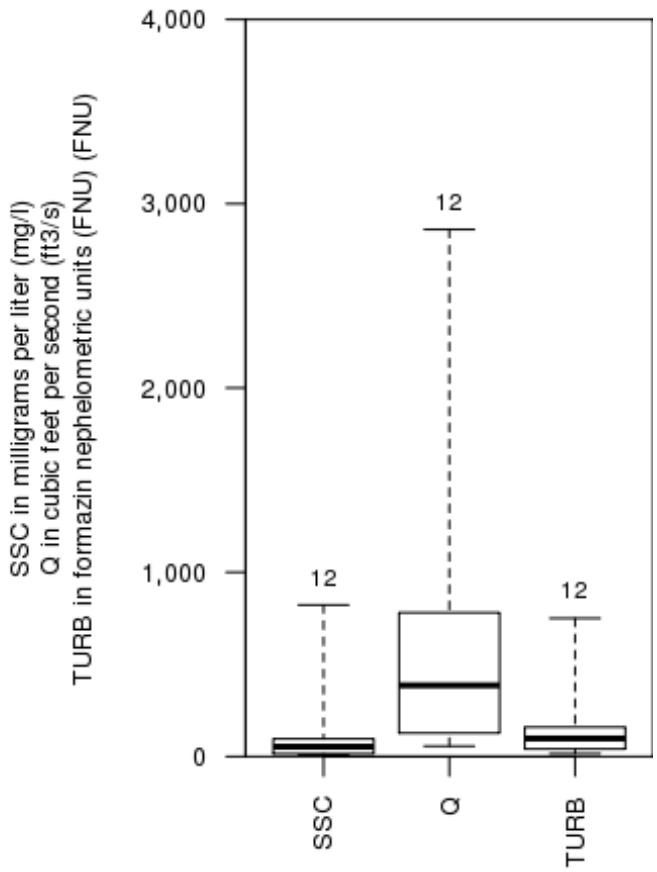
Model

SSC = + 0.0214 * Q + 1.08 * TURB - 50.4

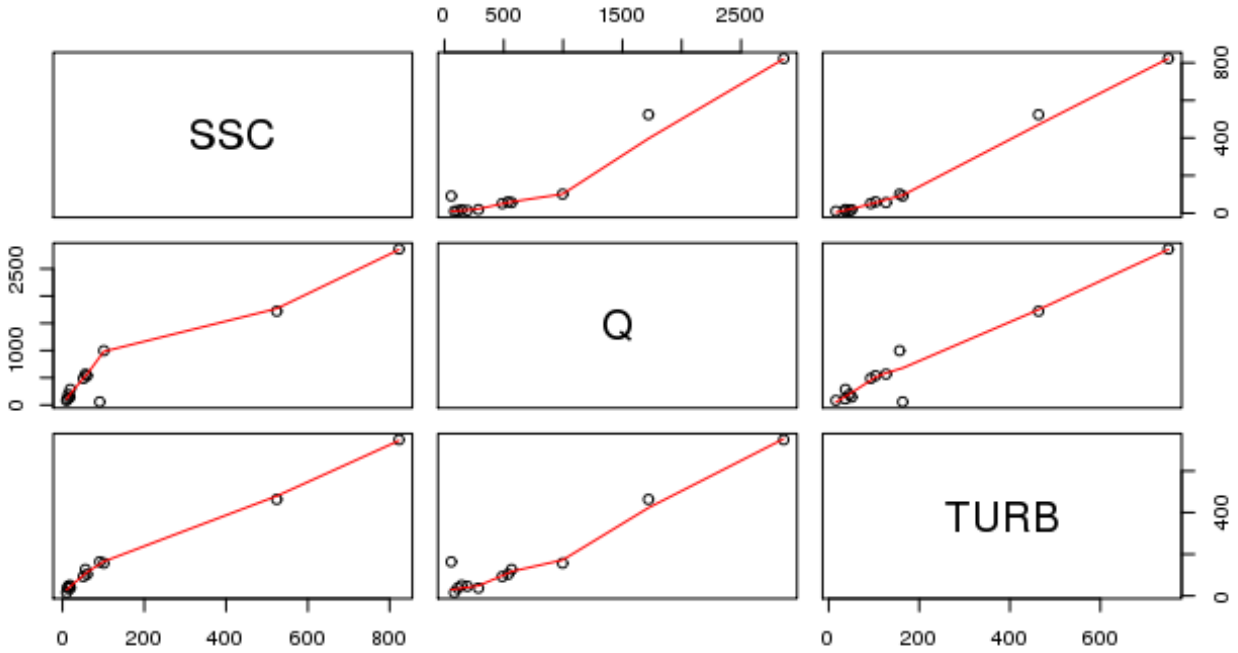
Variable Summary Statistics

	SSC	Q	TURB
Minimum	10.0	57	16.0
1st Quartile	16.0	128	41.4
Median	53.5	386	98.0
Mean	148.0	670	170.0
3rd Quartile	96.0	781	160.0
Maximum	823.0	2860	750.0

Box Plots



Exploratory Plots



Basic Model Statistics

```
Number of Observations                              12
Standard error (RMSE)                             31.9
Average Model standard percentage error (MSPE)    21.5
Coefficient of determination (R²)                0.987
Adjusted Coefficient of Determination (Adj. R²) 0.984

Variance Inflation Factors (VIF)
   Q TURB
14.4 14.4
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)     -50.4000        12.0000  -4.200 0.002320
Q                 0.0214         0.0435   0.493 0.634000
TURB              1.0800         0.1670   6.480 0.000114
```

Correlation Matrix

```
          Intercept      Q    TURB
Intercept    1.0000 -0.140 -0.0305
Q           -0.1400  1.000 -0.9640
TURB        -0.0305 -0.964  1.0000
```

