Appendix 1. Model Archival Summary for Chlorophyll Concentration at Station 06887500; Kansas River at Wamego, Kansas, during August 30, 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 August 30, 2012, through June 10, 2014.

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 YSI 6600 water-quality monitor equipped with model 6025 fChl sensor was installed during August 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: August 30, 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 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 25 concurrent measurements of sensor-measured fChl and laboratory-measured Chl collected during August 30, 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 Wamego, Kans., streamgage (U.S. Geological Survey station number 06887500) during August 30, 2012, through June 10, 2014. A limitation in model predictive ability on the upper range of the observed concentrations is likely due to the small sample size (25) and can be seen in the cross-validation plots below.

Chl concentration model:

$$Chl_E = 2.05 \times fChl + 6.55$$

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; 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.

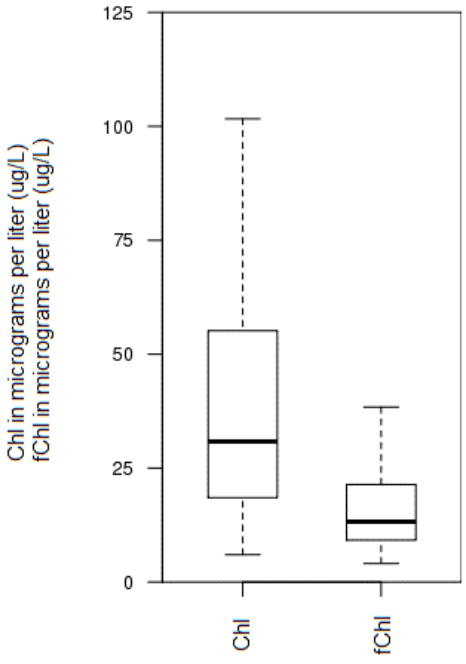
Model

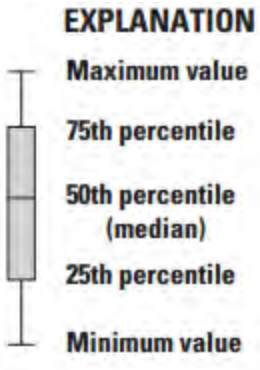
Chl_E = + 2.05 * fChl + 6.55

Variable Summary Statistics

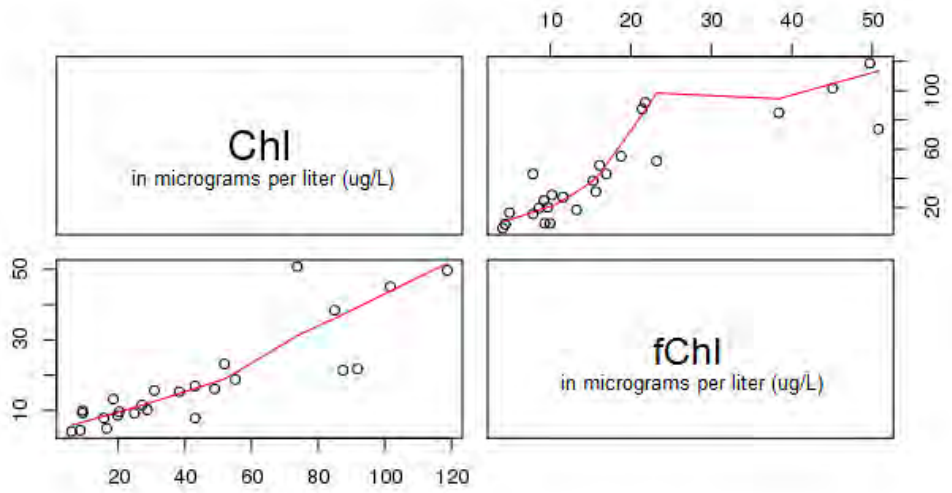
	Chl	fChl
Minimum	6.0	4.07
1st Quartile	18.5	9.20
Median	30.8	13.20
Mean	43.0	17.80
3rd Quartile	55.2	21.40
Maximum	119.0	50.80

