Appendix 2. Model Archive Summary for Total Nitrogen at Mill Creek at Johnson Drive, Shawnee, Kansas, 2015–18.

This model archive summary summarizes the total nitrogen (TN) model developed to compute 15-minute TN from January 1, 2015, to December 31, 2018. This model supersedes the previous model used from 2003 to 2009 (Rasmussen and others, 2008).

Site and Model Information

Site number: 06892513

Site name: Mill Creek at Johnson Drive, Shawnee, Kansas

Location: Latitude 39°01′45″, longitude 94°49′02″ referenced to North American Datum of 1983, in Johnson County, Kansas, Hydrologic Unit Code 10270104.

Equipment: A Yellow Springs Instruments (YSI) EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, turbidity, and pH, and a Hach Nitratax *plus* sc. The EXO2 was housed in a 4-inch metal pipe, and the Nitratax was housed in a 3-inch PVC pipe. Readings from the EXO2 and Nitratax were recorded every 15 minutes and transmitted by way of satellite hourly.

Date model was created: March 18, 2019

Model calibration data period: June 5, 2015, to October 7, 2018

Model application date: January 1, 2015, to December 31, 2018

Sampling Details

Equal-width-increment samples were collected from the downstream side of the bridge. Samples were obtained at least monthly, with the priority being during storm-runoff events, with a FISP US DH-95 depth-integrating sampler with a Teflon bottle, cap, and nozzle. Samples were analyzed at the U.S. Geological Survey (USGS) National Water Quality Laboratory (NWQL) in Lakewood, Colorado, and at the Johnson County Water Quality Laboratory (JCWQL) in Olathe, Kansas.

Model Data

All data were collected using USGS protocols and are stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019). The regression model is based on 32 concurrent measurements of total nitrogen, turbidity, dissolved nitrate plus nitrite, and total Kjeldahl nitrogen (TKN, also known as total ammonia plus organic nitrogen) collected from June 5, 2015, through October 7, 2018. Samples were collected throughout the range of continuously observed hydrologic conditions. Total nitrogen was calculated as the sum of two components: dissolved nitrate plus nitrite and TKN. The NWQL reporting limits were 0.04 milligram per liter as nitrogen (mg/L–N) for dissolved nitrate plus nitrite and 0.07 mg/L–N for TKN. The JCWQL reporting limits were 0.02 mg/L–N for dissolved nitrate plus nitrite and 0.5 mg/L–N for TKN. Zero samples had concentrations that were less than laboratory reporting limits for either component. Summary statistics and the complete model-calibration dataset are provided below. Potential outliers were identified as the data points for which the studentized residuals were greater than 3 or less than negative

3, as described by Helsel and others (2020). Values outside of that range were considered potential outliers and investigated. None of the studentized residuals for TN samples were beyond ±3.

Model Development

All continuously measured water-quality parameters and streamflow were considered as explanatory variables for estimating total nitrogen using ordinary least squares regression. A variety of models that predict TN and models that predict $\log_{10}$(TN) were evaluated. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and predicted values) compared to predicted TN were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of predicted values). This comparison led to the conclusion that the most appropriate and reliable model would be one that estimated nontransformed TN.

Turbidity and nitrate plus nitrite (NO_x) were selected as the best predictors of TN based on residual plots, relatively high adjusted coefficient of determination (adjusted R^2), and relatively low model standard percentage error (*MSPE*). Values for the aforementioned statistics and metrics were computed and are included below along with all relevant sample data and more in-depth statistical information.

Model Summary

Summary of final regression analysis for total nitrogen at USGS site number 06892513.

Total nitrogen model:

$$TN = 0.00594 \times TBY + 1.26 \times NO_x + 0.0759$$

where

TN = total nitrogen in milligrams per liter as nitrogen (mg/L–N);

TBY = turbidity, YSI EXO2, in formazin nephelometric units (FNU); and,

NO_x = sensor-measured nitrate (NO_3) plus nitrite (NO_2), in milligrams per liter as nitrogen (mg/L–N).

Turbidity and NO_x make physical and statistical sense as explanatory variables for total nitrogen. Turbidity makes sense physically because suspended solids (including some with attached TN) in the water column scatter light and increase turbidity. NO_x makes sense physically as it is a measurement of the inorganic component of total nitrogen. The relation between turbidity, NO_x, and TN makes statistical sense as the resulting model has the lowest standard error and highest adjusted R^2 values.

Previous Model

Model	Start year	End year	Model
1.0	2003	2009	$\log_{10}(TN) = 0.0006 \times \log_{10}(TBY) + 0.0483 \times \sin(\frac{2\pi D}{365}) + 0.2205 \times \cos(\frac{2\pi D}{365}) + 0.296$

where (Rasmussen and others, 2008)

TN = total nitrogen in milligrams per liter as nitrogen (mg/L–N);

TBY = turbidity, YSI model 6136, in formazin nephelometric units (FNU); and,

D = day of year, in the range of integers 1 through 365.

Total Nitrogen Record

The TN record is computed using this regression model and stored at the National Real-Time Water Quality (NRTWQ) website. Data are computed at hourly intervals. The complete water-quality record can be found at https://nrtwq.usgs.gov/ks.

Remarks

None

Computed by: Patrick Eslick

Reviewed by: Brian Klager

Model Statistics, Data, and Plots

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

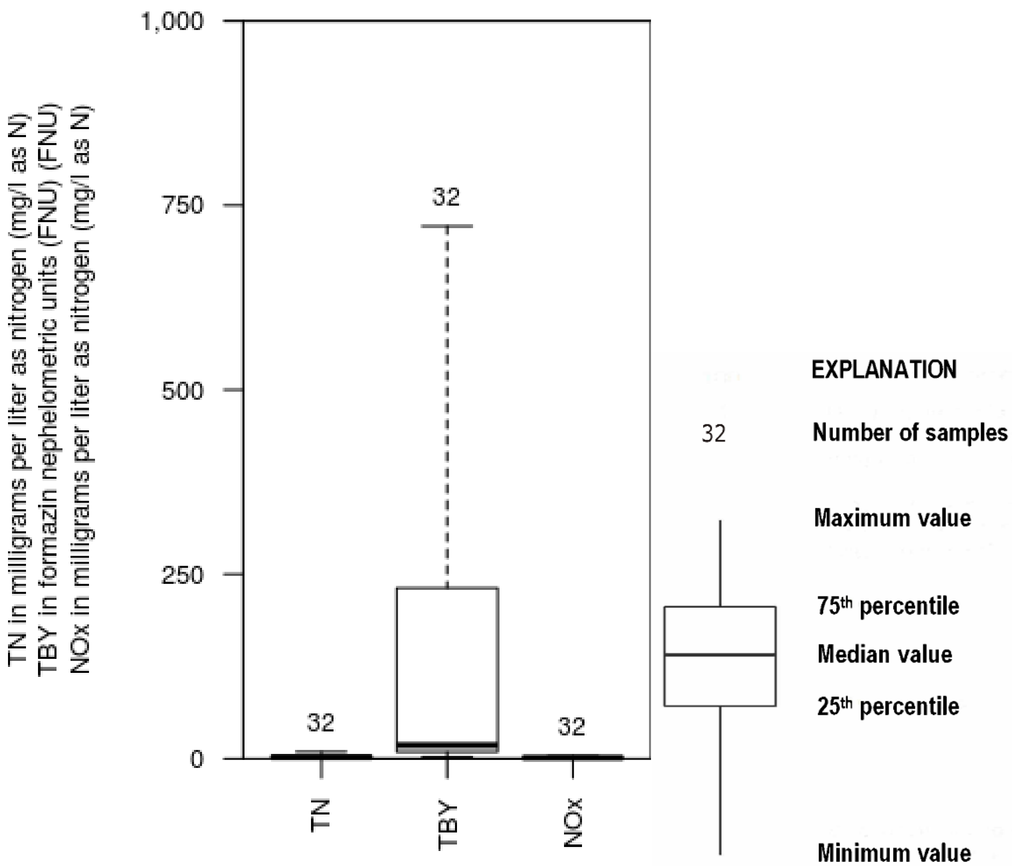
Model

$TN = + 0.00594 * TBY + 1.26 * NO_x + 0.0759$

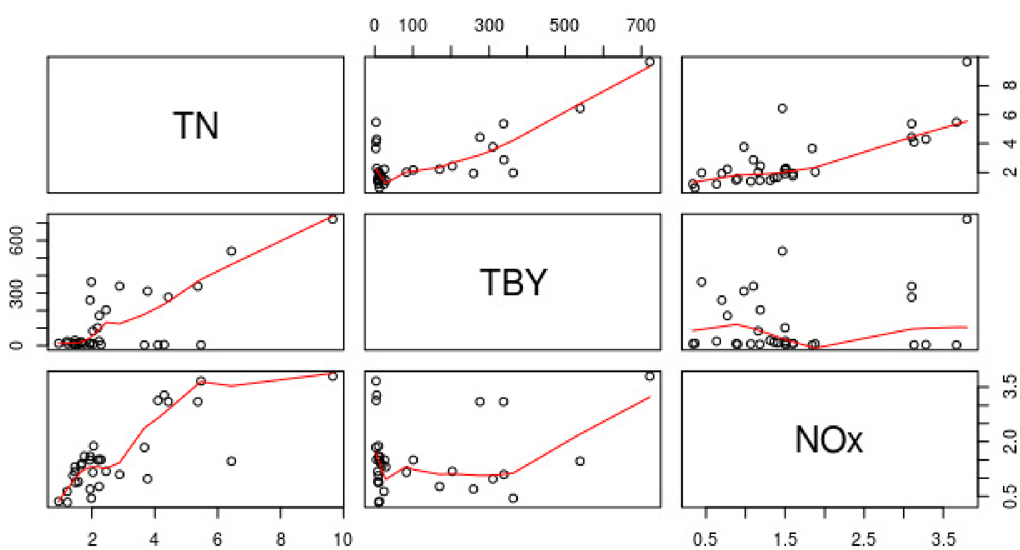
Variable Summary Statistics

```
                TN    TBY   NOx
Minimum      0.94   2.60 0.333
1st Quartile 1.61   9.25 0.940
Median       2.04  17.80 1.380
Mean         2.79 123.00 1.570
3rd Quartile 3.72 231.00 1.720
Maximum      9.66 722.00 3.800
```

Box Plots



Exploratory Plots



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

The x- and y-axis labels for a given bivariate plot are defined by the intersecting row and column labels.

Basic Model Statistics

```
Number of Observations                                  32
Standard error (RMSE)                                0.574
Average Model standard percentage error (MSPE)       20.6
Coefficient of determination (R²)                    0.912
Adjusted Coefficient of Determination (Adj. R²) 0.906

Variance Inflation Factors (VIF)
 TBY  NOx
1.05 1.05
```

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.0759	0.200	0.38	7.07e-01
TBY	0.00594	0.000583	10.20	4.37e-11
NOx	1.26	0.110	11.40	2.91e-12

Correlation Matrix

	Intercept	TBY	NOx
Intercept	1.000	-0.166	-0.787
TBY	-0.166	1.000	-0.223
NOx	-0.787	-0.223	1.000

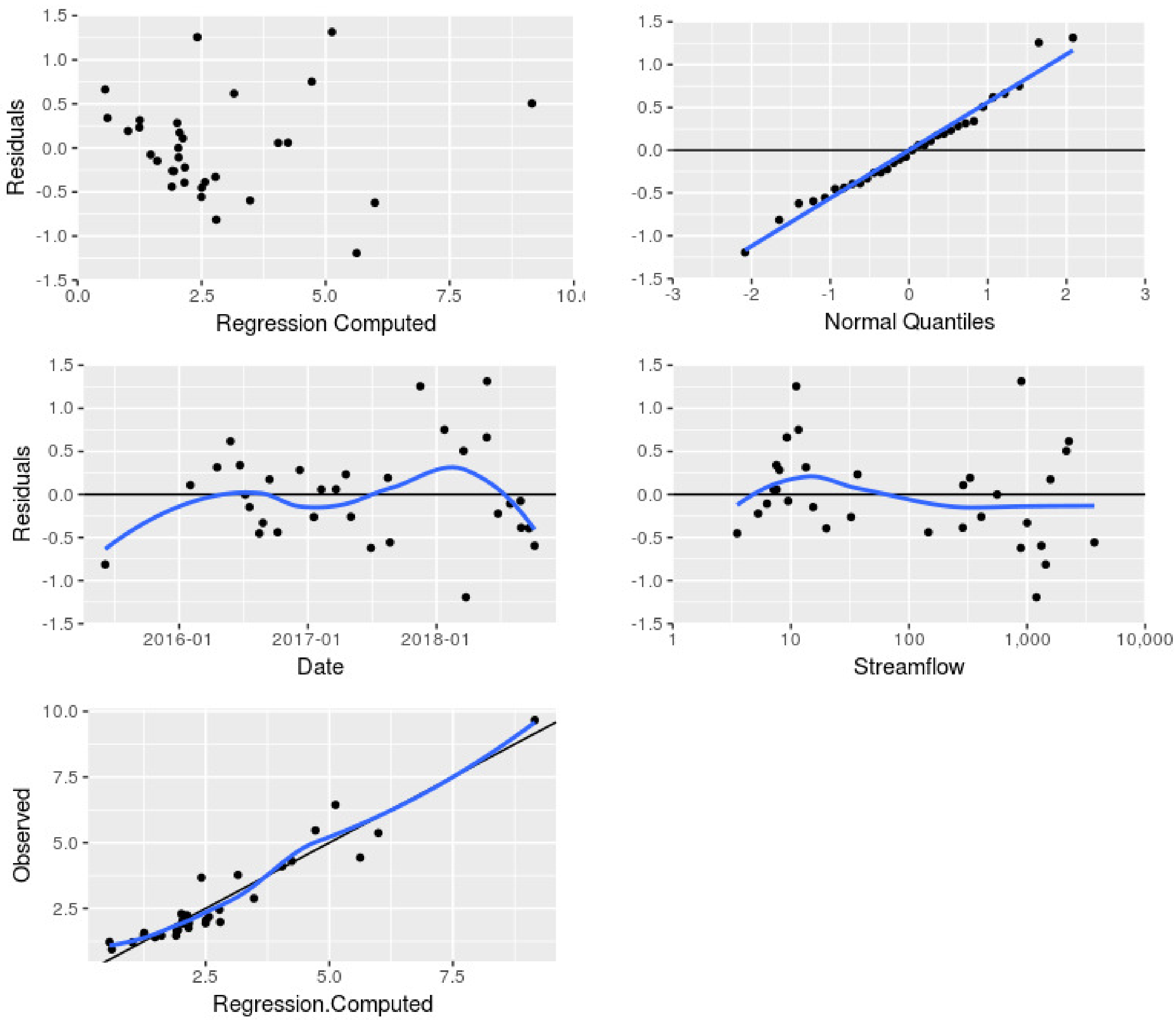
Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.281	0.262	0.612

Flagged Observations

	TN	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2015-06-05 10:10:00	1.98	2.79	-0.815	-1.55	-1.59	0.161	0.154	-0.697
2018-01-24 11:20:00	5.47	4.72	0.751	1.49	1.52	0.230	0.222	0.834
2018-03-19 13:00:00	9.66	9.16	0.505	1.21	1.22	0.469	0.429	1.140
2018-03-26 14:00:00	4.43	5.62	-1.190	-2.22	-2.39	0.121	0.227	-0.890
2018-05-25 10:50:00	6.44	5.13	1.310	2.58	2.89	0.214	0.605	1.510

Statistical Plots



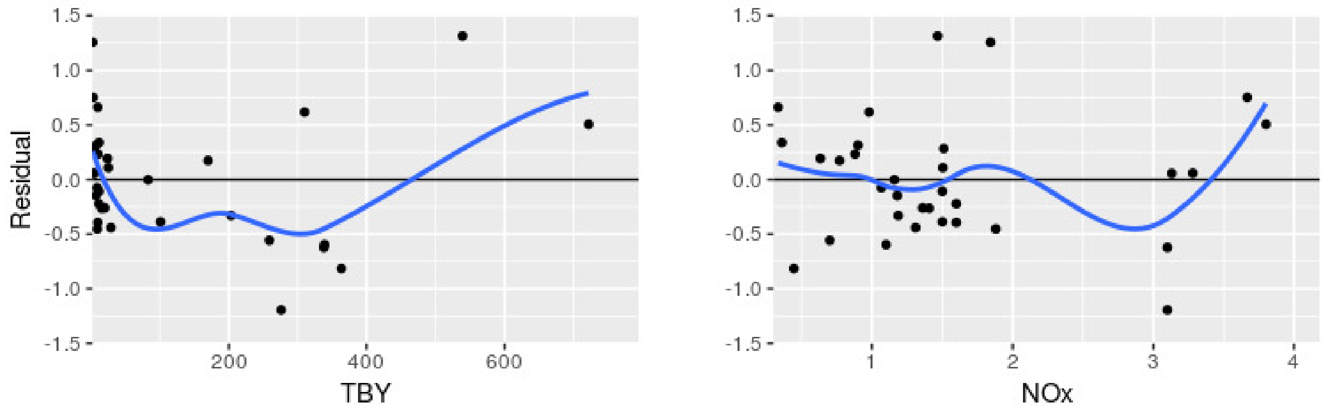
First row (left): residual TN (in mg/L–N) related to regression-computed TN (in mg/L–N).

First row (right): residual TN (in mg/L–N) related to the corresponding normal quantile (unitless) of the residual with simple linear regression indicated by the blue line.

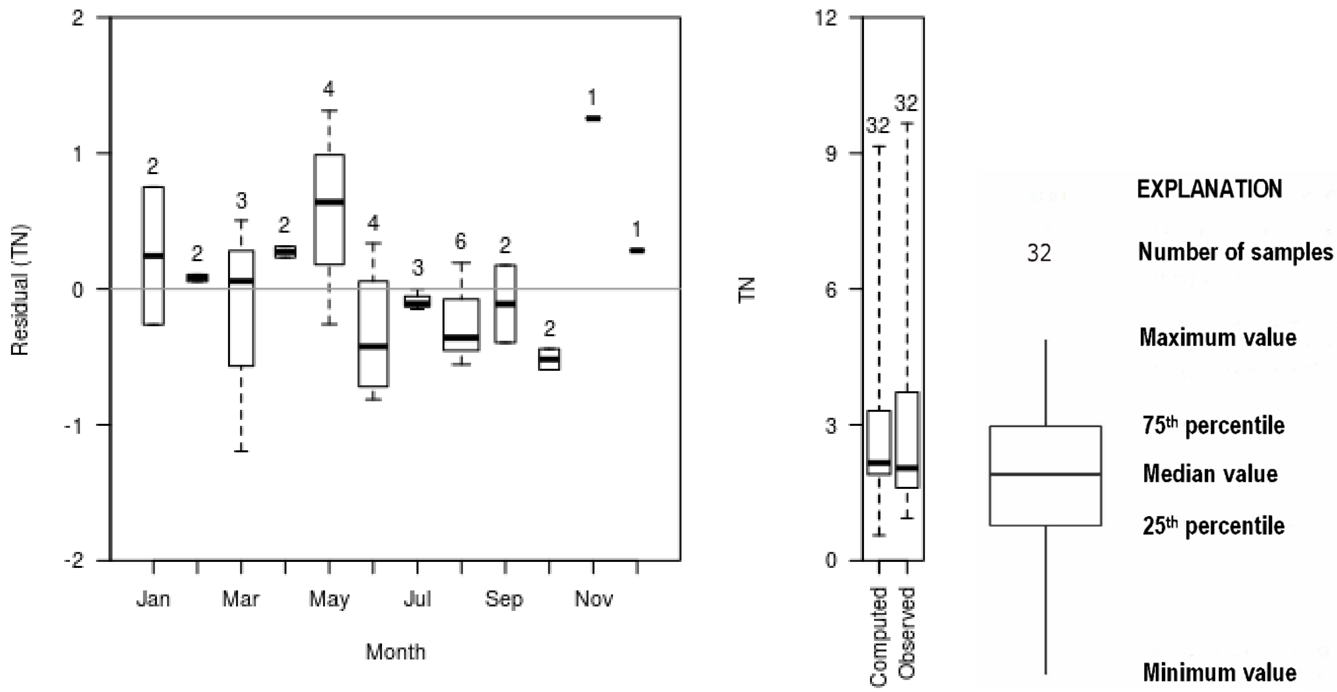
Second row (left): residual TN (in mg/L–N) related to date with local polynomial regression fitting, or locally estimated scatterplot smoothing (LOESS) indicated by the blue line.

Second row (right): residual TN (in mg/L–N) related to streamflow (in cubic feet per second) with LOESS indicated by the blue line.

Third row: observed TN (in mg/L–N) related to regression-computed TN (in mg/L–N) with LOESS indicated by the blue line.

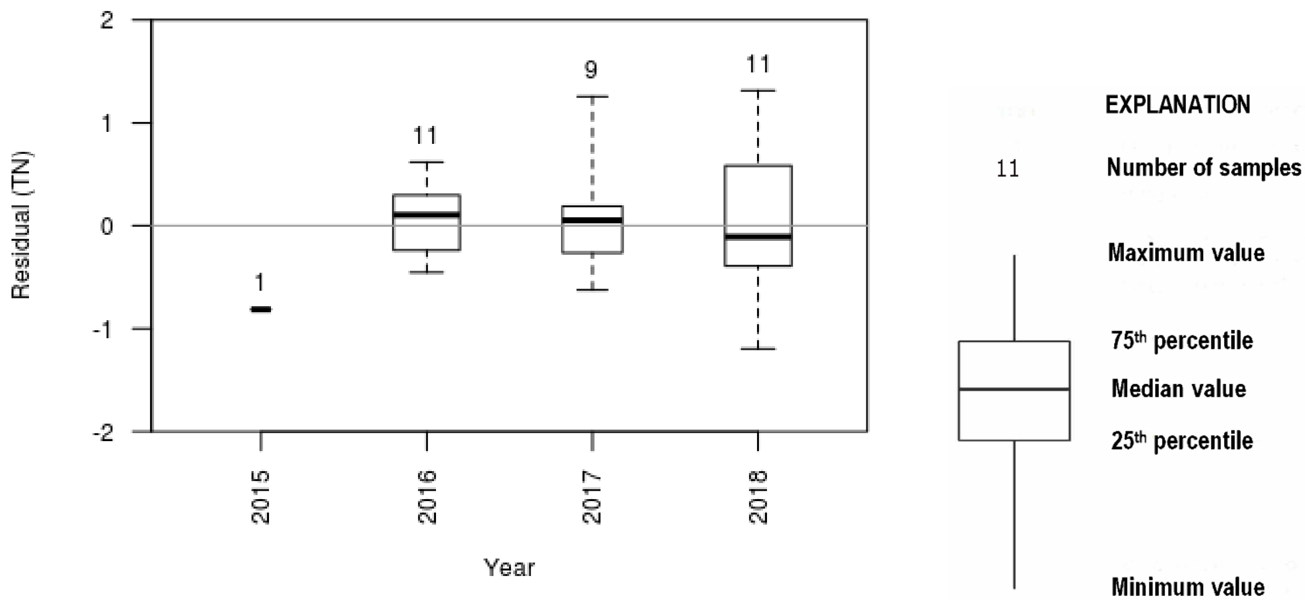


Residual TN (in mg/L–N) related to TBY (left, in FNU) and NOx (right, in mg/L–N) with LOESS indicated by the blue line.



