

Appendix 2. Model Archival Summary for Chlorophyll Concentration at Station 06892350; Kansas River at De Soto, Kansas, during July 26, 2012, through June 10, 2014

This model archival summary summarizes the chlorophyll (Chl; uncorrected for degradation products) concentration model developed to estimate Chl concentrations using YSI Model 6025 sensor-measured 15-minute chlorophyll fluorescence (fChl) data during July 26, 2012, through June 10, 2014.

Site and Model Information

Site number: 06892350

Site name: Kansas River at De Soto, Kansas

Location: Lat 38°59'00", long 94°57'52" referenced to North American Datum of 1927, in NE 1/4 SE 1/4 SE 1/4 sec.28, T.12 S., R.22 E., Leavenworth County, KS, Hydrologic Unit 10270104.

Equipment: A YSI 6600 water-quality monitor equipped with model 6025 fChl sensor was installed during July 2012 through June 2014. The monitor was housed in a 4-inch diameter galvanized steel pipe. Readings from the water-quality monitor were recorded every 15 minutes and transmitted by way of satellite, hourly. Sensor maximum was not exceeded during this time (operable range: 0 to 400 micrograms per liter).

Date model was created: August 3, 2017

Model calibration data period and application date: July 26, 2012–June 10, 2014

Model-Calibration Dataset

All data were collected using U.S. Geological Survey protocols (U.S. Geological Survey, variously dated; <https://water.usgs.gov/owq/FieldManual/>; Wagner and others, 2006) and are stored in the National Water Information System (U.S. Geological Survey, 2018) database and available to the public. Calibration of the fChl sensor was periodically checked (quarterly), and recalibrated if needed, using the YSI-recommended Rhodamine WT dye solution and YSI-provided calibration tables (Yellow Springs Instruments, 2012). If recalibration was required, calibration check data were used to apply prorated calibration drift corrections to the sensor-measured fChl data during the sensor-deployment time period. Ordinary least squares analysis was used to develop regression models between sensor-measured fChl and laboratory-measured Chl (uncorrected for degradation products) at the Kansas River at De Soto, Kansas (Kans.), site using the open-source software package “R.” This model is specific to the Kansas River at De Soto site during this study period and cannot be applied to data collected from other sites on the Kansas River or data collected from other water bodies.

The final selected regression model is based on 30 concurrent measurements of sensor-measured fChl and laboratory-measured Chl collected during July 26, 2012, through June 10, 2014. Samples were collected throughout the range of continuously observed hydrologic conditions. No samples were below laboratory detection limits. Summary statistics and the complete model-calibration dataset are provided below. Studentized residuals from the final model were inspected for values greater than three or less than negative three. Values outside of that range were considered potential outliers and were investigated. None of the samples in this dataset were deemed outliers or removed from the model calibration dataset.

Chlorophyll Sampling Details

Cross-section samples were typically collected either from the downstream side of the bridge or instream within 100 feet of the bridge. The equal-width-increment method was used, and samples typically were composited for analysis. During August 2012 through June 2014, cross-section samples were collected every 2 weeks during March through October, once a month during November through February, and during selected reservoir release and runoff events. A FISP US DH-95 or D-95 depth integrating sampler was used from the downstream side of the bridge. Additional detail on sample collection is available in Graham and others (2018). Chl (uncorrected for degradation products) was analyzed fluorometrically using U.S. Environmental Protection Agency Method 445.0 (Arar and Collins, 1997) modified for heated ethanol extraction (Sartory and Grobbelaar, 1984 and a fluorometer equipped with a flow-through cell (Knowlton, 1984) at the U.S. Geological Survey Kansas Water Science Center in Lawrence, Kans.

Model Development

Ordinary least squares regression analysis was done using R version 3.4.2 (R Core Team, 2017) with sensor-measured *fChl* as the explanatory variable for laboratory-measured *Chl* concentrations. The distribution of residuals was examined for normality, and the plots of residuals (the difference between the measured and computed values) were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of computed values).

Values for all the aforementioned statistics were computed and are included below along with all relevant sample data and more in-depth statistical information.

Model Summary

The following is a summary of final regression analysis for sensor-measured *fChl* and laboratory-measured *Chl* at the Kansas River at De Soto, Kans., streamgage (U.S. Geological Survey station number 06892350) during July 26, 2012, through June 10, 2014.