Box Plots

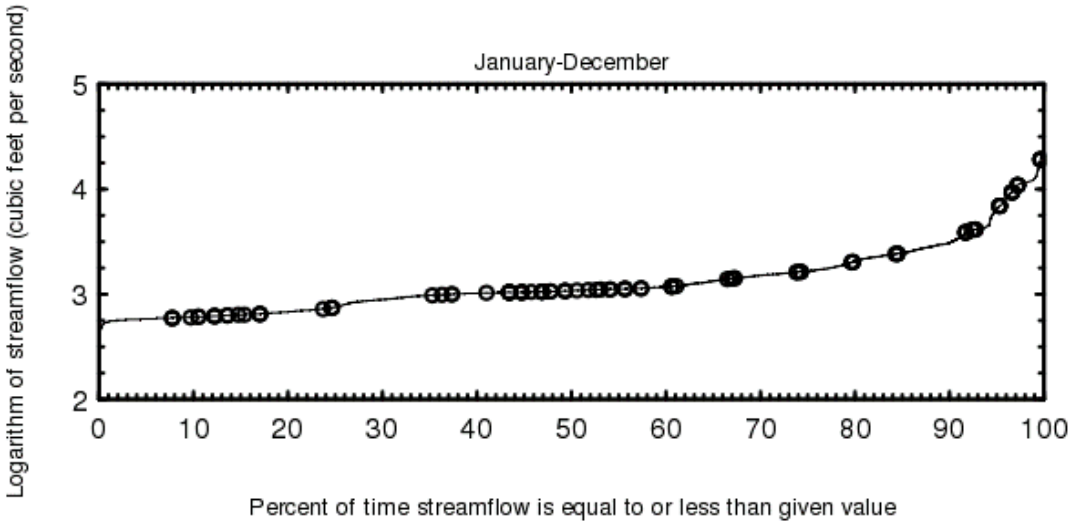


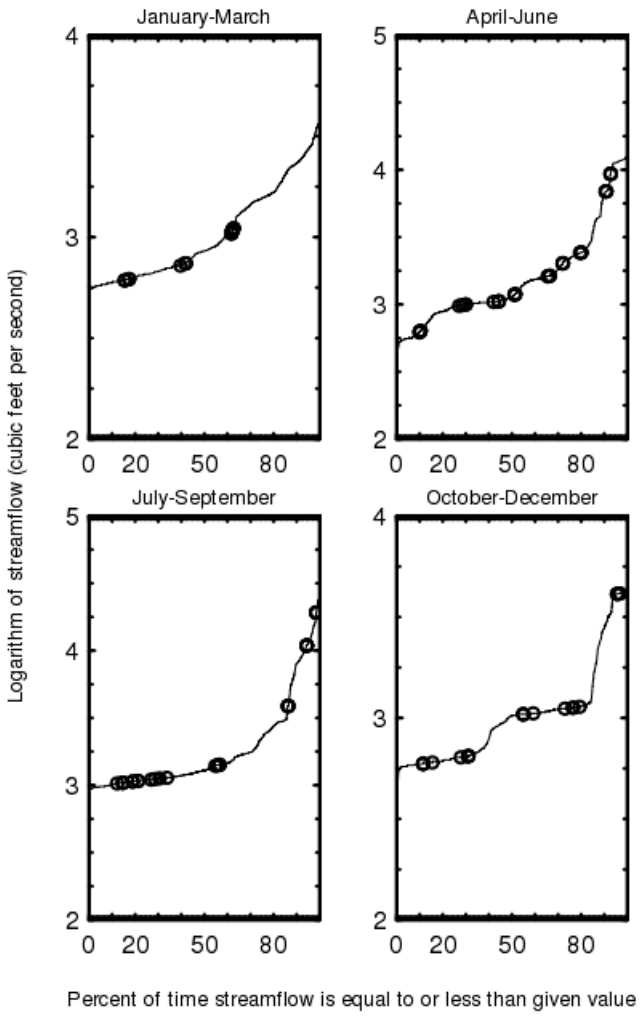


Exploratory Plots



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	25
Standard error (RMSE)	16.5
Upper Model standard percentage error (MSPE)	38.3
Lower Model standard percentage error (MSPE)	38.3
Coefficient of determination (R^2)	0.755
Adjusted Coefficient of Determination (Adj. R^2)	0.744

Explanatory Variables

	Coefficients	Standard Error	t value	Probability(>|t|)
(Intercept)	6.55	5.440	1.20	2.41e-01
fChl	2.05	0.244	8.41	1.79e-08