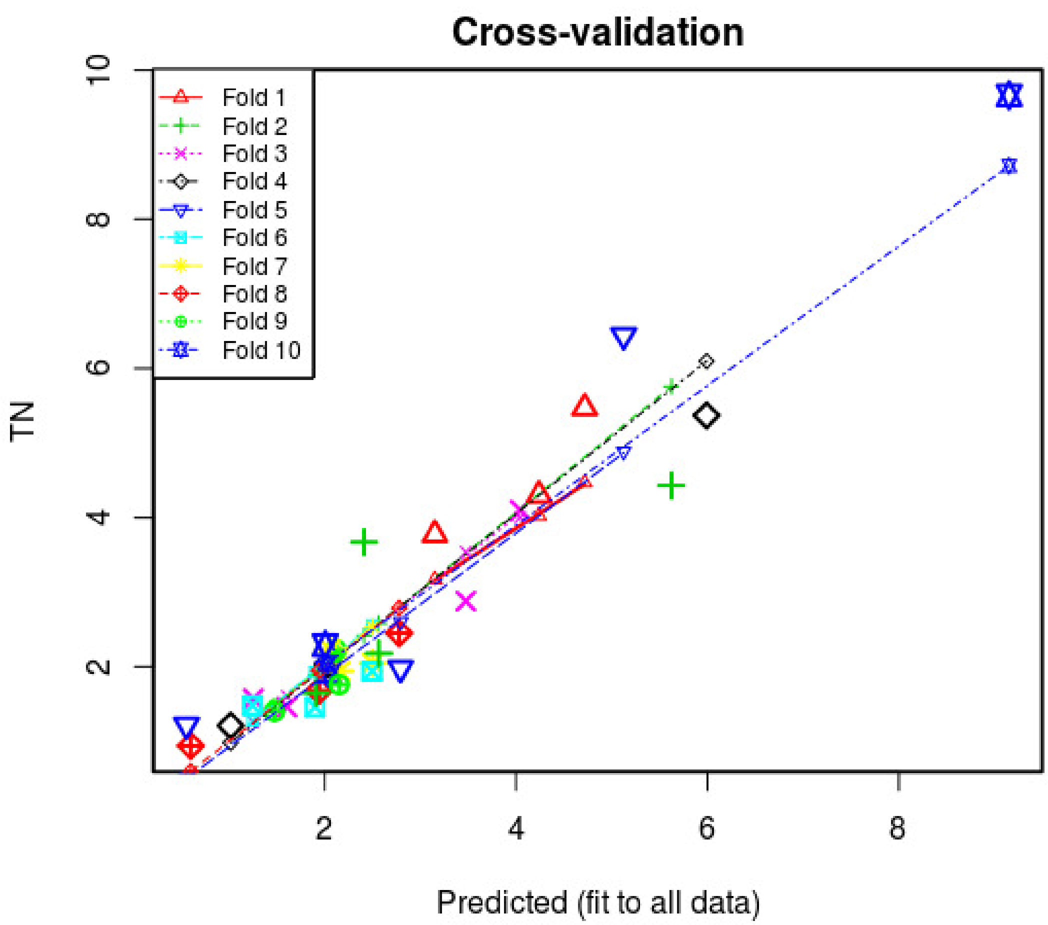
Left: residual TN (in mg/L–N) by month.

Right: TN (in mg/L–N) in regression-computed and observed values.



Residual TN (in mg/L–N) by year.

Cross Validation



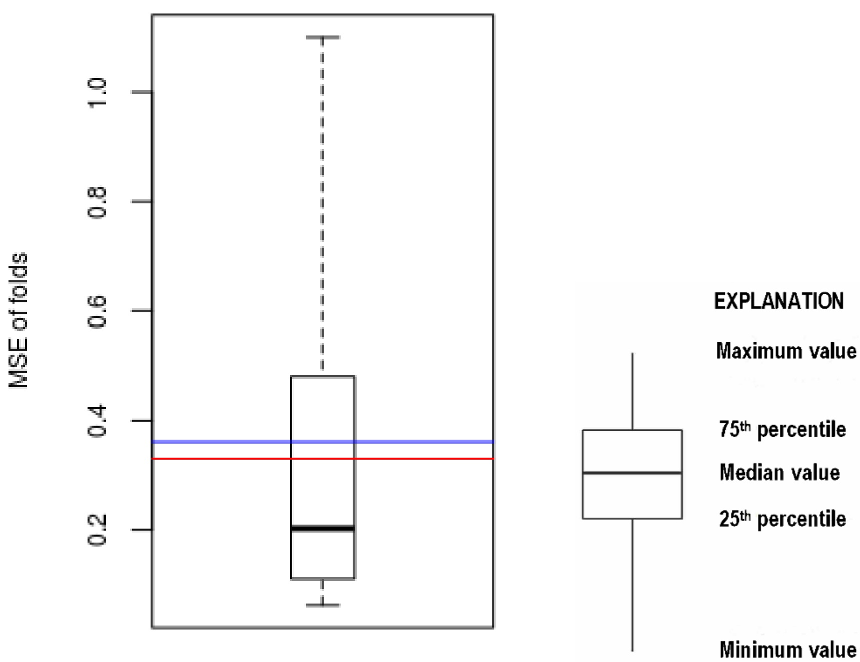
Fold: equal partition of the data (10 percent of the data)

Large symbols: observed values of data points removed in a fold

Small symbols: recomputed values of data points removed in a fold

Recomputed regression lines: adjusted regression lines with one fold removed

```
              Minimum MSE of folds:  0.0623
                 Mean MSE of folds:  0.3610
               Median MSE of folds:  0.2020
              Maximum MSE of folds:  1.1000
 (Mean MSE of folds) / (Model MSE):  1.1000
```



Red line - Model MSE (unitless)
Blue line - Mean MSE of folds (unitless)

Model-Calibration Data Set

	Date	TN	TBY	NOx	Computed TN	Residual	Normal Quantiles	Censored Values
1	2015-06-05	1.98	363	0.445	2.79	-0.815	-1.65	--
2	2016-02-02	2.23	25.3	1.5	2.12	0.108	0.276	--
3	2016-04-18	1.57	7.5	0.9	1.26	0.315	0.719	--
4	2016-05-26	3.77	310	0.98	3.15	0.618	1.07	--
5	2016-06-22	0.94	12	0.36	0.601	0.339	0.824	--
6	2016-07-07	2.03	83	1.16	2.03	-0.00155	0.0389	--
7	2016-07-19	1.46	7.3	1.18	1.61	-0.147	-0.196	--
8	2016-08-16	2.05	9.4	1.88	2.5	-0.452	-0.939	--
9	2016-08-26	2.45	203	1.19	2.78	-0.33	-0.53	--

10	2016-09-14	2.23	170	0.77	2.06	0.174	0.358	--
11	2016-10-07	1.46	29	1.31	1.9	-0.44	-0.824	--
12	2016-12-09	2.29	4.5	1.51	2.01	0.283	0.622	--
13	2017-01-18	1.68	15.7	1.41	1.94	-0.263	-0.442	--
14	2017-02-08	4.1	3.47	3.13	4.04	0.0565	0.117	--
15	2017-03-22	4.3	4.9	3.28	4.24	0.0588	0.196	--
16	2017-04-19	1.48	10.5	0.88	1.25	0.232	0.53	--
17	2017-05-03	1.65	20	1.36	1.91	-0.26	-0.358	--
18	2017-06-29	5.37	338	3.1	5.99	-0.622	-1.4	--
19	2017-08-16	1.21	24.1	0.633	1.02	0.192	0.442	--
20	2017-08-22	1.94	259	0.7	2.5	-0.557	-1.07	--
21	2017-11-16	3.67	2.6	1.84	2.41	1.26	1.65	--
22	2018-01-24	5.47	3.27	3.67	4.72	0.751	1.4	--
23	2018-03-19	9.66	722	3.8	9.16	0.505	0.939	--
24	2018-03-26	4.43	276	3.1	5.62	-1.19	-2.08	--
25	2018-05-24	1.22	10.1	0.333	0.556	0.662	1.22	--
26	2018-05-25	6.44	539	1.47	5.13	1.31	2.08	--
27	2018-06-25	1.94	11.6	1.6	2.16	-0.222	-0.276	--
28	2018-07-30	1.93	11.9	1.5	2.04	-0.108	-0.117	--
29	2018-08-28	1.4	9.1	1.07	1.48	-0.076	-0.0389	--
30	2018-08-30	2.18	101	1.5	2.57	-0.387	-0.622	--
31	2018-09-21	1.76	10.1	1.6	2.15	-0.394	-0.719	--
32	2018-10-07	2.88	339	1.1	3.48	-0.596	-1.22	--

Definitions

Cook’s D: Cook’s distance, a measure of influence (Helsel and others, 2020).

DFFITS: difference in fits, a measure of influence (Helsel and others, 2020).

leverage: a data point’s distance from the middle (mean) value in the x direction (Helsel and others, 2020).

LOESS: local polynomial regression fitting, or locally estimated scatterplot smoothing (Helsel and others, 2020).

LOWESS: locally weighted scatterplot smoothing (Cleveland, 1979; Helsel and others, 2020).

MSE: model standard error, also known as standard error of the regression (Helsel and others, 2020).

MSPE: model standard percentage error (Helsel and others, 2020).

NOx: sensor-measured nitrate plus nitrite, water, in situ, milligrams per liter as nitrogen (U.S. Geological Survey parameter code 99133).

Pr(>|t|): the probability that the independent variable has no effect on the dependent variable (Helsel and others, 2020).

RMSE: root mean square error (Helsel and others, 2020).

```
t value: Student’s t value; the coefficient divided by its associated standard error (Hel
sel and others, 2020).
```

```
TBY: Turbidity, water, unfiltered, monochrome near infra-red LED light, 780-900 nm, detec
tion angle 90 +-2.5 degrees, formazin nephelometric units (FNU) (U.S. Geological Survey p
arameter code 63680).
```

```
TN: sum of total Kjeldahl nitrogen (U.S. Geological Survey parameter code 00625, also kno
wn as total ammonia plus organic nitrogen) and dissolved nitrate plus nitrite (U.S. Geolo
gical Survey parameter code 00631), in milligrams per liter as nitrogen (mg/L-N).
```

App Version 1.0

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.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chapter A3, 458 p., https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1.]

Rasmussen, T.J., Lee, C.J., and Ziegler, A.C., 2008, Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p. [Also available at https://pubs.usgs.gov/sir/2008/5014/.]

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 1, 2019, at https://doi.org/10.5066/F7P55KJN.

Appendix 3. Model Archive Summary for *Escherichia coli* at Mill Creek at Johnson Drive, Shawnee, Kansas, 2015–18.

This model archive summary summarizes the *Escherichia coli* bacteria (ECB) model developed to compute 15-minute ECB from January 1, 2015, to December 31, 2018. This model supersedes the previous model used from 2003 to 2009 (Rasmussen and others, 2008).

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Site and Model Information

Site number: 06892513

Site name: Mill Creek at Johnson Drive, Shawnee, Kansas

Location: Latitude 39°01′45″, longitude 94°49′02″ referenced to North American Datum of 1983, in Johnson County, Kansas, Hydrologic Unit Code 10270104.

Equipment: A Yellow Springs Instruments (YSI) EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, turbidity, and pH, and a Hach Nitratax *plus* sc. The EXO2 was housed in a 4-inch metal pipe, and the Nitratax was housed in a 3-inch PVC pipe. Readings from the EXO2 and Nitratax were recorded every 15 minutes and transmitted by way of satellite hourly.

Date model was created: March 18, 2019

Model calibration data period: March 10, 2015, to October 7, 2018

Model application date: January 1, 2015, to December 31, 2018

Sampling Details

Grab samples were collected from the downstream side of the bridge. Samples were obtained at least monthly, with the priority being during storm-runoff events, with an autoclaved Nalgene plastic bottle. Samples were analyzed at the U.S. Geological Survey (USGS) Kansas Water Science Center (KSWSC) in Lawrence, Kansas; and the Johnson County Water Quality Laboratory (JCWQL) in Olathe, Kansas.

Model Data

All data were collected using USGS protocols and are stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019). The regression model is based on 34 concurrent measurements of ECB and turbidity collected from March 10, 2015, through October 7, 2018. Samples were collected throughout the range of continuously observed hydrologic conditions. Two samples analyzed by the JCWQL had densities that were less than its reporting limit of 10 most probable number per 100 milliliters (MPN/100 mL); zero samples analyzed by the KSWSC had densities that were less than its reporting limit of 1 colony forming unit per 100 milliliters (CFU/100 mL). Summary statistics and the complete model-calibration dataset are provided below. Potential outliers were identified as the data points for which both the studentized residual was greater than 3 or less than negative 3 and the Cook’s D value exceeded the outlier test criteria, as described by Helsel and others (2020). Zero ECB samples were deemed outliers.

Model Development

All continuously measured water-quality parameters and streamflow were considered as explanatory variables for estimating *Escherichia coli* using ordinary least squares regression. A variety of models that predict ECB and models that predict $\log_{10}$(ECB) were evaluated. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and predicted values) as compared to predicted ECB were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of predicted values). This comparison led to the conclusion that the most appropriate and reliable model would be one that estimated $\log_{10}$(ECB).

Turbidity was selected as the best predictor of ECB based on residual plots, relatively high coefficient of determination (R^2), and relatively low model standard percentage error (*MSPE*). Values for the aforementioned statistics and metrics were computed and are included below along with all relevant sample data and more in-depth statistical information.

Model Summary

Summary of final regression analysis for *Escherichia coli* at site number 06892513.

Escherichia coli model:

$$\log_{10}(ECB) = 1.34 \times \log_{10}(TBY) + 0.79\,,$$

where

ECB = *Escherichia coli* in colony forming units per 100 milliliters (CFU/100 mL) or in most probable number per 100 milliliters (MPN/100 mL); and,

TBY = turbidity, YSI EXO2, in formazin nephelometric units (FNU).

Turbidity makes physical and statistical sense as an explanatory variable for *E. coli*. It makes physical sense because bacterial colonies become suspended in water in the same manner that light-scattering particles, which increase turbidity, become suspended. Turbidity makes statistical sense as an explanatory variable because it resulted in a model with low standard error and high R^2 values. The model selected was the simplest model (one explanatory variable) even though some of the other models were marginally better statistically.

The log-transformed model may be retransformed to the original units so that ECB can be calculated directly. The retransformation introduces a bias in the calculated constituent. This bias may be corrected using Duan's Bias Correction Factor (BCF). For this model, the calculated BCF is 1.56. The retransformed model, accounting for BCF is:

$$ECB = 9.59 \times TBY^{1.34}.$$

Previous Models

Model	Start year	End year	Model
1.0	2003	2009	$\log_{10}(ECB) = 1.225 \times \log_{10}(TBY) + 0.508$

where (Rasmussen and others, 2008)

ECB = *Escherichia coli* in colony forming units per 100 milliliters (CFU/100 mL) or in most probable number per 100 milliliters (MPN/100 mL); and,

TBY = turbidity, YSI model 6136, in formazin nephelometric units (FNU).

Escherichia coli Record

The ECB record is computed using this regression model and stored at the National Real-Time Water Quality (NRTWQ) Web site. Data are computed at hourly intervals. The complete water-quality record can be found at https://nrtwq.usgs.gov/ks.

Remarks

None

Computed by: Patrick Eslick

Reviewed by: Brian Klager

Model Statistics, Data, and Plots

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

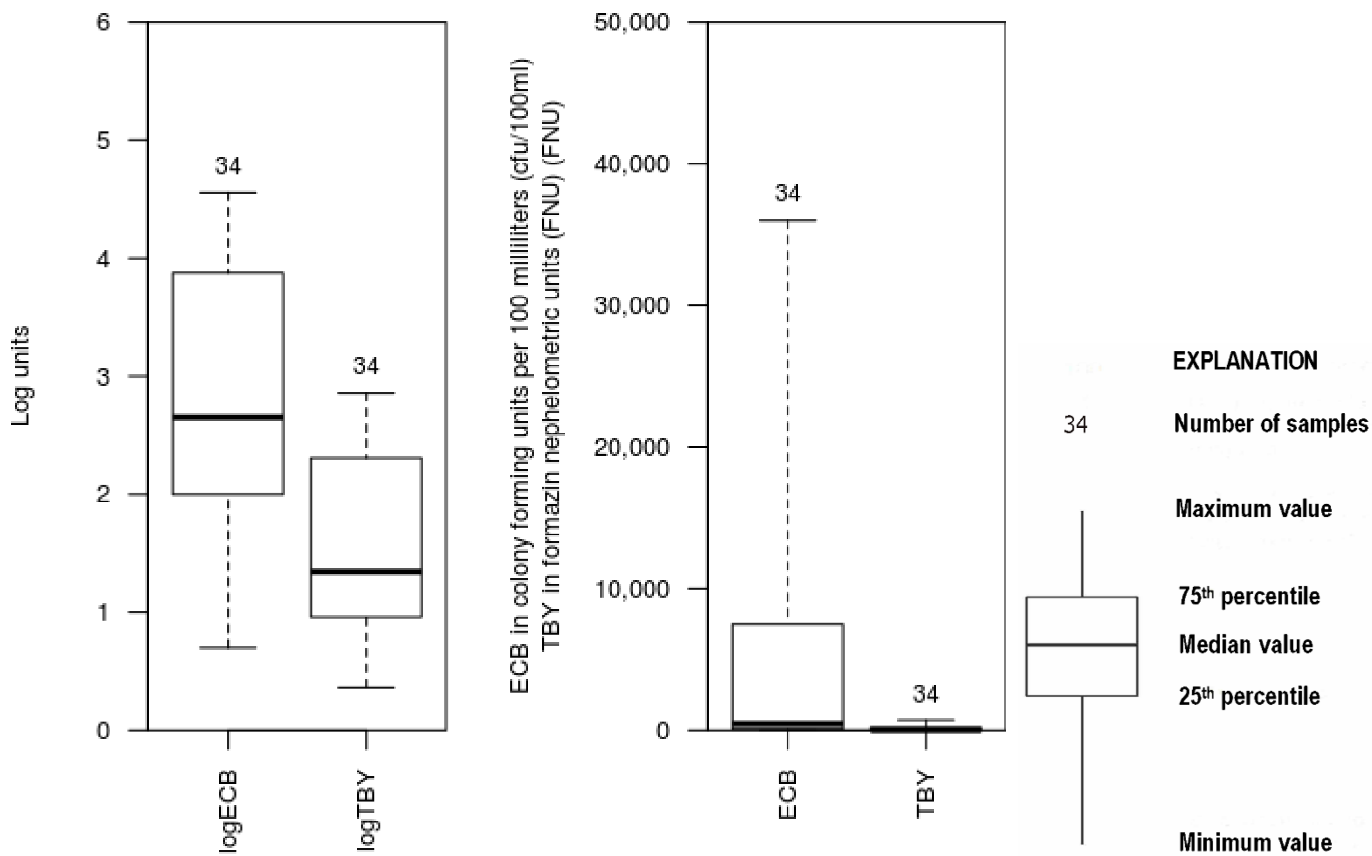
Model

log*ECB* = + 1.34 * log*TBY* + 0.79

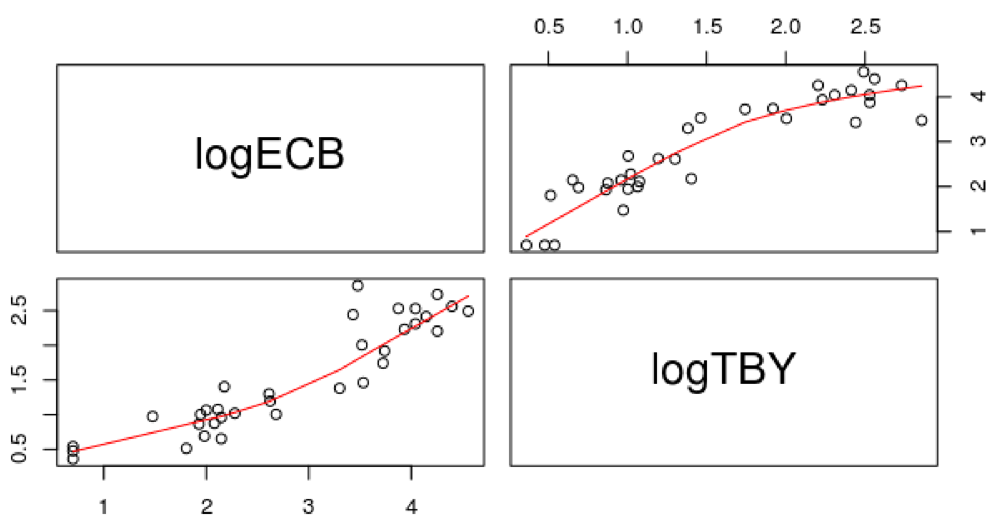
Variable Summary Statistics

```
              logECB   ECB logTBY   TBY
Minimum        0.699     5  0.362   2.3
1st Quartile   2.000   100  0.959   9.1
Median         2.650   450  1.340  22.1
Mean           2.830  5200  1.520 122.0
3rd Quartile   3.880  7500  2.310 203.0
Maximum        4.560 36000  2.860 722.0
```

Box Plots



Exploratory Plots



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

The x- and y-axis labels for a given bivariate plot are defined by the intersecting row and column labels.

