Appendix 1. Model Archival Summary for Chlorophyll Concentration at Milford Lake, May 26, June 9, July 14, July 21, and September 15, 2016

This model archival summary summarizes the laboratory-measured chlorophyll concentration (LabChl; uncorrected for degradation products) model developed to estimate LabChl concentrations at Milford Lake on May 26, June 9, July 14, July 21, and September 15, 2016. This model is specific to data collected for the purposes of this study alone and cannot be reliably applied to other data collected from Milford Lake for other studies or times, or data collected from other lakes. Model statistics and plots were developed using an internal U.S. Geological Survey R application for producing model archive summaries accessed on November 15, 2017.

Site and Model Information

Site name.—Milford Lake, Kansas

Equipment.—A Yellow Springs Instrument, Inc., EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, pH, turbidity, chlorophyll fluorescence, phycocyanin fluorescence, and fluorescent dissolved organic matter was mounted under a boat at a 0.5-meter (m) depth for spatial surveys completed on Milford Lake on May 26, June 9, July 14, July 21, and September 15, 2016. Boat speed was about 14 kilometers per hour, which provided the best balance of data quality and the ability to complete multiple representative surveys of Zone C of Milford Lake (Foster and others, 2018, fig. 1) in a timely manner. Readings from the water-quality monitor were recorded every 30 seconds. Phycocyanin was used as the single explanatory variable in the model because that model explained the most variance in LabChl, and it is consistent with the model developed by Foster and others (2017). Discrete water-quality samples for LabChl analysis were collected at multiple locations throughout Zone C of Milford Lake (Foster and others, 2018, table 1).

Date model was created.—November 15, 2017

Model calibration data periods.—May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model application dates.—May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model-Calibration Dataset

All data were collected using U.S. Geological Survey protocols (U.S. Geological Survey, variously dated) and are stored in the National Water Information System database (U.S. Geological Survey, 2018). The explanatory variable selected as input to the linear regression was phycocyanin fluorescence, in relative fluorescence units (RFU). Because most discrete samples were collected at the depth of the monitor (0.5 m), the linear relation between sensor-measured phycocyanin RFU and laboratory-measured chlorophyll could be used to compute LabChl concentrations in micrograms per liter for the spatial survey. The linear regression model was developed using the open-source software package R (version 3.2.3).

The regression model is based on 39 concurrent measurements of sensor-measured phycocyanin and laboratory-measured chlorophyll (uncorrected for degradation products) collected on May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016. No samples were below sensor- or laboratory-detection limits. Summary statistics and the complete model-calibration dataset are provided in this appendix. A total of three samples, collected on July 21, 2016, at 7:30 a.m., 11:40 a.m., and 1:00 p.m., were excluded from the dataset used to develop the regression model because the samples were collected from near-shore areas with dense surface accumulations, which are not representative of typical conditions throughout Zone C of Milford Lake. Studentized residuals from the final model were inspected for values greater than 3 or less than −3. 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

Most (about 80 percent) chlorophyll samples for laboratory analysis were collected at a 0.5-m depth (the depth of the monitor) from open-water locations. Some samples (n=8) collected during July 14, 2016, were integrated from the surface to 0.5 m; because these samples did not have undue influence on the model and were not flagged as potential outliers, they were retained in the dataset. Sample locations were not predetermined and were selected to represent the range of cyanobacterial conditions in the lake based on visual cues and continuous water-quality monitor data. Samples were analyzed for LabChl concentration at the U.S. Geological Survey Kansas Water Science Center as described in the "Methods" section of the report. Chlorophyll (uncorrected for degradation products) was analyzed fluorometrically using U.S. Environmental Protection Agency method 445.0 (Arar and Collins, 1997), modified using heated ethanol extraction (Sartory and Grobbelaar, 1984) and a fluorometer equipped with a flow-through cell (Knowlton, 1984). Additional detail on sample collection is available in the "Methods" section of the report.

Model Development

Ordinary least squares regression analysis was done using R (version 3.2.3) with sensor-measured phycocyanin RFU as the explanatory variable for laboratory-measured chlorophyll concentration. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and computed values) as compared to computed LabChl concentrations were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of computed values). Values for all regression statistics and metrics are included in this appendix 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 phycocyanin RFU and laboratory-measured chlorophyll at Milford Lake, May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016.

The LabChl concentration model is represented by the following:

$$LabChl = +30.9 * SensorPCY + 0.837$$

where

LabChl is laboratory-measured chlorophyll in micrograms per liter and

SensorPcy is sensor-measured phycocyanin in RFU.

R Output for the Relation Between Sensor-Measured Phycocyanin Relative Fluorescence Units and Laboratory-Measured Chlorophyll at Milford Lake, May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model Statistics, Data, and Plots

Definitions for terms used in this output are included at the end of this document.