Chl concentration model:

$$Chl_E = 2.29 \times fChl + 14$$

where

Chl_E = regression estimated chlorophyll (uncorrected for degradation products) concentration in micrograms per liter; and

fChl = sensor-measured chlorophyll fluorescence in micrograms per liter.

R Output for the relation between YSI model 6025 sensor-measured chlorophyll fluorescence and laboratory-measured chlorophyll concentration; 06892350; Kansas River at De Soto, Kansas

Model Statistics, Data, and Plots

Definitions for terms used in this output can be found at the end of this document.

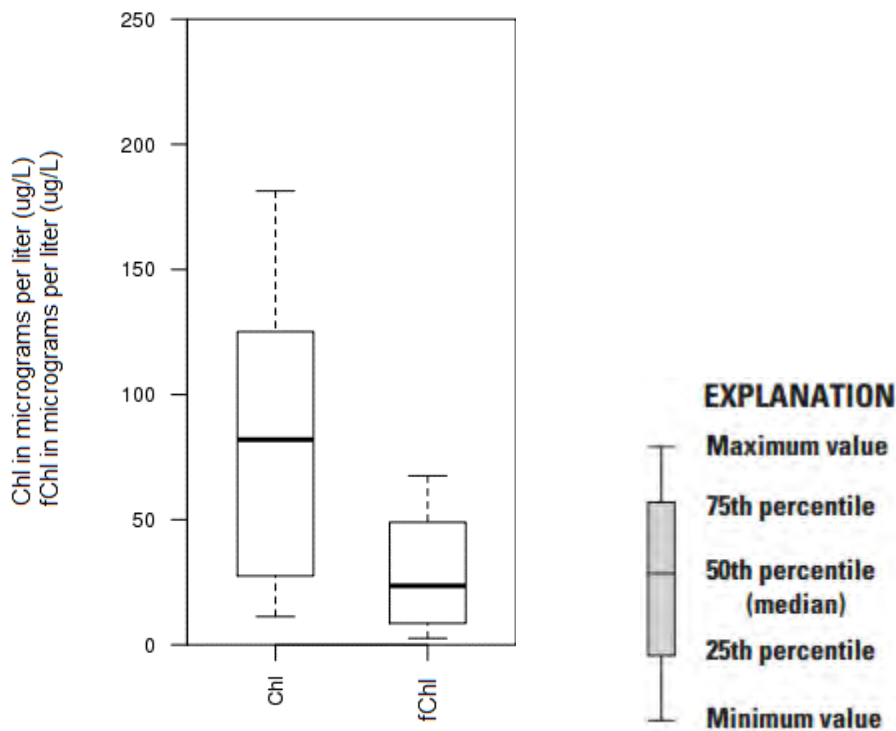
Model

$$\text{Chl}_E = + 2.29 * \text{fChl} + 14$$

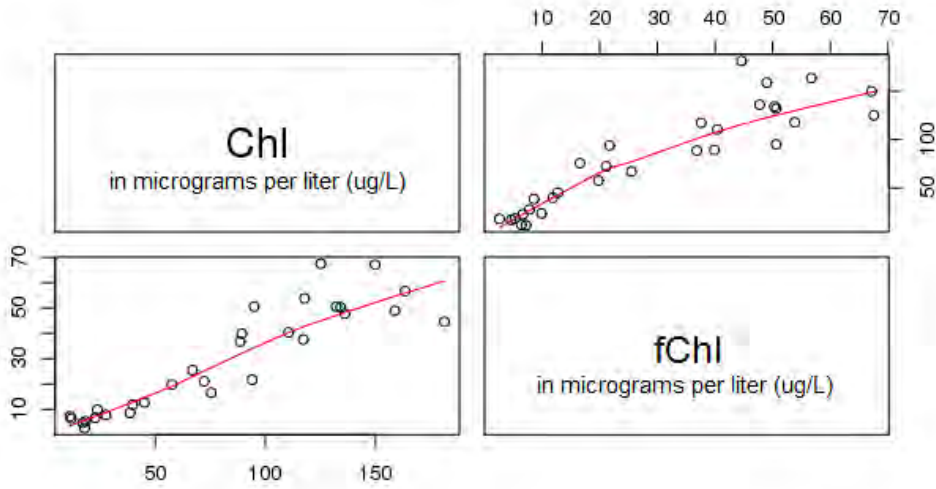
Variable Summary Statistics

	Chl	fChl
Minimum	11.3	2.63
1st Quartile	27.6	8.60
Median	82.0	23.60
Mean	81.3	29.40
3rd Quartile	125.0	49.00
Maximum	181.0	67.60