Correlation Matrix

	Intercept	fChl
Intercept	1.000	-0.796
fChl	-0.796	1.000

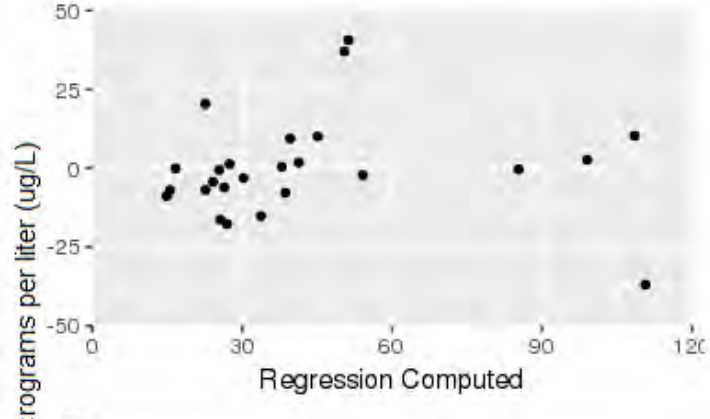
Outlier Test Criteria

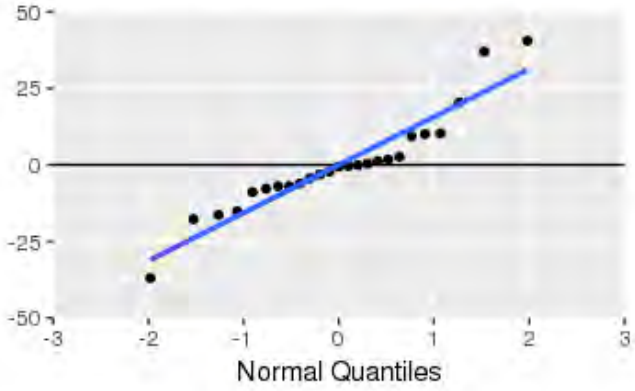
Leverage	Cook's D	DFFITS
0.120	0.106	0.400

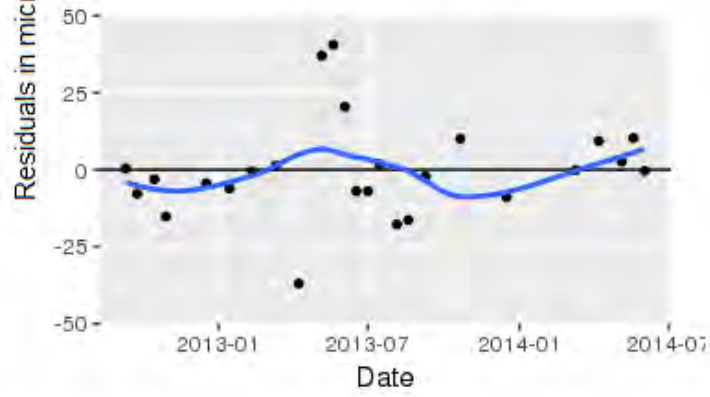
Flagged Observations

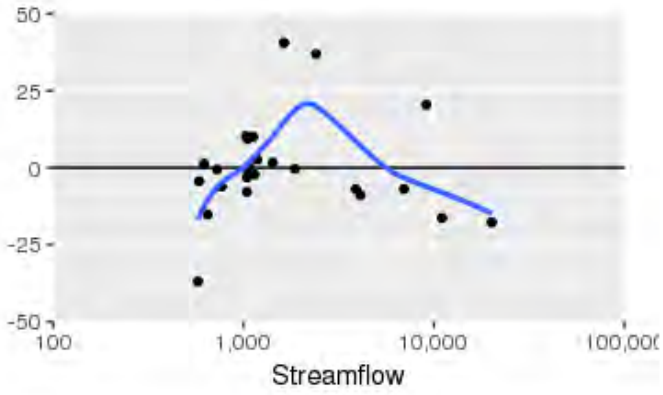
	Chl	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
15	73.7	111.0	-36.900	-2.6500	-3.1000	0.2790	1.360000	-1.93000
16	87.4	50.4	37.000	2.3000	2.5600	0.0429	0.119000	0.54300
17	91.8	51.3	40.600	2.5200	2.9000	0.0436	0.145000	0.61900
35	102.0	99.0	2.630	0.1790	0.1760	0.2040	0.004130	0.08890
36	119.0	108.0	10.300	0.7280	0.7210	0.2640	0.095100	0.43200
38	84.9	85.3	-0.386	-0.0252	-0.0247	0.1340	0.000049	-0.00968