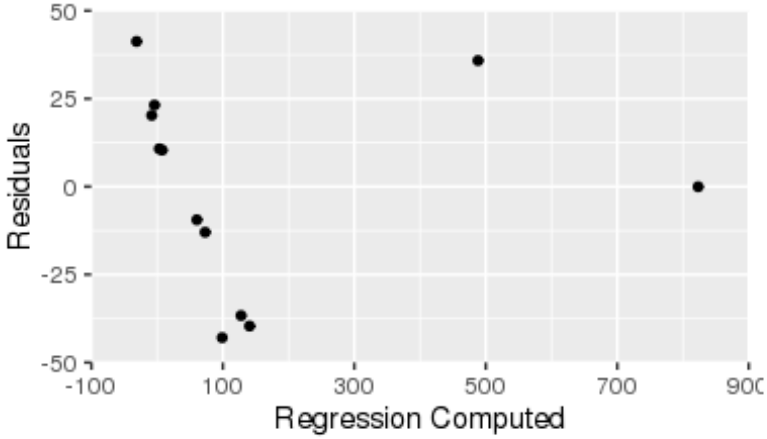
Outlier Test Criteria

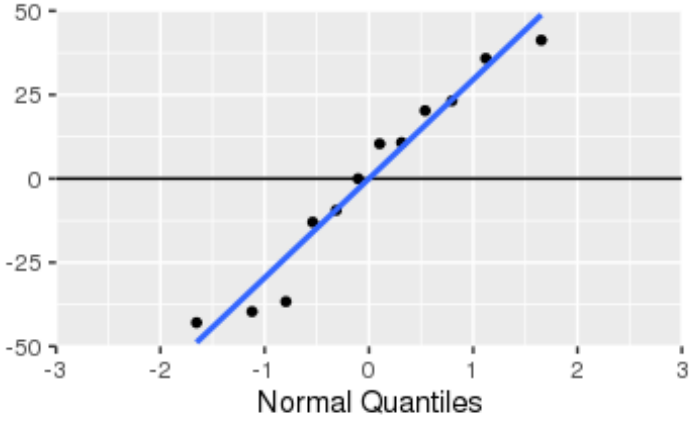
```
Leverage Cook's D   DFFITS
   0.500    0.191    0.816
```

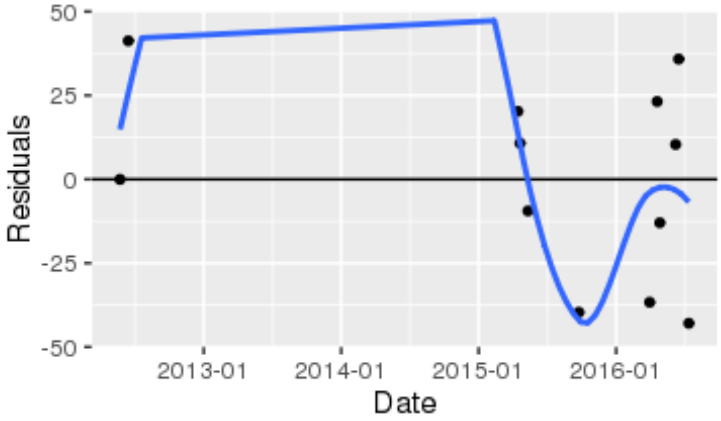
Flagged Observations

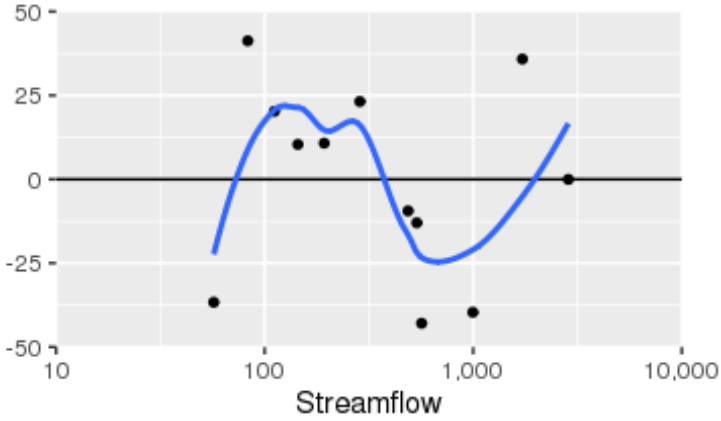
	SSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
5/24/2012 17:45	823	823	-0.0391	-0.00235	-0.00222	0.728	4.93e-06	-0.00362
9/24/2015 15:00	101	141	-39.6000	-1.53000	-1.68000	0.345	4.14e-01	-1.22000
3/30/2016 14:45	91	128	-36.7000	-2.18000	-2.99000	0.722	4.12e+00	-4.82000