Model

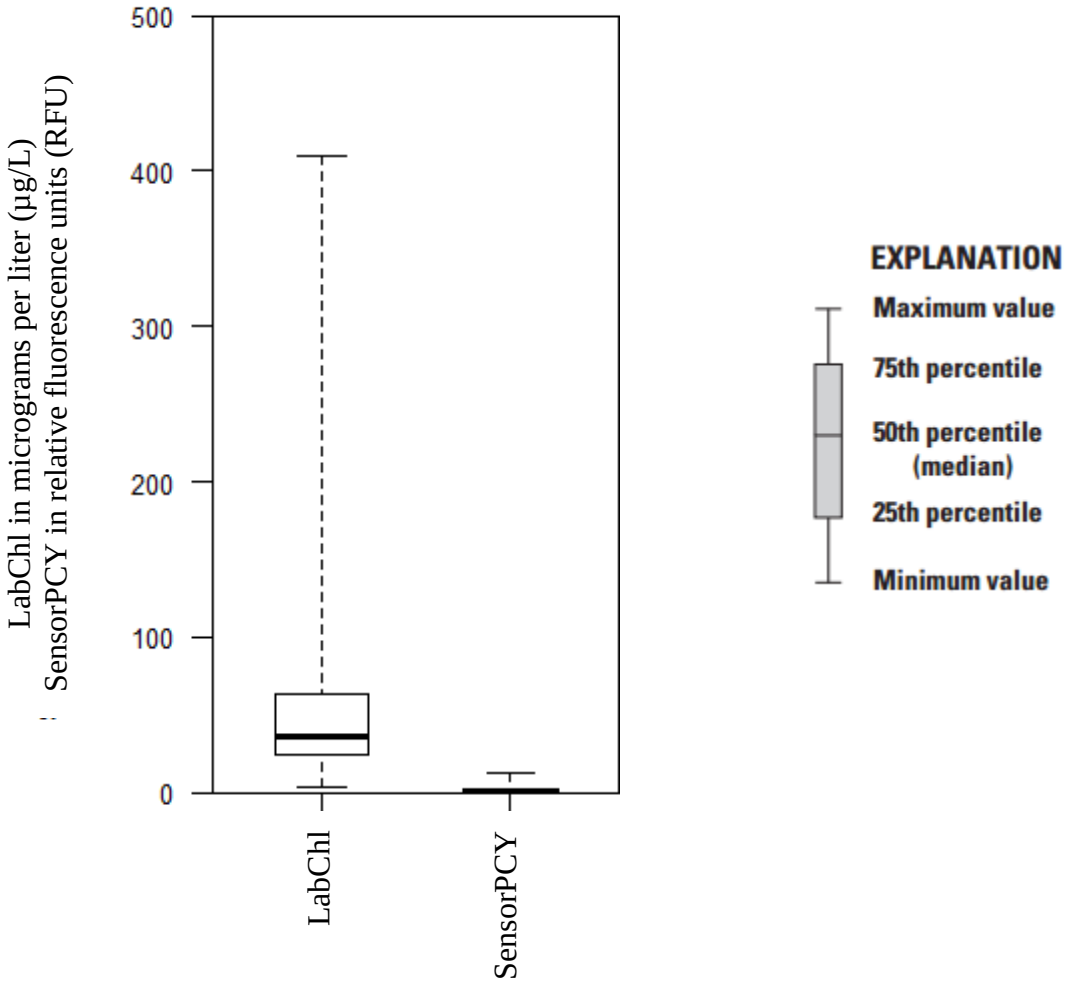
$$LabChl = +30.9 * SensorPCY + 0.837$$

Variable Summary Statistics

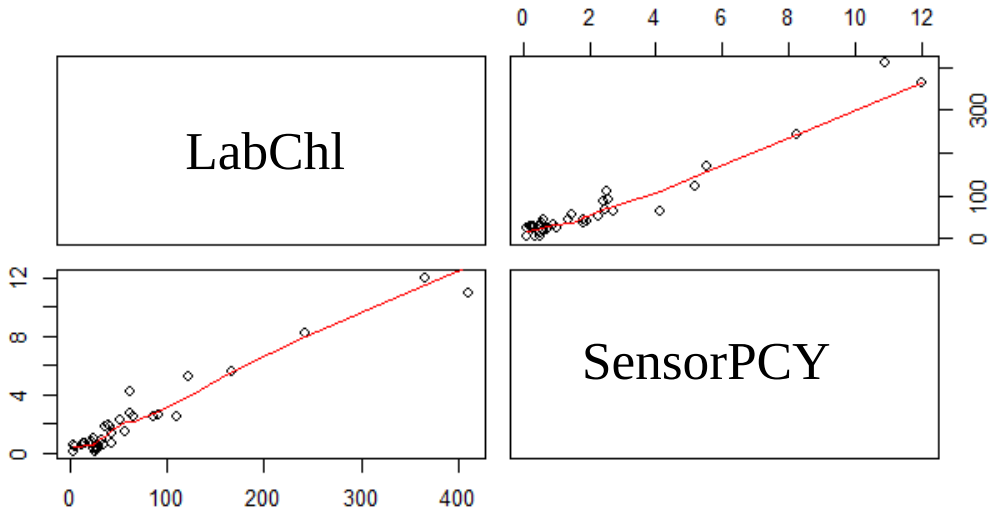
	LabChl	SensorPCY
Minimum	3.75	0.12
1st Quartile	23.50	0.51
Median	35.50	0.92

Mean	66.00	2.11
3rd Quartile	63.20	2.49
Maximum	410.00	12.00

Box Plots



Exploratory Plots



Red line shows the locally weighted scatterplot smoothing (LOWESS); LabChl, in micrograms per liter; SensorPCY, in relative fluorescence units.

Basic Model Statistics

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

Number of Observations	39
Standard error (RMSE)	22.4
Upper Model standard percentage error (MSPE)	34
Lower Model standard percentage error (MSPE)	34
Coefficient of determination (R^2)	0.938
Adjusted Coefficient of Determination (Adj. R^2)	0.937

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.837	4.52	0.185	8.54e−01
SensorPCY	30.900	1.30	23.700	5.65e−24

Correlation Matrix

	Intercept	SensorPC
Intercept	1.000	−0.608
SensorPC	-0.608	1.000

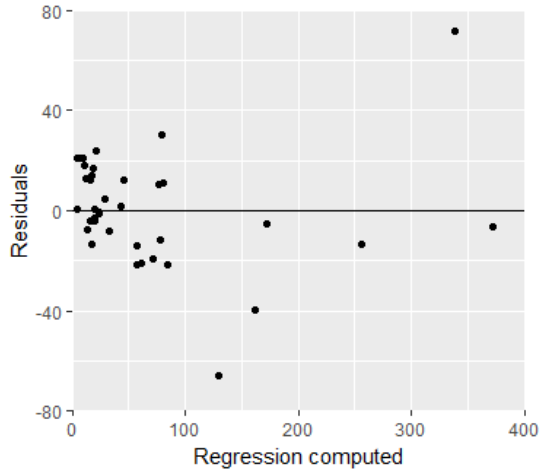
Outlier Test Criteria

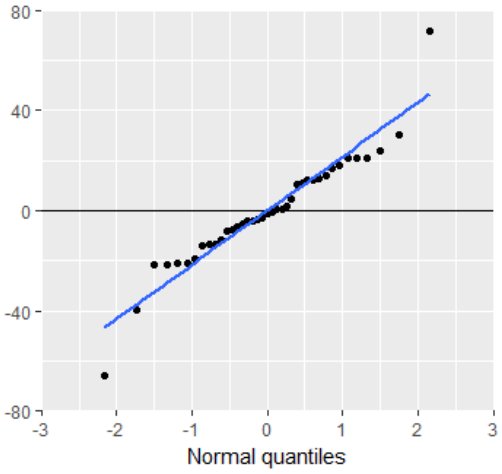
Leverage	Cook's D	DFFITS
0.154	0.194	0.453