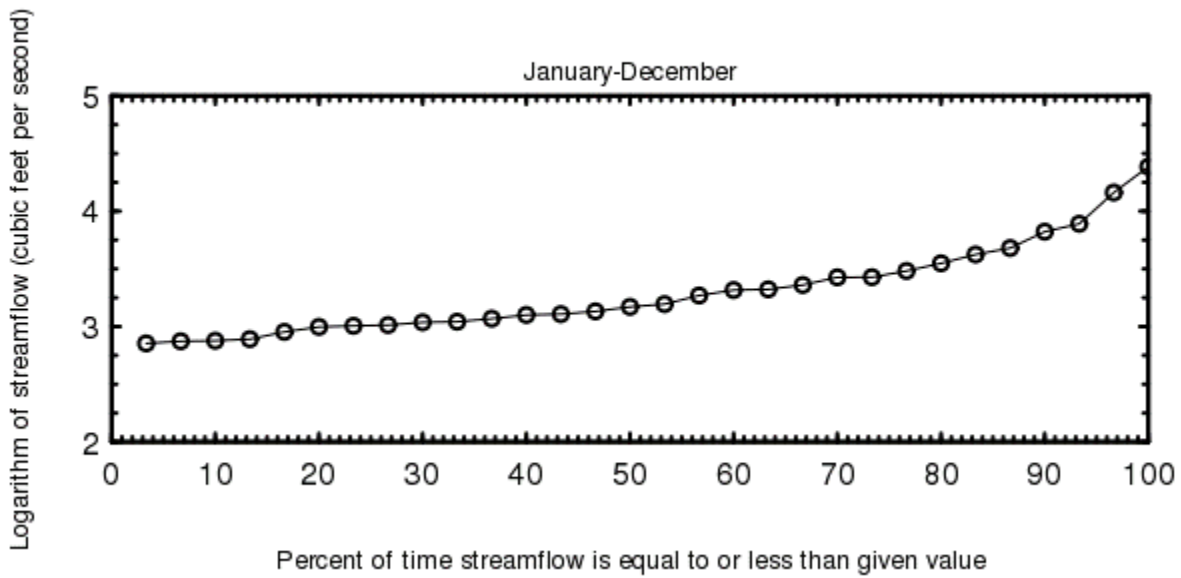
Box Plots

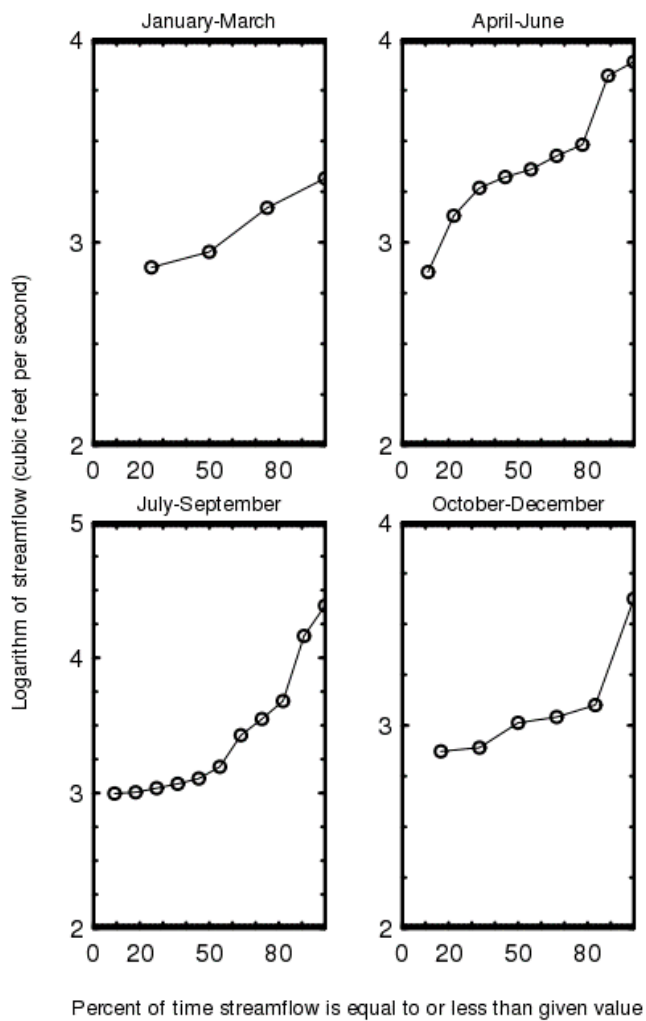


Exploratory Plots



Red line shows the locally weighted scatterplot smoothing (LOWESS).





Basic Model Statistics

For a detailed explanation of the terms used below, refer to Helsel and Hirsch (2002).

Number of Observations	30
Standard error (RMSE)	21.6
Upper Model standard percentage error (MSPE)	26.5
Lower Model standard percentage error (MSPE)	26.5
Coefficient of determination (R^2)	0.835
Adjusted Coefficient of Determination (Adj. R^2)	0.829

Explanatory Variables

	Coefficients	Standard Error	t value	Probability(> t)
(Intercept)	14.00	6.890	2.04	5.09e-02
fCh1	2.29	0.192	11.90	1.78e-12