Basic Model Statistics

Number of Observations	34
Standard error (RMSE)	0.468
Average Model standard percentage error (MSPE)	130
Coefficient of determination (R²)	0.835
Adjusted Coefficient of Determination (Adj. R²)	0.83
Bias Correction Factor (BCF)	1.56

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.79	0.179	4.4	1.12e-04
logTBY	1.34	0.105	12.7	4.52e-14

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.895
E.vars	-0.895	1.000

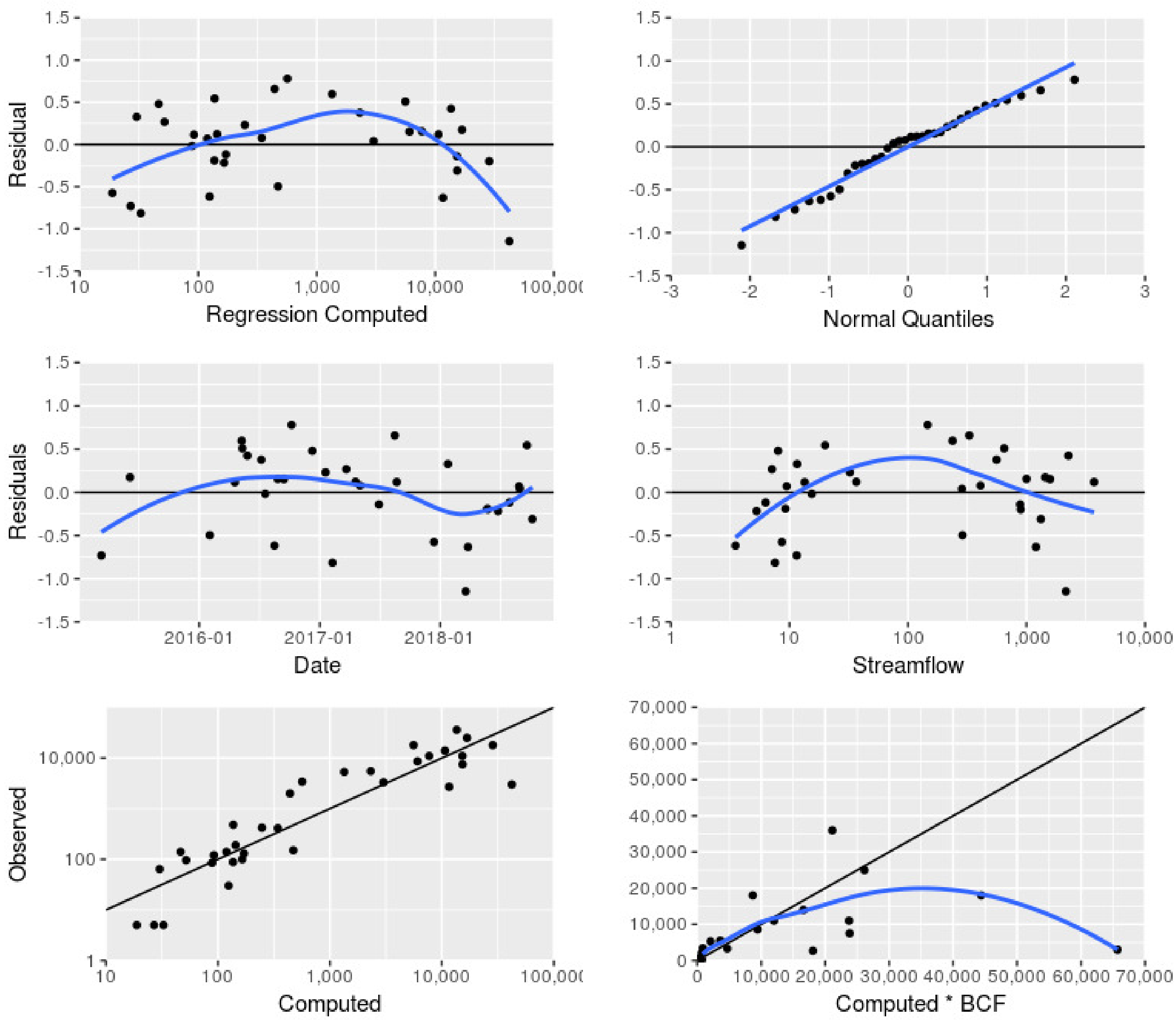
Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.176	0.194	0.485

Flagged Observations

	logECB	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2015-03-10 11:50:00	0.699	1.43	-0.731	-1.63	-1.68	0.0850	0.124	-0.512
2017-02-08 10:40:00	0.699	1.51	-0.815	-1.82	-1.89	0.0785	0.140	-0.551
2018-03-19 13:00:00	3.480	4.62	-1.150	-2.62	-2.91	0.1200	0.467	-1.070

Statistical Plots



First row (left): residual ECB (in log-space units) related to regression-computed ECB (in CFU/100 mL) with local polynomial regression fitting, or locally estimated scatterplot smoothing (LOESS) indicated by the blue line.

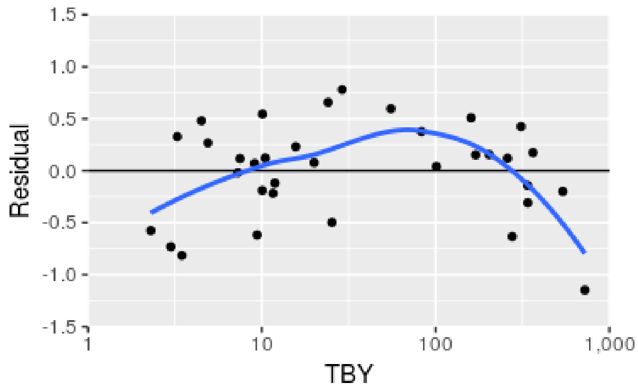
First row (right): residual ECB (in log-space units) related to the corresponding normal quantile (unitless) of the residual with simple linear regression indicated by the blue line.

Second row (left): residual ECB (in log-space units) related to date with LOESS indicated by the blue line.

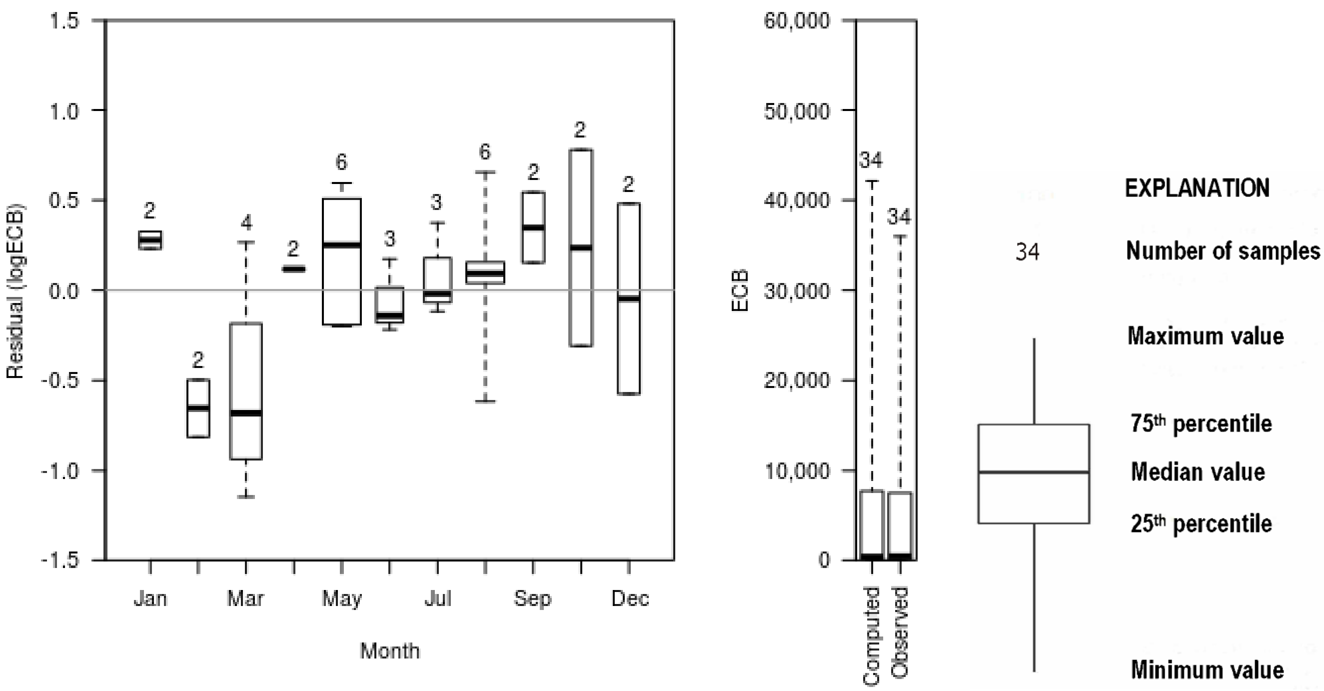
Second row (right): residual ECB (in log-space units) related to streamflow (in cubic feet per second) with LOESS indicated by the blue line.

Third row (left): observed ECB (in CFU/100 mL) related to regression-computed ECB (in CFU/100 mL).

Third row (right): observed ECB (in CFU/100 mL) related to the product of regression-computed ECB (in CFU/100 mL) and the bias correction factor with LOESS indicated by the blue line.

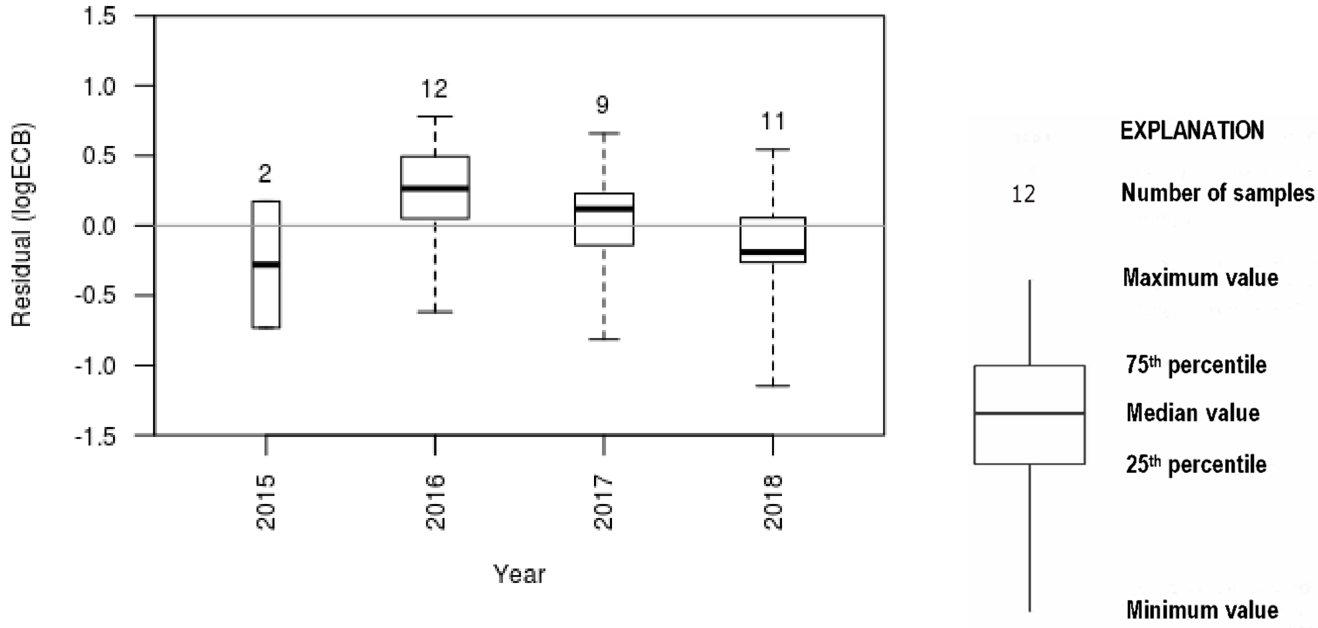


Residual ECB (in log-space units) related to TBY (in FNU) with LOESS indicated by the blue line.



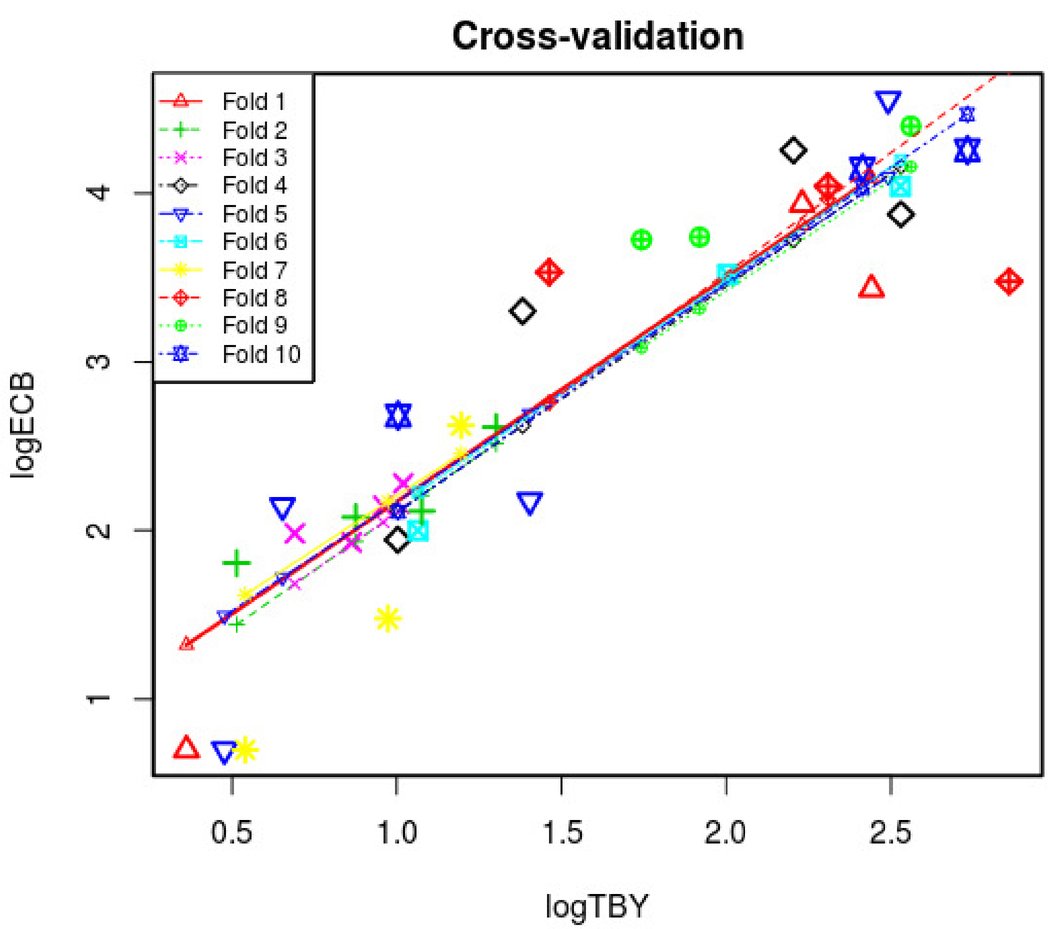
Left: residual ECB (in log-space units) by month.

Right: ECB (in CFU/100 mL) in regression-computed and observed values.

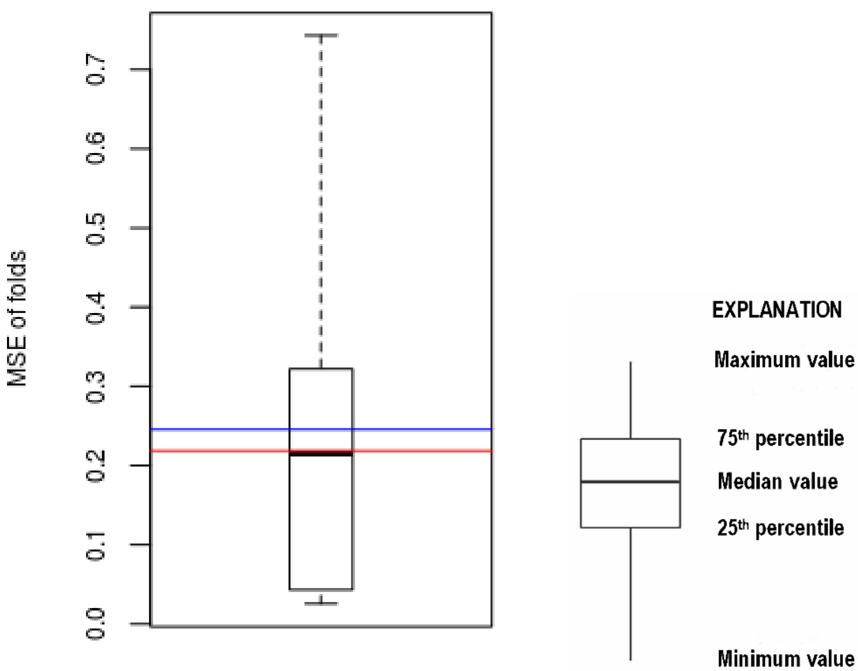


Residual ECB (in log-space units) by year.

Cross Validation



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

Large symbols: observed values of data points removed in a fold

Small symbols: recomputed values of data points removed in a fold

Recomputed regression lines: adjusted regression lines with one fold removed


              Minimum MSE of folds:  0.0253
                 Mean MSE of folds:  0.2450
               Median MSE of folds:  0.2160
              Maximum MSE of folds:  0.7430
 (Mean MSE of folds) / (Model MSE):  1.1200
```




```
Red line - Model MSE (unitless)
Blue line - Mean MSE of folds (unitless)
```

Model-Calibration Data Set

	Date	logECB	logTBY	ECB	TBY	Computed logECB	Computed ECB	Residual	Normal Quantiles	Censored Values
1	2015-03-10	0.699	0.477	5	3	1.43	41.9	-0.731	-1.43	--
2	2015-06-05	4.4	2.56	25000	363	4.22	26100	0.173	0.415	--
3	2016-02-02	2.18	1.4	150	25.3	2.67	734	-0.497	-0.867	--
4	2016-04-18	2.08	0.875	120	7.5	1.96	143	0.116	0.0367	--
5	2016-05-09	3.72	1.74	5300	55.3	3.13	2090	0.596	1.43	--
6	2016-05-11	4.26	2.2	18000	160	3.75	8700	0.508	1.11	--
7	2016-05-26	4.56	2.49	36000	310	4.13	21100	0.424	0.867	--
8	2016-07-07	3.74	1.92	5500	83	3.36	3610	0.376	0.765	--
9	2016-07-19	1.93	0.863	85	7.3	1.95	138	-0.0184	-0.259	--

10	2016-08-16	1.48	0.973	30	9.4	2.1	194	-0.618	-1.11	--
11	2016-08-26	4.04	2.31	11000	203	3.89	12000	0.155	0.336	--
12	2016-09-14	3.93	2.23	8600	170	3.78	9440	0.152	0.259	--
13	2016-10-07	3.53	1.46	3400	29	2.75	880	0.78	2.11	--
14	2016-12-09	2.15	0.653	140	4.5	1.67	72.2	0.48	0.979	--
15	2017-01-18	2.62	1.19	420	15.7	2.39	385	0.23	0.496	--
16	2017-02-08	0.699	0.54	5	3.47	1.51	50.9	-0.815	-1.68	--
17	2017-03-22	1.98	0.69	96	4.9	1.72	81	0.267	0.581	--
18	2017-04-19	2.28	1.02	190	10.5	2.16	224	0.121	0.184	--
19	2017-05-03	2.61	1.3	410	20	2.54	534	0.0776	-0.0367	--
20	2017-06-29	4.04	2.53	11000	338	4.18	23700	-0.141	-0.415	--
21	2017-08-16	3.3	1.38	2000	24.1	2.64	686	0.657	1.68	--
22	2017-08-22	4.15	2.41	14000	259	4.03	16600	0.119	0.11	--
23	2017-12-13	0.699	0.362	5	2.3	1.27	29.3	-0.576	-0.979	--
24	2018-01-24	1.81	0.514	64	3.27	1.48	47	0.327	0.67	--
25	2018-03-19	3.48	2.86	3000	722	4.62	65700	-1.15	-2.11	--
26	2018-03-26	3.43	2.44	2700	276	4.06	18100	-0.633	-1.25	--
27	2018-05-24	1.94	1	88	10.1	2.14	213	-0.191	-0.496	--
28	2018-05-25	4.26	2.73	18000	539	4.45	44400	-0.199	-0.581	--
29	2018-06-25	2	1.06	100	11.6	2.22	257	-0.218	-0.67	--
30	2018-07-30	2.11	1.08	130	11.9	2.23	266	-0.119	-0.336	--
31	2018-08-28	2.15	0.959	140	9.1	2.08	186	0.0698	-0.11	--
32	2018-08-30	3.52	2	3300	101	3.48	4690	0.0397	-0.184	--
33	2018-09-21	2.68	1	480	10.1	2.14	214	0.544	1.25	--
34	2018-10-07	3.88	2.53	7500	339	4.18	23800	-0.309	-0.765	--

Definitions

Cook’s D: Cook’s distance, a measure of influence (Helsel and others, 2020).

DFFITS: difference in fits, a measure of influence (Helsel and others, 2020).

E.vars: explanatory variables.

ECB: Escherichia coli, modified m-TEC MF method, water, colony forming units per 100 milliliters (U.S. Geological Survey parameter code 90902), or Escherichia coli, defined substrate test method (DSTM), water, most probable number per 100 milliliters (U.S. Geological Survey parameter code 50468).

Leverage: a data point’s distance from the middle (mean) value in the x direction (Helsel and others, 2020).

LOESS: local polynomial regression fitting, or locally estimated scatterplot smoothing (Helsel and others, 2020).

LOWESS: locally weighted scatterplot smoothing (Cleveland, 1979; Helsel and others, 2020).

MSE: model standard error, also known as standard error of the regression (Helsel and others, 2020).

```
MSPE: model standard percentage error (Helsel and others, 2020).

Pr(>|t|): the probability that the independent variable has no effect on the dependent va
riable (Helsel and others, 2020).

RMSE: root mean square error (Helsel and others, 2020).

t value: Student’s t value; the coefficient divided by its associated standard error (Hel
sel and others, 2020).