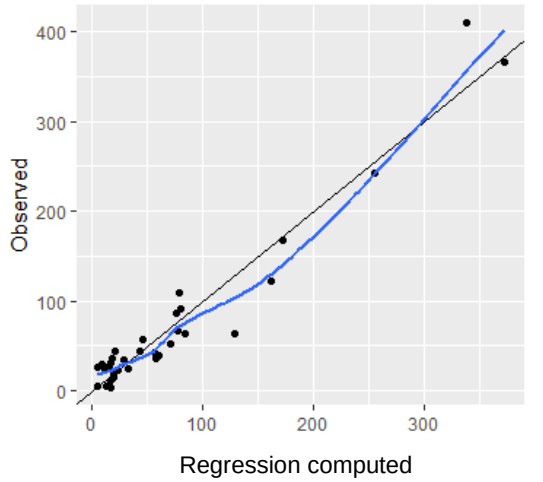
Flagged Observations (Observations that Exceed One of the Test Criteria, Helsel and Hirsch, 2002)

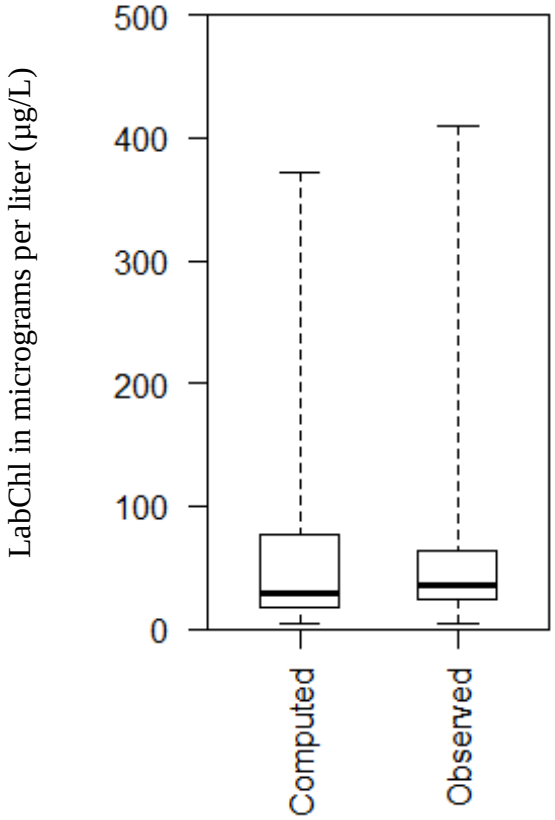
	LabChl	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
21	122.0	161	−39.90	−1.840	−1.900	0.0579	0.1040	−0.471
22	63.2	129	−65.70	−2.990	−3.390	0.0397	0.1850	−0.690
23	410.0	338	71.70	3.790	4.780	0.2880	2.9100	3.040
24	242.0	255	−13.10	−0.637	−0.632	0.1530	0.0367	−0.269
28	366.0	372	−6.37	−0.355	−0.351	0.3580	0.0351	−0.262

Statistical Plots (LabChl, in micrograms per liter)