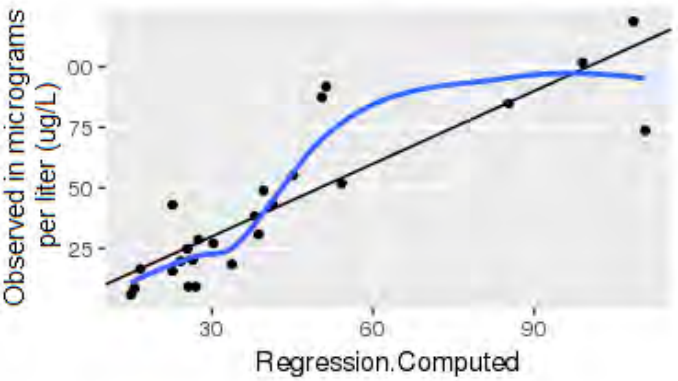
Statistical Plots

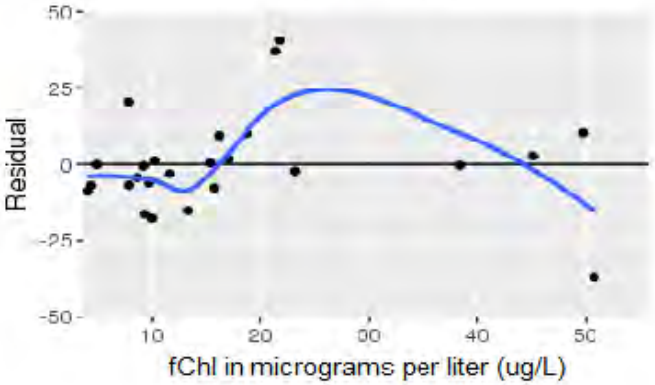


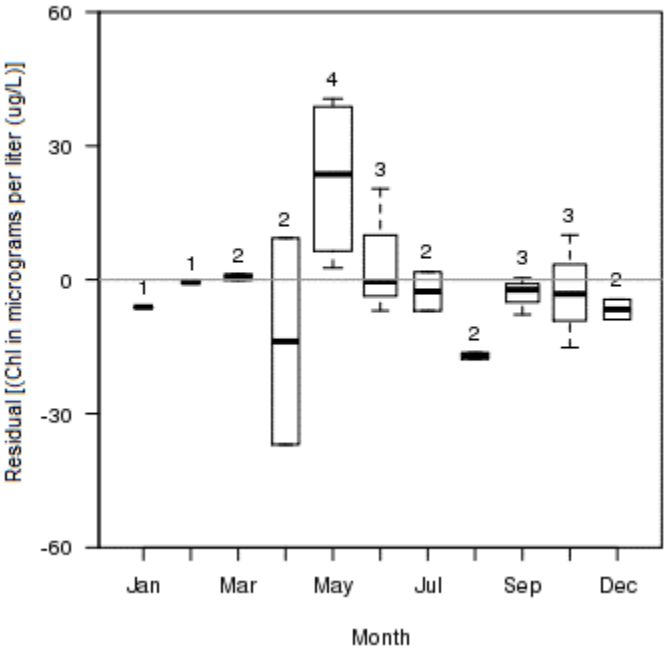


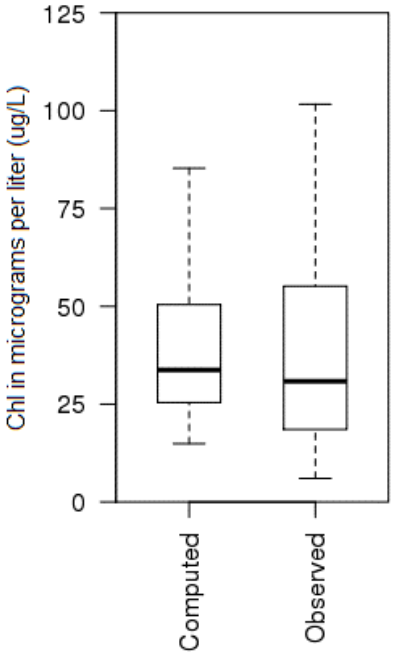


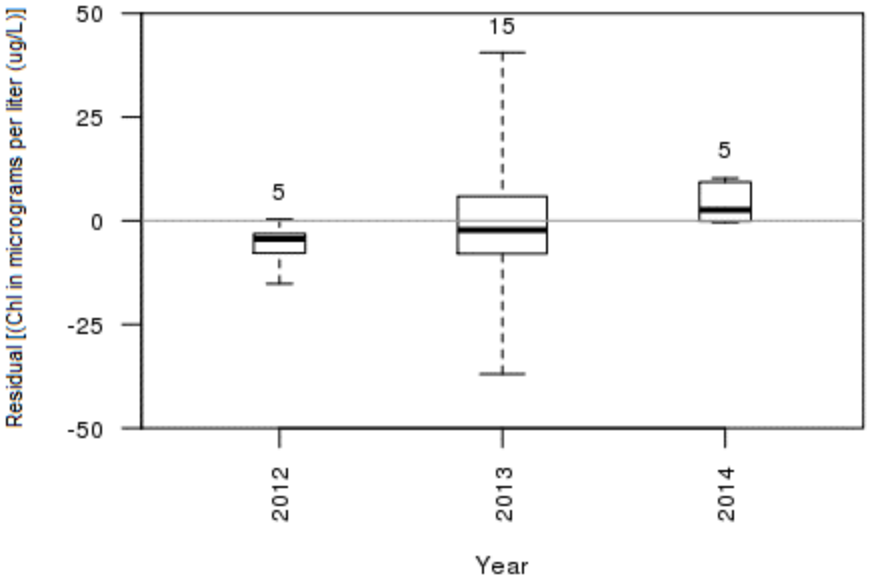




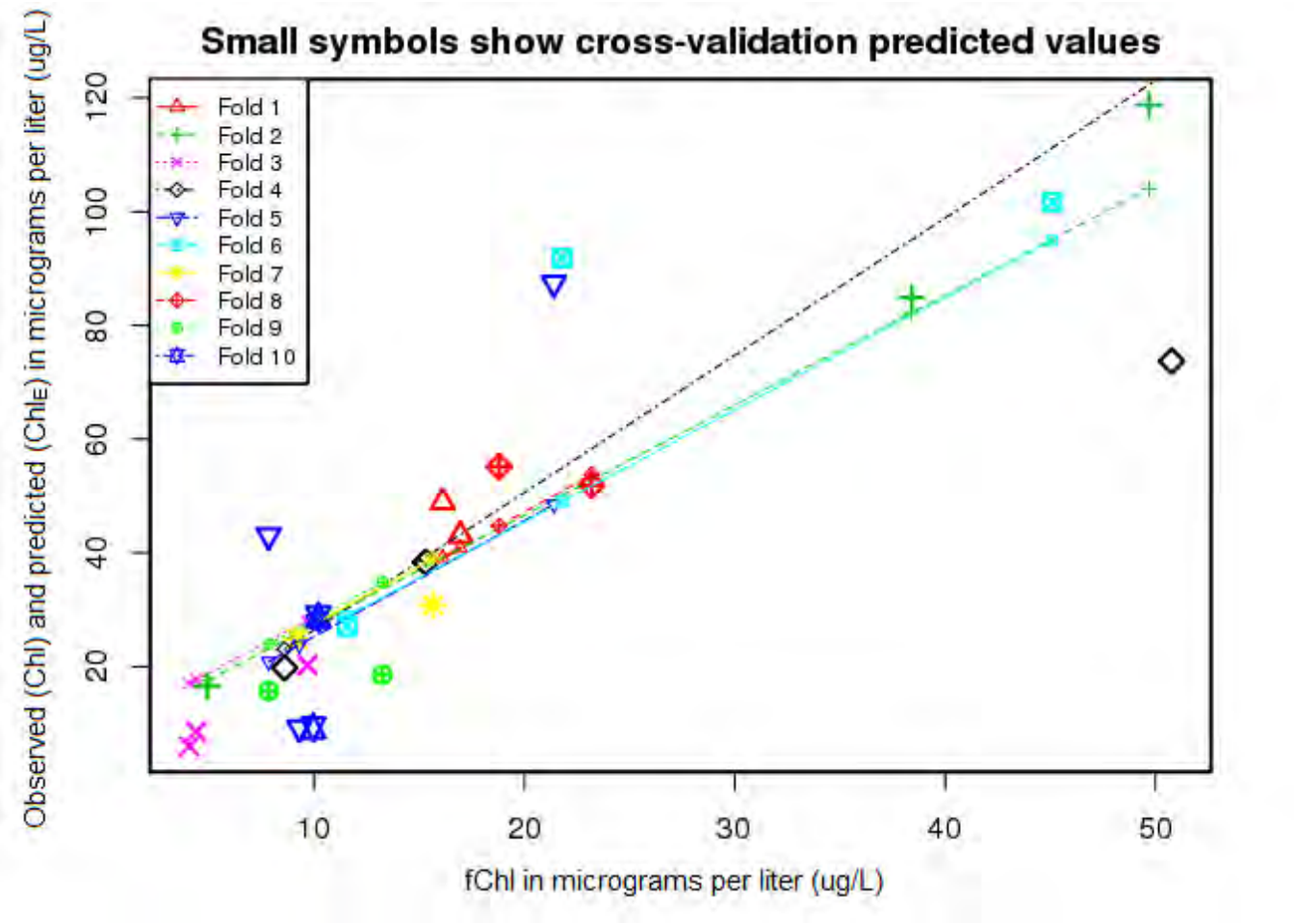



Residual [(Chl in micrograms per liter (ug/L)]
60
30
0
-30
-60
1
1
2
2
4
3
2
2
3
3
2
Jan
Mar
May
Jul
Sep
Dec
Month


Chl in micrograms per liter (ug/L)
125
100
75
50
25
0
Computed
Observed


Residual [(Chl in micrograms per liter (ug/L)]
50
25
0
-25
-50
5
15
5
2012
2013
2014
Year

Cross Validation



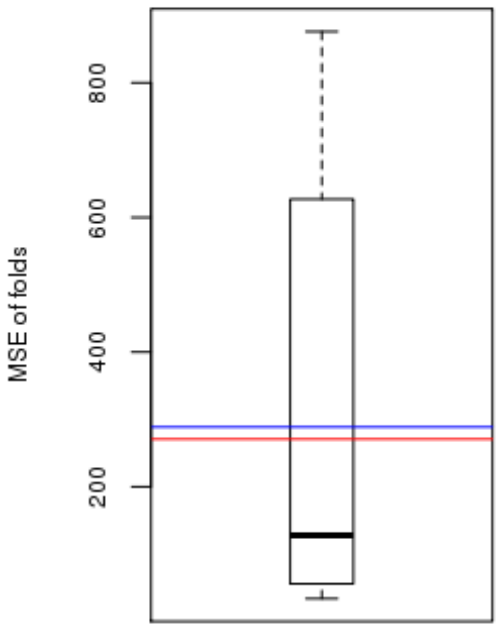

```
Fold - equal partition of the data (10 percent of the data)

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:   33.80