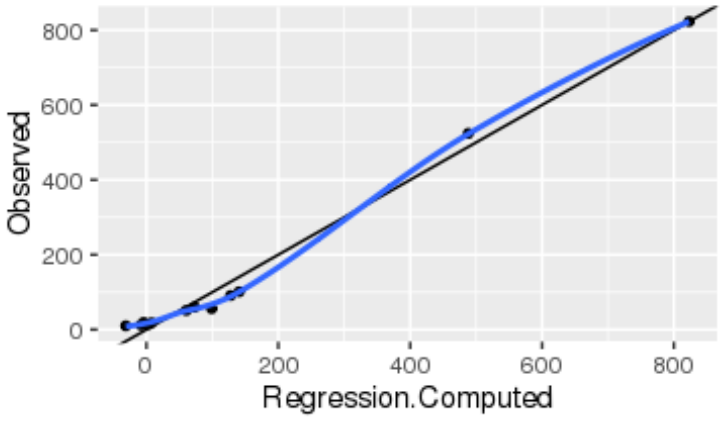
Statistical Plots

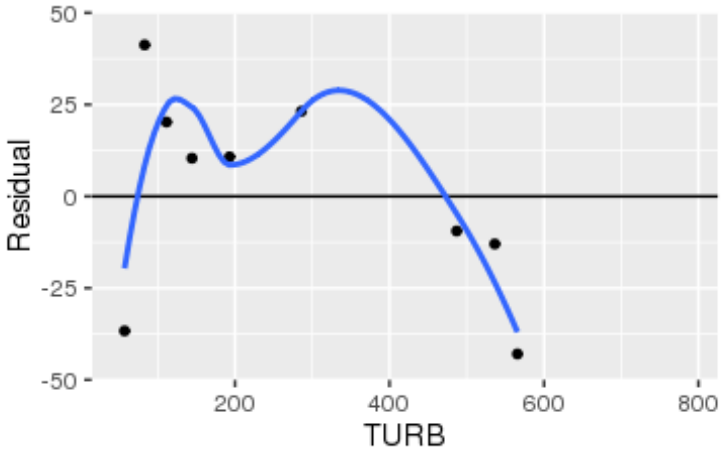


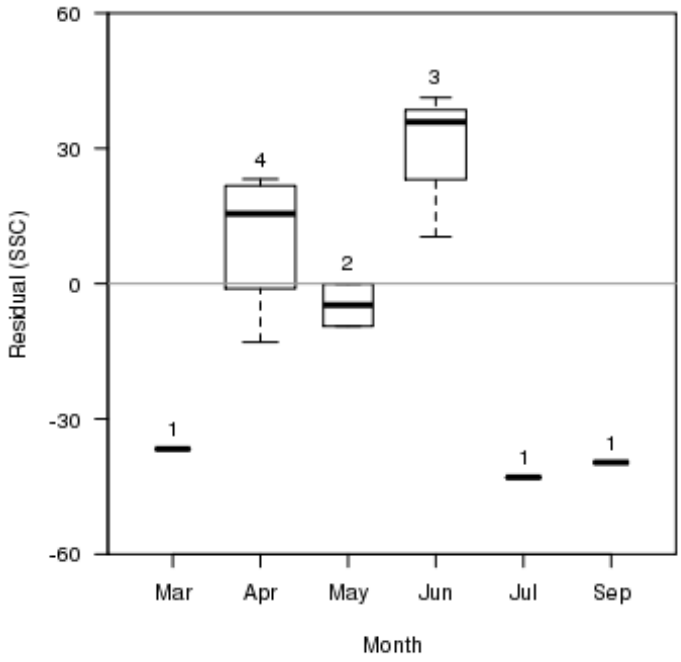


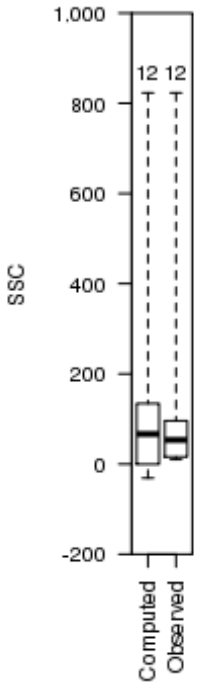


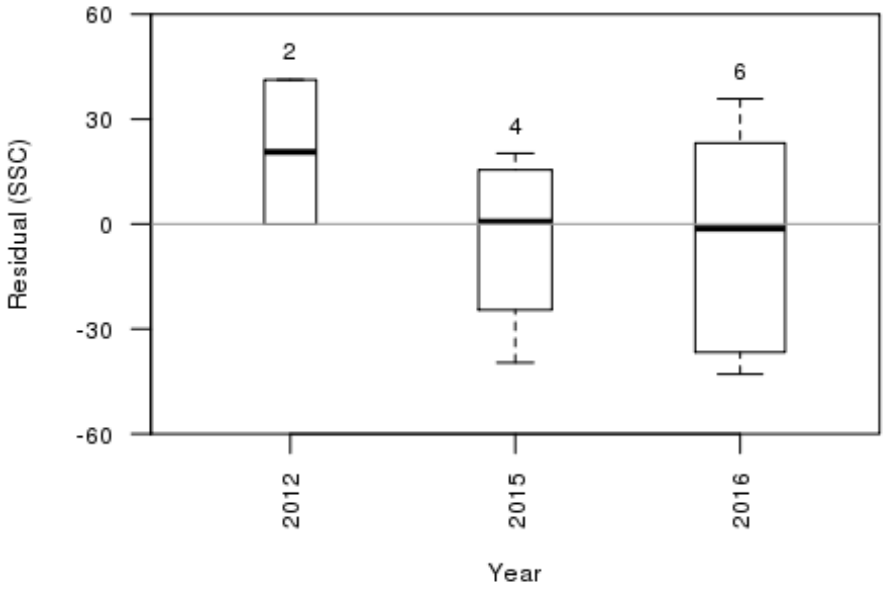




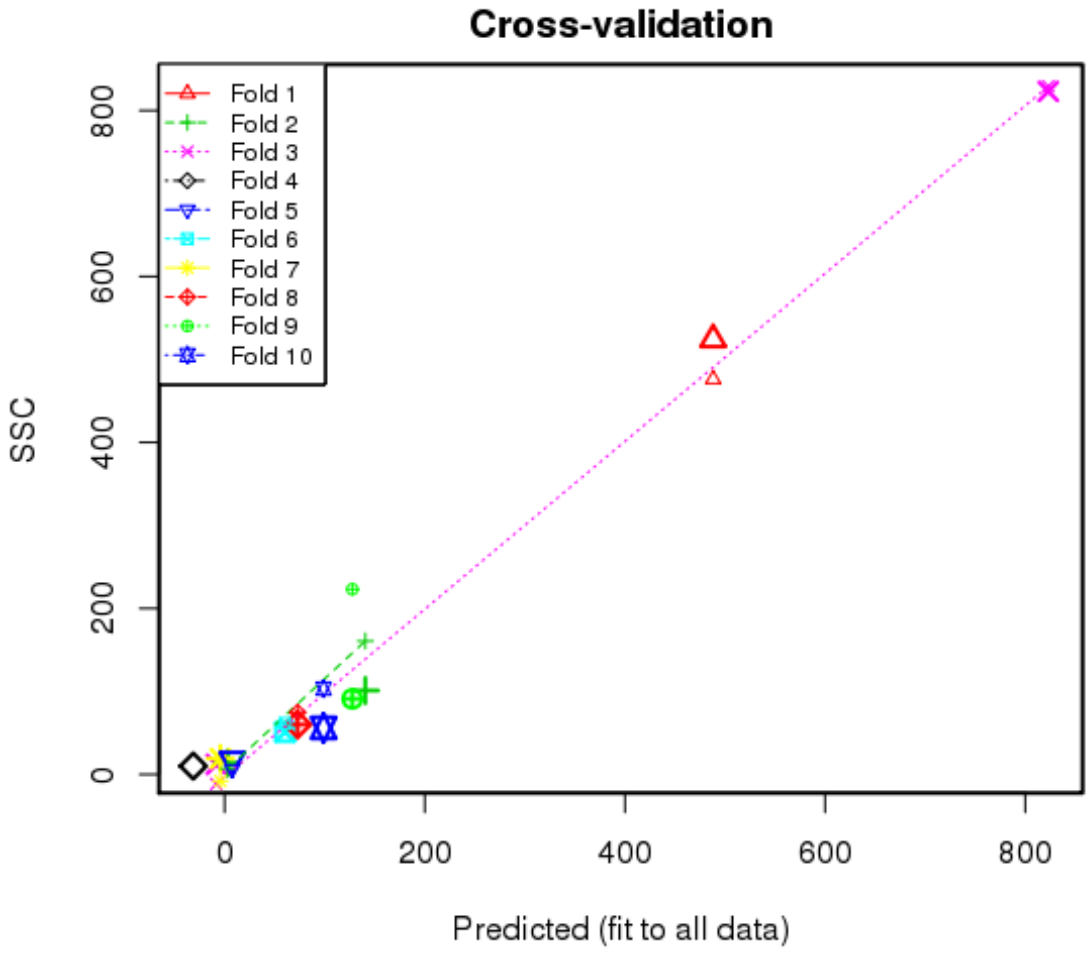

Residual
50
25
0
-25
-50
200
400
600
800
TURB


Residual (SSC)
60
30
0
-30
-60
1
4
2
3
1
1
Mar
Apr
May
Jun
Jul
Sep
Month


SSC
1,000
800
600
400
200
0
-200
12
12
Computed
Observed

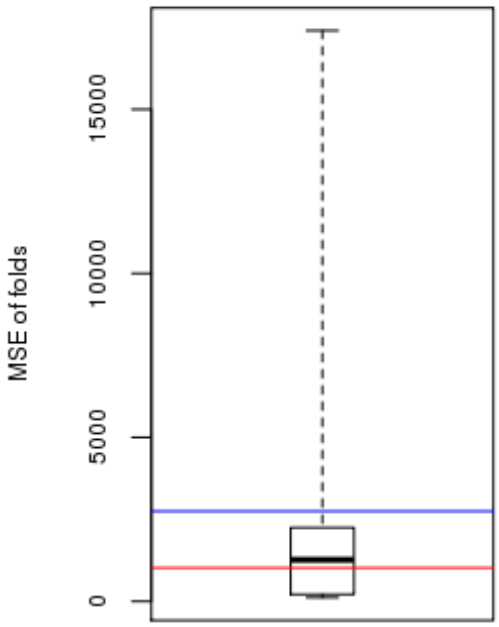


Cross Validation




```
Minimum MSE of folds:    114.00
   Mean MSE of folds:   2750.00
```

```
              Median MSE of folds:   1270.00
             Maximum MSE of folds:  17400.00
(Mean MSE of folds) / (Model MSE):      2.69
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Data Set

	Date	SSC	Q	TURB	Computed SSC	Residual	Normal Quantiles	Censored Values
1	2012-05-24	823	2860	750	823	-0.0391	-0.103	--
2	2012-06-15	10	83	16	-31.3	41.3	1.65	--
3	2015-04-15	12	111	36.7	-8.25	20.3	0.539	--
4	2015-04-21	14	193	45.7	3.24	10.8	0.313	--
5	2015-05-12	51	487	92.7	60.4	-9.44	-0.313	--
6	2015-09-24	101	996	157	141	-39.6	-1.12	--
7	2016-03-30	91	57	163	128	-36.7	-0.796	--
8	2016-04-19	19	286	37	-4.19	23.2	0.796	--
9	2016-04-26	60	536	103	73	-13	-0.539	--
10	2016-06-07	18	144	50.7	7.61	10.4	0.103	--
11	2016-06-15	524	1720	463	488	35.8	1.12	--
12	2016-07-12	56	566	127	98.9	-42.9	-1.65	--

Definitions