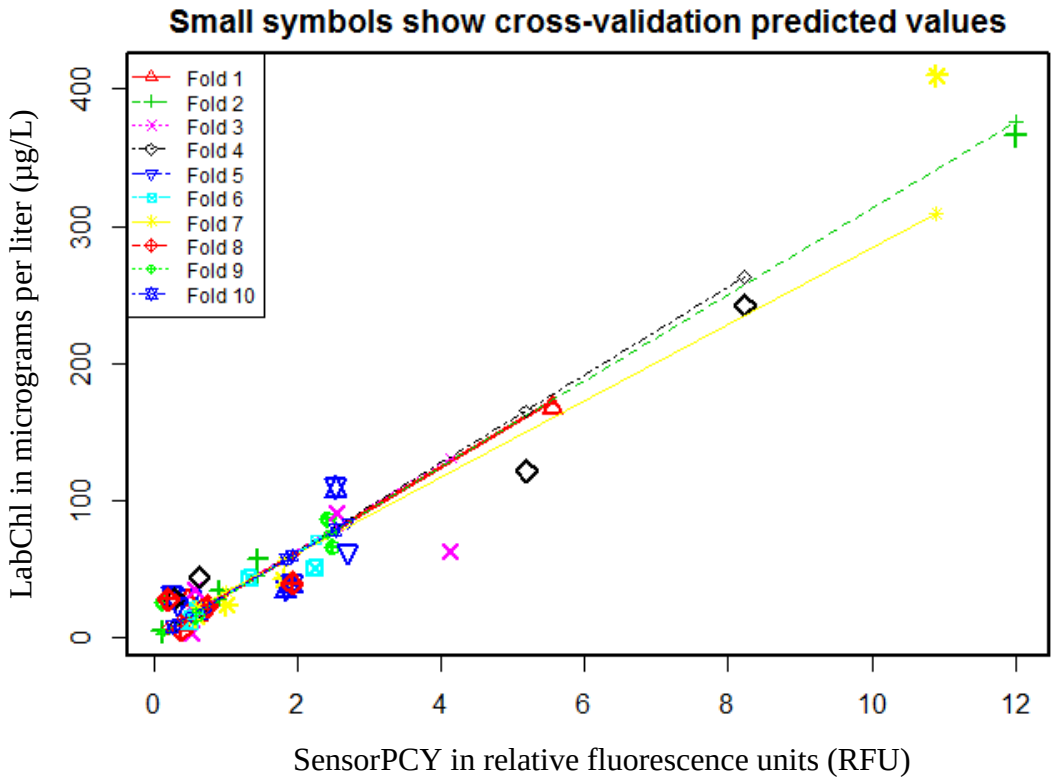




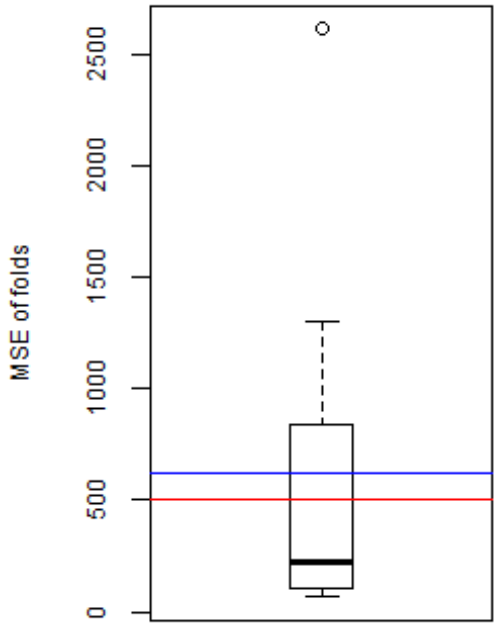



LabChl in micrograms per liter (µg/L)
500
400
300
200
100
0
Computed
Observed

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:    65.80

                Mean MSE of folds:   620.00

              Median MSE of folds:   226.00

             Maximum MSE of folds:  2610.00

(Mean MSE of folds) / (Model MSE):     1.24
```




```
Red line—Model MSE

Blue line—Mean MSE of folds
```

Model-Calibration Dataset

	Date	Time	LabChl	SensorPCY	Computed LabChl	Residual	Normal Quantiles	Censored Values
1	5/26/2016	0910	30.7	0.51	16.6	14.1	0.774	--
2	5/26/2016	0930	27.7	0.47	15.4	12.4	0.612	--
3	5/26/2016	1020	30.2	0.26	8.88	21.3	1.33	--
4	5/26/2016	1130	3.75	0.53	17.2	-13.5	-0.774	--
5	5/26/2016	1150	4.98	0.12	4.55	0.43	0.128	--
6	5/26/2016	1200	5.58	0.39	12.9	-7.32	-0.464	--
7	5/26/2016	1250	25.7	0.12	4.55	21.1	1.07	--
8	5/26/2016	1340	28.4	0.3	10.1	18.3	0.961	--
9	6/9/2016	0940	24.3	0.34	11.4	12.9	0.691	--

10	6/9/2016	1010	28.3	0.2	7.02	21.2	1.19	--
11	7/14/2016	1745	36.3	1.84	57.8	−21.5	−1.33	--
12	7/14/2016	1720	42.7	1.81	56.8	−14.1	−0.864	--
13	7/14/2016	1715	39.4	1.93	60.5	−21.1	−1.19	--
14	7/14/2016	1705	39.7	1.93	60.5	−20.8	−1.07	--
15	7/14/2016	1630	12	0.5	16.3	−4.31	−0.259	--
16	7/14/2016	1540	16.9	0.61	19.7	−2.81	−0.064	--
17	7/14/2016	1535	15.2	0.6	19.4	−4.2	−0.193	--
18	7/14/2016	1415	86.6	2.43	76	10.6	0.394	--
19	7/14/2016	1350	110	2.53	79.1	30.5	1.74	--
20	7/14/2016	1345	91.3	2.56	80	11.3	0.464	--
21	7/14/2016	1235	122	5.19	161	−39.9	−1.74	--
22	7/14/2016	1230	63.2	4.14	129	−65.7	−2.16	--
23	7/21/2016	0850	410	10.9	338	71.7	2.16	--
24	7/21/2016	0930	242	8.23	255	−13.1	−0.691	--
25	7/21/2016	1030	66.5	2.49	77.9	−11.4	−0.612	--
26	7/21/2016	1050	51.6	2.26	70.8	−19.2	−0.961	--
27	7/21/2016	1210	167	5.55	173	−5.23	−0.325	--
28	7/21/2016	1420	366	12	372	−6.37	−0.394	--
29	7/21/2016	1450	62.9	2.71	84.7	−21.8	−1.5	--
30	9/15/2016	0900	35.5	0.57	18.5	17	0.864	--
31	9/15/2016	0940	34.1	0.92	29.3	4.8	0.325	--
32	9/15/2016	1020	44.3	1.35	42.6	1.7	0.259	--
33	9/15/2016	1050	57.8	1.45	45.7	12.1	0.536	--
34	9/15/2016	1100	22.4	0.73	23.4	−1.02	0	--
35	9/15/2016	1210	44.4	0.64	20.6	23.8	1.5	--