Correlation Matrix

	Intercept	fCh1
Intercept	1.00	-0.82
fCh1	-0.82	1.00

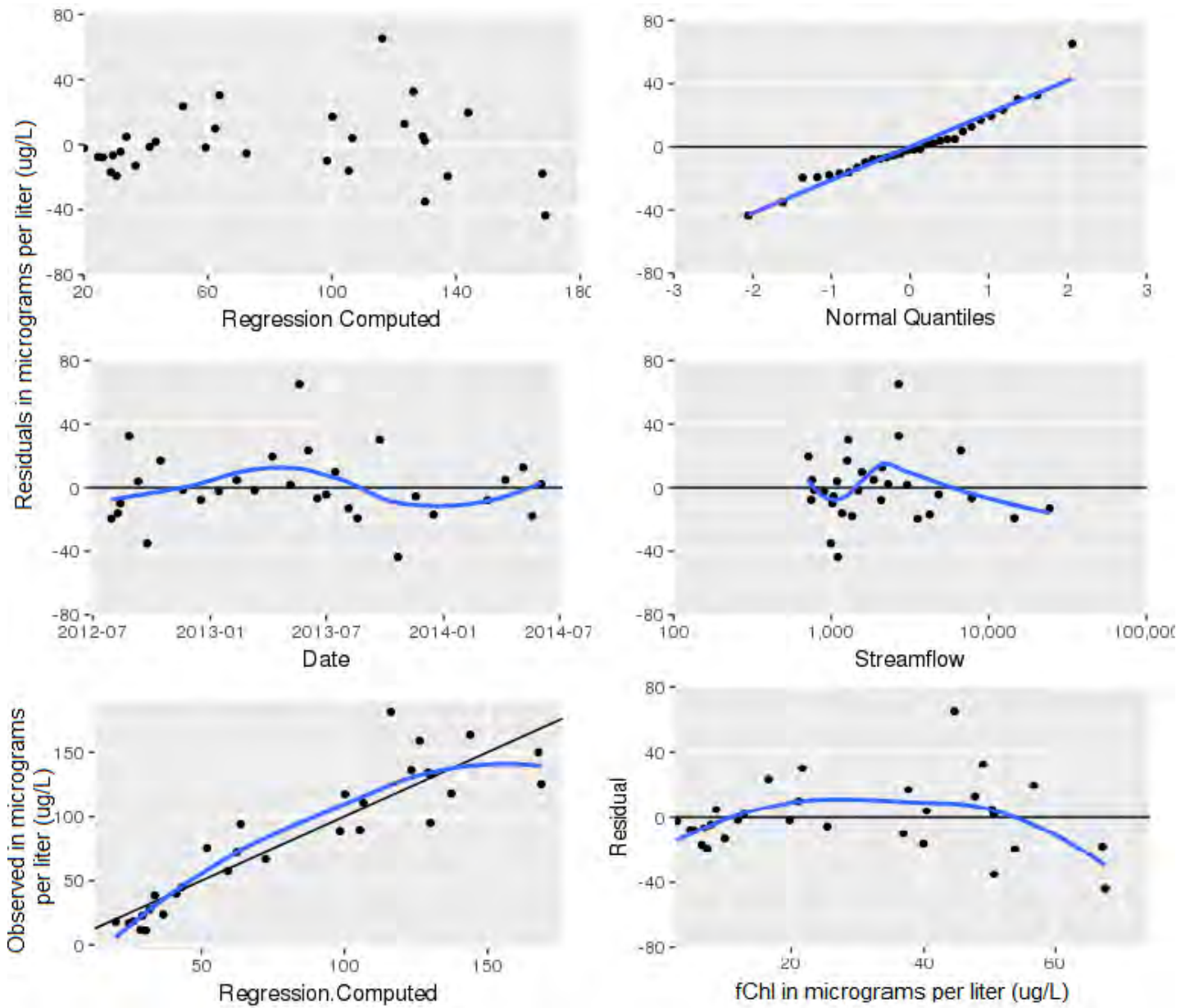
Outlier Test Criteria

