Appendix 1. Model Archive for Total Carbon Concentration at U.S. Geological Survey Station 022889001: Tamiami Canal 11 Mile Road to Monroe Station, Florida

This model archive summary summarizes the total carbon (TC) (includes total particulate carbon [TPC], dissolved organic carbon [DOC], and dissolved inorganic carbon [DIC]) model developed to compute 15-minute TC concentrations and loads from September 2015 to October 2017.

The purpose of this model is to calculate continuous TC along Tamiami Canal at 11 Mile Road to Monroe Station (022889001). Station 022889001 represents outflow at 12 culverts along U.S. Highway 41 between 11 Mile Road and Monroe Station; culverts are located under bridges 97 to 108. Data will be used as an input to a larger carbon budget study for Sweetwater Strand.

Site and Model Information

Site number: 022889001

Site name: Tamiami Canal 11 Mile Road to Monroe Station, Florida

Location: lat 25°51'05" N., long 80°58'50" W., referenced to the North American Datum of 1927, Collier County, Florida, hydrologic unit 03090204

Equipment: A YSI EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, turbidity, and fluorescence of chromophoric (colored) dissolved organic matter (fDOM, a proxy for DOC). The monitor is housed in an 8-inch polyvinyl chloride (PVC) pipe off the end of the structure into the stream at bridge 105 (station number 02288900) along the Tamiami Trail. Readings from the monitor were recorded every 15 minutes and transmitted hourly by way of satellite. The model applies only to this site (022889001) and specified time period (September 9, 2015–October 3, 2017).

Model number: 1.0

Date model was created: April 23, 2018

Model calibration data period: September 10, 2015–October 3, 2017

Model application date: September 9, 2015–October 3, 2017

Computed by: Amanda Booth, Caribbean-Florida Water Science Center

Reviewed by: Mandy Stone, Kansas Water Science Center, and Michael Rosen, U.S. Geological Survey (USGS) Hydro-Eco Interactions Branch

Approved by: David Sumner, Caribbean-Florida Water Science Center

Model Data

All data were collected using USGS protocols and are stored in the National Water Information System (NWIS) database (U.S. Geological Survey, 2019) with the exception of DIC, which is included in this appendix. The methods used follow USGS guidance in Rasmussen and others (2009).

Specific conductance (SC) was published at 02288900. The regression model is based on 22 concurrent measurements of DOC, DIC, and TPC samples and concomitant SC measurements collected from September 10, 2015, to October 3, 2017. Samples were collected throughout the range of observed hydrologic conditions. Summary statistics and the complete model-calibration data are provided in the dataset. All data were retained in the model; no data were removed.

Discrete Sample Data

Point samples were collected at the location and depth of the water-quality sensors at 02288900. Samples used in model development were analyzed for TPC, DOC, and DIC by the USGS National Water Quality Lab (NWQL). DIC is not an approved method at the NWQL. Data were collected under laboratory information management system proposal #CL15025. Replicates were collected at 02288900 on March 10, 2016, June 14, 2017, and October 3, 2017. Blanks were collected at 02288900 on May 4, 2017. All replicates were sequential unless otherwise noted.

Absolute differences between replicate pairs of DOC ranged from −0.48 to −0.01 milligram per liter (mg/L). Relative percent differences ranged from −5 to 0 percent for DOC. The DOC blank was <0.23 mg/L (the detection limit). DIC replicate pairs varied by −0.2 to 0.6 mg/L. Relative percent differences from the replicate samples ranged from 0 to 2 percent. The DIC blank was 0.3 mg/L.

TPC sequential replicate pairs varied by 0.048 to 0.148 mg/L. Relative percent differences from the sequential replicate samples ranged from 10 to 39 percent. One split replicate for TPC was done on October 3, 2017; the difference was 0.032 mg/L, with a relative percent difference of −7 percent. TPC blanks were below the detection limit.

Surrogate Data

Specific conductance at 02288900 ranged from 90 microsiemens per centimeter at 25 degrees Celsius (µS/cm at 25°C) on June 7, 2017, to 624 µS/cm at 25°C on May 13, 2017. The SC sensor was maintained, and the data were computed using Wagner and others (2006).