TBY: Turbidity, water, unfiltered, monochrome near infra-red LED light, 780-900 nm, detec
tion angle 90 +-2.5 degrees, formazin nephelometric units (FNU) (U.S. Geological Survey p
arameter code 63680).
```

App Version 1.0

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.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chapter A3, 458 p., https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1.]

Rasmussen, T.J., Lee, C.J., and Ziegler, A.C., 2008, Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p. [Also available at https://pubs.usgs.gov/sir/2008/5014/.]

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 1, 2019, at https://doi.org/10.5066/F7P55KJN.

Appendix 4. Model Archive Summary for Total Suspended Solids at Mill Creek at Johnson Drive, Shawnee, Kansas, 2015–18.

This model archive summary summarizes the total suspended solids (TSS) model developed to compute 15-minute TSS from January 1, 2015, to December 31, 2018. This model supersedes the model used from 2003 to 2009 (Rasmussen and others, 2008).

Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Site and Model Information

Site number: 06892513

Site name: Mill Creek at Johnson Drive, Shawnee, Kansas

Location: Latitude 39°01′45″, longitude 94°49′02″ referenced to North American Datum of 1983, in Johnson County, Kansas, Hydrologic Unit Code 10270104.

Equipment: A Yellow Springs Instruments (YSI) EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, turbidity, and pH, and a Hach Nitratax *plus* sc. The EXO2 was housed in a 4-inch metal pipe, and the Nitratax was housed in a 3-inch PVC pipe. Readings from the EXO2 and Nitratax were recorded every 15 minutes and transmitted by way of satellite hourly.

Date model was created: March 18, 2019

Model calibration data period: March 10, 2015, to October 7, 2018

Model application date: January 1, 2015, to December 31, 2018

Sampling Details

Equal-width-increment samples were collected from the downstream side of the bridge. Samples were obtained at least monthly, with the priority being during storm-runoff events, with a FISP US DH-95 depth-integrating sampler with a Teflon bottle, cap, and nozzle. Samples were analyzed at the U.S. Geological Survey (USGS) National Water Quality Laboratory (NWQL) in Lakewood, Colorado, and at the Johnson County Water Quality Laboratory (JCWQL) in Olathe, Kansas.

Model Data

All data were collected using USGS protocols and are stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019). The regression model is based on 36 concurrent measurements of total suspended solids and turbidity collected from March 10, 2015, through October 7, 2018. Samples were collected throughout the range of continuously observed hydrologic conditions. Four samples analyzed by the NWQL had concentrations that were less than its reporting limit of 15 milligrams per liter (mg/L); zero samples analyzed by the JCWQL had concentrations that were less than its reporting limit of 1 mg/L. Summary statistics and the complete model-calibration dataset are provided below. Potential outliers were identified as the data points for which both the studentized residual was greater than 3 or less than negative 3 and the Cook’s D value exceeded the outlier test criteria, as described by Helsel and others (2020). Zero TSS samples were deemed outliers.

Model Development

All continuously measured water-quality parameters and streamflow were considered as explanatory variables for estimating total suspended solids using ordinary least squares (OLS) regression. A variety of models that predict TSS and models that predict $\log_{10}$(TSS) were evaluated. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and predicted values) compared to predicted TSS were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of predicted values). This comparison led to the conclusion that the most appropriate and reliable model would be one that estimated $\log_{10}$(TSS).

Turbidity was selected as the best predictor of TSS based on residual plots, relatively high coefficient of determination (R^2), and relatively low model standard percentage error (*MSPE*). Values for all aforementioned statistics and metrics were computed and are included below along with all relevant sample data and more in-depth statistical information.

Model Summary

Summary of final regression analysis for total suspended solids at USGS site number 06892513.

Total suspended solids model:

$$\log_{10}(TSS) = 1.05 \times \log_{10}(TBY) + 0.339,$$

where

TSS = total suspended solids in milligrams per liter (mg/L); and,

TBY = Turbidity, YSI EXO2, in formazin nephelometric units (FNU).

Turbidity makes physical and statistical sense as an explanatory variable for total suspended solids. It makes sense physically because the particles that comprise the suspended solids scatter light, which affects turbidity. Turbidity makes statistical sense as an explanatory variable because it resulted in a model with low standard error and high R^2 values. The model selected was the simplest model (one explanatory variable) and the best, statistically.

The log-transformed model may be retransformed to the original units so that TSS can be calculated directly. The retransformation introduces a bias in the calculated constituent. This bias may be corrected using Duan's Bias Correction Factor (BCF). For this model, the calculated BCF is 1.12. The retransformed model, accounting for BCF is:

$$TSS = 2.44 \times TBY^{1.05}.$$

Previous Models

Model	Start year	End year	Model
1.0	2003	2009	$\log_{10}(TSS) = 0.985 \times \log_{10}(TBY) + 0.242$

where (Rasmussen and others, 2008)

TSS = total suspended solids in milligrams per liter (mg/L); and,

TBY = Turbidity, YSI model 6136, in formazin nephelometric units (FNU).

Total Suspended Solids Record

The TSS record is computed using this regression model and stored at the National Real-Time Water Quality (NRTWQ) Web site. Data are computed at hourly intervals. The complete water-quality record can be found at https://nrtwq.usgs.gov/ks.

Remarks

Because more than 5 percent of the data points were censored (4 of 36, 11 percent), a non-parametric model was also computed. The Theil-Sen method computes slope as the median of all pairwise slopes between observations and is not strongly influenced by the presence of outliers (Helsel, 2005). The intercept is calculated to ensure the fitted line goes through the median point (Helsel and others, 2020). The Thiel-Sen method does not require the assumption of a specific distribution. The non-parametric model computed using this method was

$$\log_{10}(TSS_{T-S}) = 1.064 \times \log_{10}(TBY) + 0.307,$$

where

TSS_{T-S} = total suspended solids in milligrams per liter (mg/L), estimated using the Thiel-Sen method; and,

TBY = Turbidity, YSI EXO2, in formazin nephelometric units (FNU).

The non-parametric model had coefficients similar to those calculated using the OLS method.

Computed by: Patrick Eslick

Reviewed by: Brian Klager

Model Statistics, Data, and Plots

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

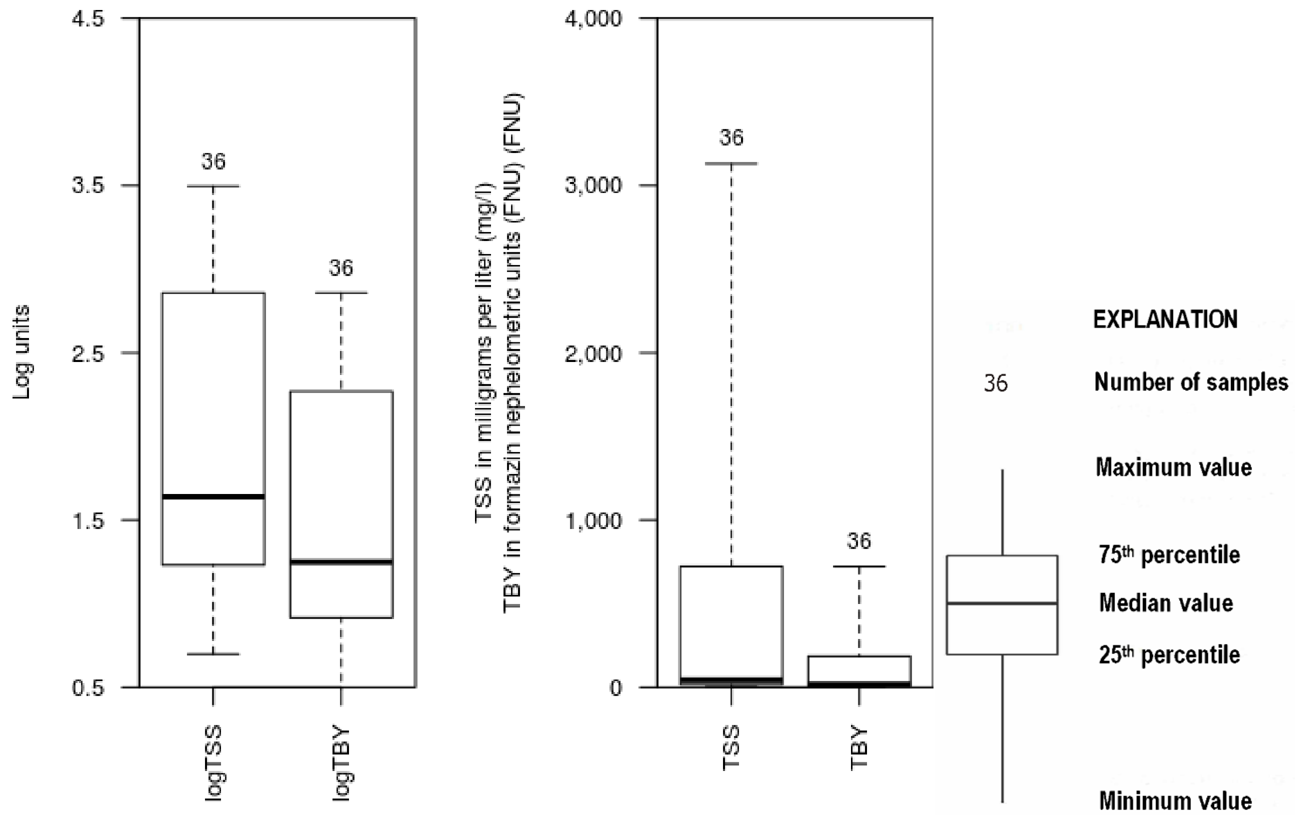
Model

log*TSS* = + 1.05 * log*TBY* + 0.339

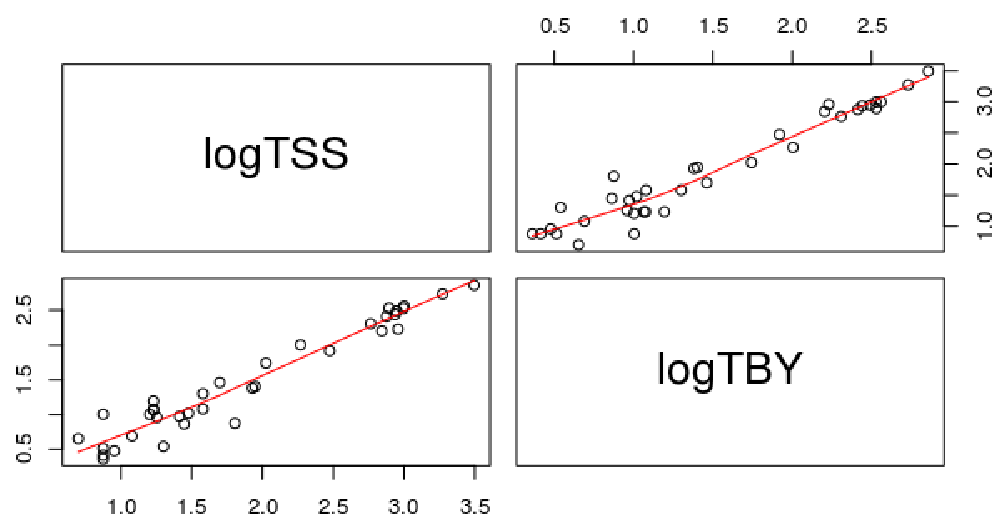
Variable Summary Statistics

```
               logTSS  TSS logTBY   TBY
Minimum         0.699    5  0.362   2.3
1st Quartile    1.230   17  0.917   8.3
Median          1.640   44  1.250  17.8
Mean            1.900  379  1.480 115.0
3rd Quartile    2.860  723  2.270 187.0
Maximum         3.500 3130  2.860 722.0
```

Box Plots



Exploratory Plots



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

The x- and y-axis labels for a given bivariate plot are defined by the intersecting row and column labels.

Basic Model Statistics

Number of Observations	36
Standard error (RMSE)	0.212
Average Model standard percentage error (MSPE)	50.7
Coefficient of determination (R^2)	0.939
Adjusted Coefficient of Determination (Adj. R^2)	0.937
Bias Correction Factor (BCF)	1.12

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.339	0.0769	4.41	9.89e-05
logTBY	1.05	0.0462	22.80	3.41e-22

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.888
E.vars	-0.888	1.000

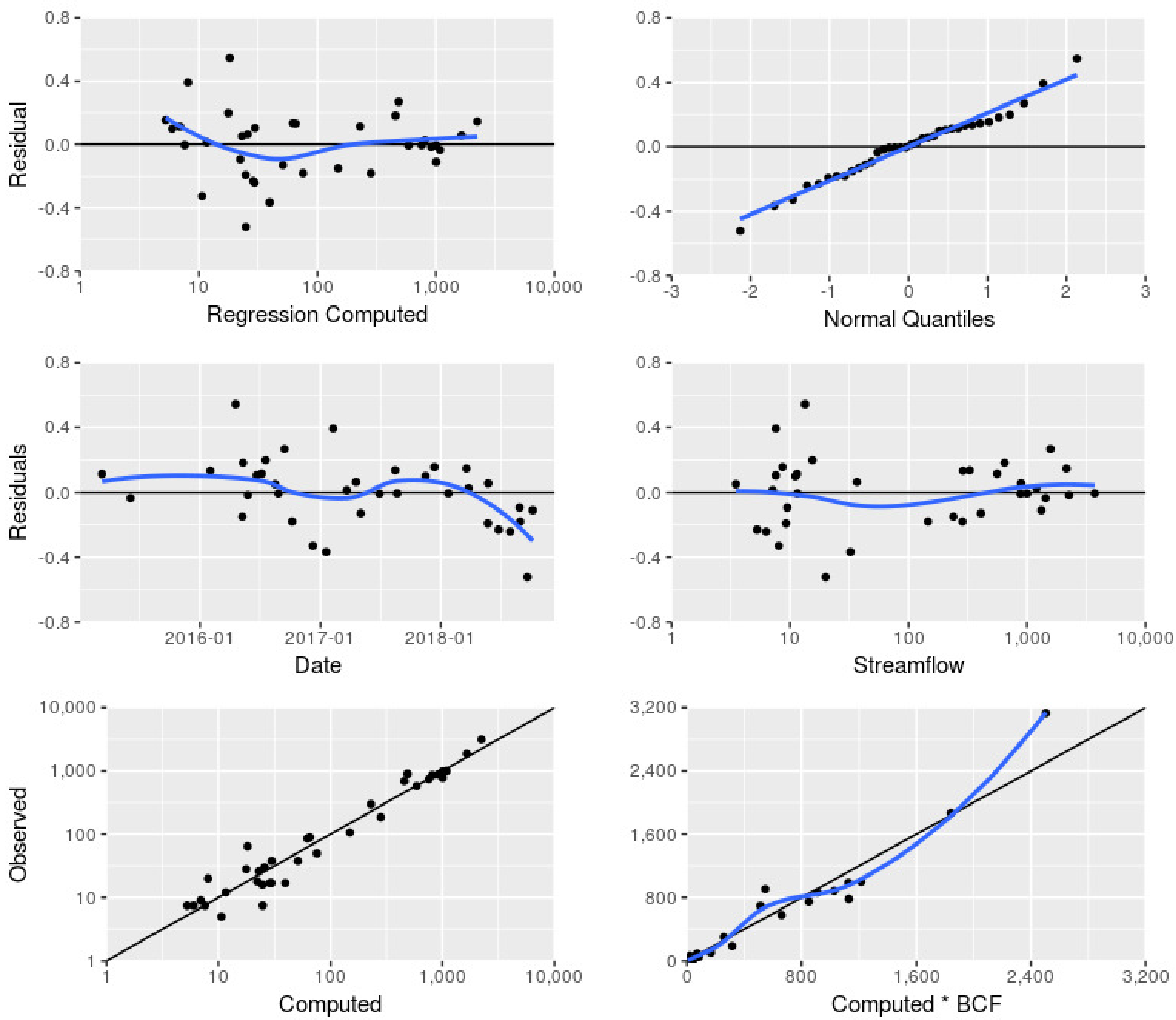
Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.167	0.194	0.471

Flagged Observations

	logTSS	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2016-04-18 10:15:00	1.810	1.260	0.546	2.64	2.91	0.0452	0.164	0.633
2017-02-08 10:40:00	1.300	0.908	0.393	1.93	2.01	0.0698	0.139	0.551
2018-09-21 10:10:00	0.875	1.400	-0.522	-2.51	-2.74	0.0385	0.126	-0.549

Statistical Plots



First row (left): residual TSS (in log-space units) related to regression-computed TSS (in mg/L) with local polynomial regression fitting, or locally estimated scatterplot smoothing (LOESS) indicated by the blue line.

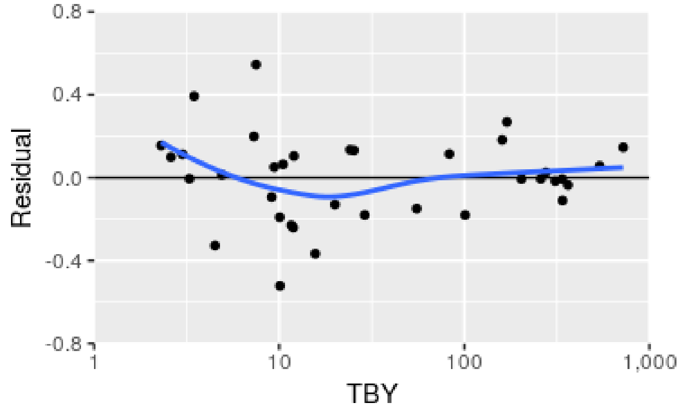
First row (right): residual TSS (in log-space units) related to the corresponding normal quantile (unitless) of the residual with simple linear regression indicated by the blue line.

Second row (left): residual TSS (in log-space units) related to date with LOESS indicated by the blue line.

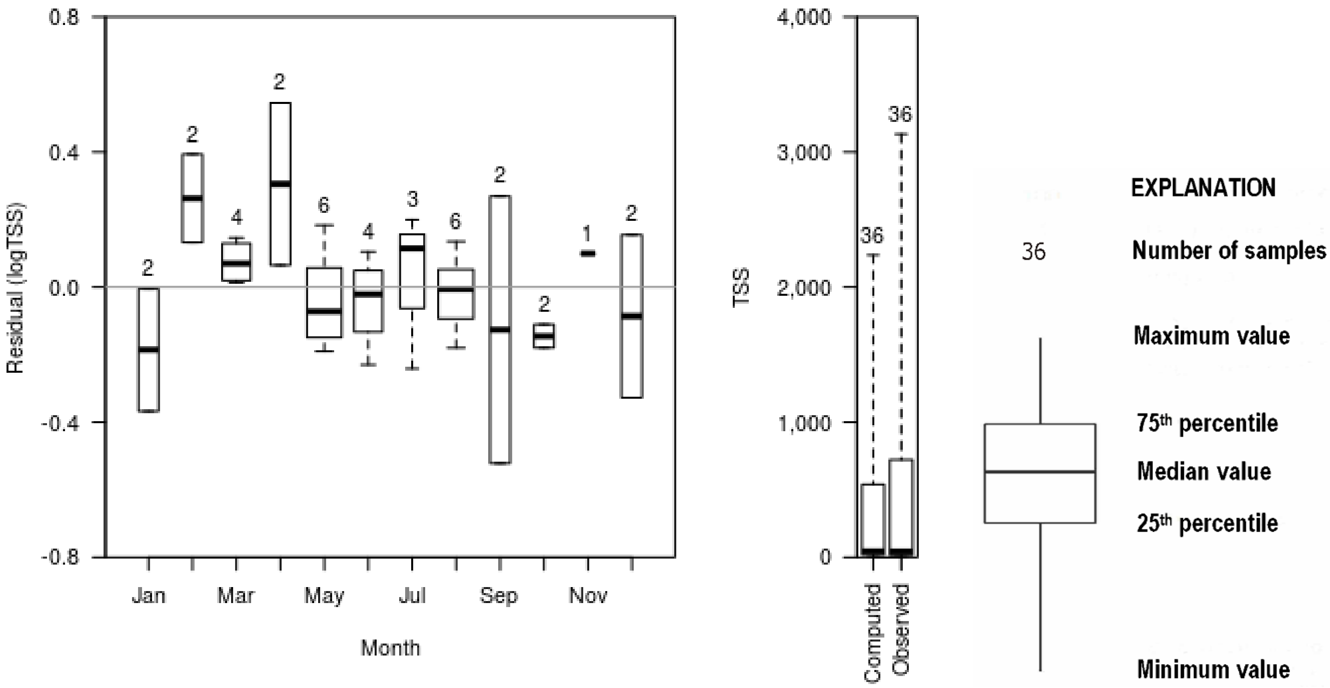
Second row (right): residual TSS (in log-space units) related to streamflow (in cubic feet per second) with LOESS indicated by the blue line.

Third row (left): observed TSS (in mg/L) related to regression-computed TSS (in mg/L).

Third row (right): observed TSS (in mg/L) related to the product of regression-computed TSS (in mg/L) and the bias correction factor with LOESS indicated by the blue line.

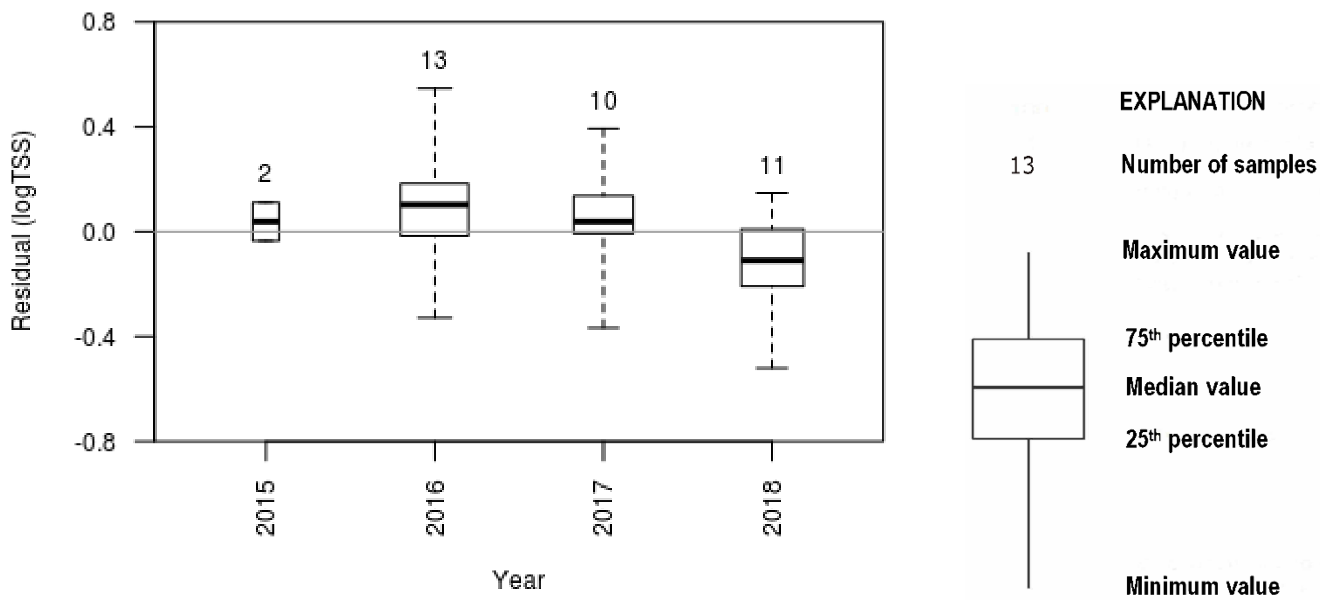


Residual TSS (in log-space units) related to TBY (in FNU) with LOESS indicated by the blue line.



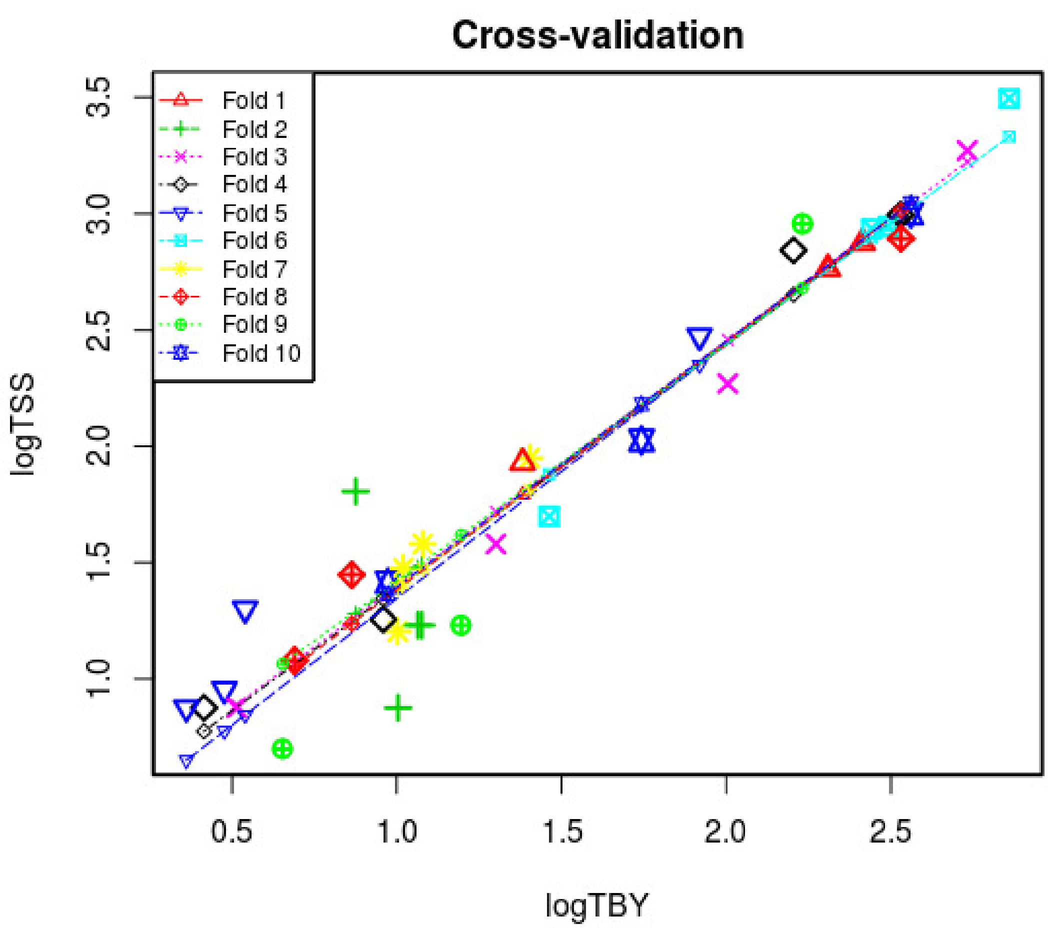
Left: residual TSS (in log-space units) by month.

Right: TSS (in mg/L) in regression-computed and observed values.

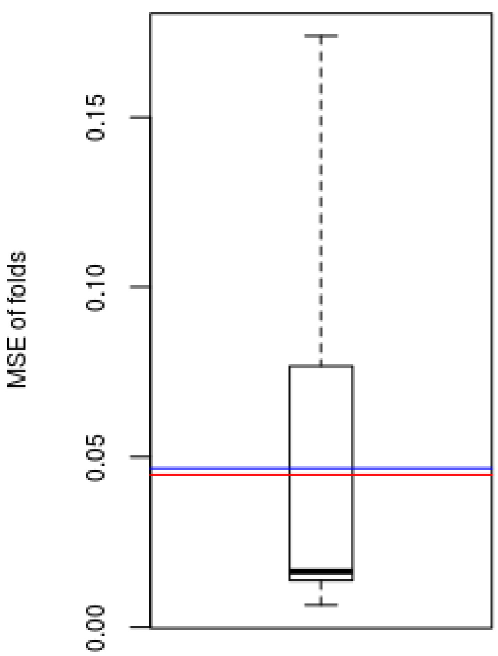


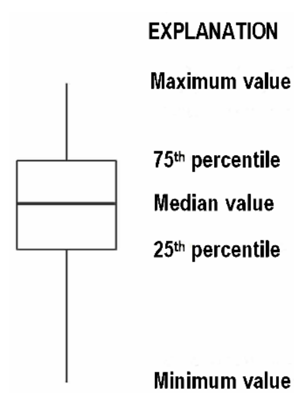
Residual TSS (in log-space units) by year.

Cross Validation



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

Large symbols: observed values of data points removed in a fold

Small symbols: recomputed values of data points removed in a fold

Recomputed regression lines: adjusted regression lines with one fold removed


              Minimum MSE of folds:  0.00642
                 Mean MSE of folds:  0.04680
               Median MSE of folds:  0.01630
              Maximum MSE of folds:  0.17400
 (Mean MSE of folds) / (Model MSE):  1.04000
```






```
Red line - Model MSE (unitless)

Blue line - Mean MSE of folds (unitless)
```