```
SSC: Suspended sediment concentration (SSC) in mg/l (80154)
Q: Stream flow, mean. daily in ft3/s (00060)
TURB: Turbidity in FNU (63680)
```

App Version 1.0

Appendix 2

Model Statistics, Data, and Plots

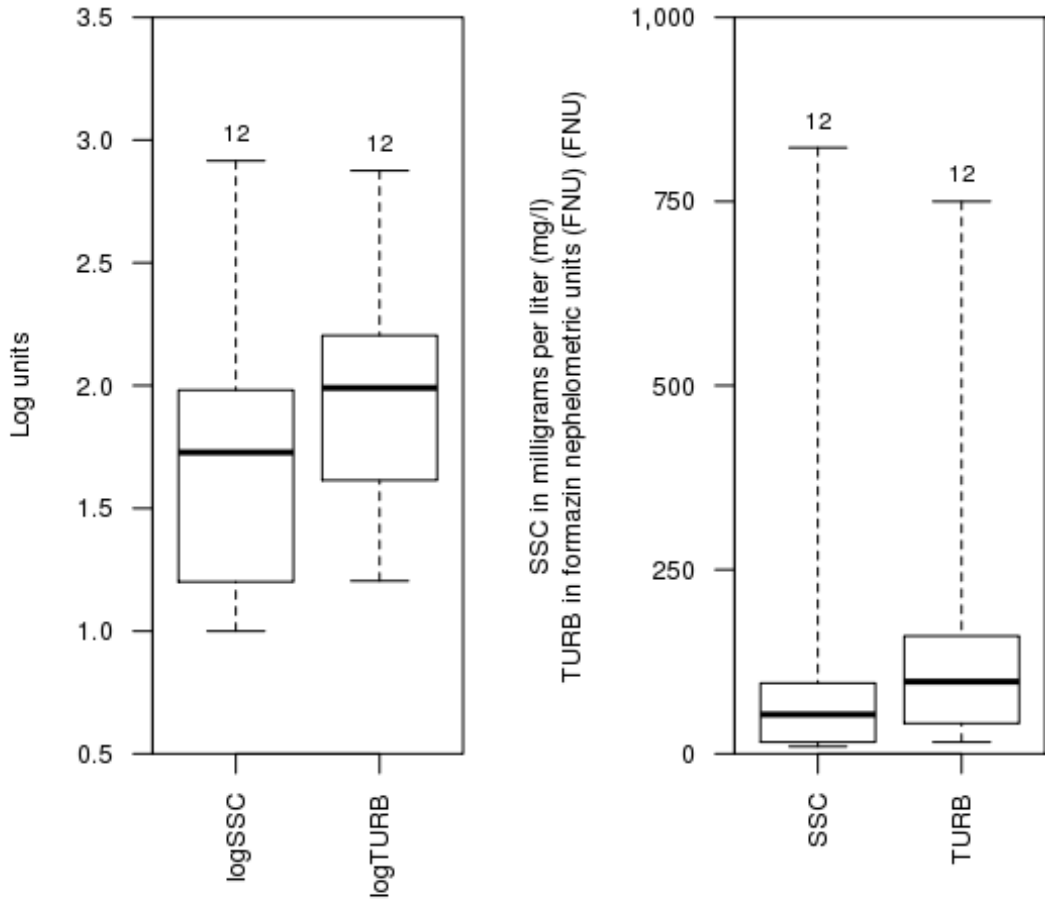
Model

logSSC = + 1.27 * logTURB - 0.802

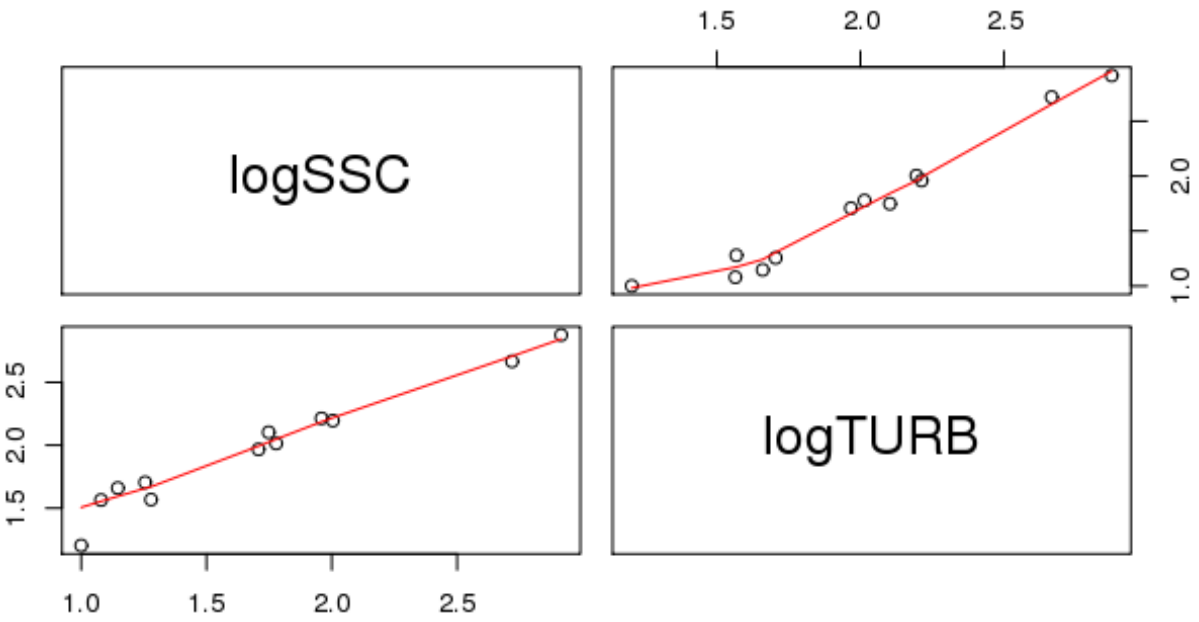
Variable Summary Statistics

	logSSC	SSC	logTURB	TURB
Minimum	1.00	10.0	1.20	16.0
1st Quartile	1.20	16.0	1.61	41.3
Median	1.73	53.5	1.99	98.0
Mean	1.72	148.0	1.98	170.0
3rd Quartile	1.98	96.0	2.20	160.0
Maximum	2.92	823.0	2.88	750.0

Box Plots



Exploratory Plots



Basic Model Statistics

```
Number of Observations                              12
Standard error (RMSE)                            0.131
Average Model standard percentage error (MSPE)    30.6
Coefficient of determination (R²)                 0.96
Adjusted Coefficient of Determination (Adj. R²) 0.956
Bias Correction Factor (BCF)                      1.04
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)       -0.802         0.1680   -4.78 7.41e-04
logTURB            1.270         0.0825   15.40 2.68e-08
```

Correlation Matrix

