Appendix 2. Model Archive for Total Carbon Concentration at U.S. Geological Survey Station 022909471: Loop Road Culverts Monroe Station to Florida Trail, 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 (June 2015–September 2017) and loads (September 2015–September 2017).

The purpose of this model is to compute TC concentrations at Loop Road between Monroe Station and the Florida Trail (022909471) during the study period. Station 02290947 represents outflow at 40 culvert locations along Loop Road. Data will be used as an input to a larger carbon budget study for Sweetwater Strand.

Site and Model Information

Site number: 022909471

Site name: Loop Road culverts Monroe Station to Florida Trail, Florida

Location: lat 25°47'18.86" N., long 81°05'59.72" W., referenced to North American Datum of 1983, Monroe County, Florida, hydrologic unit 03090204

Equipment: A YSI EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, turbidity, and a fluorescence of chromophoric dissolved organic matter (fDOM) sensor. The monitor was housed in a 6-inch polyvinyl chloride (PVC) pipe attached to the upstream side wall of Sweetwater Strand bridge on Loop Road (USGS 02290947). Readings from the YSI EXO2 were recorded every 15 minutes and recorded internally to the YSI. The model applies only to this site (022909471) and specified time period (June 16, 2015–October 5, 2017).

Model number: 1.0

Date model was created: April 24, 2018

Model calibration data period: June 16, 2015–October 5, 2017

Model application date: June 16, 2015–October 5, 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 as referenced in Rasmussen and others (2009).

Specific conductance (SC) was published at 02290947. The regression model is based on 21 concurrent measurements of DOC, DIC, and TPC samples and concomitant SC measurements collected from June 16, 2015, to October 5, 2017. Samples were collected throughout the range of observed hydrologic conditions. Summary statistics and the complete model-calibration data are provided in the dataset.

Discrete Sample Data

Point samples were collected at the location and depth of the water-quality sensors at 02290947. 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 02290947 on March 22, 2016, July 17, 2016, and October 5, 2017. Blanks were collected at 02290947 on November 3, 2016, February 22, 2017, and July 17, 2017. All replicates were sequential unless otherwise noted.

Absolute differences between replicate pairs of DOC ranged from −0.70 to −0.09 milligrams per liter (mg/L). Relative percent differences ranged from −5 to −1 percent for DOC. The DOC blank was <0.23 mg/L (the detection limit) on November 3, 2016, 0.31 mg/L on February 22, 2017, and 0.49 mg/L on July 17, 2017. DIC replicate pairs varied by −3 to 1.1 mg/L. Relative percent differences from the replicate samples ranged from −8 to 2 percent. The DIC blanks ranged between 0.2 and 0.3 mg/L.

TPC replicate pairs varied by −0.002 to 0.050 mg/L. Relative percent differences from the sequential replicate samples ranged from −1 to 23 percent. All TPC blanks were below the detection limit.

Surrogate Data

SC at 02290947 ranged from 69 microsiemens at 25 degrees Celsius (µS/cm at 25°C) on June 6, 2017, to 625 µS/cm at 25°C on April 20, 2015. 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. A variety of models that predict TC were evaluated. 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 indepth 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 concentration at site number 022909471.

Total carbon concentration-based model:

$$TC = 0.117(SC) + 12$$

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

Total.Carbon = 0.117 * *Specific.Conductance* + 12

Variable Summary Statistics

	TC	SC
Minimum	35.0	256
1st Quartile	48.6	310
Median	53.0	374
Mean	56.2	379
3rd Quartile	66.7	455
Maximum	70.7	516