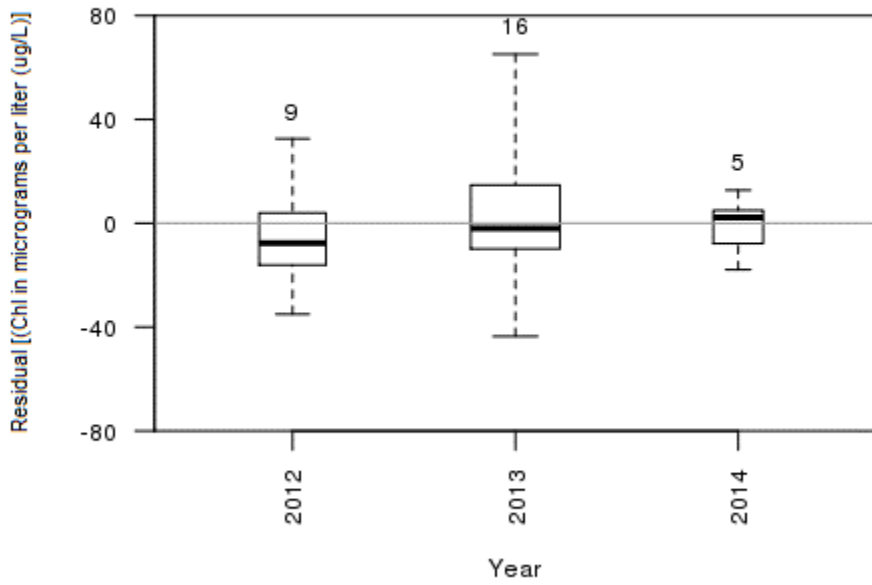
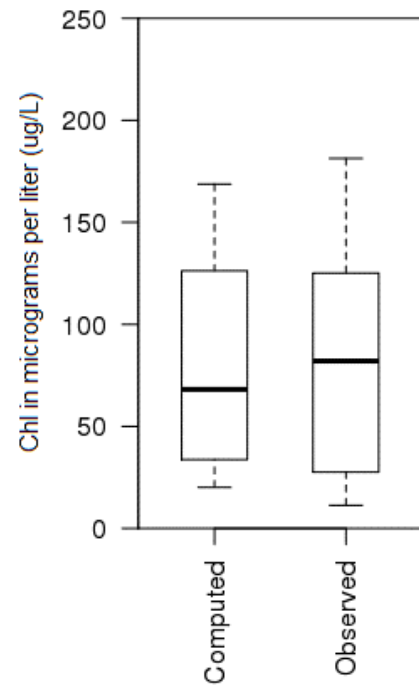
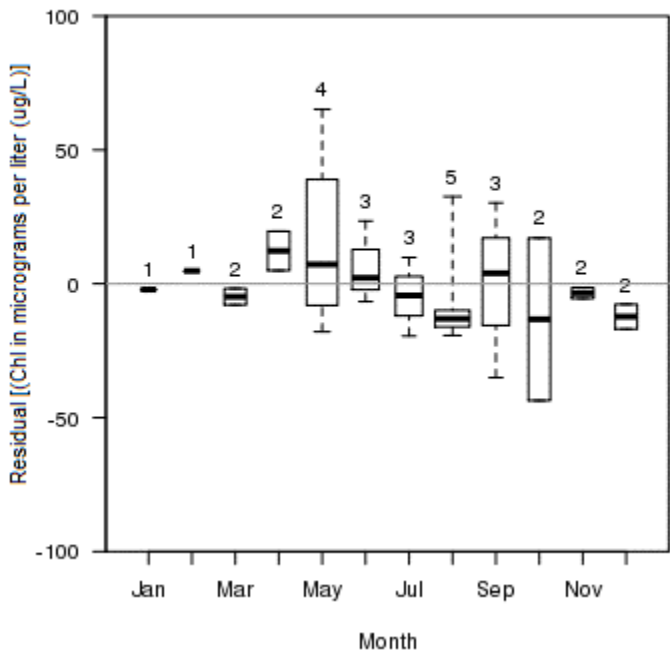
Leverage	Cook's D	DFFITS
0.100	0.106	0.365