Model Development

Regression analysis was done using Microsoft Excel and the USGS Surrogate Analysis and Index Developer (SAID) tool (Domanski and others, 2015) by examining

continuously measured water-quality data as explanatory variables for estimating TC concentration. The distribution of residuals was examined for normality. Scatter plots of residuals (the difference between the measured and predicted values) compared to predicted TC were examined for homoscedasticity. SC was selected as the best predictor of TC based on residual plots, relatively high adjusted coefficient of determination (adjusted R^2), and relatively low model standard percentage error (MSPE). Values for all of the aforementioned statistics and metrics were computed and are included below, along with all relevant sample data and more in-depth statistical information. When discharge equaled zero, a discharge value (Q) of 0.001 was entered in order for the program to create the graphics.

Model Summary

Summary of final regression analysis for TC at site number 022909471.

Total carbon concentration-based model:

$$TC = 0.129(SC) + 6.62$$

where

TC = total carbon (dissolved organic carbon, dissolved inorganic carbon, and particulate carbon), in milligrams per liter (mg/L);

and

SC = specific conductance, in microsiemens per centimeter at 25 degrees (µS/cm at 25°C).

The use of SC as a variable is supported statistically and correlates strongly with DIC (Curtis and Adams, 1995; Monteiro and others, 2014).

Model Statistics, Data, and Plots

Model

$$TC = 0.129 * SC + 6.62$$

Variable Summary Statistics

	TC	SC
Minimum	31.2	203
1st Quartile	38.8	236
Median	47.3	310
Mean	49.8	334
3rd Quartile	56.1	387
Maximum	81.0	576

Box Plots

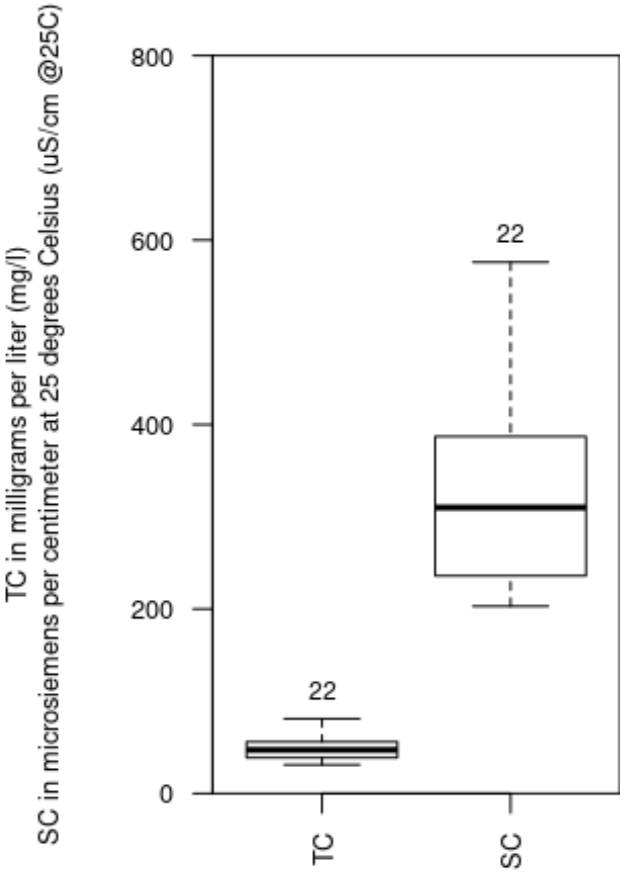


Figure 1.1. Total carbon (TC) in milligrams per liter and specific conductance (SC) in microsiemens per centimeter at 25 degrees Celsius during discrete sampling events.