Box Plots

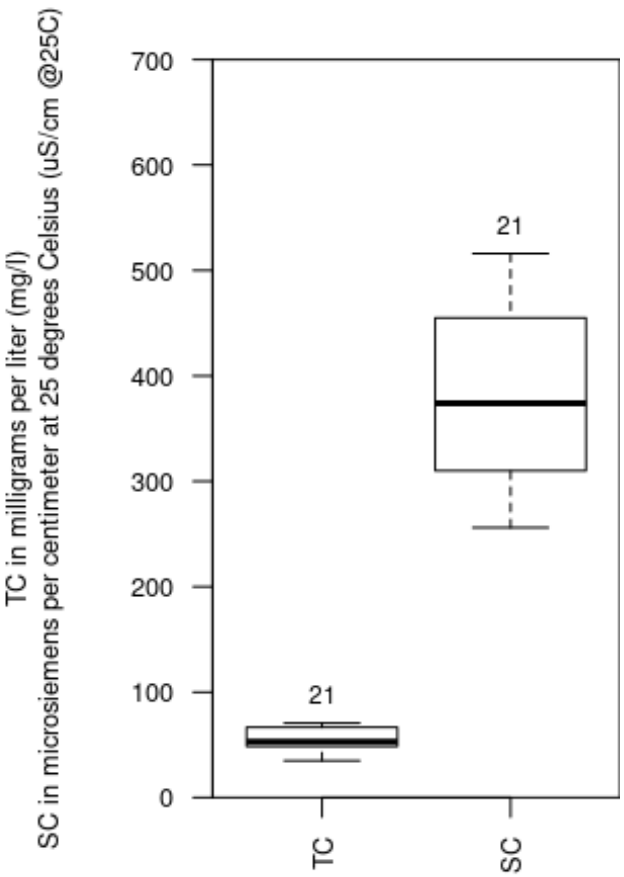


Figure 2.1. Total carbon (TC) in milligrams per liter and specific conductance (SC) in microsiemens per centimeter at 25 degrees Celsius during discrete sampling events.

Exploratory Plots

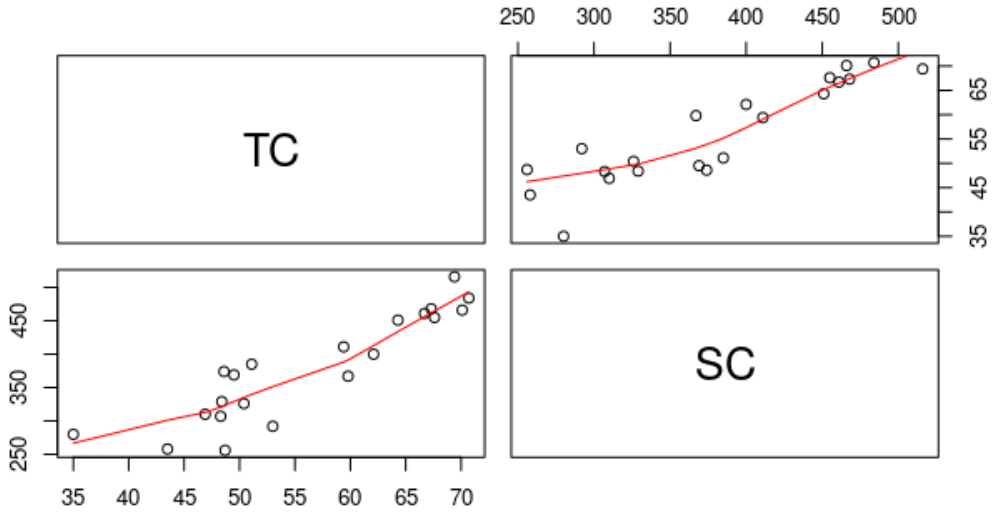


Figure 2.2. Comparison of total carbon (TC) and specific conductance (SC).

Basic Model Statistics

Number of Observations	21
Standard error (RMSE)	4.48
Average Model standard percentage error (MSPE)	7.97
Coefficient of determination (R^2)	0.819
Adjusted Coefficient of Determination (Adj. R^2)	0.81

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	12.000	4.8700	2.46	2.38e-02
SC	0.117	0.0126	9.29	1.71e-08

Correlation Matrix

	Intercept	E.vars
Intercept	1.00	-0.98
E.vars	-0.98	1.00

Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.286	0.193	0.617

Flagged Observations

	TC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
9/20/2016 10:00	35.0	44.6	-9.64	-2.30	-2.63	0.125	0.378	-0.996
10/5/2017 9:30	48.7	41.8	6.86	1.68	1.77	0.167	0.283	0.793