Flagged Observations

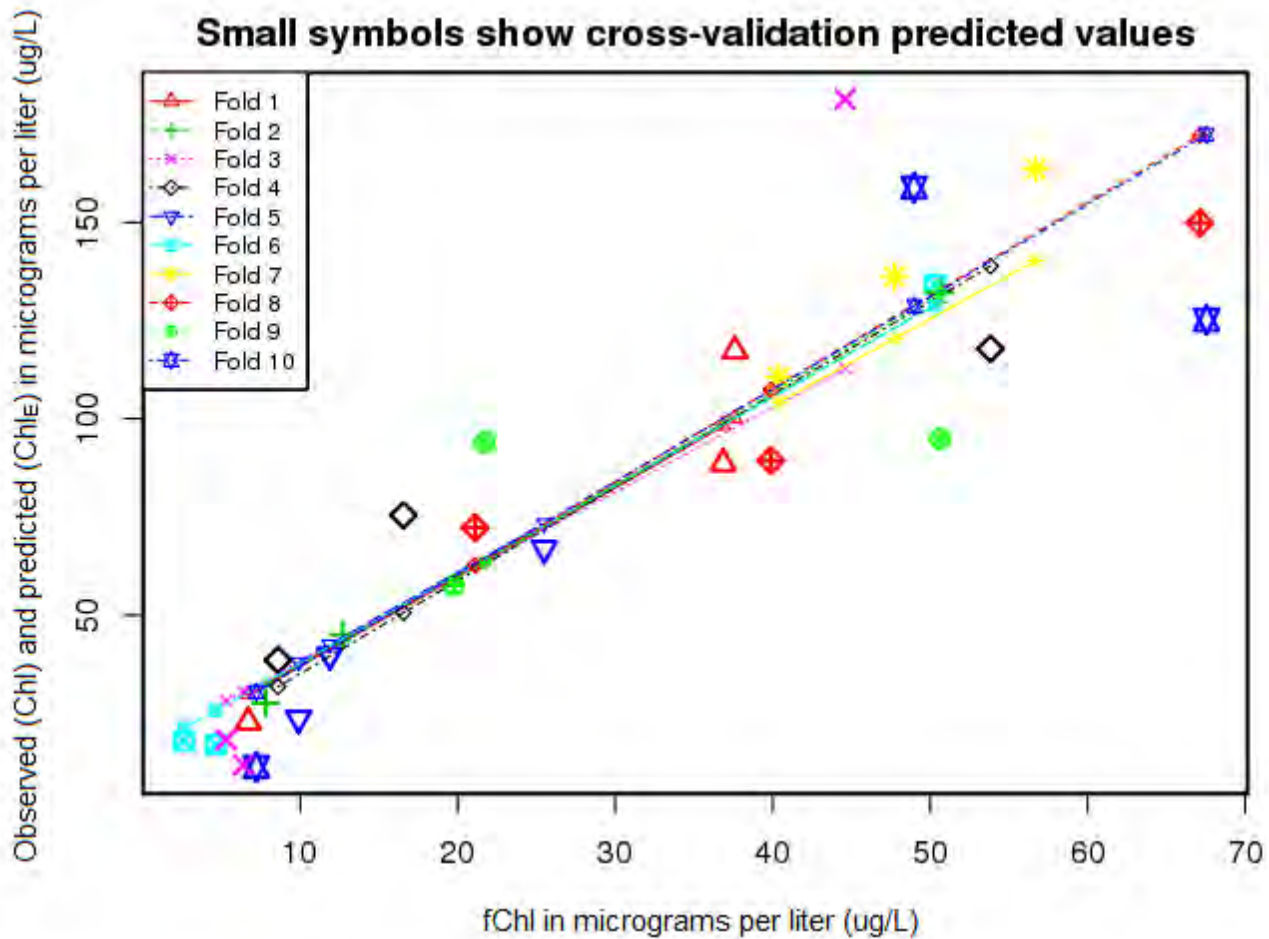
	Ch1	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
5	159.0	126	32.6	1.560	1.610	0.0639	0.0833	0.420
7	94.9	130	-35.0	-1.680	-1.740	0.0692	0.1050	-0.476
19	181.0	116	65.2	3.100	3.760	0.0518	0.2630	0.879
30	125.0	169	-43.6	-2.190	-2.370	0.1490	0.4210	-0.991
42	150.0	168	-17.9	-0.899	-0.896	0.1470	0.0695	-0.371