Model-Calibration Data Set

	Date	logTSS	logTBY	TSS	TBY	Computed logTSS	Computed TSS	Residual	Normal Quantiles	Censored Values
1	2015-03-10	0.954	0.477	9	3	0.841	7.78	0.113	0.545	--
2	2015-06-05	3	2.56	1000	363	3.04	1220	-0.0354	-0.391	--
3	2016-02-02	1.95	1.4	89	25.3	1.82	73.5	0.132	0.714	--
4	2016-04-18	1.81	0.875	64	7.5	1.26	20.4	0.546	2.13	--
5	2016-05-09	2.03	1.74	106	55.3	2.17	167	-0.149	-0.714	--
6	2016-05-11	2.84	2.2	696	160	2.66	512	0.182	1.14	--
7	2016-05-26	2.95	2.49	884	310	2.96	1030	-0.0163	-0.317	--
8	2016-06-22	1.58	1.08	38	12	1.48	33.5	0.104	0.466	--

9	2016-07-07	2.47	1.92	298	83	2.36	257	0.114	0.627	--
10	2016-07-19	1.45	0.863	28	7.3	1.25	19.8	0.199	1.28	--
11	2016-08-16	1.41	0.973	26	9.4	1.36	25.9	0.0511	0.174	--
12	2016-08-26	2.76	2.31	580	203	2.77	659	-0.00647	-0.174	--
13	2016-09-14	2.96	2.23	906	170	2.69	546	0.269	1.46	--
14	2016-10-07	1.7	1.46	50	29	1.88	84.8	-0.18	-0.807	--
15	2016-12-09	0.699	0.653	5	4.5	1.03	11.9	-0.328	-1.46	--
16	2017-01-18	1.23	1.19	17	15.7	1.6	44.3	-0.367	-1.7	--
17	2017-02-08	1.3	0.54	20	3.47	0.908	9.05	0.393	1.7	--
18	2017-03-22	1.08	0.69	12	4.9	1.07	13	0.0133	0.0346	--
19	2017-04-19	1.48	1.02	30	10.5	1.41	29	0.0641	0.317	--
20	2017-05-03	1.58	1.3	38	20	1.71	57.3	-0.129	-0.627	--
21	2017-06-29	2.99	2.53	988	338	3	1130	-0.00758	-0.245	--
22	2017-08-16	1.93	1.38	85	24.1	1.79	69.8	0.135	0.807	--
23	2017-08-22	2.88	2.41	750	259	2.88	851	-0.00551	-0.104	--
24	2017-11-16	0.875	0.415	7.5	2.6	0.776	6.69	0.0991	0.391	< 15
25	2017-12-13	0.875	0.362	7.5	2.3	0.72	5.88	0.155	1.02	< 15
26	2018-01-24	0.875	0.514	7.5	3.27	0.88	8.51	-0.00535	-0.0346	< 15
27	2018-03-19	3.5	2.86	3130	722	3.35	2500	0.146	0.907	--
28	2018-03-26	2.94	2.44	862	276	2.91	910	0.0259	0.104	--
29	2018-05-24	1.2	1	16	10.1	1.4	27.8	-0.191	-1.02	--
30	2018-05-25	3.27	2.73	1870	539	3.22	1840	0.0561	0.245	--
31	2018-06-25	1.23	1.06	17	11.6	1.46	32.3	-0.23	-1.14	--
32	2018-07-30	1.23	1.08	17	11.9	1.47	33.2	-0.241	-1.28	--
33	2018-08-28	1.26	0.959	18	9.1	1.35	25	-0.0937	-0.466	--
34	2018-08-30	2.27	2	186	101	2.45	316	-0.18	-0.907	--
35	2018-09-21	0.875	1	7.5	10.1	1.4	27.9	-0.522	-2.13	< 15
36	2018-10-07	2.89	2.53	782	339	3	1130	-0.11	-0.545	--

Definitions

Cook’s D: Cook’s distance, a measure of influence (Helsel and others, 2020).

DFFITS: difference in fits, a measure of influence (Helsel and others, 2020).

E.vars: explanatory variables.

Leverage: a data point’s distance from the middle (mean) value in the x direction (Helsel and others, 2020).

LOESS: local polynomial regression fitting, or locally estimated scatterplot smoothing (Helsel and others, 2020).

LOWESS: locally weighted scatterplot smoothing (Cleveland, 1979; Helsel and others, 2020).

MSE: model standard error, also known as standard error of the regression (Helsel and others, 2020).

MSPE: model standard percentage error (Helsel and others, 2020).

```
Pr(>|t|): the probability that the independent variable has no effect on the dependent variable (Helsel and others, 2020).

RMSE: root mean square error (Helsel and others, 2020).

t value: Student’s t value; the coefficient divided by its associated standard error (Helsel and others, 2020).

TBY: turbidity, water, unfiltered, monochrome near infra-red LED light, 780-900 nm, detection angle 90 +-2.5 degrees, formazin nephelometric units (FNU) (U.S. Geological Survey parameter code 63680).

TSS: total suspended solids, water, unfiltered, milligrams per liter (U.S. Geological Survey parameter code 00530).
```

App Version 1.0

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.

Helsel, D. R., 2005, Nondetects and data analysis: statistics for censored environmental data: Hoboken, New Jersey, John Wiley & Sons, Inc., 250 p.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chapter A3, 458 p., https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1.]

Rasmussen, T.J., Lee, C.J., and Ziegler, A.C., 2008, Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p. [Also available at https://pubs.usgs.gov/sir/2008/5014/.]

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 1, 2019, at https://doi.org/10.5066/F7P55KJN.

Appendix 5. Model Archive Summary for Suspended Sediment Concentration at Mill Creek at Johnson Drive, Shawnee, Kansas, 2015–18.

This model archive summary summarizes the suspended sediment (SS) concentration model developed to compute 15-minute SS from January 1, 2015, to December 31, 2018. This model supersedes the model used from 2003 to 2009 (Rasmussen and others, 2008).

Site and Model Information

Site number: 06892513

Site name: Mill Creek at Johnson Drive, Shawnee, Kansas

Location: Latitude 39°01′45″, longitude 94°49′02″ referenced to North American Datum of 1983, in Johnson County, Kansas, Hydrologic Unit Code 10270104.

Equipment: A Yellow Springs Instruments (YSI) EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, turbidity, and pH, and a Hach Nitratax *plus* sc. The EXO2 was housed in a 4-inch metal pipe, and the Nitratax was housed in a 3-inch PVC pipe. Readings from the EXO2 and Nitratax were recorded every 15 minutes and transmitted by way of satellite hourly.

Date model was created: March 18, 2019

Model calibration data period: March 10, 2015, to October 7, 2018

Model application date: January 1, 2015, to December 31, 2018

Sampling Details

Equal-width-increment samples were collected from the downstream side of the bridge. Samples were obtained at least monthly, with the priority being during storm-runoff events, with a FISP US DH-95 depth-integrating sampler with a Teflon bottle, cap, and nozzle. Samples were analyzed for suspended sediment concentration at the U.S. Geological Survey (USGS) Sediment Laboratory in Iowa City, Iowa.

Model Data

All data were collected using USGS protocols and are stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019). The regression model is based on 35 concurrent measurements of suspended sediment concentration and turbidity collected from March 10, 2015, through October 7, 2018. Samples were collected throughout the range of continuously observed hydrologic conditions. No samples had concentrations that were below the laboratory reporting limit of 1 milligram per liter (mg/L). Summary statistics and the complete model-calibration dataset are provided below. Potential outliers were identified as the data points for which both the studentized residual was greater than 3 or less than negative 3 and the Cook’s D value exceeded the outlier test criteria, as described by Helsel and others (2020). The sample collected December 13, 2017, was deemed an outlier but was not removed from the dataset; this sample’s SS value was 1 mg/L, which resulted in a $\log_{10}$(SS) value of zero. The $\log_{10}$(SS) value was set to 0.01 and the model was recalculated; this sample remained an outlier.

Model Development

All continuously measured water-quality parameters and streamflow were considered as explanatory variables for estimating suspended sediment concentration using ordinary least squares regression. A variety of models that predict SS and models that predict $\log_{10}$(SS) were evaluated. The distribution of residuals was examined for normality, and plots of residuals (the difference between the measured and predicted values) as compared to predicted SS were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of predicted values). This comparison led to the conclusion that the most appropriate and reliable model would be one that estimated $\log_{10}$(SS).

Turbidity was selected as the best predictor of SS based on residual plots, relatively high coefficient of determination (R^2), and relatively low model standard percentage error (*MSPE*). Values for the aforementioned statistics and metrics were computed and are included below along with all relevant sample data and more in-depth statistical information.

Model Summary

Summary of final regression analysis for suspended sediment concentration at site number 06892513.

Suspended sediment concentration model:

$$\log_{10}(SS) = 1.09 \times \log(TBY) + 0.345$$

where

SS = suspended sediment concentration in milligrams per liter (mg/L); and,

TBY = turbidity, YSI EXO2, in formazin nephelometric units (FNU).

Turbidity makes physical and statistical sense as an explanatory variable for SS. Turbidity makes sense physically because suspended sediment is composed of particles that scatter light in water. The relation between turbidity and SS can vary given varying concentrations of organic suspended particles that increase turbidity, but these are not included in the SS analysis. The model selected was the simplest model (one explanatory variable) and the best, statistically.

The log-transformed model may be retransformed to the original units so that SS can be calculated directly. The retransformation introduces a bias in the calculated constituent. This bias may be corrected using Duan's Bias Correction Factor (BCF). For this model, the calculated BCF is 1.18. The retransformed model, accounting for BCF is:

$$SS = 2.61 \times TBY^{1.09}.$$

Previous Models

Model	Start year	End year	Model
1.0	2003	2009	$\log_{10}(SS) = 1.026 \times \log_{10}(TBY) + 0.144$

where (Rasmussen and others, 2008)

SS = suspended sediment concentration in milligrams per liter (mg/L); and

TBY = turbidity, YSI model 6136, in formazin nephelometric units (FNU).

Suspended Sediment Concentration Record

The SS record is computed using this regression model and stored at the National Real-Time Water Quality (NRTWQ) Web site. Data are computed at hourly intervals. The complete water-quality record can be found at https://nrtwq.usgs.gov/ks.

Remarks

None

Computed by: Patrick Eslick

Reviewed by: Brian Klager

Model Statistics, Data, and Plots

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

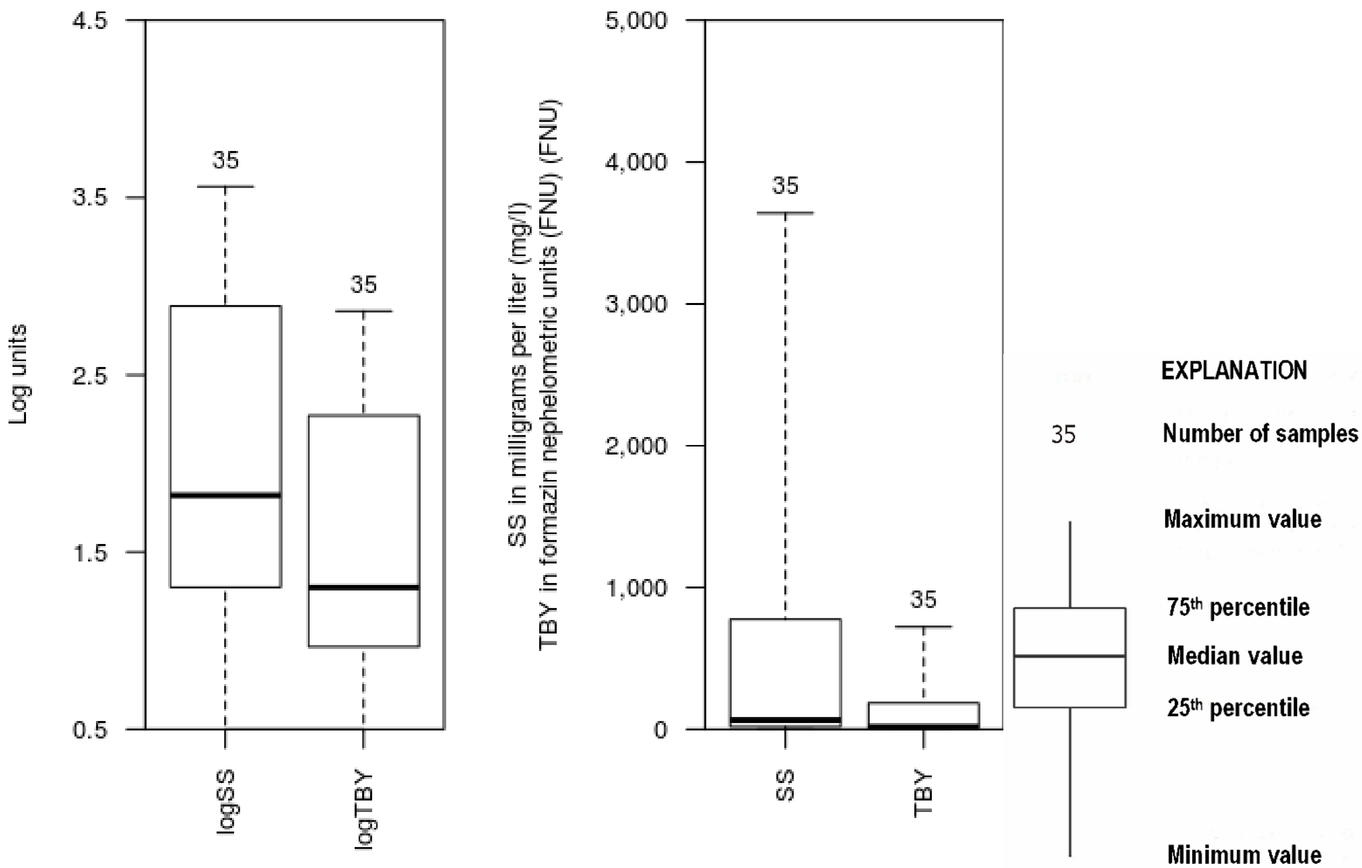
Model

log*SS* = + 1.09 * log*TBY* + 0.345

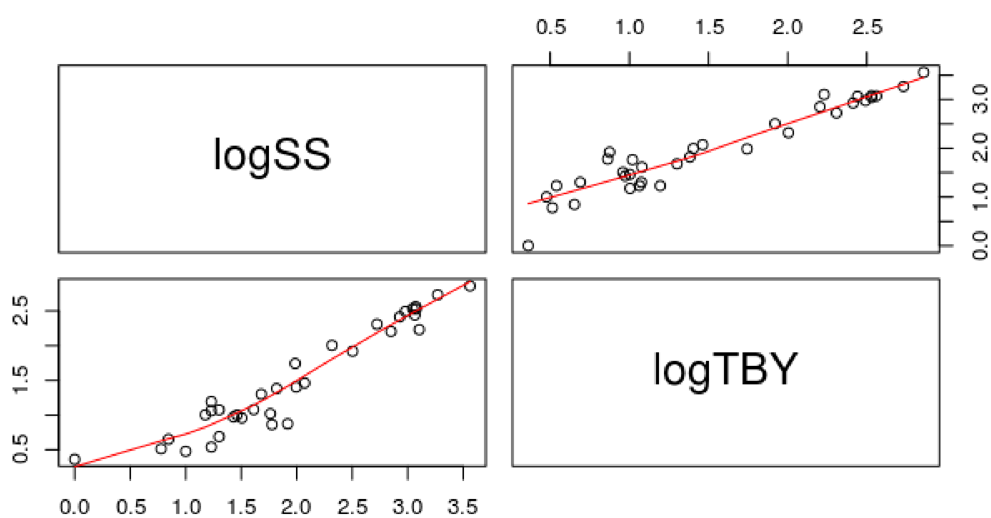
Variable Summary Statistics

```
             logSS   SS logTBY   TBY
Minimum       0.00    1  0.362   2.3
1st Quartile  1.30   20  0.959   9.1
Median        1.82   66  1.300  20.0
Mean          1.99  453  1.510 119.0
3rd Quartile  2.93  847  2.310 203.0
Maximum       3.56 3640  2.860 722.0
```

Box Plots



Exploratory Plots



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

The x- and y-axis labels for a given bivariate plot are defined by the intersecting row and column labels.