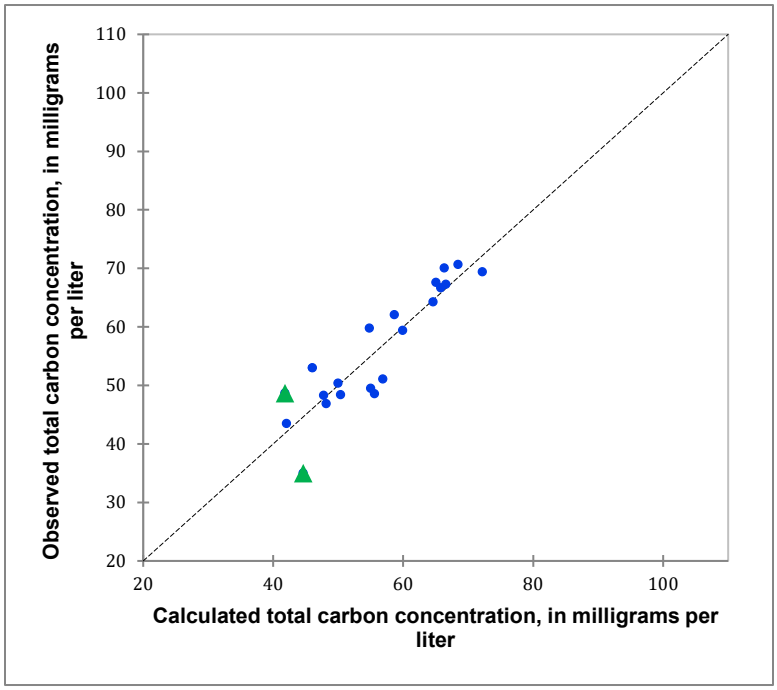


Figure 2.3. Relation between observed and calculated total carbon concentration in milligrams per liter; flagged observations are highlighted by a green triangle.