```
          Intercept E.vars
Intercept     1.000 -0.974
E.vars       -0.974  1.000
```

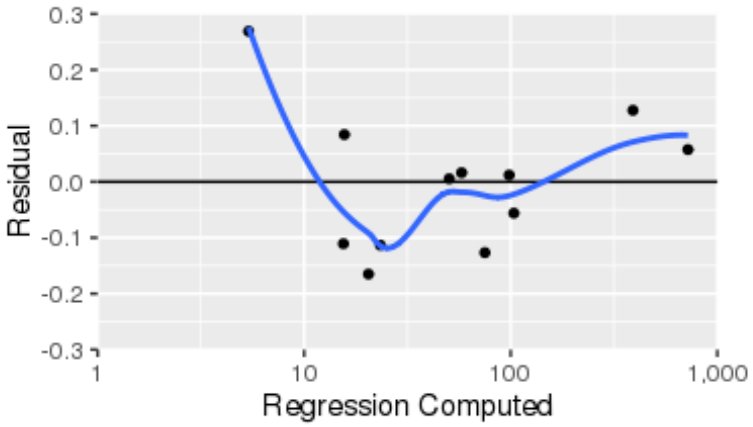
Outlier Test Criteria

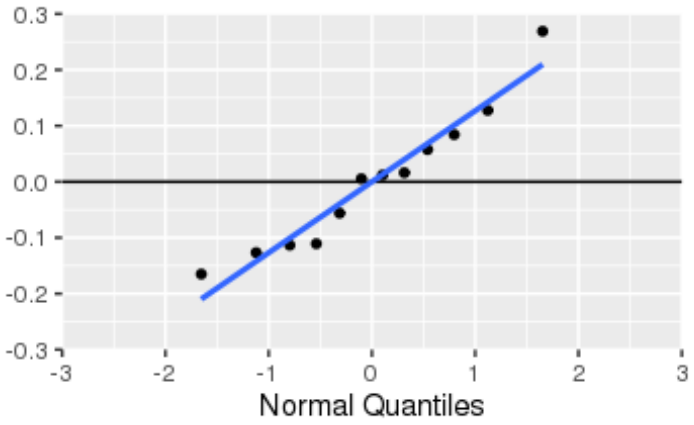
```
Leverage Cook's D   DFFITS
   0.250    0.106    0.577
```

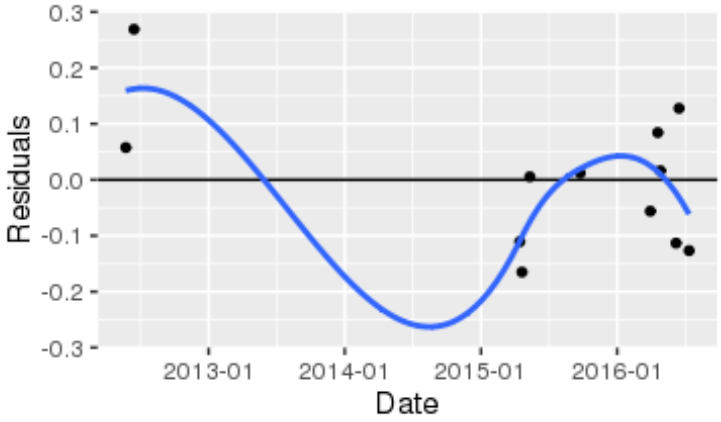
Flagged Observations

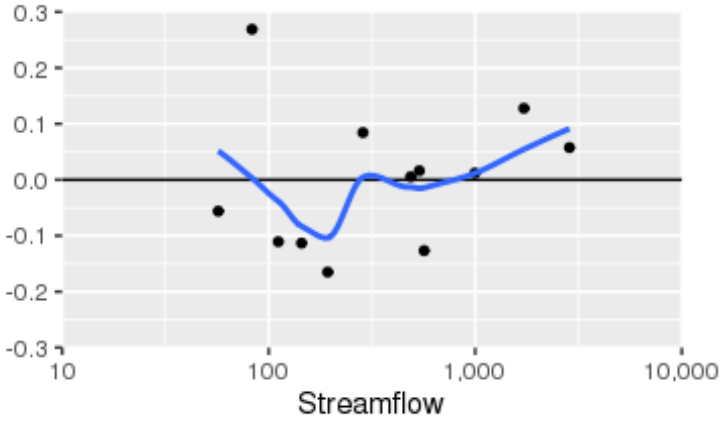
	logSSC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
5/24/2012 17:45	2.92	2.860	0.0575	0.569	0.549	0.404	0.110	0.452
6/15/2012 13:15	1.00	0.731	0.2690	2.500	3.860	0.322	1.480	2.660
4/21/2015 17:45	1.15	1.310	-0.1650	-1.350	-1.410	0.124	0.128	-0.531