Statistical Plots





Cross Validation



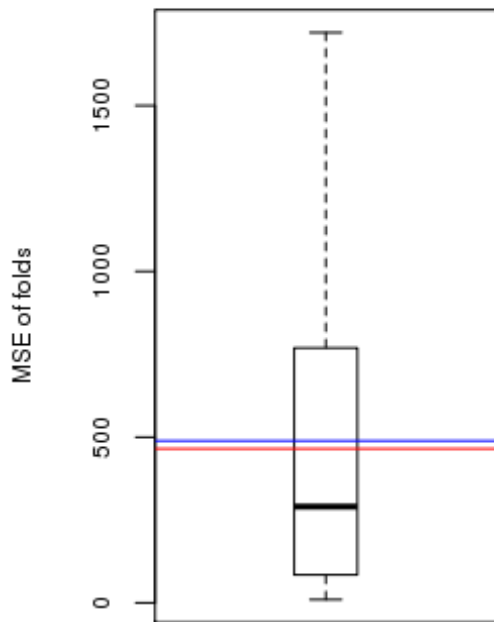
Fold - equal partition of the data (10 percent of the data)

Large symbols - observed value of a data point removed in a fold

Small symbols - recomputed value of a data point removed in a fold

Recomputed regression lines - adjusted regression line with one fold removed

Minimum MSE of folds:	9.82
Mean MSE of folds:	489.00
Median MSE of folds:	290.00
Maximum MSE of folds:	1720.00
(Mean MSE of folds) / (Model MSE):	1.05



Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Data Set

	Date	Ch1	fCh1	Computed Ch1 _E	Residual	Normal Quantiles	Censored Values
0							
1	7/30/2012	118	53.8	137	-19.5	-1.37	--
2	8/9/2012	89.3	39.9	105	-16.1	-0.777	--
3	8/13/2012	88.5	36.8	98.4	-9.9	-0.569	--
4	8/27/2012	159	49	126	32.6	1.62	--
5	9/10/2012	111	40.4	107	3.97	0.383	--
6	9/24/2012	94.9	50.6	130	-35	-1.62	--
7	10/15/2012	117	37.6	100	17.1	0.895	--
10	11/19/2012	39.8	11.9	41.3	-1.44	0.125	--
11	12/17/2012	17.1	4.65	24.7	-7.64	-0.383	--
12	1/14/2013	17.9	2.63	20.1	-2.18	-0.0415	--
13	2/11/2013	38.5	8.6	33.7	4.8	0.474	--
14	3/11/2013	57.6	19.8	59.3	-1.77	0.0415	--
15	4/8/2013	164	56.7	144	19.6	1.03	--