Basic Model Statistics

Number of Observations	35
Standard error (RMSE)	0.258
Average Model standard percentage error (MSPE)	62.9
Coefficient of determination (R^2)	0.915
Adjusted Coefficient of Determination (Adj. R^2)	0.912
Bias Correction Factor (BCF)	1.18

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	0.345	0.0975	3.54	1.21e-03
logTBY	1.09	0.0578	18.80	3.20e-19

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.895
E.vars	-0.895	1.000

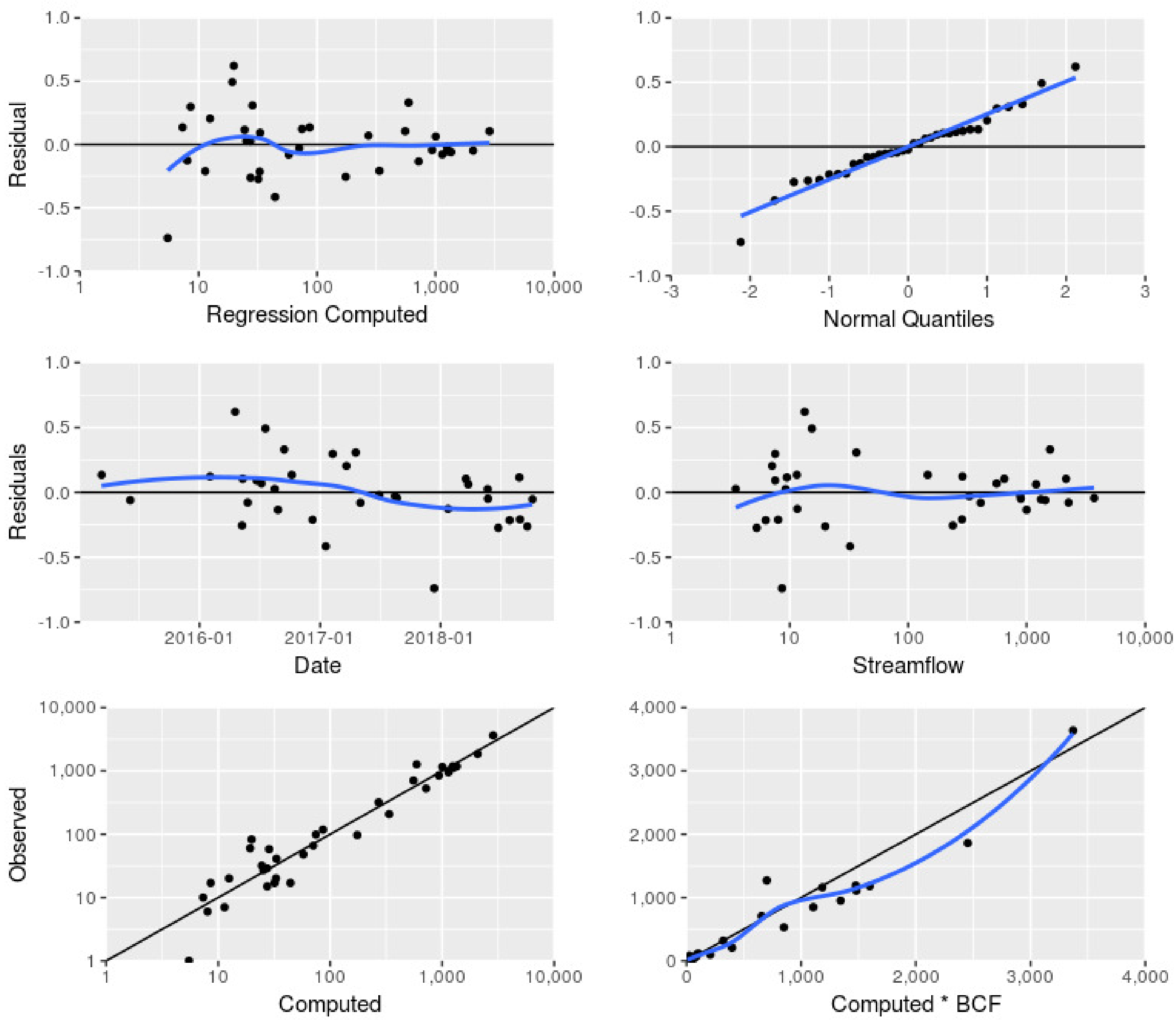
Outlier Test Criteria

Leverage	Cook's D	DFFITS
0.171	0.194	0.478

Flagged Observations

	logSS	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
2016-04-18 10:15:00	1.92	1.300	0.621	2.47	2.70	0.0489	0.157	0.612
2017-12-13 10:30:00	0.00	0.739	-0.739	-3.02	-3.49	0.0949	0.477	-1.130

Statistical Plots



First row (left): residual SS (in log-space units) related to regression-computed SS (in mg/L) with local polynomial regression fitting, or locally estimated scatterplot smoothing (LOESS) indicated by the blue line.

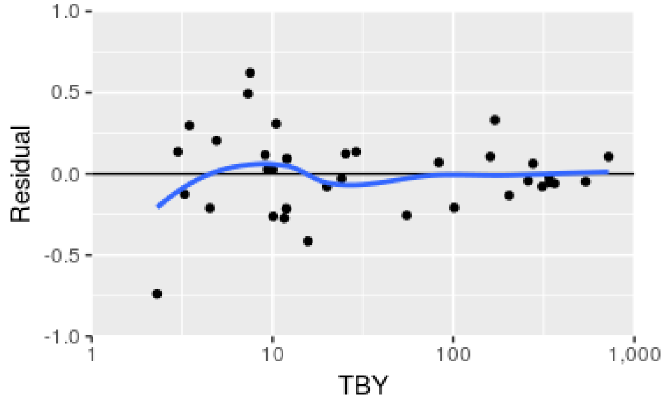
First row (right): residual SS (in log-space units) related to the corresponding normal quantile (unitless) of the residual with simple linear regression indicated by the blue line.

Second row (left): residual SS (in log-space units) related to date with LOESS indicated by the blue line.

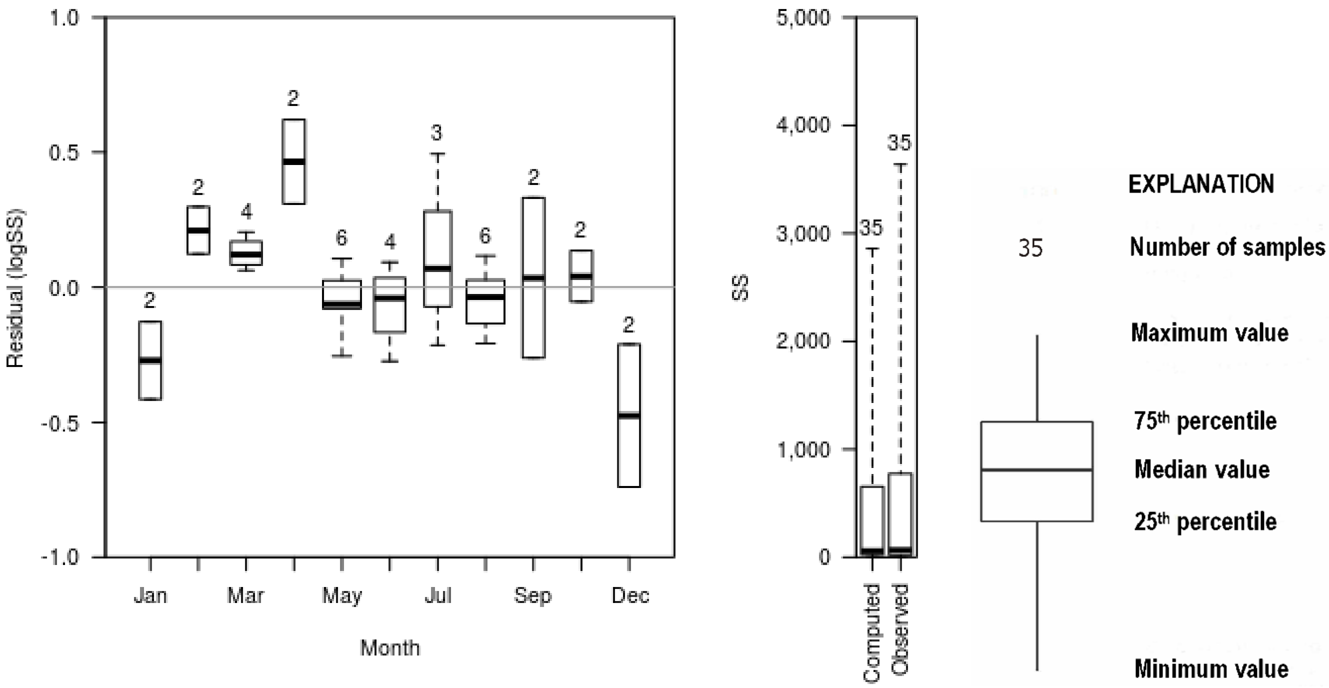
Second row (right): residual SS (in log-space units) related to streamflow (in cubic feet per second) with LOESS indicated by the blue line.

Third row (left): observed SS (in mg/L) related to regression-computed SS (in mg/L).

Third row (right): observed SS (in mg/L) related to the product of regression-computed SS (in mg/L) and the bias correction factor with LOESS indicated by the blue line.

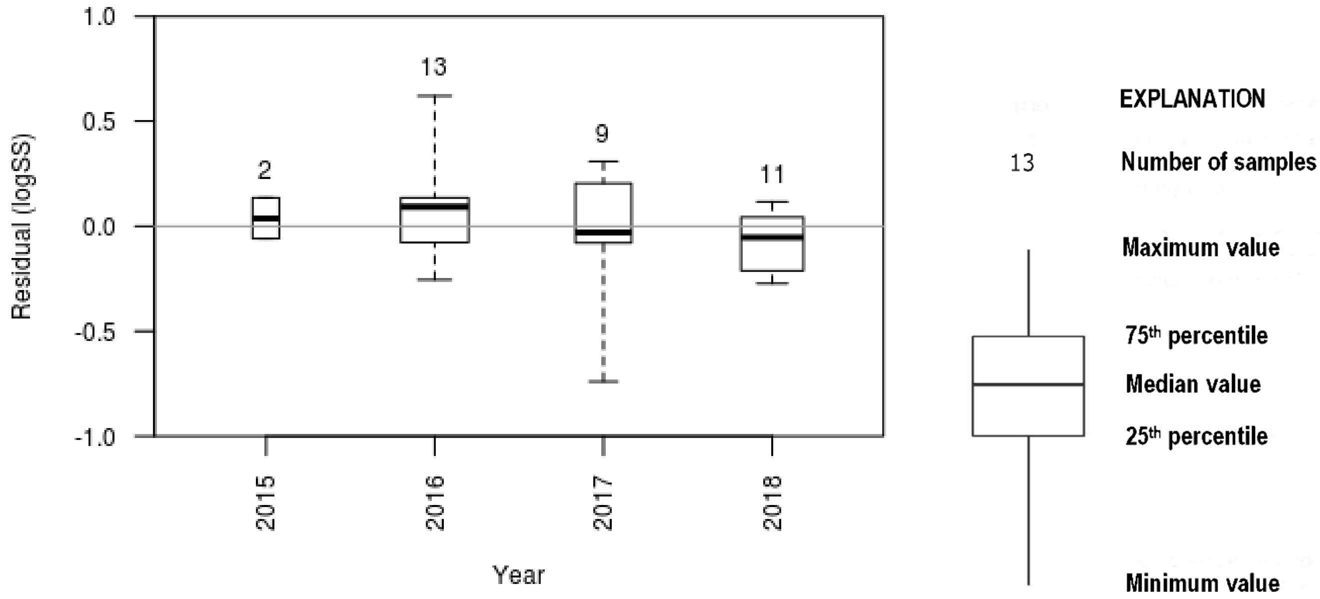


Residual SS (in log-space units) related to TBY (in FNU) with LOESS indicated by the blue line.



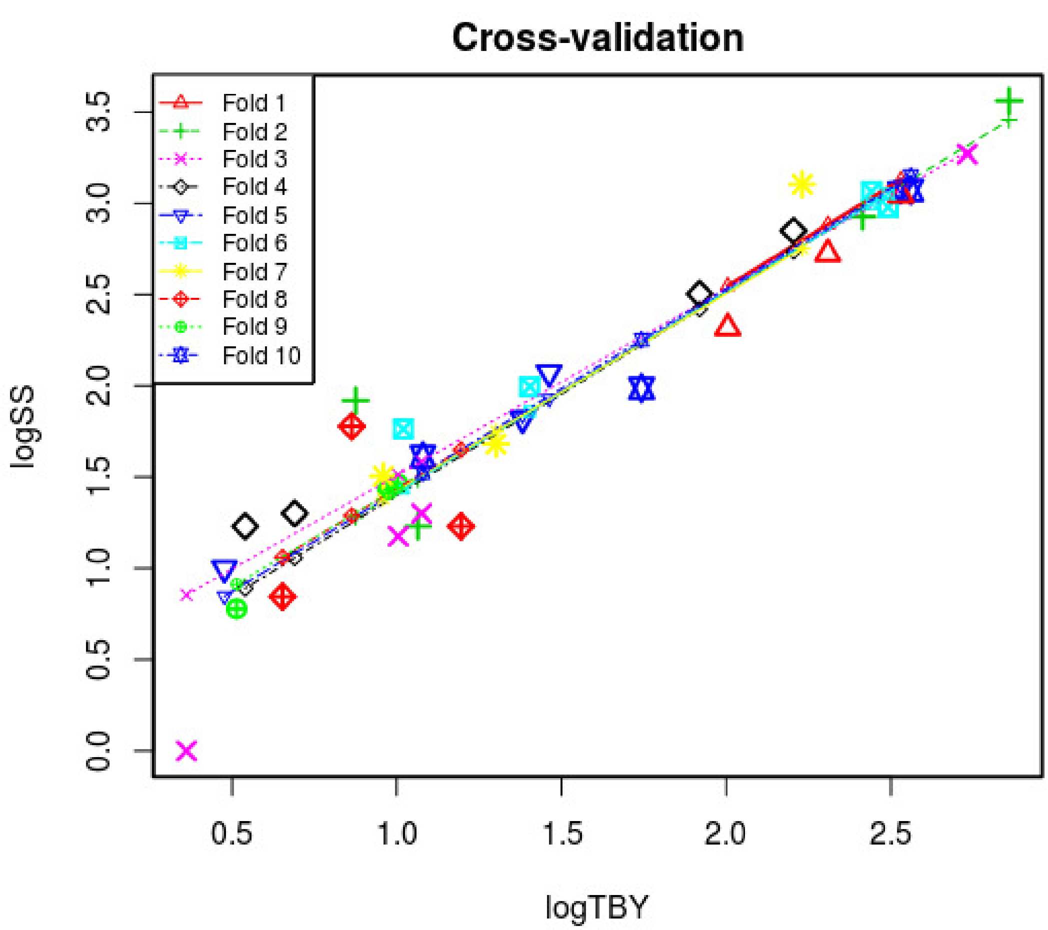
Left: residual SS (in log-space units) by month.

Right: SS (in mg/L) in regression-computed and observed values.

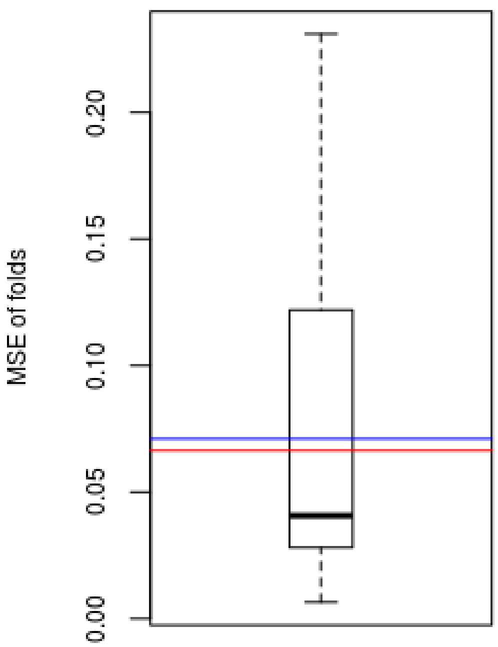


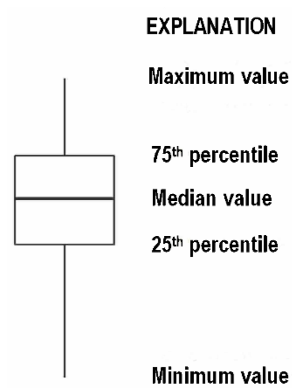
Residual SS (in log-space units) by year.

Cross Validation



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

Large symbols: observed values of data points removed in a fold

Small symbols: recomputed values of data points removed in a fold

Recomputed regression lines: adjusted regression lines with one fold removed


              Minimum MSE of folds:  0.00637
                 Mean MSE of folds:  0.07110
               Median MSE of folds:  0.04060
              Maximum MSE of folds:  0.23100
 (Mean MSE of folds) / (Model MSE):  1.07000
```






```
Red line - Model MSE (unitless)
Blue line - Mean MSE of folds (unitless)
```

Model-Calibration Data Set

	Date	logSS	logTBY	SS	TBY	Computed logSS	Computed SS	Residual	Normal Quantiles	Censored Values
1	2015-03-10	1	0.477	10	3	0.865	8.64	0.135	0.887	--
2	2015-06-05	3.07	2.56	1180	363	3.13	1600	-0.0596	-0.364	--
3	2016-02-02	2	1.4	99	25.3	1.87	88.1	0.123	0.692	--
4	2016-04-18	1.92	0.875	83	7.5	1.3	23.4	0.621	2.12	--
5	2016-05-09	1.99	1.74	97	55.3	2.24	206	-0.255	-1.12	--
6	2016-05-11	2.85	2.2	707	160	2.74	654	0.106	0.521	--
7	2016-05-26	2.98	2.49	951	310	3.06	1340	-0.0783	-0.441	--
8	2016-06-22	1.61	1.08	41	12	1.52	39.1	0.093	0.364	--
9	2016-07-07	2.5	1.92	319	83	2.43	320	0.0701	0.289	--

10	2016-07-19	1.78	0.863	60	7.3	1.28	22.7	0.493	1.69	--
11	2016-08-16	1.43	0.973	27	9.4	1.4	29.9	0.027	0.143	--
12	2016-08-26	2.72	2.31	529	203	2.86	849	-0.134	-0.692	--
13	2016-09-14	3.1	2.23	1270	170	2.77	699	0.331	1.45	--
14	2016-10-07	2.07	1.46	118	29	1.94	102	0.135	0.786	--
15	2016-12-09	0.845	0.653	7	4.5	1.06	13.4	-0.211	-0.887	--
16	2017-01-18	1.23	1.19	17	15.7	1.65	52.2	-0.415	-1.69	--
17	2017-02-08	1.23	0.54	17	3.47	0.933	10.1	0.297	1.12	--
18	2017-03-22	1.3	0.69	20	4.9	1.1	14.7	0.205	0.998	--
19	2017-04-19	1.76	1.02	58	10.5	1.46	33.7	0.308	1.27	--
20	2017-05-03	1.68	1.3	48	20	1.76	68.1	-0.0799	-0.521	--
21	2017-06-29	3.08	2.53	1190	338	3.1	1480	-0.0218	0	--
22	2017-08-16	1.82	1.38	66	24.1	1.85	83.4	-0.0298	-0.0713	--
23	2017-08-22	2.93	2.41	847	259	2.97	1110	-0.0436	-0.143	--
24	2017-12-13	0	0.362	1	2.3	0.739	6.47	-0.739	-2.12	--
25	2018-01-24	0.778	0.514	6	3.27	0.905	9.48	-0.127	-0.605	--
26	2018-03-19	3.56	2.86	3640	722	3.46	3370	0.105	0.441	--
27	2018-03-26	3.06	2.44	1160	276	3	1180	0.0629	0.215	--
28	2018-05-24	1.46	1	29	10.1	1.44	32.3	0.0256	0.0713	--
29	2018-05-25	3.27	2.73	1860	539	3.32	2450	-0.0484	-0.215	--
30	2018-06-25	1.23	1.06	17	11.6	1.5	37.6	-0.273	-1.45	--
31	2018-07-30	1.3	1.08	20	11.9	1.52	38.7	-0.215	-0.998	--
32	2018-08-28	1.51	0.959	32	9.1	1.39	28.9	0.116	0.605	--
33	2018-08-30	2.32	2	208	101	2.53	397	-0.208	-0.786	--
34	2018-09-21	1.18	1	15	10.1	1.44	32.4	-0.262	-1.27	--
35	2018-10-07	3.05	2.53	1110	339	3.1	1480	-0.0534	-0.289	--

Definitions

Cook’s D: Cook’s distance, a measure of influence (Helsel and others, 2020).

DFFITS: difference in fits, a measure of influence (Helsel and others, 2020).

E.vars: explanatory variables.

Leverage: a data point’s distance from the middle (mean) value in the x direction (Helsel and others, 2020).

LOESS: local polynomial regression fitting, or locally estimated scatterplot smoothing (Helsel and others, 2020).

LOWESS: locally weighted scatterplot smoothing (Cleveland, 1979; Helsel and others, 2020).

MSE: model standard error, also known as standard error of the regression (Helsel and others, 2020).

MSPE: model standard percentage error (Helsel and others, 2020).

Pr(>|t|): the probability that the independent variable has no effect on the dependent variable (Helsel and others, 2020).

```
RMSE: root mean square error (Helsel and others, 2020).

SS: suspended sediment concentration, milligrams per liter (U.S. Geological Survey parame
ter code 80154).

t value: Student’s t value; the coefficient divided by its associated standard error (Hel
sel and others, 2020).