6/15/2016 17:00	2.72	2.590	0.1280	1.140	1.160	0.272	0.244	0.711

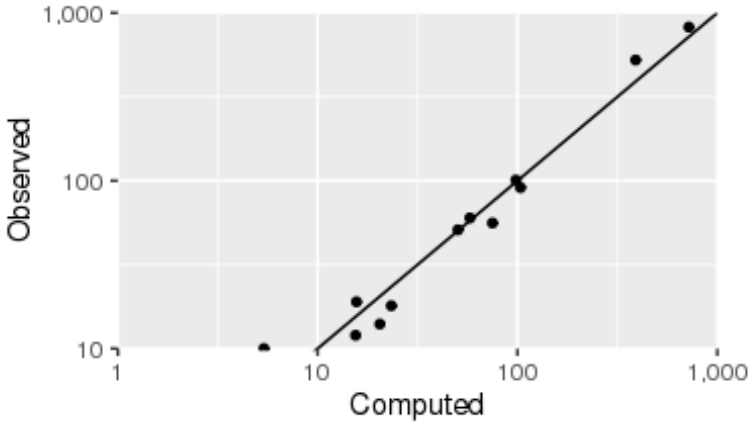
Statistical Plots

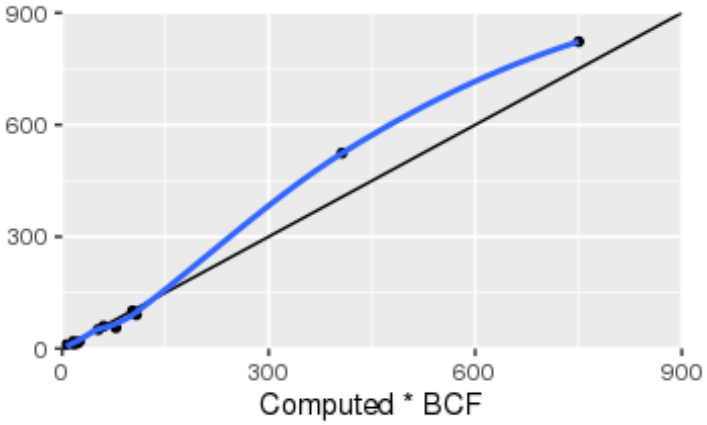


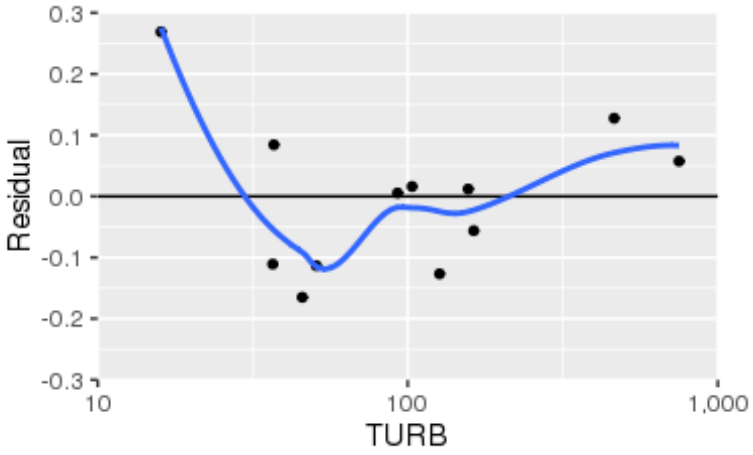


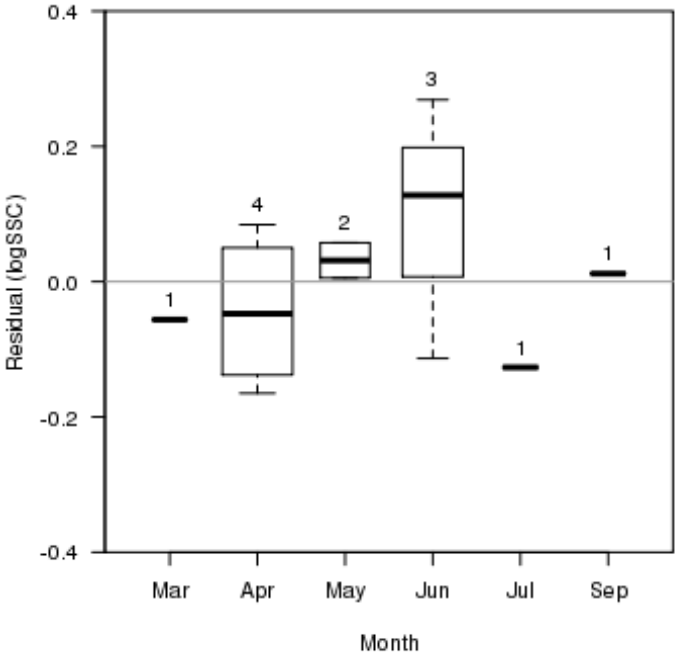


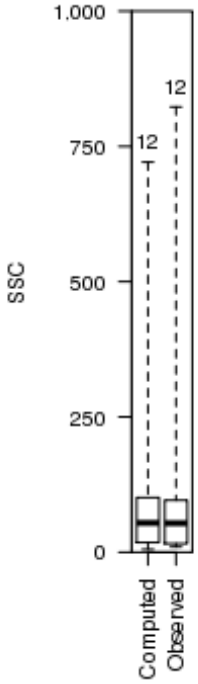


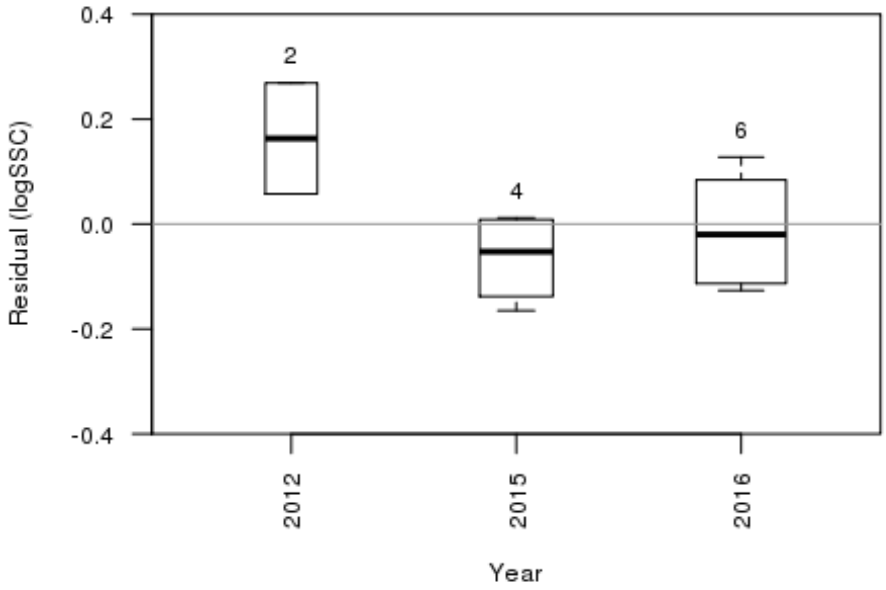




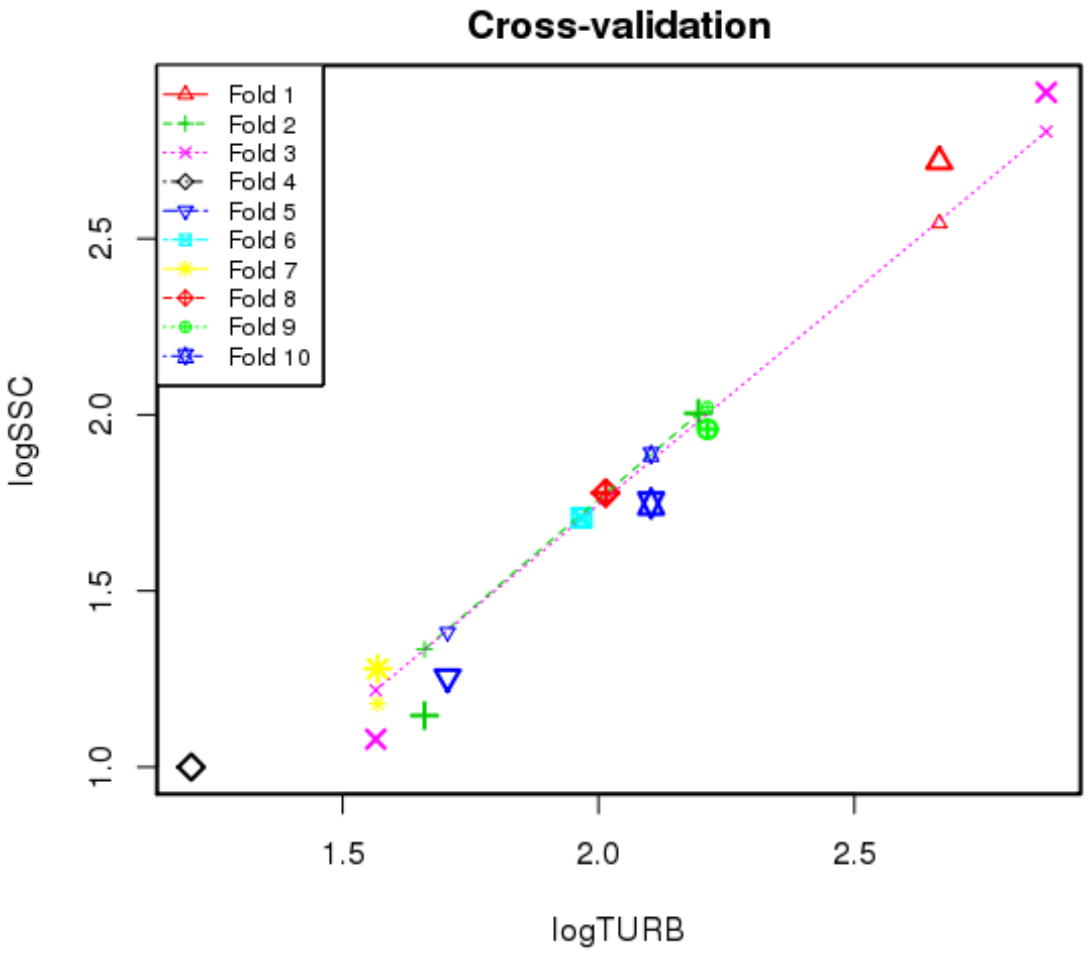

0.3
0.2
0.1
0.0
-0.1
-0.2
-0.3
Residual
10
100
1,000
TURB


0.4
0.2
0.0
-0.2
-0.4
Residual (logSSC)
1
4
2
3
1
1
Mar
Apr
May
Jun
Jul
Sep
Month


1,000
750
500
250
0
SSC
12
12
Computed
Observed

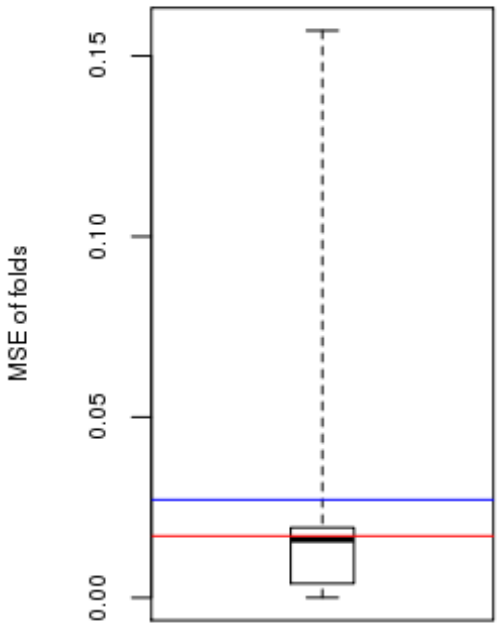


Cross Validation