17	5/6/2013	45	12.7	43.2	1.81	0.209	--
18	5/20/2013	181	44.6	116	65.2	2.06	--
19	6/3/2013	75.4	16.6	52	23.5	1.18	--
20	6/17/2013	22.8	6.7	29.4	-6.64	-0.295	--
21	7/1/2013	27.6	7.8	31.9	-4.35	-0.125	--
22	7/15/2013	72.2	21.1	62.4	9.85	0.669	--
24	8/5/2013	23.7	9.9	36.7	-13	-0.669	--
26	8/19/2013	11.3	7.22	30.6	-19.3	-1.18	--
28	9/23/2013	94	21.7	63.7	30.2	1.37	--
29	10/21/2013	125	67.6	169	-43.6	-2.06	--
30	11/18/2013	67	25.5	72.4	-5.45	-0.209	--
32	12/16/2013	11.8	6.42	28.7	-16.9	-0.895	--
36	3/10/2014	18.3	5.3	26.2	-7.86	-0.474	--
38	4/7/2014	134	50.3	129	4.96	0.569	--
40	5/5/2014	136	47.7	123	12.7	0.777	--
41	5/19/2014	150	67.1	168	-17.9	-1.03	--
43	6/2/2014	132	50.6	130	2.23	0.295	--

Definitions

Ch1: Chlorophyll a, fluorometric method, uncorrected, micrograms per liter (parameter code 32217).

Ch1_E: Estimated chlorophyll concentration, micrograms per liter.

Cook's D: Cook's distance (Helsel and Hirsch, 2002).

DIFFITS: Difference in fits statistic (Helsel and Hirsch, 2002).

Leverage: An outlier's measure in the x direction (Helsel and Hirsch, 2002).

LOWESS: Locally weighted scatterplot smoothing (Cleveland, 1979; Helsel and Hirsch, 2002).

MSE: Model standard error (Helsel and Hirsch, 2002).

MSPE: Model standard percentage error (Helsel and Hirsch, 2002).

Probability(>|t|): The probability that the independent variable has no effect on the dependent variable (Helsel and Hirsch, 2002).

RMSE: Root mean square error (Helsel and Hirsch, 2002).

fCh1: Chlorophyll, total, water, fluorometric, 650-700 nanometers, in situ sensor, micrograms per liter (parameter code 62361).

t value: Student's t value; the coefficient divided by its associated standard error (Helsel and Hirsch, 2002).

References Cited

- Arar, E.J., and Collins, G.B., 1997, Method 445.0—In vitro determination of chlorophyll a and pheophytin a in marine and freshwater algae by fluorescence (rev. 1.2): U.S. Environmental Protection Agency, Office of Research and Development, 22 p.
- Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: *Journal of the American Statistical Association*, v. 74, p. 829-836.
- Graham, J.L., Foster, G.M., Williams, T.J., Mahoney, M.D., May, M.R., and Loftin, K.A., 2018, Water-quality conditions with an emphasis on cyanobacteria and associated toxins and taste-and-odor compounds in the Kansas River, Kansas, July 2012 through September 2016: U.S. Geological Survey Scientific Investigations Report 2018-5089, 55 p., <https://doi.org/10.3133/sir20185089>.
- Helsel, D.R., and Hirsch, R.M., 2002, Statistical methods in water resources—Hydrologic analysis and interpretation: U.S. Geological Survey Techniques of Water-Resources Investigations, book 4, chap. A3, 510 p.
- Knowlton, M.F., 1984, Flow-through microcuvette for fluorometric determination of chlorophyll: *Water Resources Bulletin*, v. 20, p. 1198-1205.
- R Core Team, 2017, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing. [Also available at <https://www.R-project.org/>.]
- Sartory, D.P., and Grobbelaar, J.U., 1984, Extraction of chloro-phyll a from freshwater phytoplankton for spectrophotometric analysis: *Hydrobiologia*, v. 114, no. 3, p. 177-187.
- U.S. Geological Survey, variously dated, National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A9 [variously paged].
- U.S. Geological Survey, 2018, National Water Information System—Web interface: accessed January 2017 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. + 8 attachments, accessed April 10, 2006, at <https://pubs.water.usgs.gov/tm1d3>.
- Yellow Springs Instruments, 2012, 6-Series Multiparameter Water Quality Sondes—User manual: Yellow Springs Instruments, 379 p.