Exploratory Plots

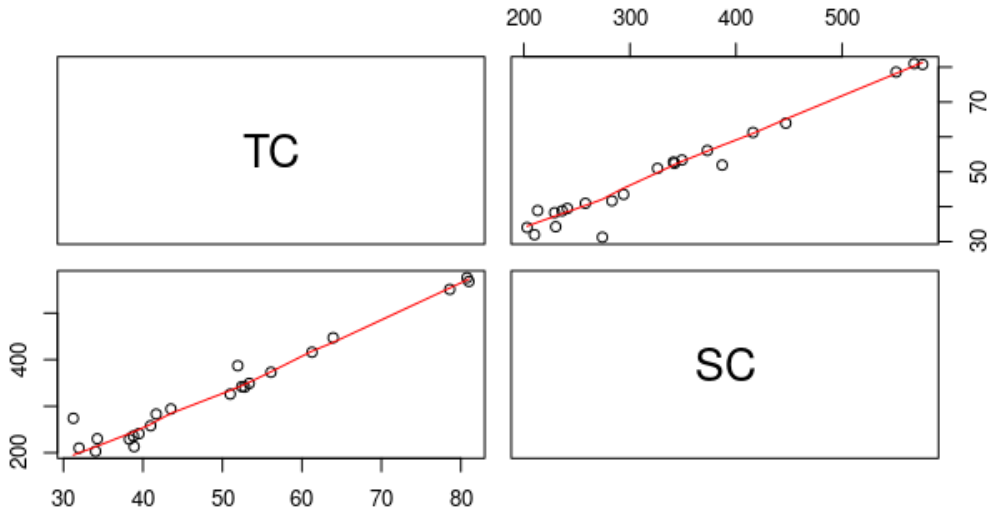


Figure 1.2. Comparison of total carbon (TC) and specific conductance (SC).

Basic Model Statistics

```
Number of Observations                              22
Standard error (RMSE)                             3.19
Average Model standard percentage error (MSPE)     6.4
Coefficient of determination (R²)                0.959
Adjusted Coefficient of Determination (Adj. R²) 0.957
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)        6.620        2.11000    3.14 5.12e-03
SC                 0.129        0.00597   21.70 2.28e-15
```

Correlation Matrix

```
          Intercept E.vars
Intercept     1.000 -0.946
E.vars       -0.946  1.000
```

Outlier Test Criteria

```
Leverage Cook's D   DFFITS
   0.273    0.193    0.603
```

Flagged Observations

	TC	Estimate Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
7/11/2016 9:45	31.2	42.1	-10.8	-3.5	-5.47	0.058	0.377	-1.36

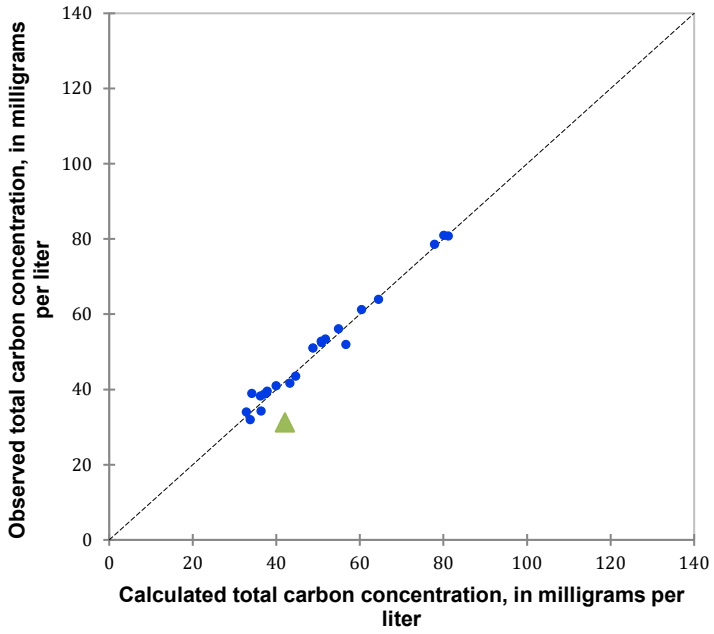


Figure 1.3. Relation between observed and calculated total carbon concentration in milligrams per liter; flagged observations are highlighted by a green triangle.