TBY: Turbidity, water, unfiltered, monochrome near infra-red LED light, 780-900 nm, detec
tion angle 90 +-2.5 degrees, formazin nephelometric units (FNU) (U.S. Geological Survey p
arameter code 63680).
```

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.

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chapter A3, 458 p., https://doi.org/10.3133/tm4a3. [Supersedes USGS Techniques of Water-Resources Investigations, book 4, chapter A3, version 1.1.]

Rasmussen, T.J., Lee, C.J., and Ziegler, A.C., 2008, Estimation of constituent concentrations, loads, and yields in streams of Johnson County, northeast Kansas, using continuous water-quality monitoring and regression models, October 2002 through December 2006: U.S. Geological Survey Scientific Investigations Report 2008–5014, 103 p. [Also available at https://pubs.usgs.gov/sir/2008/5014/.]

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 1, 2019, at https://doi.org/10.5066/F7P55KJN.

Appendix 6. Historical and Project Data Comparison

Discrete water-quality data from sites in Johnson County where at least one sample was collected by the U.S. Geological Survey and analyzed for at least one nutrient since 1994 are displayed in figure 6.1. Sites that represent effluent from wastewater treatment facilities were excluded from comparisons in this appendix and from figure 6.1. There are 111 sites with historical nutrient data collected before this study began in 2015; 8 of those sites also had data collected during this study. Historical data labeled as routine samples are generally comparable to the low-flow samples collected in the current study although preceding rainfall conditions are unknown for the historical routine samples. Historical and current-study data are stored in the USGS National Water Information System (NWIS) database (U.S. Geological Survey, 2019).

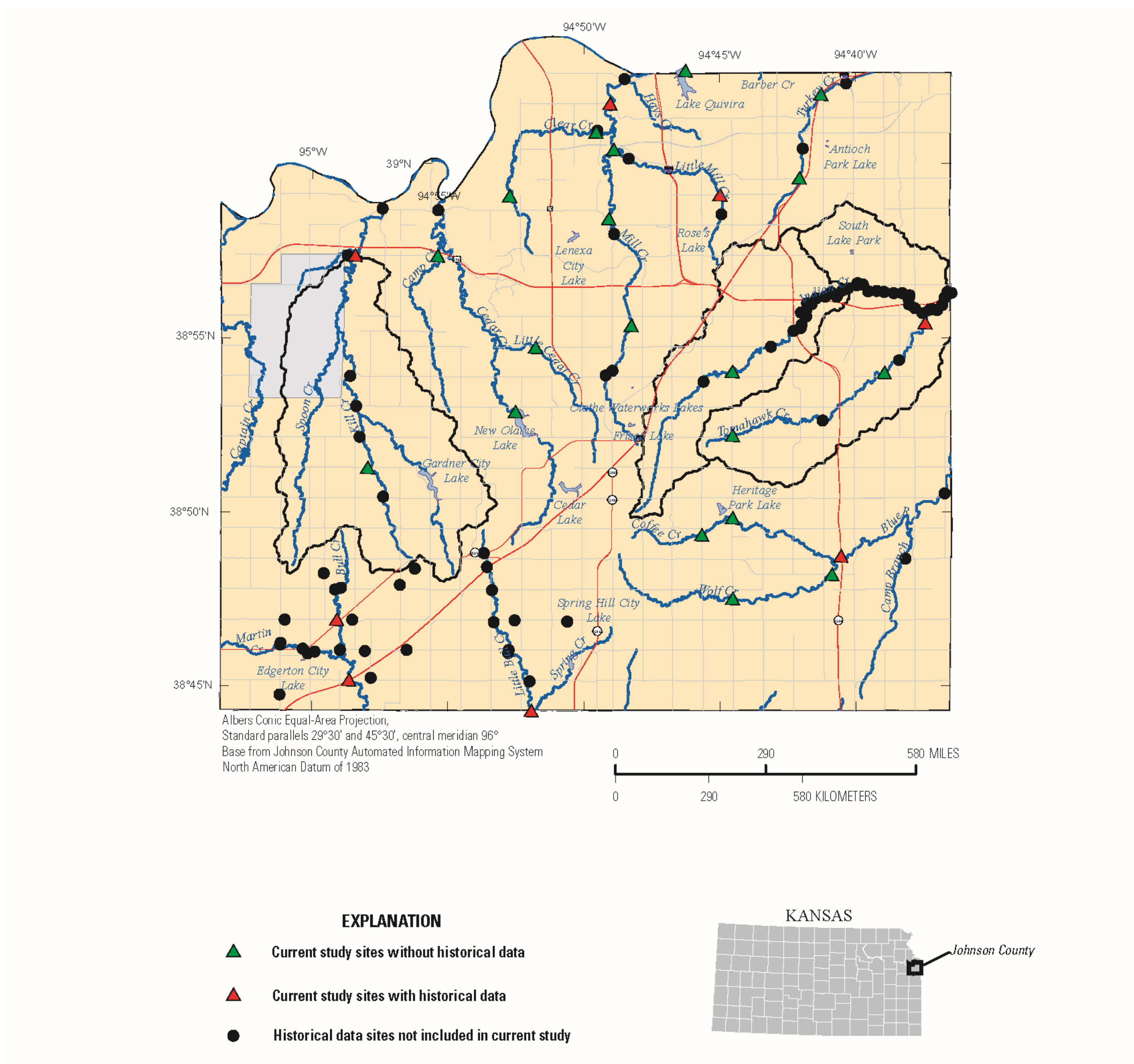


Figure 6.1. Water-quality data sites with historical data, data collected in the current study, or both, Johnson County, Kansas.

Blue River Watershed

There were three sites in the Blue River watershed with historical nutrient data (table 6.1). Fifty-one historical samples were collected between 2003 and 2011, 32 of which were storm-event samples and 19 of which were routine samples. In the current study, there were 95 storm-event and 20 routine (low-flow) samples collected in this watershed (sites 23–27, fig. 1 and table 2 in main body of report). No differences between historical and current-study data were apparent in either storm-event samples (fig. 6.2*A*) or in routine samples (fig. 6.2*B*), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.1. Water-quality data sites with historical data and (or) data collected in the current study, Blue River watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Blue River nr Stanley, KS	06893080	0	3	Yes
Blue River at Kenneth Road, Overland Park, KS	06893100	32	13	No
Camp Branch at 175th Street, Johnson Co, KS	384840094381100	0	3	No

A. Storm-event samples by nutrient

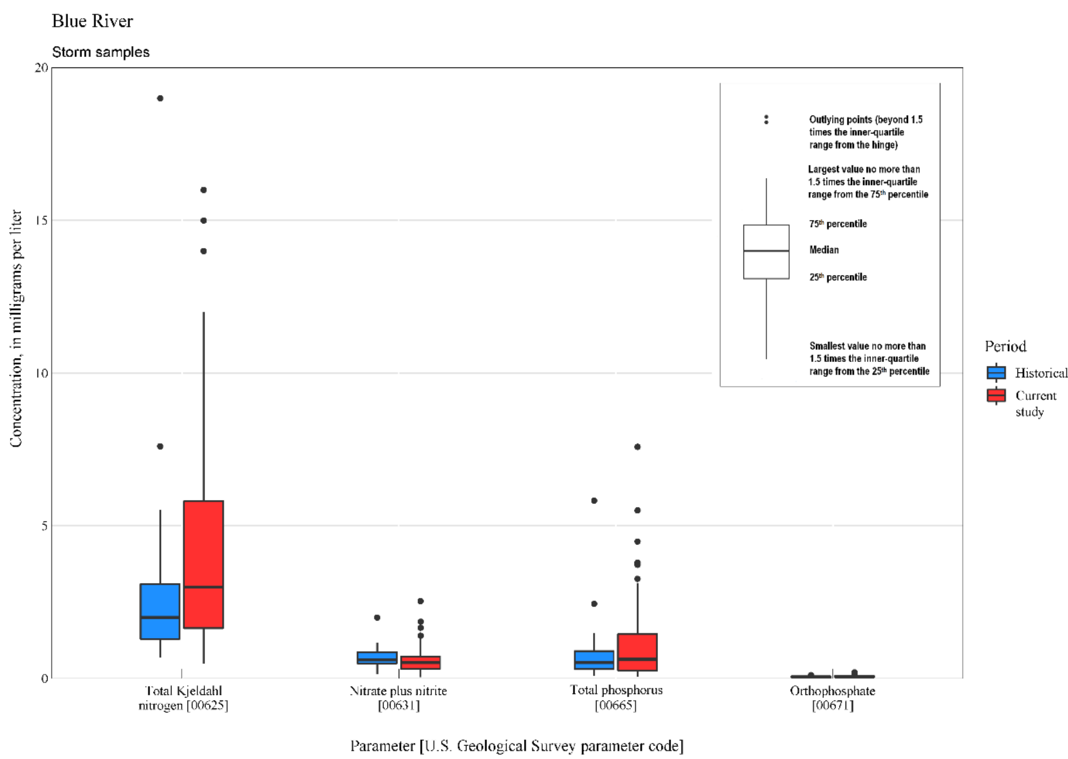


Figure 6.2. Comparison of historical and current-study nutrient data in the Blue River watershed. *A*. Storm-event samples. *B*. Routine samples.

B. **Routine samples by nutrient**

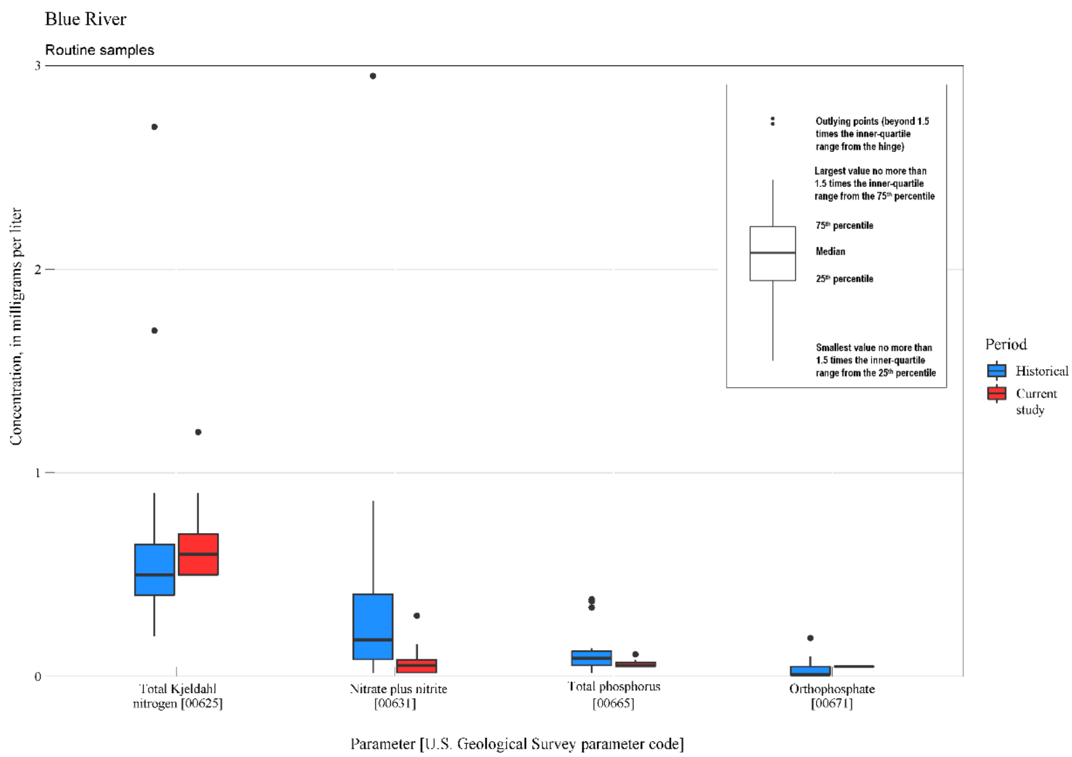


Figure 6.2. Comparison of historical and current-study nutrient data in the Blue River watershed. *A*. Storm-event samples. *B*. Routine samples.

Bull/Little Bull Creek Watershed

There were 19 sites in the (Big) Bull/Little Bull Creek watershed with historical nutrient data (table 6.2). Eighty-two historical samples were collected between 1994 and 2010, all of which were routine samples. In the current study, there were 56 storm-event and 12 routine (low-flow) samples collected in this watershed (sites 6–8, fig. 1 and table 2 in main body of report). No differences between historical and current-study data were apparent in routine samples (fig. 6.3), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.2. Water-quality data sites with historical data and (or) data collected in the current study, Bull/Little Bull Creek watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Big Bull C 1 mile US of Big Bull Gage, KS (BB5)	384603094585700	0	5	No
Big Bull C at 191st St, Edgerton, KS	384656094590400	0	4	Yes
Big Bull C nr Edgerton, KS	06914950	0	9	Yes
Big Bull E Fk Trib nr Gardner, KS (BB12)	384750094585500	0	3	No
Big Bull Tr nr Edgerton, KS	06914948	0	5	No

Big Bull Trib Blw Gardner, KS (BB9)	384823094561200	0	5	No
Big Bull Trib nr Edgerton, KS (BB2)	384515094574900	0	4	No
Big Bull Trib nr Edgerton, KS (BB3)	384603094563000	0	4	No
Big Bull Trib nr Edgerton, KS (BB4)	384601094580200	0	4	No
Big Bull Trib nr Gardner, KS (BB8)	384755094564500	0	5	No
Big Bull W Fk nr Gardner, KS (BB10)	384747094590800	0	3	No
Big Bull W Fk nr Gardner, KS (BB11)	384815094593300	0	4	No
Martin C Ab Southlake nr Edgerton, KS (MC5)	384446095011000	0	4	No
Martin C at Edgerton, KS (MC1)	384605095001900	0	4	No
Martin C nr Edgerton, KS (MC 7)	384557095000800	0	4	No
Martin C nr Edgerton, KS (MC4)	384612095010900	0	4	No
Martin C Trib 2 nr Edgerton, KS (MC3)	384615095010800	0	4	No
Martin C Trib nr Edgerton, KS (MC2)	384655095005900	0	3	No
Martin C Trib nr Santa Fe Lake at Edgerton, KS	384600094595300	0	4	No

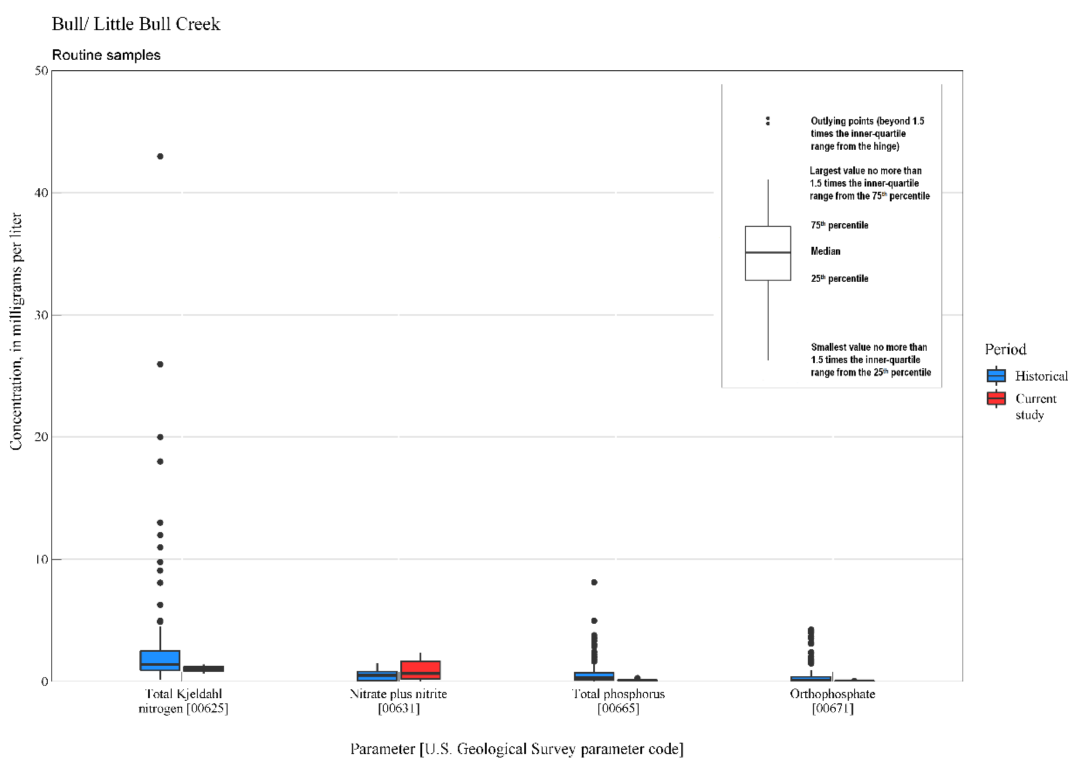


Figure 6.3. Comparison of historical and current-study nutrient data in the Bull/Little Bull Creek watershed in routine samples.

Tomahawk/Indian Creek Watershed

There were 40 sites in the Tomahawk/Indian Creek watershed with historical nutrient data (table 6.3). Two hundred twenty-two historical samples were collected between 2002 and 2014, 102 of which were storm-event samples and 120 of which were routine samples (including 6 snowmelt samples). In the current study, there were 65 storm-event and 13 routine (low-flow) samples collected in this watershed (sites 19–22, fig. 1 and table 2 in main body of report). Routine historical data displayed wider variability in nitrate plus nitrite than current-study samples (fig. 6.4*B*) but no other differences between historical and current-study data were apparent in either storm-event samples (fig. 6.4*A*) or in routine samples (fig. 6.4*B*), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.3. Water-quality data sites with historical data and (or) data collected in the current study, Tomahawk/Indian Creek watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Indian C 0.02 mi W of Roe Ave, Overland Park, KS	385618094383500	0	1	No
Indian C 0.05 mi N of College Blvd, Leawood, KS	385549094371600	0	1	No
Indian C 0.25 mi E of Hwy 69, Overland Park, KS	385524094415900	0	1	No
Indian C 0.25 mi N of Hwy 50, Overland Park, KS	385618094405300	0	1	No
Indian C at 111 St, Johnson Co, KS	385518094420100	0	2	No
Indian C at 119th St, Overland Park, KS	385446094430700	12	9[a]	No
Indian C at Black Bob Rd, Johnson Co, KS	385345094453600	0	2	No
Indian C at College Blvd, Johnson Co, KS	385520094420000	11	10[a]	No
Indian C at Hwy 50, Overland Park, KS	385559094380200	0	1	No