Statistical Plots

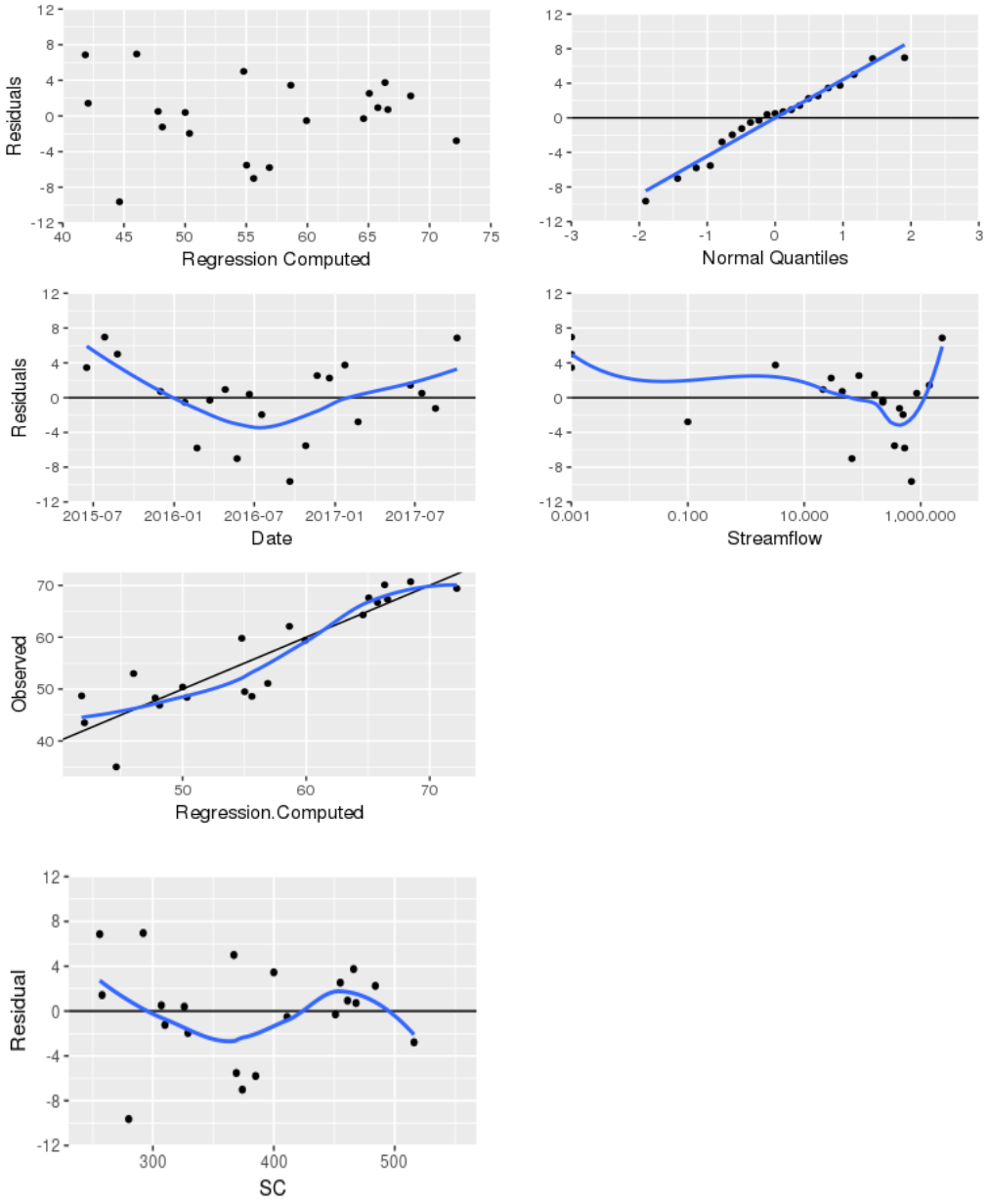


Figure 2.4. Residual and observed versus computed plots.

(SC, specific conductance)

A. B.

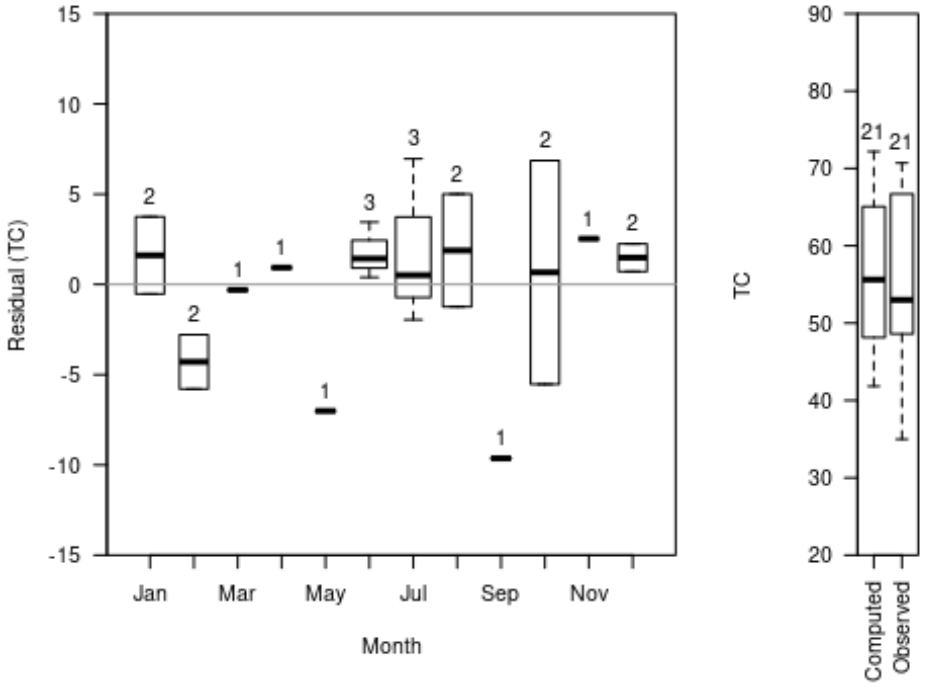


Figure 2.5. Seasonal variation in residuals of *A*, total carbon concentration (TC), and *B,* computed and observed total carbon (TC) concentration.

