

Appendix 3. Model Archival Summary for Chlorophyll Concentration at Station 06887500; Kansas River at Wamego, Kansas, during June 10, 2014, through September 30, 2016

This model archival summary summarizes the chlorophyll (Chl; uncorrected for degradation products) concentration model developed to estimate Chl concentrations using Xylem YSI EXO2 sensor-measured 15-minute chlorophyll fluorescence (fChl) data during June 10, 2014, through September 30, 2016.

Site and Model Information

Site number: 06887500

Site name: Kansas River at Wamego, Kansas

Location: Lat 39°11'54", long 96°18'19" referenced to North American Datum of 1927, in SW 1/4 NW 1/4 SE 1/4 sec.9, T.10 S., R.10 E., Pottawatomie County, KS, Hydrologic Unit 10270102.

Equipment: A Xylem YSI EXO2 water-quality monitor equipped with a fChl sensor was installed during June 2014 through September 2016. 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 ($\mu\text{g/L}$)).

Date model was created: August 3, 2017

Model calibration data period and application date: June 10, 2014–September 30, 2016

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 Wamego, Kansas (Kans.), site using the open-source software package “R.” This model is specific to the Kansas River at Wamego 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 43 concurrent measurements of sensor-measured fChl and laboratory-measured Chl collected during June 10, 2014, through September 30, 2016. 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. One of the sensor-measured fChl samples, from May 18, 2015, was deemed an outlier and removed from the model calibration dataset. The sensor-measured fChl value of 79.98 $\mu\text{g/L}$ was likely biased high due to turbidity interference from a recent reservoir release. The corresponding laboratory-measured Chl was 7.24 $\mu\text{g/L}$. This was the only data pair in which sensor-measured fChl was greater than laboratory-measured Chl.

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 June 2014 through September 2016, 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 Wamego, Kans., streamgauge (USGS station number 06887500) during June 10, 2014, through September 30, 2016.

Chl concentration-based model:

$$Chl_E = 1.9 \times fChl + 4.12$$

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 Xylem YSI EXO2 sensor-measured chlorophyll fluorescence and laboratory-measured chlorophyll concentration; 06887500; Kansas River at Wamego, Kansas