```
 36 9/15/2016     1310       20.7      0.62       20    0.682     0.193       --
 37 9/15/2016     1340         24      1.02     32.4    -8.39    -0.536       --
 38 9/15/2016     1410       23.5      0.75       24   -0.539     0.064       --
 39 9/15/2016     1430       15.8       0.6     19.4     -3.6    -0.128       --
-- = value was not censored
```

Definitions

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).

LabChl Chlorophyll, fluorometric method, uncorrected, micrograms per liter (NWIS parameter code 32217).

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).

SensorPCY in Phycocyanins (cyanobacteria), water, in situ, fluorometric method, excitation at 590 ± 15 nm, emission at 685 +-20 nm, relative fluorescence units (NWIS parameter code 32321).

***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): Washington D.C., 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, no. 368, p. 829–836. [Also available at https://doi.org/10.2307/2286407.]

Foster, G.M., Graham, J.L., Stiles, T.C., Boyer, M.G., King, L.R., and Loftin, K.A., 2017, Spatial variability of harmful algal blooms in Milford Lake, Kansas, July and August 2015: U.S. Geological Survey Scientific Investigations Report 2016–5168, 45 p., accessed December 4, 2017 at https://doi.org/10.3133/sir20165168.

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., accessed December 4, 2017 at https://pubs.usgs.gov/twri/twri4a3/.

Knowlton, M.F., 1984, Flow-through microcuvette for fluorometric determination of chlorophyll—Water resources bulletin: Journal of the American Water Resources Association, v. 20, no. 5, p. 1198–1205. [Also available at https://doi.org/10.1111/j.1752-1688.1984.tb04763.x.]

Sartory, D.P., and Grobbelaar, J.U., 1984, Extraction of chlorophyll *a* from freshwater phytoplankton for spectrophotometric analysis: Hydrobiologia, v. 114, no. 3, p. 177–187.

U.S. Geological Survey, 2018, National Water Information System—Web interface: accessed December 4, 2017 at https://doi.org/10.5066/F7P55KJN.

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, accessed December 4, 2017 at https://pubs.water.usgs.gov/twri9A.

Appendix 2. Model Archival Summary for Total Microcystin Concentration at Milford Lake, May 26, June 9, July 14, July 21, and September 15, 2016

This model archival summary summarizes the laboratory-measured total microcystin concentration (LabMC) model developed to estimate LabMC concentrations at Milford Lake on May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016. This model is specific to data collected for the purposes of this study alone and cannot be reliably applied to other data collected from Milford Lake for other studies or times, or data collected from other lakes. Model statistics and plots were developed using an internal U.S. Geological Survey R application for producing model archive summaries accessed on November 15, 2017.

Site and Model Information

Site name.—Milford Lake, Kansas

Equipment.—A Yellow Springs Instrument, Inc., EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, pH, turbidity, chlorophyll fluorescence, phycocyanin fluorescence, and fluorescent dissolved organic matter was mounted under a boat at a 0.5-meter (m) depth for spatial surveys completed on Milford Lake on May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016. Boat speed was about 14 kilometers per hour, which provided the best balance of data quality and the ability to complete multiple representative surveys of Zone C of Milford Lake (Foster and others, 2018, fig. 1) in a timely manner. Readings from the water-quality monitor were recorded every 30 seconds. Phycocyanin was used as the single explanatory variable in the model because that model explained the most variance in LabMC, and it is consistent with the model developed by Foster and others