                 Mean MSE of folds:  289.00
               Median MSE of folds:  128.00
              Maximum MSE of folds:  876.00
 (Mean MSE of folds) / (Model MSE):    1.07
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Data Set

	Date	Chl	fChl	Computed Chl_E	Residual	Normal Quantiles	Censored Values
5	9/10/2012	38.3	15.3	37.9	0.417	0.303	--
6	9/24/2012	30.8	15.6	38.6	-7.79	-0.765	--
7	10/15/2012	27.1	11.6	30.3	-3.14	-0.2	--
8	10/29/2012	18.5	13.2	33.7	-15.2	-1.07	--
10	12/17/2012	19.8	8.6	24.2	-4.38	-0.303	--
11	1/14/2013	20.3	9.68	26.4	-6.13	-0.409	--
12	2/11/2013	24.8	9.2	25.4	-0.563	0	--
13	3/11/2013	28.7	10.2	27.5	1.27	0.409	--
14	4/8/2013	73.7	50.8	111	-36.9	-1.98	--
15	5/6/2013	87.4	21.4	50.4	37	1.53	--
16	5/20/2013	91.8	21.8	51.3	40.6	1.98	--
17	6/3/2013	43	7.84	22.6	20.4	1.26	--
18	6/17/2013	15.7	7.85	22.6	-6.91	-0.52	--

19	7/1/2013	8.6	4.4	15.6	-6.97	-0.637	--
20	7/15/2013	43	17	41.3	1.74	0.52	--
21	8/5/2013	9.23	9.96	27	-17.7	-1.53	--
23	8/19/2013	9.27	9.3	25.6	-16.3	-1.26	--
24	9/9/2013	51.9	23.2	54.1	-2.23	-0.0996	--
26	10/21/2013	55.2	18.8	45.1	10	0.906	--
29	12/16/2013	6	4.07	14.9	-8.89	-0.906	--
32	3/10/2014	16.6	4.93	16.7	-0.106	0.2	--
33	4/7/2014	48.9	16.1	39.6	9.35	0.765	--
34	5/5/2014	102	45.1	99	2.63	0.637	--
35	5/19/2014	119	49.7	108	10.3	1.07	--
37	6/2/2014	84.9	38.4	85.3	-0.386	0.0996	--

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: 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 spectrophotomet¬ric 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.