Model Statistics, Data, and Plots

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

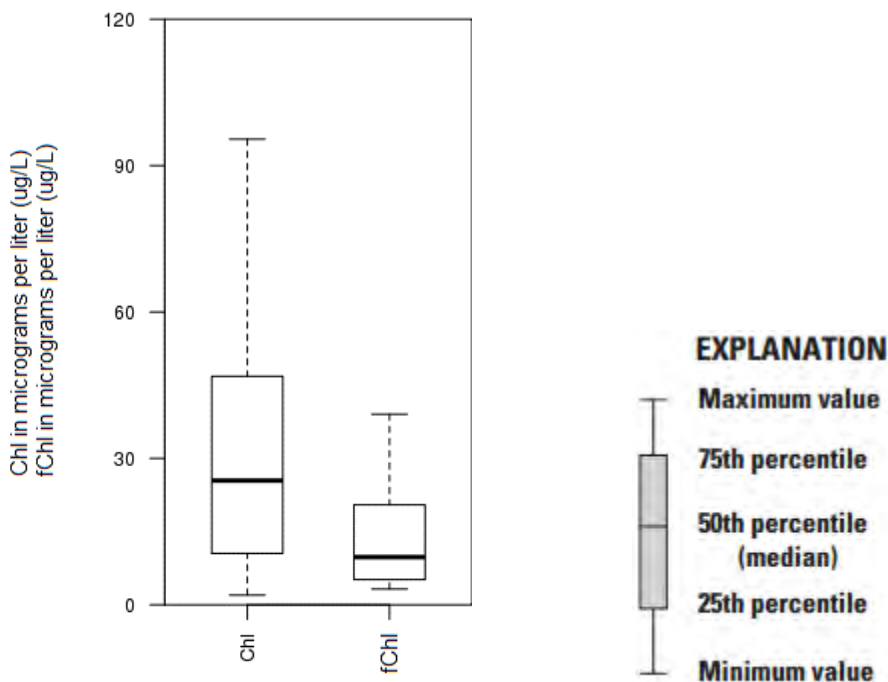
Model

$$\text{Chl}_E = + 1.9 * \text{fChl} + 4.12$$

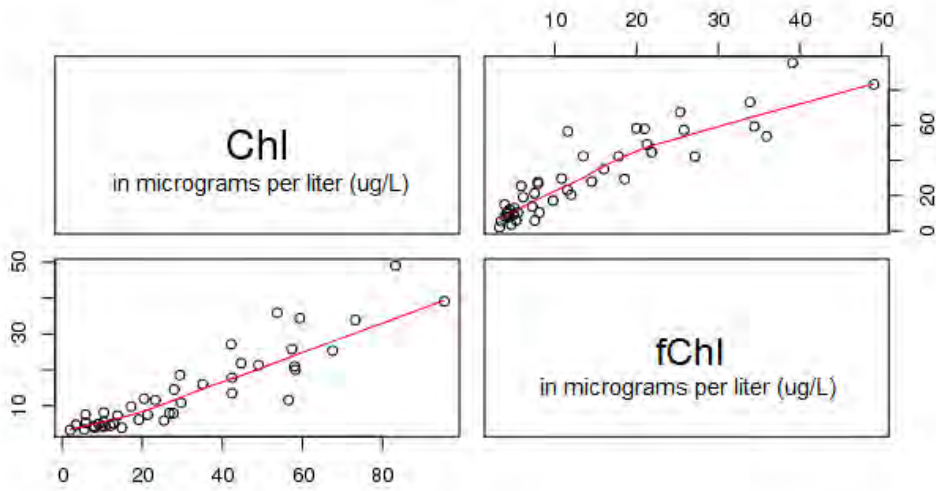
Variable Summary Statistics

	Chl	fChl
Minimum	1.99	3.24
1st Quartile	10.40	5.08
Median	25.50	9.78
Mean	30.90	14.10
3rd Quartile	49.00	21.00
Maximum	95.40	49.00