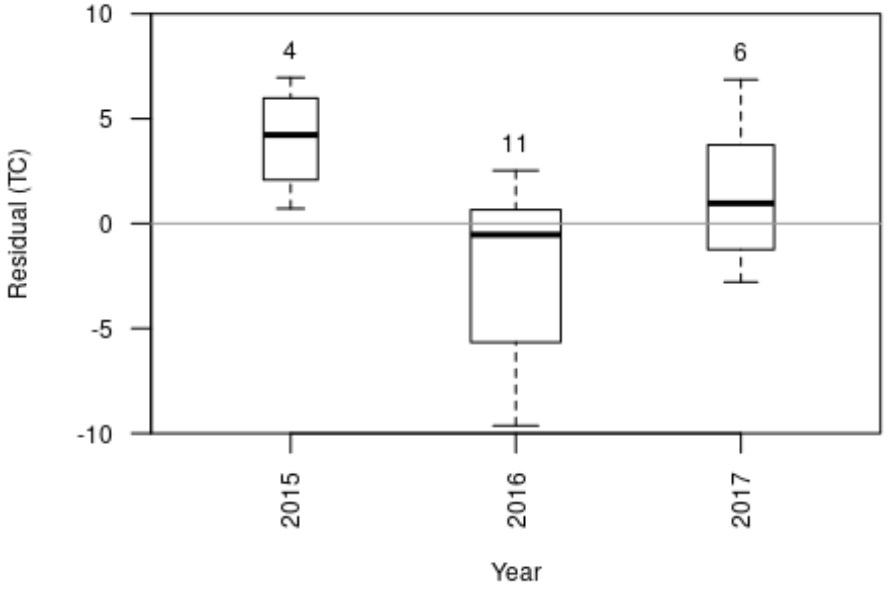
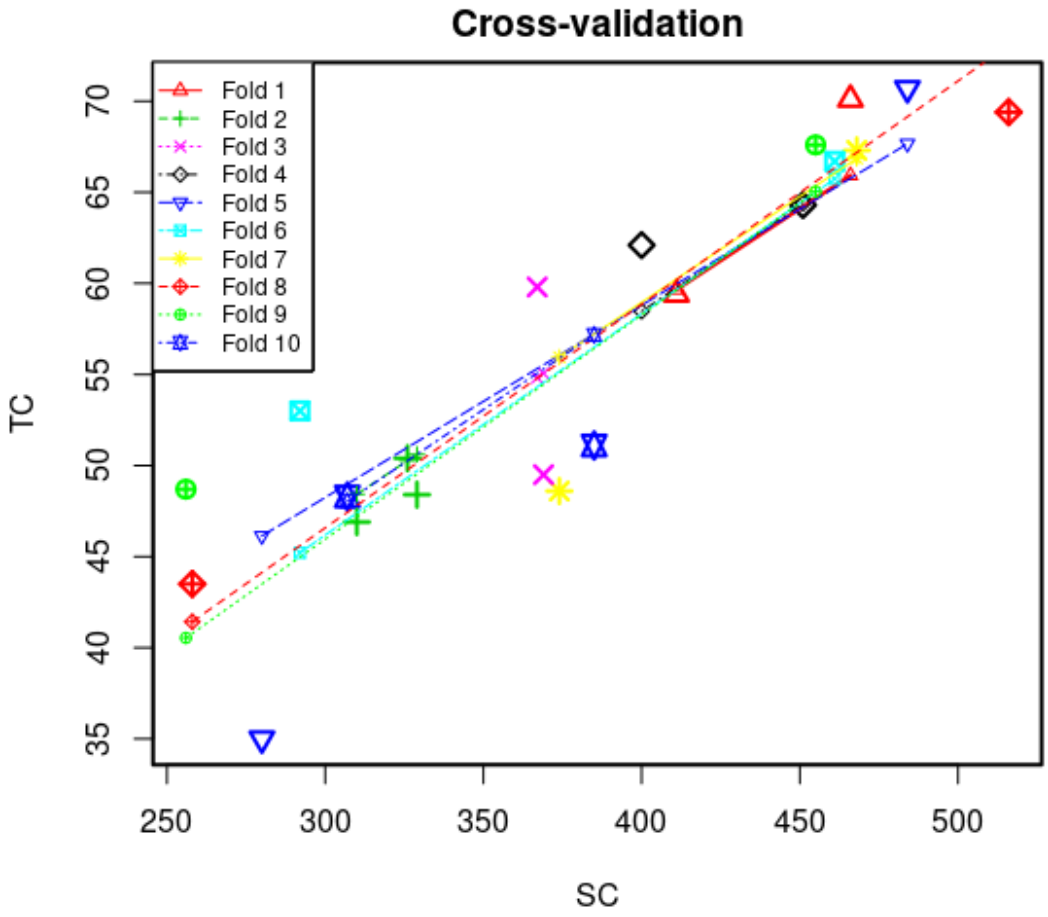


Figure 2.6. Annual variation in residuals.