(2017). Discrete water-quality samples for LabMC analysis were collected at multiple locations throughout Zone C of Milford Lake (Foster and others, 2018, table 1).

Date model was created.—January 11, 2017

Model calibration data period.—May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model application date.—May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model-Calibration Dataset

All data were collected using U.S. Geological Survey protocols (U.S. Geological Survey, variously dated) and are stored in the National Water Information System database (U.S. Geological Survey, 2018). The explanatory variable selected as input to the linear regression was phycocyanin, in relative fluorescence units (RFU). Because most discrete samples were collected at the depth of the monitor (0.5 m), the linear relation between sensor-measured phycocyanin RFU and laboratory-measured total microcystin could be used to compute LabMC concentrations in micrograms per liter for the spatial survey. The linear regression model was developed using the open-source software package R (version 3.2.3).

The regression model is based on 39 concurrent measurements of sensor-measured phycocyanin and LabMC collected on May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016. A total of eight samples were below laboratory-detection limit (<0.10 microgram per liter [µg/L]) and were replaced with 0.05 µg/L for model development. Summary statistics and the complete model-calibration dataset are provided in this appendix. A total of three samples, collected on July 21, 2016, at 7:30 a.m., 11:40 a.m., and 1:00 p.m., were excluded from the dataset used to develop the regression model because the samples were collected from near-shore areas with dense surface accumulations, which are not representative of typical conditions throughout Zone C of Milford Lake. Studentized residuals from the final model were inspected for values greater than 3 or less than −3. Flagged observations were considered potential outliers and were

investigated. None of the samples in the flagged observations dataset were deemed outliers or removed from the model calibration dataset.

Total Microcystin Sampling

Most (about 80 percent) total microcystin samples for laboratory analysis were collected at a 0.5-m depth (the depth of the monitor) from open-water locations. Some samples (*n*=8) collected during July 14, 2016, were integrated from the surface to 0.5 m; because these samples did not have undue influence on the model and were not flagged as potential outliers, they were retained in the dataset. Sample locations were not predetermined and were selected to represent the range of cyanobacterial conditions in the lake based on visual cues and continuous water-quality monitor data. Samples were analyzed for total microcystin concentration using enzyme-linked immunosorbent assay (ELISA) at the U.S. Geological Survey Organic Geochemistry Research Laboratory as described in Foster and others (2017). Additional detail on sample collection is available in the "Methods" section of the report.

Model Development

Ordinary least squares regression analysis was done using R (version 3.2.3) with sensor-measured phycocyanin RFU as the explanatory variable for laboratory-measured total microcystin concentrations. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and computed values) as compared to computed LabMC concentrations were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of computed values). Values for all regression statistics and metrics are included in this appendix along with all relevant sample data and more in-depth statistical information.