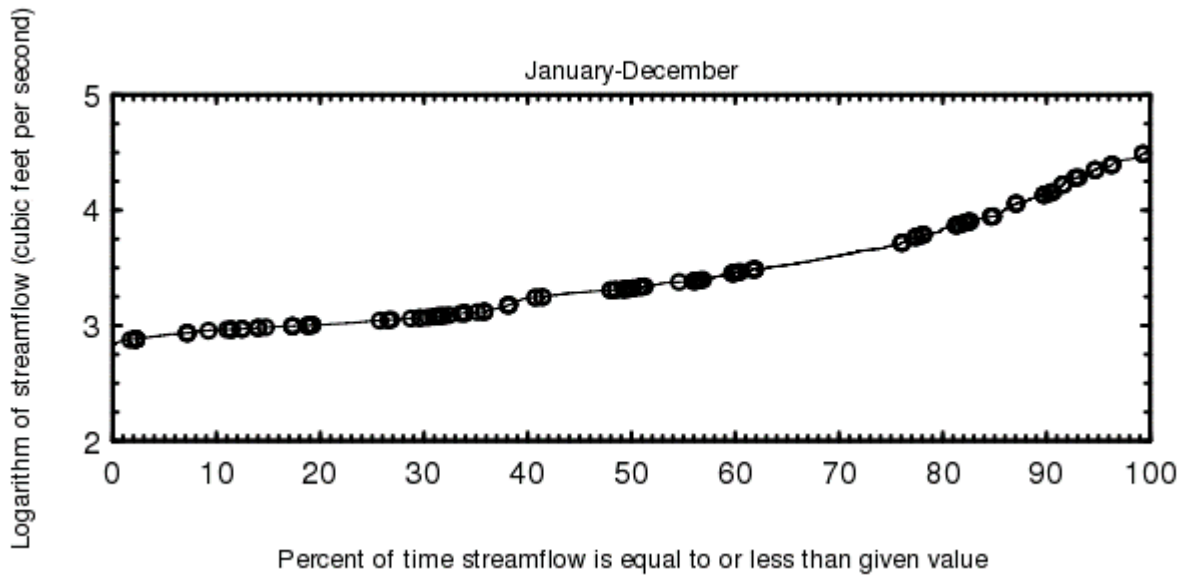
Box Plots

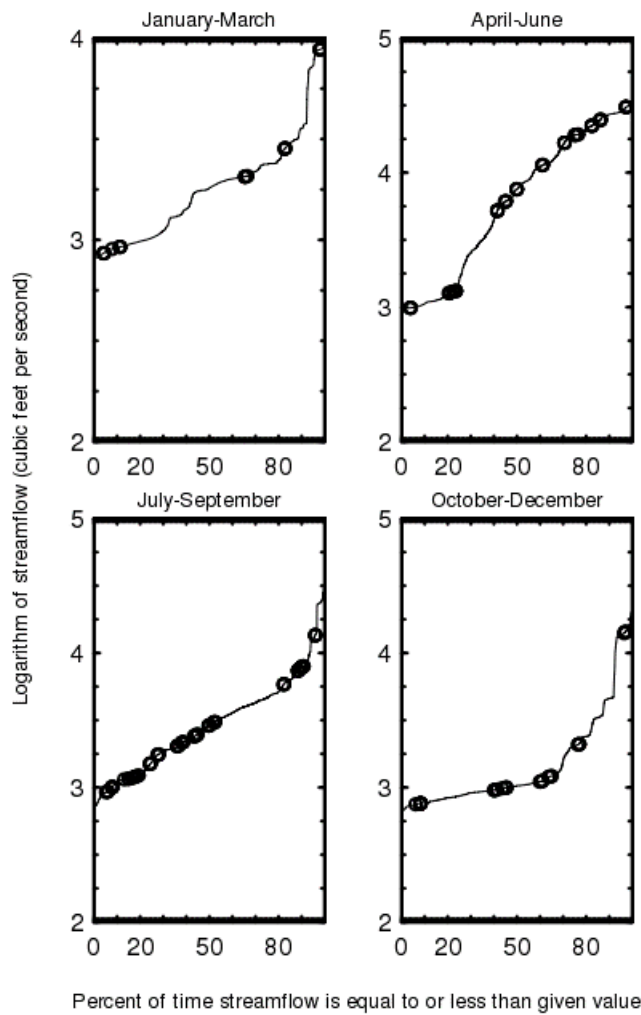


Exploratory Plots



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	43
Standard error (RMSE)	9.93
Upper Model standard percentage error (MSPE)	32.1
Lower Model standard percentage error (MSPE)	32.1
Coefficient of determination (R^2)	0.829
Adjusted Coefficient of Determination (Adj. R^2)	0.825

Explanatory Variables

	Coefficients	Standard Error	t value	Probability(> t)
(Intercept)	4.12	2.430	1.7	9.74e-02
fCh1	1.90	0.135	14.1	2.48e-17