(TC, total carbon concentration)

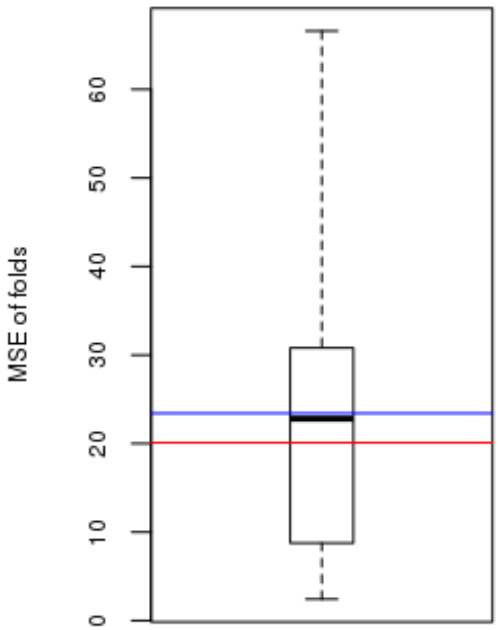
Cross Validation




```
             Minimum MSE of folds:   2.42
                Mean MSE of folds:  23.40
              Median MSE of folds:  22.80
             Maximum MSE of folds:  66.60
(Mean MSE of folds) / (Model MSE):   1.17
```

Figure 2.7. 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 2-8. 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
6/16/2015 10:00	10.7	51.0	0.432	62.1
7/27/2015 12:00	15.4	37.3	0.279	53.0
8/25/2015 12:02	15.4	44.3	0.173	59.8
12/1/2015 10:52	11.4	55.7	0.227	67.3

1/25/2016 9:45	8.8	50.5	0.095	59.4
2/22/2016 12:45	8.6	42.3	0.177	51.1
3/22/2016 8:30	10.5	53.6	0.154	64.3
4/26/2016 9:45	11.4	54.9	0.319	66.7
5/23/2016 11:45	9.5	38.9	0.142	48.6
6/20/2016 13:00	10.2	40.0	0.168	50.4
7/18/2016 11:00	9.3	38.8	0.293	48.4
9/20/2016 10:00	9.9	24.8	0.304	35.0
10/26/2016 10:45	9.0	40.1	0.390	49.5
11/21/2016 11:30	9.5	57.8	0.196	67.6
12/19/2016 12:30	10.3	59.9	0.484	70.7
1/23/2017 11:30	9.7	59.4	0.998	70.1
2/22/2017 10:15	9.2	58.4	1.827	69.4
6/21/2017 10:00	15.4	27.9	0.222	43.5
7/17/2017 8:00	13.8	34.3	0.222	48.3
8/17/2017 8:15	11.9	34.4	0.601	46.9
10/5/2017 9:30	10.8	37.6	0.343	48.7

	Date	TC	SC	Computed TC	Residual	Normal Quantiles	Censored Values
1	2015-06-16	62.1	400	58.6	3.45	0.782	--
2	2015-07-27	53.0	292	46	6.96	1.91	--
3	2015-08-25	59.8	367	54.8	5.01	1.16	--
4	2015-12-01	67.3	468	66.6	0.716	0.119	--
5	2016-01-25	59.4	411	59.9	-0.531	-0.362	--
6	2016-02-22	51.1	385	56.9	-5.8	-1.16	--
7	2016-03-22	64.3	451	64.6	-0.3	-0.239	--
8	2016-04-26	66.7	461	65.8	0.933	0.239	--
9	2016-05-23	48.6	374	55.6	-7.01	-1.44	--
10	2016-06-20	50.4	326	50	0.392	-0.119	--
11	2016-07-18	48.4	329	50.4	-1.96	-0.631	--
12	2016-09-20	35.0	280	44.6	-9.64	-1.91	--
13	2016-10-26	49.5	369	55	-5.53	-0.955	--
14	2016-11-21	67.6	455	65.1	2.53	0.631	--
15	2016-12-19	70.7	484	68.5	2.25	0.492	--
16	2017-01-23	70.1	466	66.4	3.75	0.955	--
17	2017-02-22	69.4	516	72.2	-2.79	-0.782	--
18	2017-06-21	43.5	258	42.1	1.43	0.362	--
19	2017-07-17	48.3	307	47.8	0.51	0	--
20	2017-08-17	46.9	310	48.1	-1.24	-0.492	--
21	2017-10-05	48.7	256	41.8	6.86	1.44	--

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 more than 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 within 5 percent on average. 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
particulate 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