The model is not sensitive at lower LabMC concentrations. When phycocyanin fluorescence is less than 0.74 RFU, the model outputs negative values. Various types of models and transformations, including approximate maximum likelihood estimation (AMLE) for censored data, were explored and did not

substantially improve model fit at the low end. Regression-estimated LabMC concentrations should be censored to exclude negative values as described in the "Methods" section of the report. The model developed for Foster and others (2018) was considered appropriate to meet study objectives; however, this model should not be used outside of the scope of Foster and others (2018).

Model Summary

The following is a summary of final regression analysis for sensor-measured phycocyanin RFU and laboratory-measured total microcystin at Milford Lake, May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016.

The LabMC concentration model is represented by the following:

$$LabMC = +8.60 * SensorPCY - 6.38$$

where

LabMC is laboratory-measured total microcystin in micrograms per liter (µg/L) and

SensorPCY is sensor-measured phycocyanin in RFU.

R Output for the Relation Between Sensor-Measured Phycocyanin Relative Fluorescence Units and Laboratory-Measured Total Microcystin at Milford Lake, May 26, 2016; June 9, 2016; July 14, 2016; July 21, 2016; and September 15, 2016

Model Statistics, Data, and Plots

Definitions for terms used in this output are included at the end of this document.

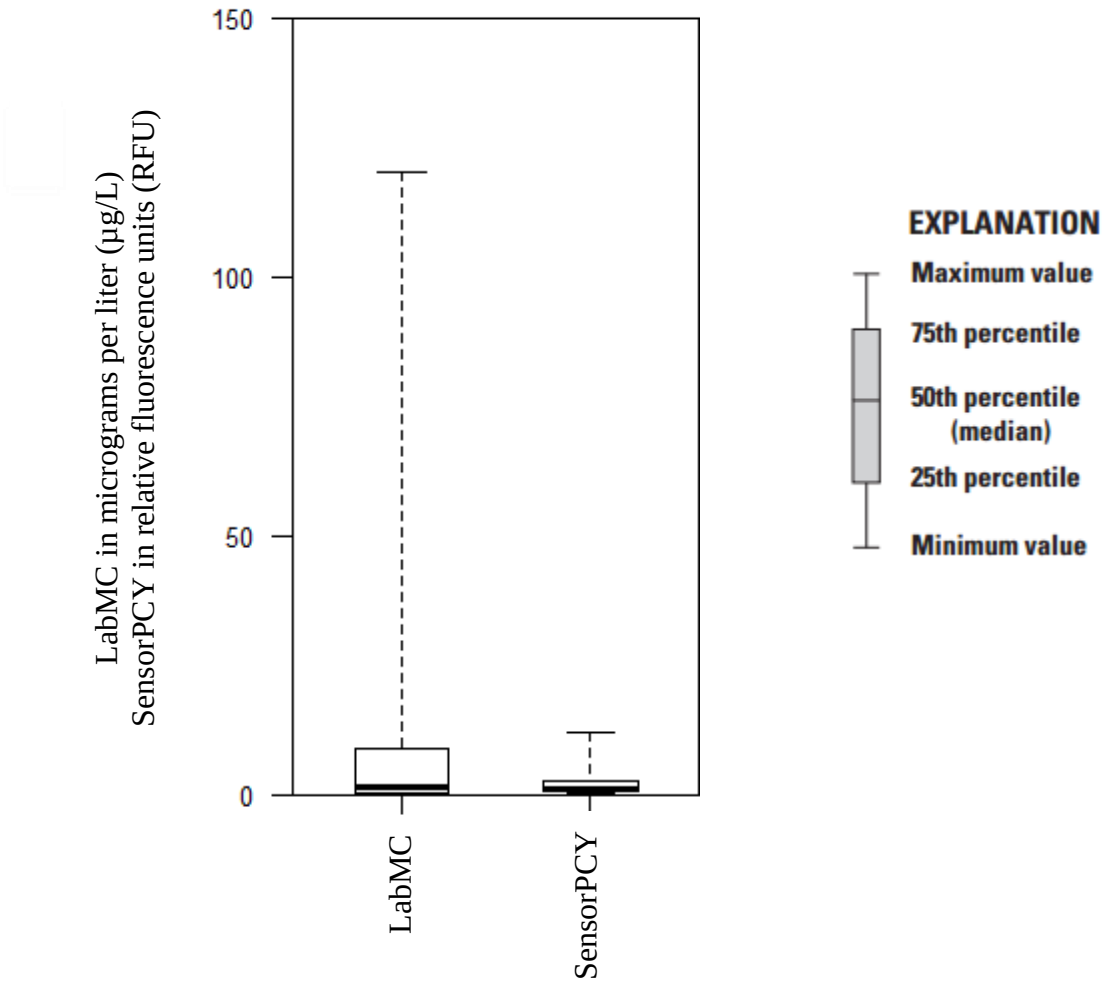
Model

$$LabMC = +8.6 * SensorPCY - 6.38$$

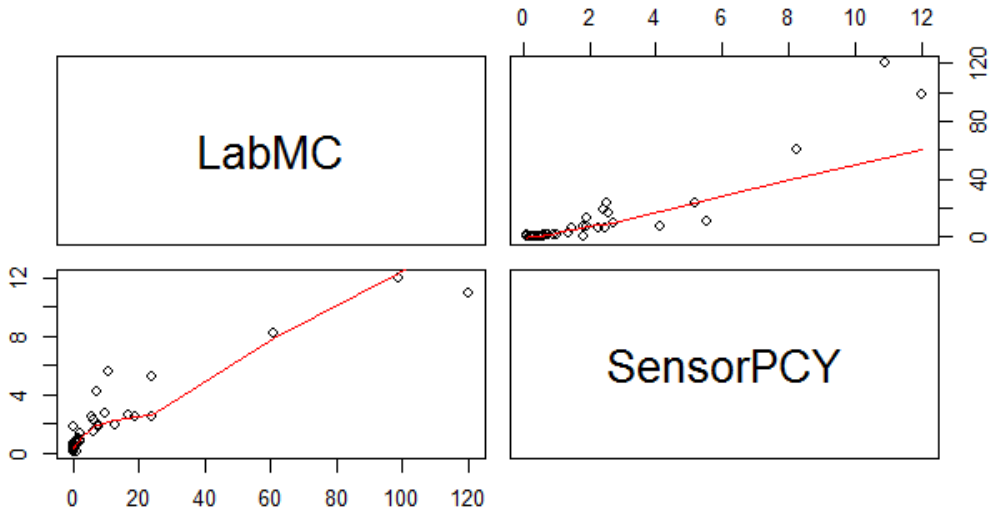
Variable Summary Statistics

	LabMC	SensorPCY
Minimum	0.05	0.12
1st Quartile	0.26	0.51
Median	1.50	0.92
Mean	11.70	2.11
3rd Quartile	9.60	2.49
Maximum	120.00	12.00

Box Plots



Exploratory Plots



Red line shows the locally weighted scatterplot smoothing (LOWESS); LabChl, in micrograms per liter; SensorPCY, in relative fluorescence units.

Basic Model Statistics

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

Number of Observations	39
Standard error (RMSE)	9.41
Upper Model standard percentage error (MSPE)	80.3
Lower Model standard percentage error (MSPE)	80.3
Coefficient of determination (R^2)	0.869
Adjusted Coefficient of Determination (Adj. R^2)	0.866

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-6.38	1.900	−3.36	1.83e−03
SensorPCY	8.60	0.548	15.70	6.33e−18

Correlation Matrix

	Intercept	SensorPCY
Intercept	1.000	−0.608
SensorPCY	-0.608	1.000

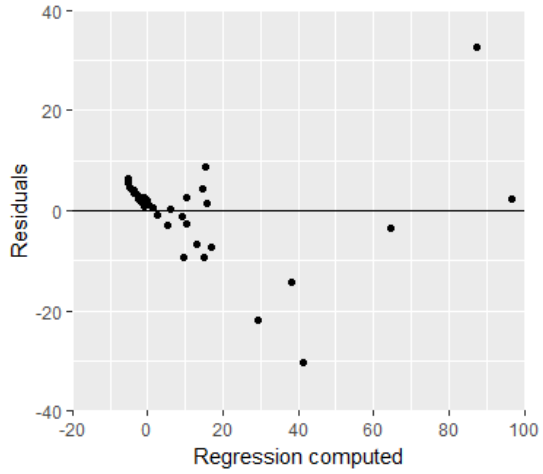
Outlier Test Criteria

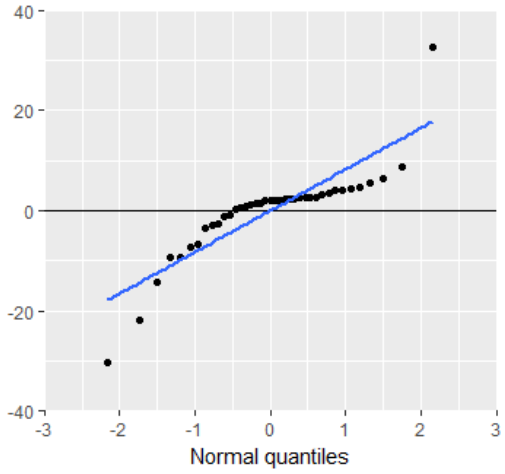
Leverage	Cook's D	DFFITS
0.154	0.194	0.453

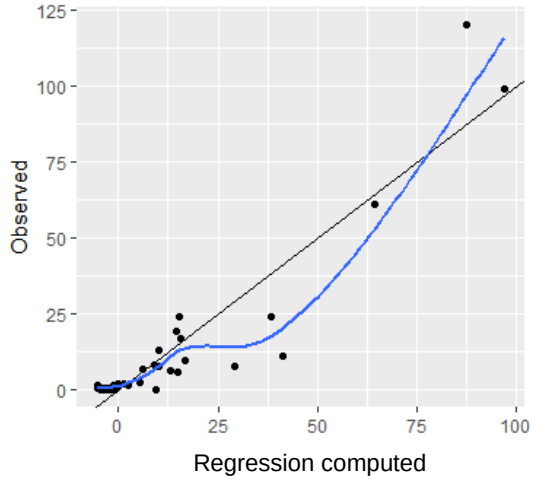
Flagged Observations (Observations that Exceed One of the Test Criteria, Helsel and Hirsch, 2002)

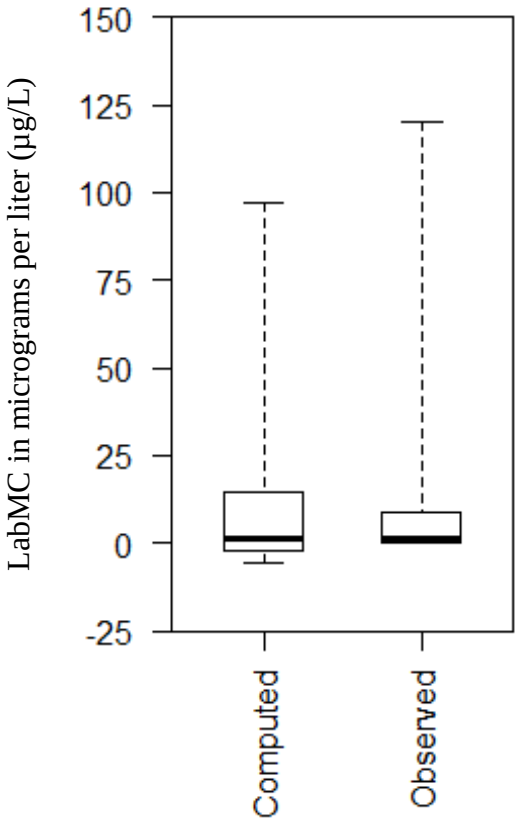
	LabMC	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
22	7.5	29.2	−21.70	−2.350	−2.52	0.0397	0.114	−0.512
23	120.0	87.3	32.70	4.110	5.51	0.2880	3.420	3.500
27	11.0	41.3	−30.30	−3.330	−3.93	0.0659	0.392	−1.040
28	99.0	96.8	2.21	0.293	0.29	0.3580	0.024	0.216

Statistical Plots (LabMC, in micrograms per liter)

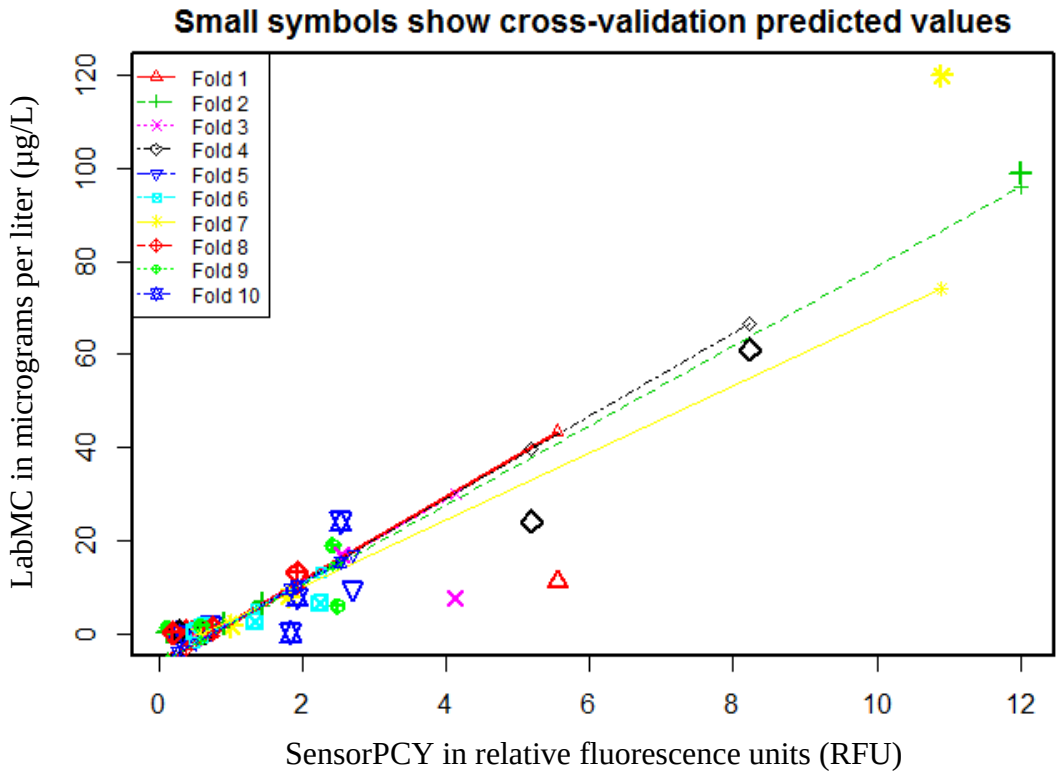




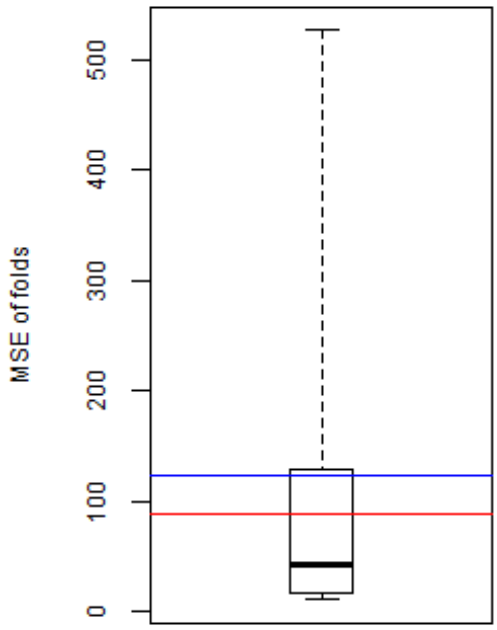



LabMC in micrograms per liter (µg/L)
150
125
100
75
50
25
0
-25
Computed
Observed

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:   10.70

                Mean MSE of folds:  123.00

              Median MSE of folds:   43.20

             Maximum MSE of folds:  526.00

(Mean MSE of folds) / (Model MSE):    1.39
```




```
Red line—Model MSE

Blue line—Mean MSE of folds
```