Correlation Matrix

	Intercept	fCh1
Intercept	1.000	-0.782
fCh1	-0.782	1.000

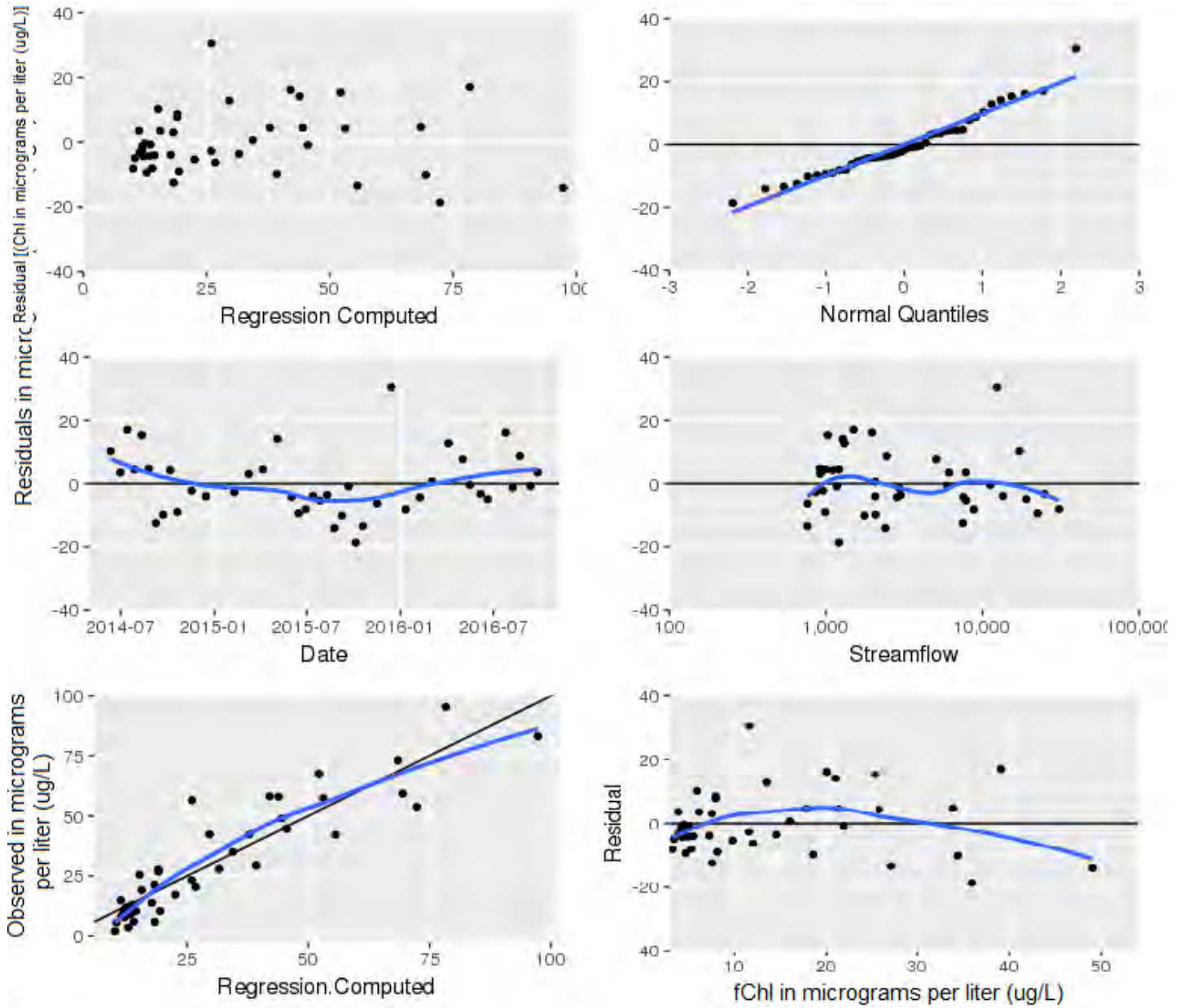
Outlier Test Criteria