Statistical Plots

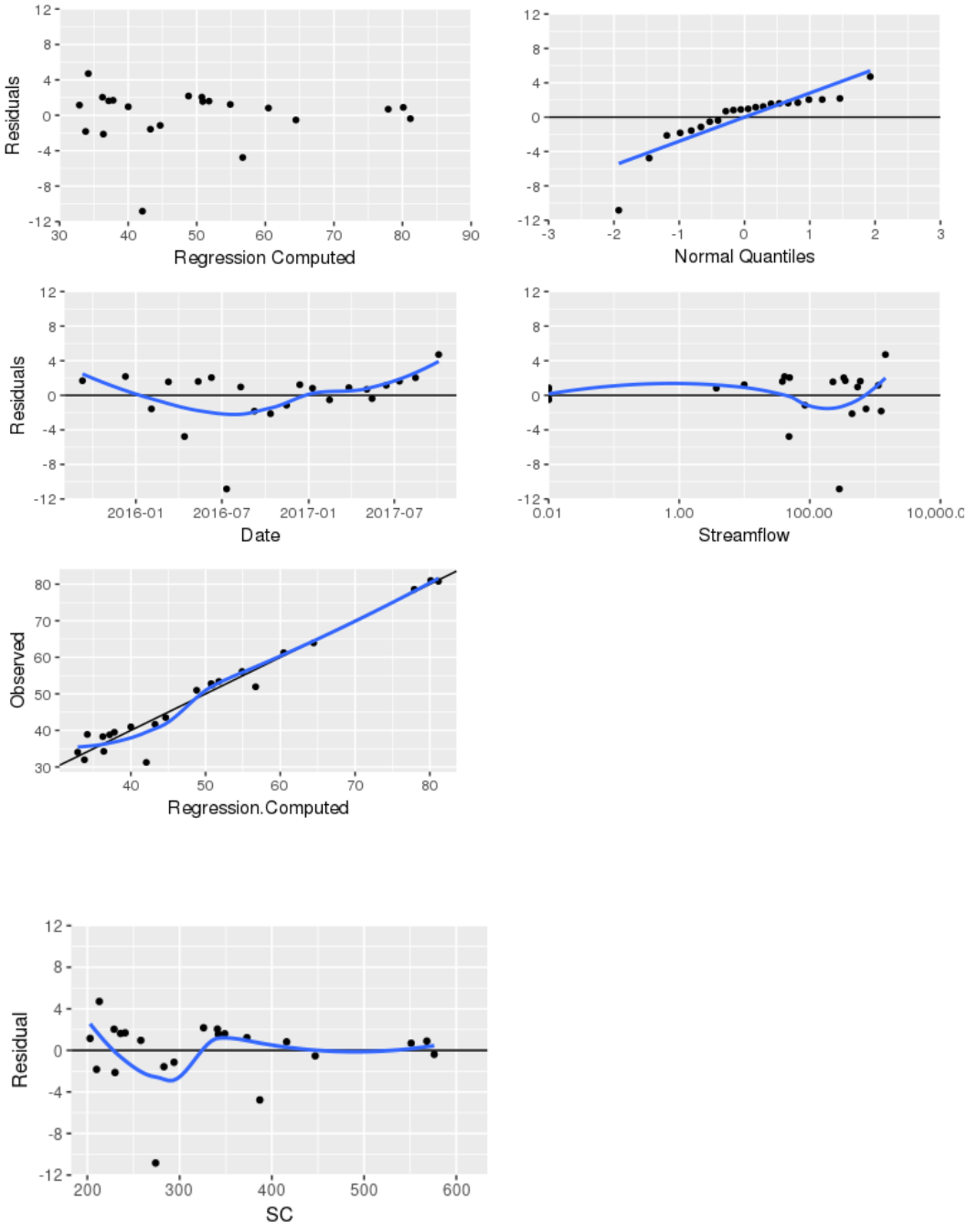


Figure 1.4. Residual and observed versus computed plots.

(SC, specific conductance)