Model-Calibration Dataset

	Date	Time	LabMC	SensorPCY	Computed LabMC	Residual	Normal Quantiles	Censored Values
1	5/26/2016	0910	0.05	0.51	-1.99	2.04	-0.064	X
2	5/26/2016	0930	0.05	0.47	-2.34	2.39	0.325	X
3	5/26/2016	1020	0.05	0.26	-4.14	4.19	0.864	X
4	5/26/2016	1130	0.36	0.53	-1.82	2.18	0.128	--
5	5/26/2016	1150	0.22	0.12	-5.34	5.56	1.33	--
6	5/26/2016	1200	0.26	0.39	-3.02	3.28	0.691	--
7	5/26/2016	1250	1.2	0.12	-5.34	6.54	1.5	--
8	5/26/2016	1340	0.43	0.3	-3.8	4.23	0.961	--

9	6/9/2016	0940	0.05	0.34	−3.45	3.5	0.774	X
10	6/9/2016	1010	0.05	0.2	−4.66	4.71	1.19	X
11	7/14/2016	1745	0.05	1.84	9.44	−9.39	−1.33	X
12	7/14/2016	1720	7.9	1.81	9.18	−1.28	−0.612	--
13	7/14/2016	1715	7.7	1.93	10.2	−2.52	−0.691	--
14	7/14/2016	1705	13	1.93	10.2	2.78	0.612	--
15	7/14/2016	1630	0.49	0.5	−2.08	2.57	0.464	--
16	7/14/2016	1540	1.1	0.61	−1.13	2.23	0.259	--
17	7/14/2016	1535	0.96	0.6	−1.22	2.18	0.064	--
18	7/14/2016	1415	19	2.43	14.5	4.49	1.07	--
19	7/14/2016	1350	24	2.53	15.4	8.63	1.74	--
20	7/14/2016	1345	17	2.56	15.6	1.37	−0.193	--
21	7/14/2016	1235	24	5.19	38.2	-14.2	−1.5	--
22	7/14/2016	1230	7.5	4.14	29.2	−21.7	−1.74	--
23	7/21/2016	0850	120	10.9	87.3	32.7	2.16	--
24	7/21/2016	0930	61	8.23	64.4	−3.38	−0.864	--
25	7/21/2016	1030	5.8	2.49	15	−9.23	−1.19	--
26	7/21/2016	1050	6.4	2.26	13.1	−6.65	−0.961	--
27	7/21/2016	1210	11	5.55	41.3	−30.3	−2.16	--
28	7/21/2016	1420	99	12	96.8	2.21	0.193	--
29	7/21/2016	1450	9.6	2.71	16.9	−7.32	−1.07	--
30	9/15/2016	0900	0.05	0.57	−1.48	1.53	−0.128	X
31	9/15/2016	0940	2.1	0.92	1.53	0.567	−0.394	--
32	9/15/2016	1020	2.4	1.35	5.23	−2.83	−0.774	--
33	9/15/2016	1050	6.5	1.45	6.09	0.411	−0.464	--
34	9/15/2016	1100	2	0.73	−0.101	2.1	0	--

```
 35 9/15/2016      1210        0.05      0.64   -0.875    0.925    -0.325       X
 36 9/15/2016      1310         1.5      0.62    -1.05     2.55     0.394       --
 37 9/15/2016      1340         1.5      1.02     2.39   -0.892    -0.536       --
 38 9/15/2016      1410         1.3      0.75   0.0711     1.23    -0.259       --
 39 9/15/2016      1430         1.5       0.6    -1.22     2.72     0.536       --
-- = value was not censored
X = value was censored
```

Definitions

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).

LabMC Total microcystins plus nodularins, unfiltered water, freeze/thaw extraction, ADDA specific enzyme-linked immunosorbent assay, recoverable, micrograms per liter (NWIS parameter code 89011).

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).

SensorPCY in Phycocyanins (cyanobacteria), water, in situ, fluorometric method, excitation at 590 ±15 nm, emission at 685 ±20 nm, relative fluorescence units (NWIS parameter code 32321).

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

References Cited

Cleveland, W.S., 1979, Robust locally weighted regression and smoothing scatterplots: Journal of the American Statistical Association, v. 74, no. 368, p. 829–836. [Also available at https://doi.org/10.2307/2286407.]

Foster, G.M., Graham, J.L., Stiles, T.C., Boyer, M.G., King, L.R., and Loftin, K.A., 2017, Spatial variability of harmful algal blooms in Milford Lake, Kansas, July and August 2015: U.S. Geological Survey Scientific Investigations Report 2016–5168, 45 p., accessed December 4, 2017 at https://doi.org/10.3133/sir20165168.

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., accessed December 4, 2017 at https://pubs.usgs.gov/twri/twri4a3/.

U.S. Geological Survey, 2018, National Water Information System—Web interface: accessed December 4, 2017 at https://doi.org/10.5066/F7P55KJN.

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, accessed December 4, 2017 at https://pubs.water.usgs.gov/twri9A.