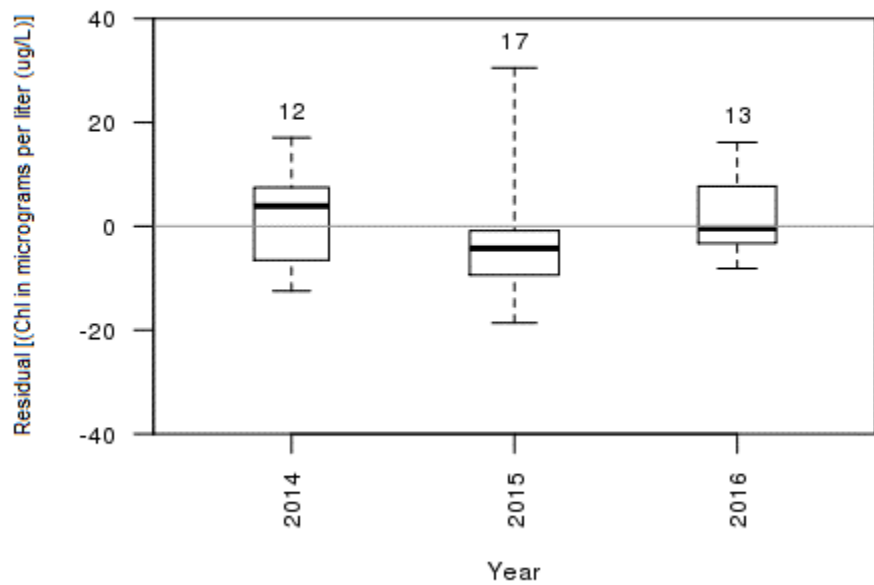
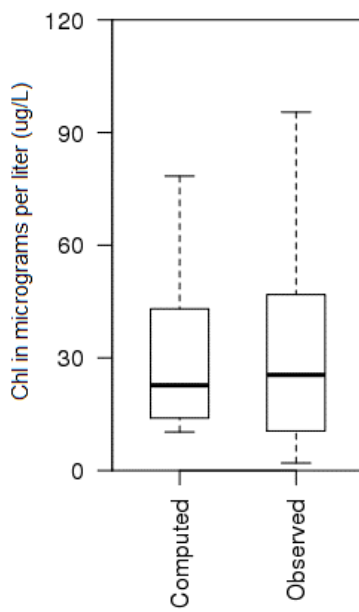
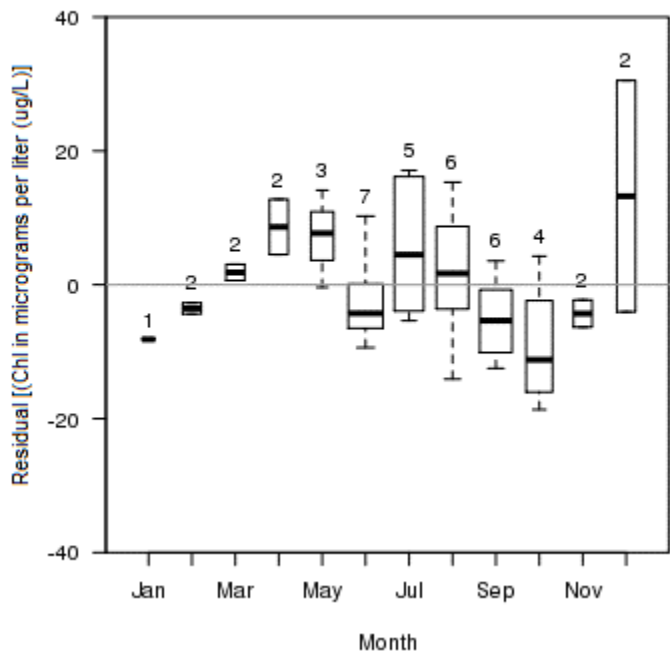
Leverage	Cook's D	DFFITS
0.0698	0.1056	0.3050