A. B.

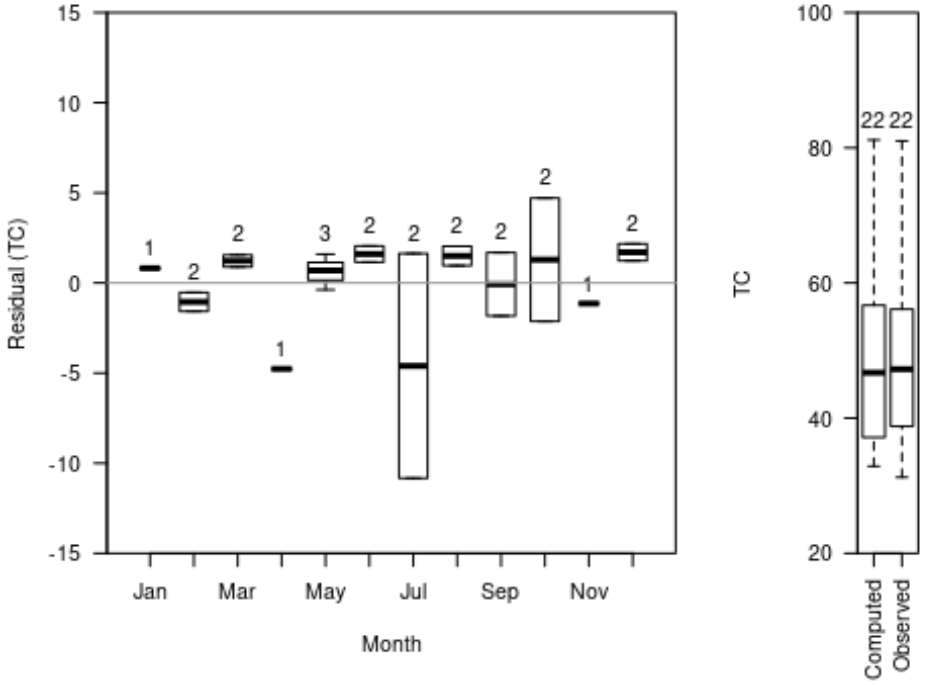


Figure 1.5. Seasonal variation in residuals of *A,* total carbon concentration (TC), and *B*, computed and observed total carbon (TC) concentration.

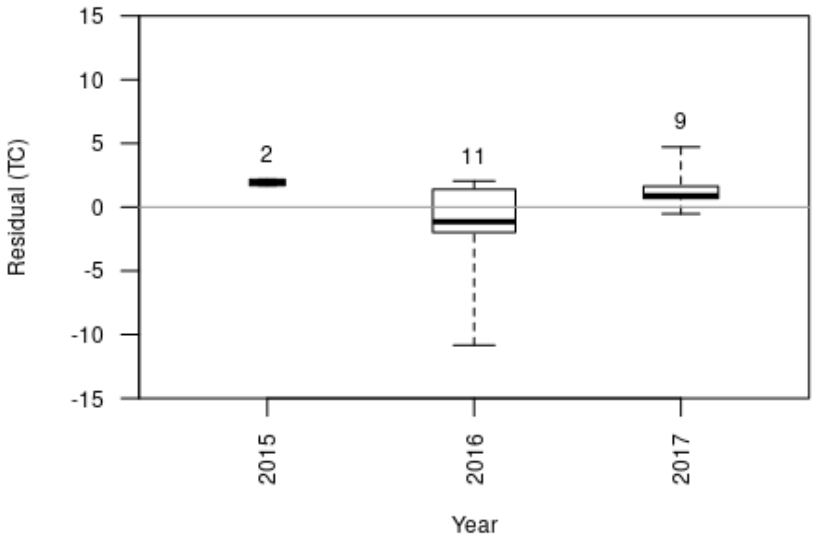
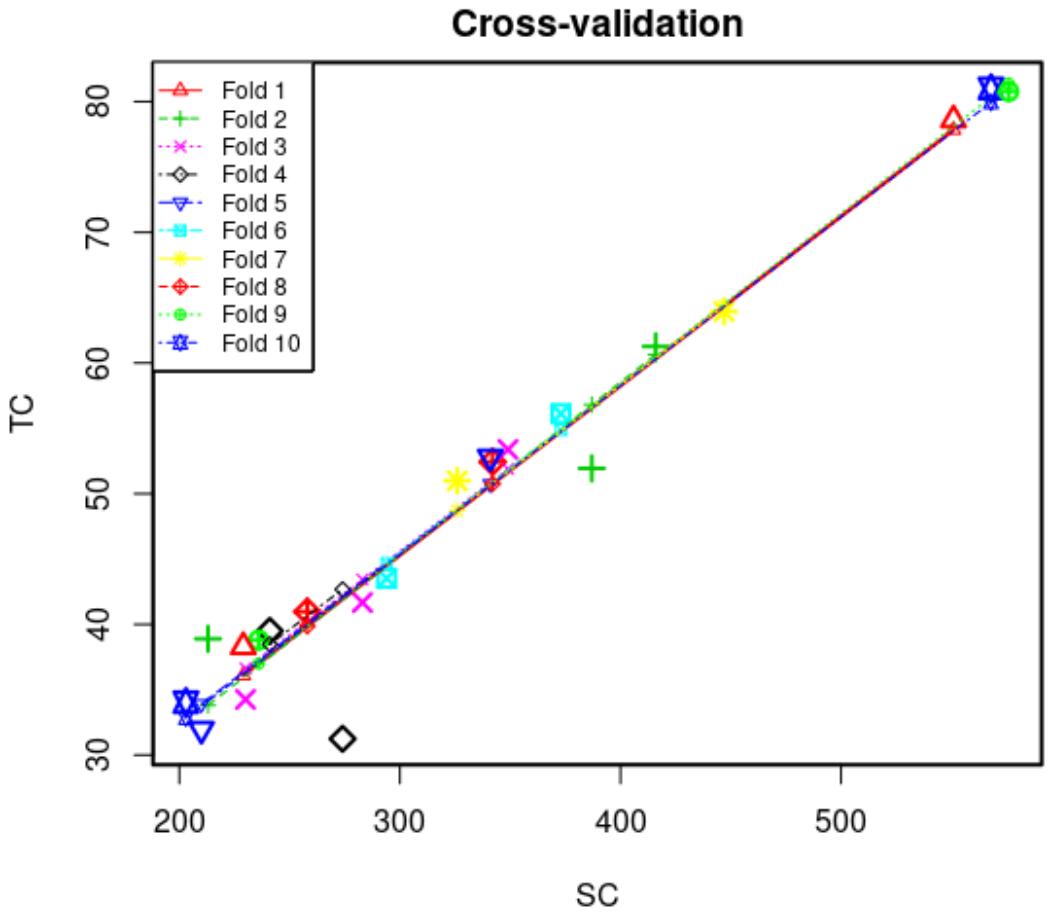