```
Minimum MSE of folds:  3.47e-05
   Mean MSE of folds:  2.71e-02
```

```
              Median MSE of folds:  1.61e-02
             Maximum MSE of folds:  1.57e-01
(Mean MSE of folds) / (Model MSE):  1.59e+00
```




```
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	2012-05-24	2.92	2.88	823	750	2.86	750	0.0575	0.539	--
2	2012-06-15	1	1.2	10	16	0.731	5.6	0.269	1.65	--
3	2015-04-15	1.08	1.56	12	36.7	1.19	16.1	-0.111	-0.539	--
4	2015-04-21	1.15	1.66	14	45.7	1.31	21.3	-0.165	-1.65	--
5	2015-05-12	1.71	1.97	51	92.7	1.7	52.4	0.0054	-0.103	--
6	2015-09-24	2	2.2	101	157	1.99	102	0.0119	0.103	--
7	2016-03-30	1.96	2.21	91	163	2.02	108	-0.0561	-0.313	--
8	2016-04-19	1.28	1.57	19	37	1.19	16.3	0.0843	0.796	--
9	2016-04-26	1.78	2.01	60	103	1.76	60.2	0.0161	0.313	--
10	2016-06-07	1.26	1.71	18	50.7	1.37	24.3	-0.113	-0.796	--
11	2016-06-15	2.72	2.67	524	463	2.59	406	0.128	1.12	--
12	2016-07-12	1.75	2.1	56	127	1.87	78	-0.127	-1.12	--

Definitions

```
SSC: Suspended sediment concentration (SSC) in mg/l (80154)
TURB: Turbidity in FNU (63680)
```

App Version 1.0