Flagged Observations

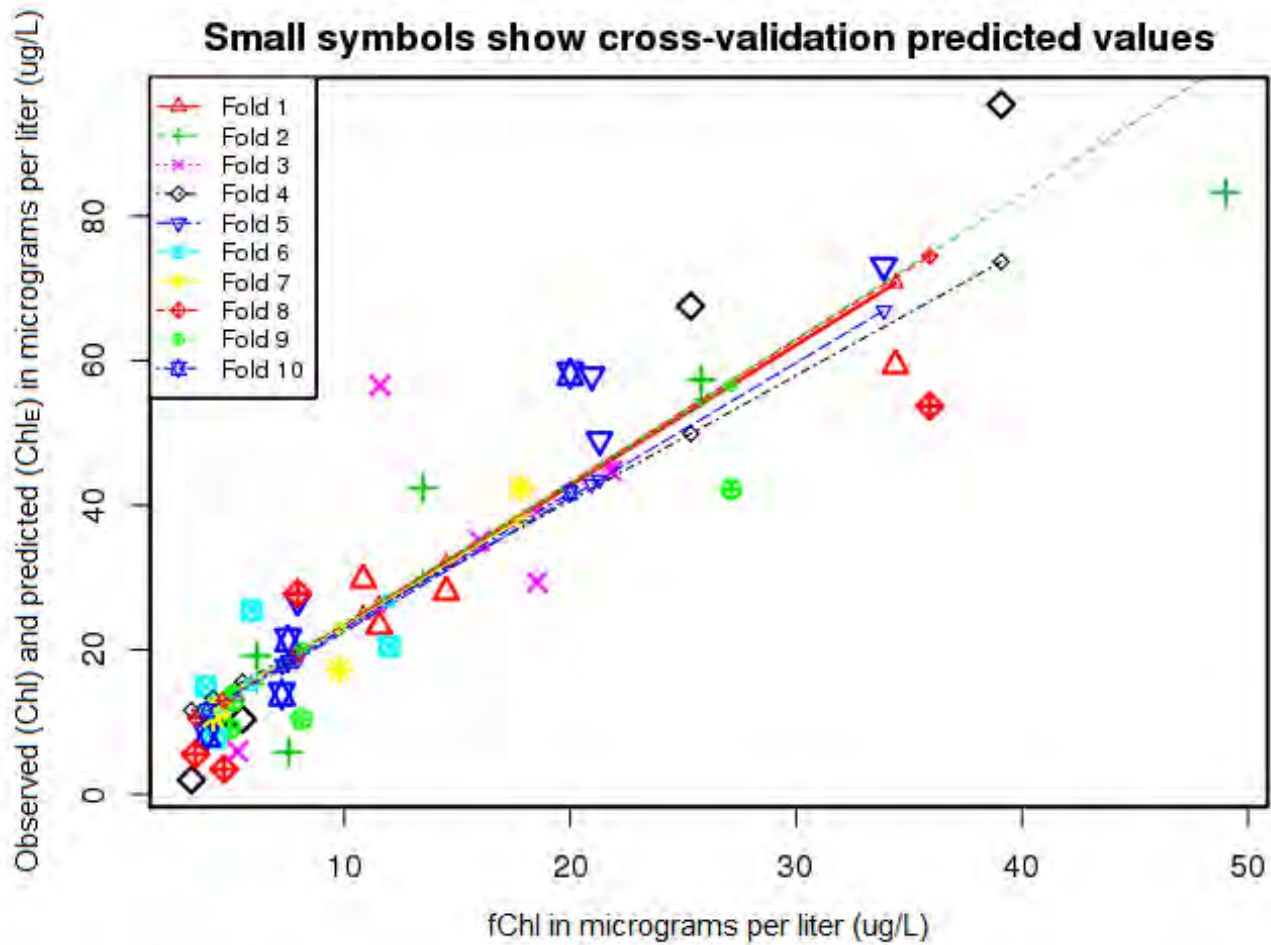
	Ch1	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
3	95.4	78.4	17.00	1.85	1.900	0.1380	0.2730	0.762
6	67.5	52.3	15.20	1.57	1.600	0.0465	0.0602	0.354
7	73.2	68.5	4.63	0.49	0.486	0.0952	0.0127	0.158
26	83.2	97.3	-14.10	-1.64	-1.670	0.2480	0.4400	-0.959
28	59.3	69.5	-10.20	-1.08	-1.080	0.0990	0.0641	-0.359
30	53.7	72.4	-18.70	-2.00	-2.070	0.1110	0.2480	-0.732
32	42.2	55.7	-13.50	-1.40	-1.410	0.0545	0.0561	-0.339
35	56.5	26.1	30.40	3.10	3.490	0.0244	0.1200	0.553

Statistical Plots





Cross Validation



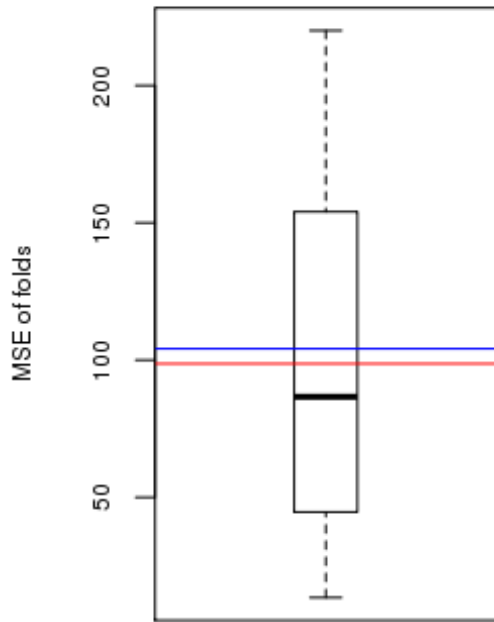
Fold - equal partition of the data (10 percent of the data)