Figure 1.6. Annual variation in residuals.

(TC, total carbon concentration)

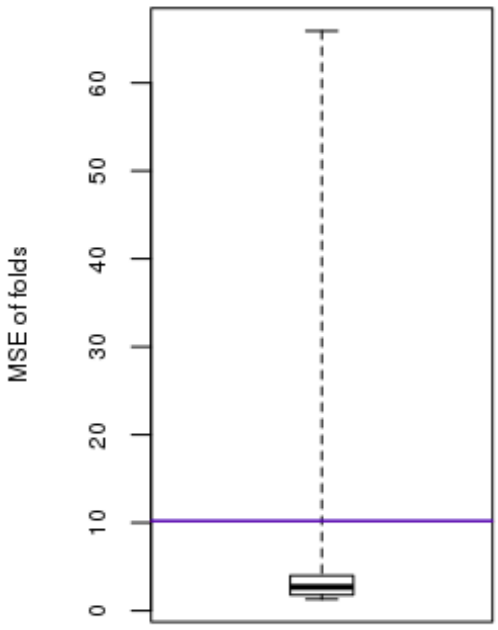
Cross Validation




```
             Minimum MSE of folds:   1.31
                Mean MSE of folds:  10.20
              Median MSE of folds:   2.70
             Maximum MSE of folds:  65.80
(Mean MSE of folds) / (Model MSE):   1.00
```

Figure 1-7. Cross validation plot.

(MSE, mean standard of error; TC, total carbon concentration; SC, specific conductance)




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Figure 1.8. Mean standard of error (MSE) of folds boxplot.

Model-Calibration Dataset

(EST, Eastern Standard Time; mg/L, milligrams per liter)

Date and time (EST)	Dissolved organic carbon in mg/L, (00681)	Dissolved inorganic carbon in mg/L, (00691)	Total particulate carbon in mg/L, (00694)	Total carbon in mg/L
09/10/2015 09:30	9.81	29.40	0.29	39.5
12/10/2015 11:25	12.1	38.23	0.67	51.0
02/03/2016 12:00	7.76	33.55	0.36	41.7
03/10/2016 12:00	13.4	38.54	0.54	52.4

04/13/2016 10:00	16.1	34.99	0.86	51.9
05/12/2016 09:45	14.1	38.57	0.70	53.4
06/09/2016 11:15	17.66	34.28	0.75	52.7
07/11/2016 09:45	9.91	20.91	0.42	31.2
08/10/2016 11:30	10.3	30.18	0.45	41.0
09/08/2016 11:15	8.63	22.93	0.41	32.0
10/12/2016 12:45	8.63	25.25	0.38	34.3
11/15/2016 12:15	11.89	30.91	0.72	43.5
12/13/2016 12:30	12.93	42.48	0.70	56.1
01/09/2017 13:00	12.92	47.91	0.44	61.3
02/14/2017 12:00	12.27	49.67	2.00	63.9
03/27/2017 10:30	9.24	69.66	2.11	81.0
05/04/2017 11:15	8.44	61.61	8.55	78.6
05/15/2017 13:45	8.07	64.43	8.32	80.8
06/14/2017 10:45	9.53	23.99	0.53	34.0
07/12/2017 12:30	12.44	25.94	0.42	38.8
08/15/2017 10:15	12.48	25.32	0.50	38.3
10/03/2017 10:30	11.13	27.20	0.57	38.9