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: 13.50
Mean MSE of folds: 104.00
Median MSE of folds: 86.60
Maximum MSE of folds: 220.00
(Mean MSE of folds) / (Model MSE): 1.05



Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Data Set

θ	Date	Ch1	fCh1	Computed Ch1 _E	Residual	Normal Quantiles	Censored Values
13	6/11/2014	25.5	5.9	15.3	10.1	1.02	--
1	6/30/2014	19.2	6.13	15.8	3.38	0.355	--
2	7/14/2014	95.4	39.1	78.4	17	1.79	--
3	7/28/2014	49	21.3	44.6	4.4	0.549	--
5	8/11/2014	67.5	25.3	52.3	15.2	1.38	--
6	8/25/2014	73.2	33.9	68.5	4.63	0.689	--
7	9/8/2014	5.85	7.55	18.5	-12.6	-1.38	--
8	9/22/2014	29.4	18.5	39.3	-9.95	-1.13	--
9	10/6/2014	57.4	25.8	53.2	4.19	0.482	--
10	10/20/2014	10.4	8.12	19.6	-9.14	-0.931	--
11	11/17/2014	9.7	4.2	12.1	-2.4	-0.0581	--
12	12/15/2014	10.4	5.5	14.6	-4.18	-0.355	--
14	2/9/2015	23.2	11.6	26.1	-2.82	-0.116	--

15	3/9/2015	21.3	7.52	18.4	2.93	0.294	--
16	4/6/2015	42.4	17.8	38	4.41	0.617	--
17	5/4/2015	58	21	44	14	1.25	--
19	6/1/2015	9.13	4.95	13.5	-4.4	-0.418	--
20	6/15/2015	3.48	4.68	13	-9.54	-1.02	--
21	6/29/2015	5.95	5.28	14.2	-8.21	-0.765	--
22	7/13/2015	13.8	7.24	17.9	-4.05	-0.294	--
23	7/27/2015	17.2	9.78	22.7	-5.5	-0.617	--
24	8/10/2015	28	14.5	31.7	-3.73	-0.234	--
25	8/24/2015	83.2	49	97.3	-14.1	-1.79	--
27	9/8/2015	59.3	34.4	69.5	-10.2	-1.25	--
28	9/21/2015	44.7	21.8	45.6	-0.952	0.0581	--
29	10/5/2015	53.7	35.9	72.4	-18.7	-2.2	--
31	10/19/2015	42.2	27.1	55.7	-13.5	-1.55	--
33	11/16/2015	20.5	12	26.9	-6.45	-0.689	--
34	12/14/2015	56.5	11.6	26.1	30.4	2.2	--
35	1/11/2016	1.99	3.24	10.3	-8.29	-0.845	--
36	2/8/2016	7.8	4.3	12.3	-4.49	-0.482	--
37	3/3/2016	35.1	16	34.5	0.528	0.234	--
38	4/4/2016	42.4	13.5	29.7	12.7	1.13	--
40	5/2/2016	26.8	7.94	19.2	7.58	0.845	--
41	5/16/2016	12.1	4.48	12.6	-0.536	0.175	--
42	6/6/2016	8.15	3.92	11.6	-3.42	-0.175	--
43	6/20/2016	5.53	3.42	10.6	-5.09	-0.549	--
45	7/11/2016	29.7	10.8	24.7	4.98	0.765	--
46	7/25/2016	58.2	20	42.1	16.1	1.55	--
47	8/8/2016	10.7	4.2	12.1	-1.42	0	--
49	8/22/2016	27.8	7.94	19.2	8.58	0.931	--
51	9/12/2016	12.9	5.08	13.8	-0.877	0.116	--
52	9/26/2016	14.9	3.88	11.5	3.4	0.418	--

Definitions

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

Chl_E: Estimated chlorophyll concentration, micrograms per liter.

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

DFFITS: 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).

fChl: Chlorophylls, water, in situ, fluorometric method, excitation at 470 +-15 nm, emission at 685 +-20 nm, micrograms per liter (parameter code 32318).

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, EXO user manual-Advanced water quality monitoring platform: Yellow Springs Instruments, 137 p.