	Date	TC	SC	Computed TC	Residual	Normal Quantiles	Censored Values
1	2015-09-10	39.5	241	37.8	1.69	0.816	--
2	2015-12-10	51.0	326	48.8	2.18	1.46	--
3	2016-02-03	41.7	283	43.2	-1.57	-0.816	--
4	2016-03-10	52.4	342	50.9	1.56	0.406	--
5	2016-04-13	51.9	387	56.7	-4.76	-1.46	--
6	2016-05-12	53.4	349	51.8	1.6	0.532	--
7	2016-06-09	52.7	341	50.7	1.96	0.986	--
8	2016-07-11	31.2	274	42.1	-10.8	-1.93	--
9	2016-08-10	41.0	258	40	0.967	0.0565	--
10	2016-09-08	32.0	210	33.8	-1.82	-0.986	--
11	2016-10-12	34.3	230	36.4	-2.12	-1.19	--
12	2016-11-15	43.5	294	44.7	-1.14	-0.667	--
13	2016-12-13	56.1	373	54.9	1.23	0.286	--
14	2017-01-09	61.3	416	60.4	0.821	-0.17	--
15	2017-02-14	63.9	447	64.5	-0.528	-0.532	--
16	2017-03-27	81.0	568	80.1	0.886	-0.0565	--
17	2017-05-04	78.6	551	77.9	0.686	-0.286	--
18	2017-05-15	80.8	576	81.2	-0.355	-0.406	--
19	2017-06-14	34.0	203	32.9	1.16	0.17	--
20	2017-07-12	38.8	236	37.2	1.64	0.667	--
21	2017-08-15	38.3	229	36.3	2.04	1.19	--
22	2017-10-03	38.9	213	34.2	4.72	1.93	--

Model Limitations

Error in the model can be attributed to several factors, including those related to SC data. There is error associated in the calibration of the standards; corrections were only applied when the instrument was 3 percent for SC. Another limitation to this model is in the assumption that the data at the fixed location are representative of the entire study area. The sensor profiles showed that the mean cross section and the data at 02288900 were not always equivelant, however, they were always within 8 percent. An additional source of model error is from the discrete data analysis.

Definitions

```
TC: Inorganic carbon in mg/l (00691), Organic carbon in mg/l (00681),Total pa
rticulate carbon, mg/l (00694)
SC: Specific conductance in uS/cm @25C (00095)
```

References Cited

Curtis, P.J., and Adams, H.E., 1995, Dissolved organic matter quantity and quality from freshwater and saltwater lakes in east-central Alberta: Biogeochemistry, v. 30, p. 59–76, accessed April 24, 2019, at https://link.springer.com/article/10.1007/BF02181040.

Domanski, M.M., Straub, T.D., and Landers, M.N., 2015, Surrogate analysis and index developer (SAID) tool (version 1.0, September 2015): U.S. Geological Survey Open-File Report 2015–1177, 38 p., accessed October 1, 2017, at https://pubs.er.usgs.gov/publication/ofr20151177.

Monteiro, M.T., Oliveira, S.M., Luizao, F.J., Candido, L.A., Ishida, F.Y., and Tomasella, J., 2014, Dissolved organic carbon concentration and its relationship to electrical conductivity in the waters of a stream in forested Amazonian blackwater catchment: Plant Ecology and Diversity, v. 4, p. 205–213, accessed December 1, 2018, at https://doi.org/10.1080/17550874.2013.820223.

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, 53 p., accessed October 1, 2017, at https://doi.org/10.3133/tm3C4.

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed May 7, 2019, at https://doi.org/10.5066/F7P55KJN.

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., accessed October 1, 2017, at http://pubs.usgs.gov/tm/2006/tm1D3/.

App Version 1.0