Indian C at Hwy 69, Overland Park, KS	06893270	0	3	No
Indian C at Indian C Pkwy, Overland Park, KS	385608094380300	11	9[a]	No
Indian C at Nall Ave, Overland Park, KS	385620094385700	0	1	No
Indian C at Overland Park, KS	06893300	12	8[a]	No
Indian C at State Line Rd, Leawood, KS	06893390	45	20[a]	No
Indian C nr 103rd St, Overland Park, KS	385633094394400	0	1	No
Indian C nr 105th St, Overland Park, KS	385614094380000	0	1	No
Indian C nr 106th Terr, Overland Park, KS	385609094412600	0	1	No
Indian C nr 109th St, Overland Park, KS	385559094413700	0	1	No
Indian C nr 110th St, Leawood, KS	385552094374900	0	1	No
Indian C nr 111 St, Johnson Co, KS	385550094371100	0	2	No

Indian C nr Antioch Rd, Overland Park, KS	385612094410600	0	1	No
Indian C nr Brookwood Ave, Leawood, KS	385549094370000	0	1	No
Indian C nr Conser St, Overland Park, KS	385623094402900	0	1	No
Indian C nr Country Club Dr, Overland Park, KS	385612094404100	0	1	No
Indian C nr Farley St, Overland Park, KS	385600094414700	0	1	No
Indian C nr Hwy 50, Leawood, KS	385610094364300	0	1	No
Indian C nr Indian C Dr, Overland Park, KS	385622094391300	0	1	No
Indian C nr Lamar Ave, Overland Park, KS	385621094393200	0	1	No
Indian C nr Lee Blvd, Leawood, KS	385557094364900	0	1	No
Indian C nr Mastin St, Overland Park, KS	385548094414900	0	1	No
Indian C nr Metcalf Ave, Overland Park, KS	385632094395800	0	1	No
Indian C nr Mission Rd, Leawood, KS	385544094373200	0	1	No
Indian C nr Roe Ave, Overland Park, KS	385618094381700	0	1	No
Indian C Trib nr 104th St, Leawood, KS	385620094363700	0	1	No
Indian C Trib nr 109th St, Overland Park, KS	385545094420100	0	1	No
Indian C Trib nr Antioch Rd, Overland Park, KS	385619094405400	0	1	No
Indian C Trib nr Metcalf Ave, Overland Park, KS	385635094395500	0	1	No
Tomahawk C at Antioch Rd, Johnson Co, KS	385238094411300	0	2	No
Tomahawk C nr 111th St, Johnson Co, KS	385539094372100	0	4	Yes
Tomahawk C nr Overland Park, KS	06893350	11	21[a]	No

[a] Includes one snowmelt sample.

A. Storm-event samples by nutrient

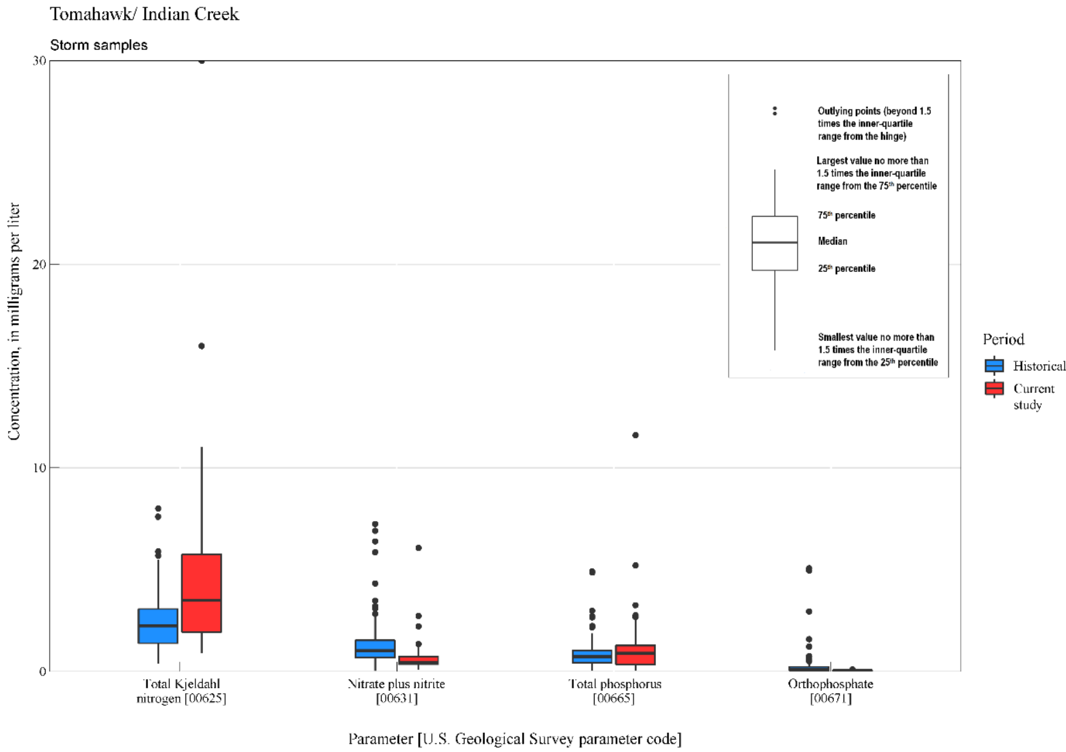


Figure 6.4. Comparison of historical and current-study nutrient data in the Tomahawk/Indian Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

B. **Routine samples by nutrient**

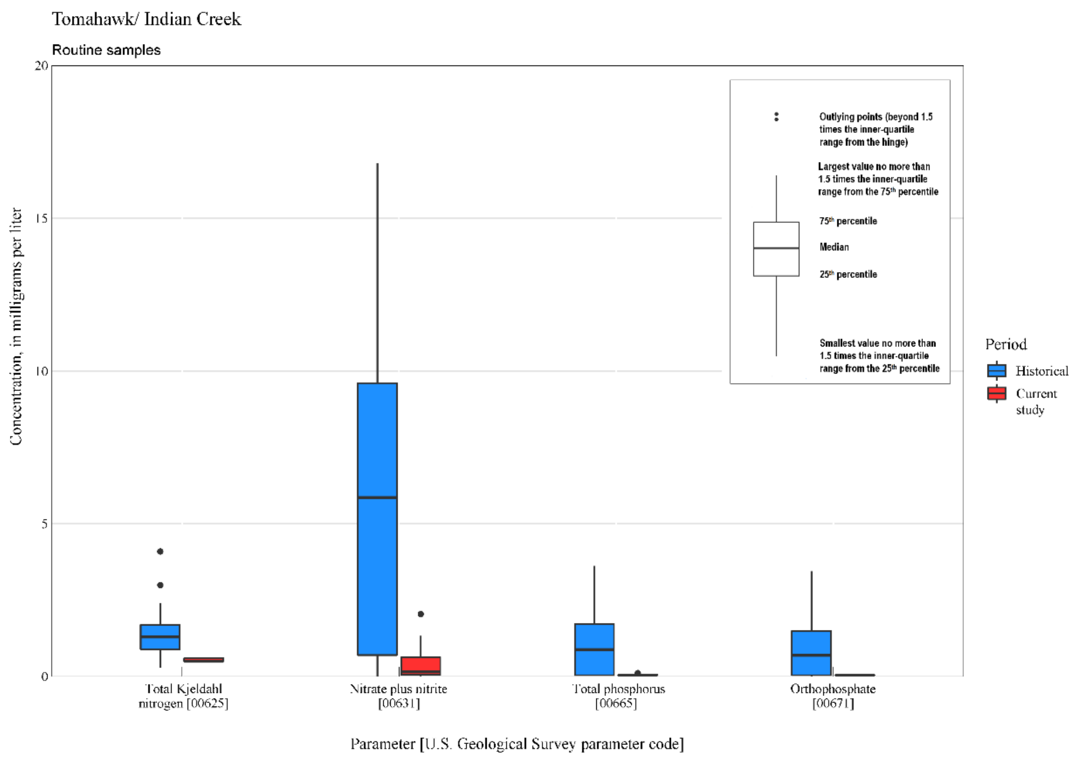


Figure 6.4. Comparison of historical and current-study nutrient data in the Tomahawk/Indian Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

Kill Creek Watershed

There were seven sites in the Kill Creek watershed with historical nutrient data (table 6.4). Forty-three historical samples were collected between 2001 and 2010, 21 of which were storm-event samples and 22 of which were routine samples. Historical samples at Kill Creek at 95th Street were collected between 2003 and 2010. In the current study, there were 28 storm-event and 7 routine (low-flow) samples collected in this watershed (sites 1–2, fig. 1 and table 2 in main body of report). Routine current-study data displayed wider variability in nitrate plus nitrite, phosphorus, and orthophosphate than historical samples (fig. 6.5*B*) but no other differences between historical and current-study data were apparent in either storm-event samples (fig. 6.5*A*) or in routine samples (fig. 6.5*B*), and statistically significant determinations were not possible due to the compositions of the datasets. At Kill Creek at 95th Street, the current-study data generally displayed wider variability than historical samples (fig. 6.6), but no other differences between historical and current-study data were apparent in either storm-event samples (fig. 6.6*A*) or in routine samples (fig. 6.6*B*), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.4. Water-quality data sites with historical data and (or) data collected in the current study, Kill Creek watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Kill C at 127th St at DeSoto, KS	06892359	0	2	No

Kill C at 135 St, Johnson Co, KS	385303094582300	0	1	No
Kill C at 143 St, Johnson Co, KS	385210094581500	0	1	No
Kill C at 159 St, Johnson Co, KS	385027094572300	0	1	No
Kill C at 83 St, Johnson Co, KS	385844094572500	0	2	No
Kill C at 83rd St, KS	385723094584200	0	3	No
Kill C at 95th St nr DeSoto, KS	06892360	21	12	Yes

***A.* Storm-event samples by nutrient**

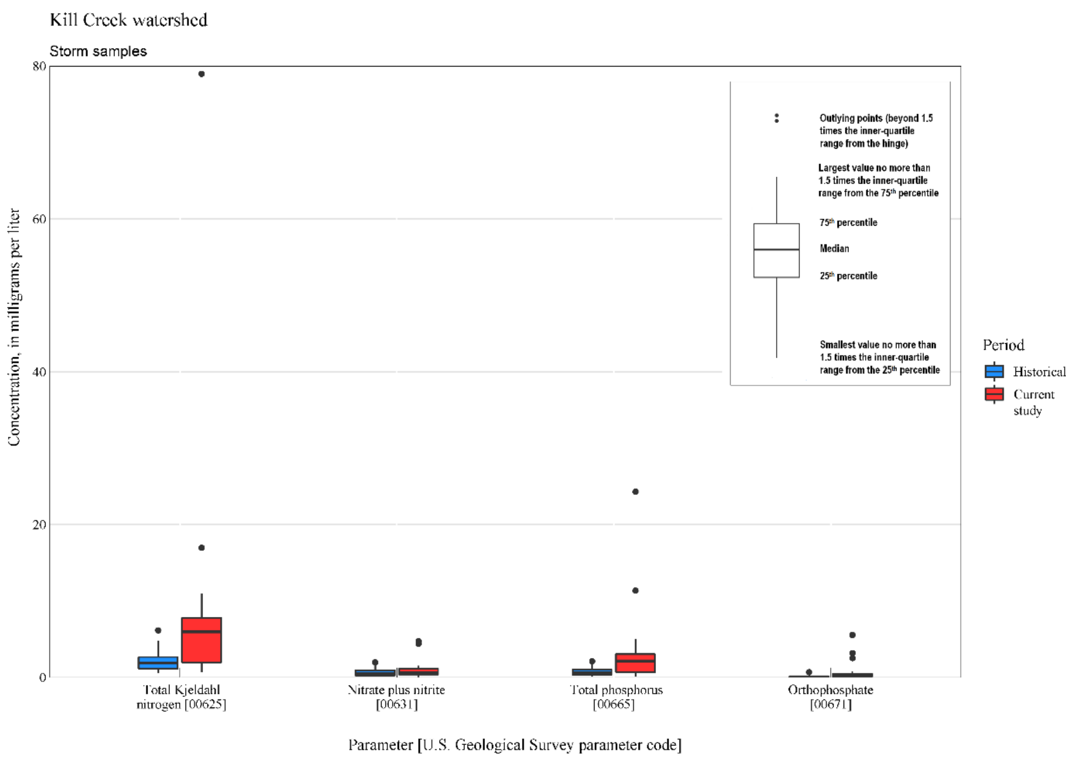


Figure 6.5. Comparison of historical and current-study nutrient data in the Kill Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

B. Routine samples by nutrient

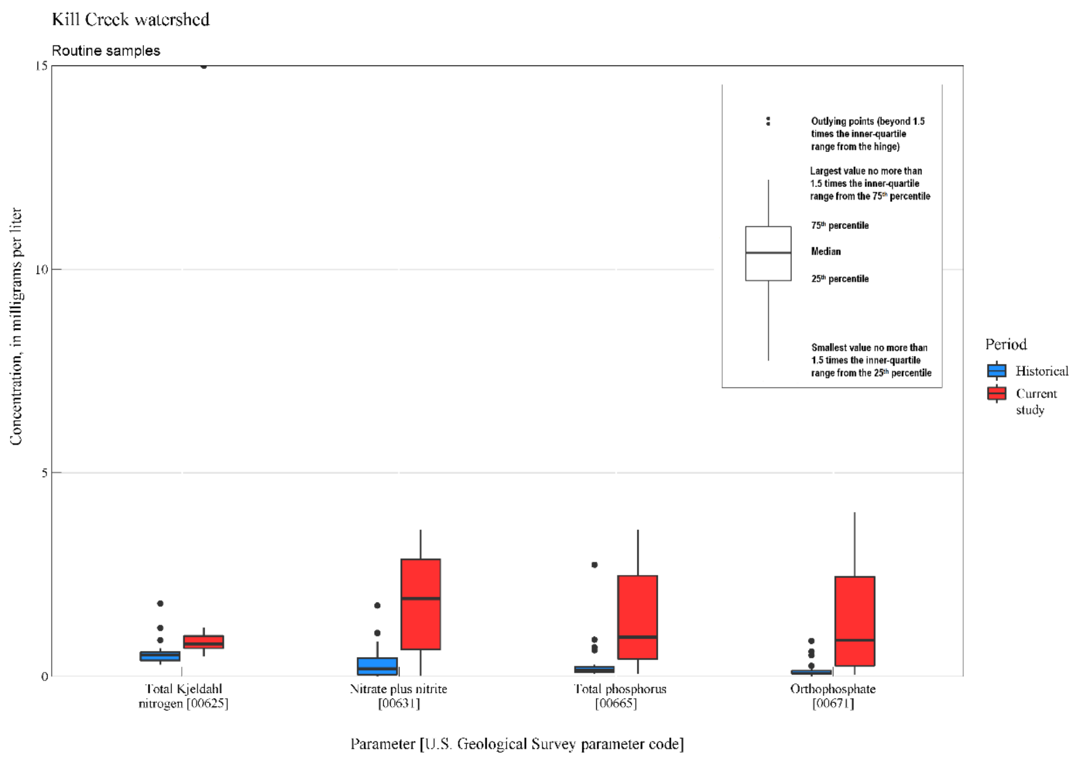


Figure 6.5. Comparison of historical and current-study nutrient data in the Kill Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

A. **Storm-event samples by nutrient**

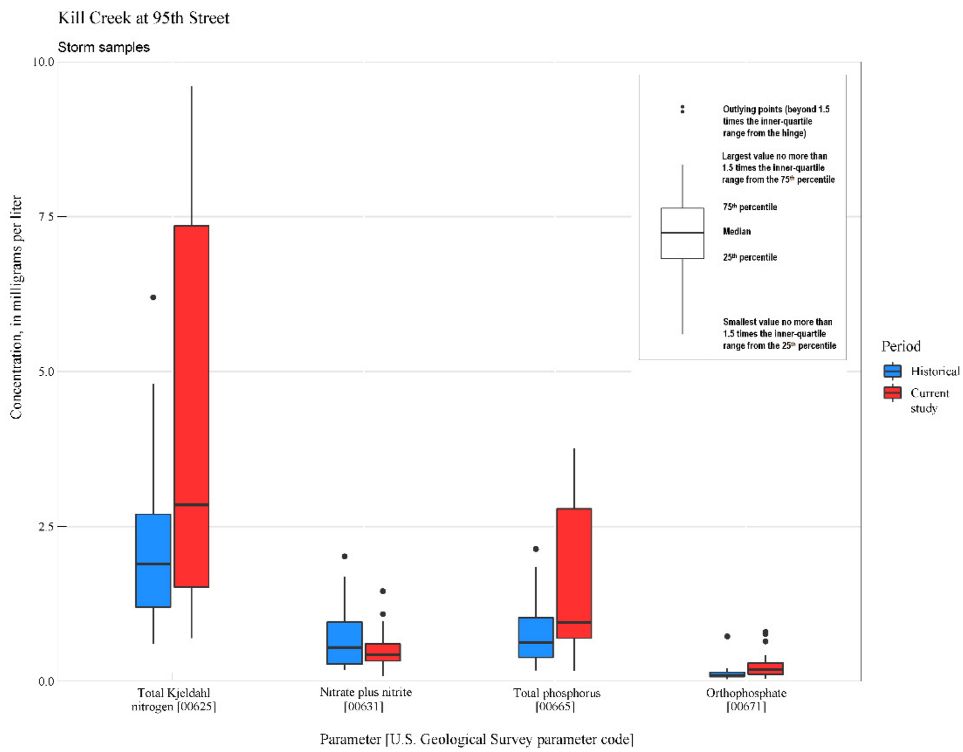


Figure 6.6. Comparison of historical and current-study nutrient data at Kill Creek at 95th Street. *A*. Storm-event samples. *B*. Routine samples.

B. Routine samples by nutrient

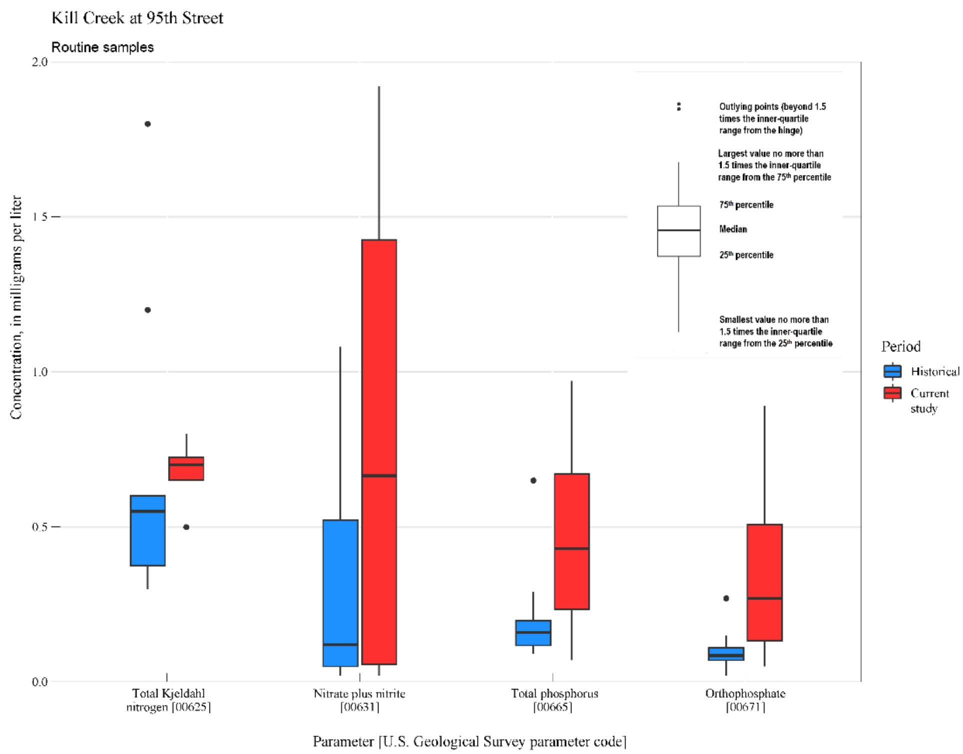


Figure 6.6. Comparison of historical and current-study nutrient data at Kill Creek at 95th Street. *A*. Storm-event samples. *B*. Routine samples.

Clear/Little Mill/Mill Creek Watershed

There were nine sites in the Clear/Little Mill/Mill Creek watershed with historical nutrient data (table 6.5). Fifty-two historical samples were collected between 2001 and 2010, 21 of which were storm-event samples and 31 of which were routine samples. Historical samples at Mill Creek at Johnson Drive were collected between 2002 and 2010. In the current study, there were 148 storm-event and 73 routine (low-flow) samples collected in this watershed (sites 9–15, fig. 1 and table 2 in main body of report). No differences between historical and current-study data were apparent in either storm-event samples or in routine samples across the watershed and at Mill Creek at Johnson Drive (figs. 6.7, 6.8), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.5. Water-quality data sites with historical data and (or) data collected in the current study, Clear/Little Mill/Mill Creek watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Clear C at 63 St, Johnson Co, KS	390058094493000	0	1	No
L Mill at 79th St, Lenexa, KS	385908094445900	0	2	Yes
L Mill C at W 84th Terr, Lenexa, KS	385834094445600	0	1	No
L Mill C at Warwick Ln, Shawnee, KS	390010094482100	0	2	No

Mill C at 127 St, Johnson Co, KS	385356094491200	0	4	No
Mill C at 87th St Ln, Johnson Co, KS	385800094485300	0	4	No
Mill C at Johnson Drive, Shawnee, KS	06892513	21	12	Yes
Mill C at Wilder Rd, KS	390227094483000	0	3	No
Mill C nr Northgate Rd, Johnson Co, KS	385404094485700	0	2	No

***A.* Storm-event samples by nutrient**

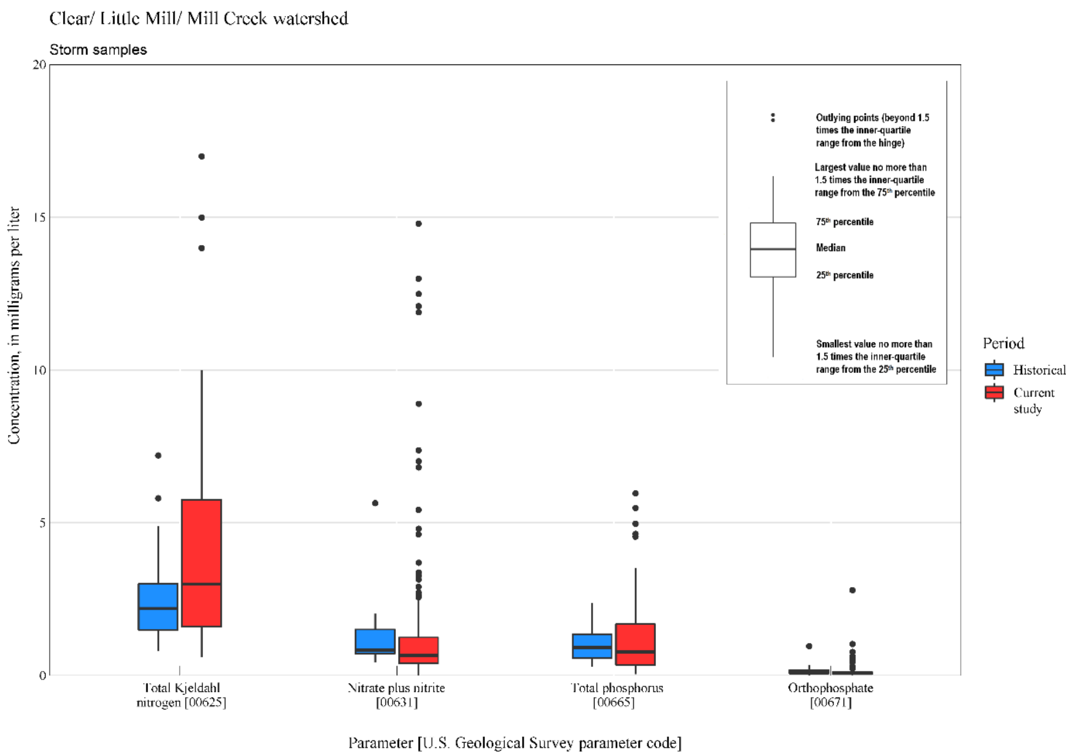


Figure 6.7. Comparison of historical and current-study nutrient data in the Clear/Little Mill/Mill Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

B. Routine samples by nutrient

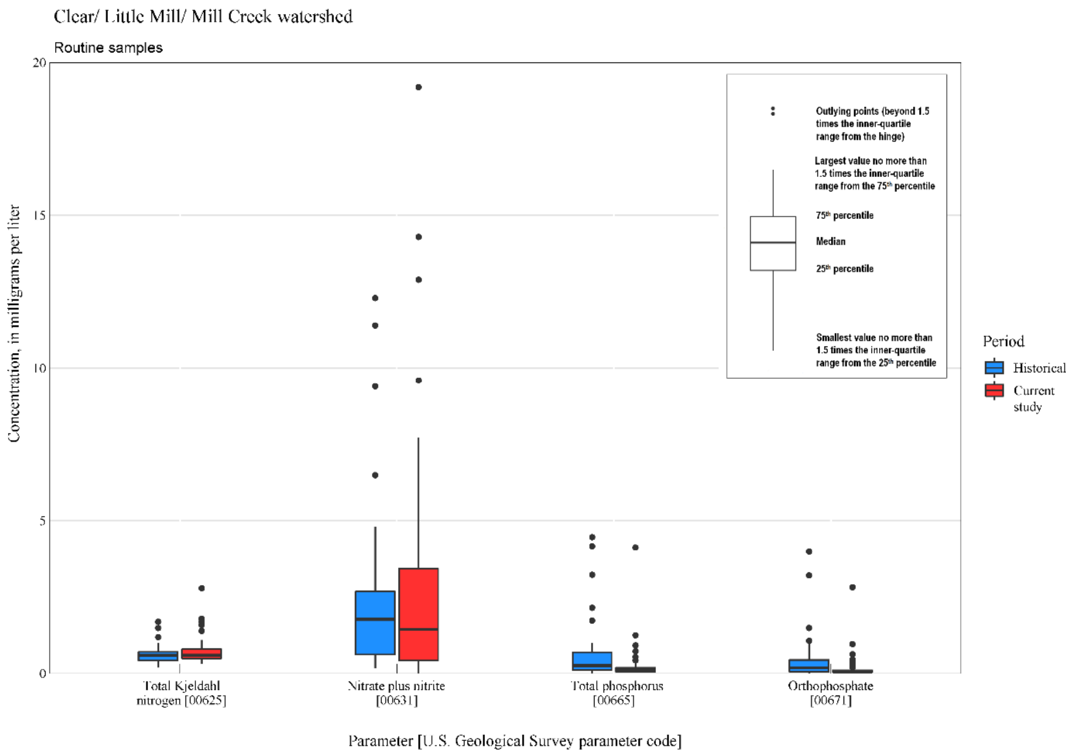


Figure 6.7. Comparison of historical and current-study nutrient data in the Clear/Little Mill/Mill Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

A. Storm-event samples by nutrient

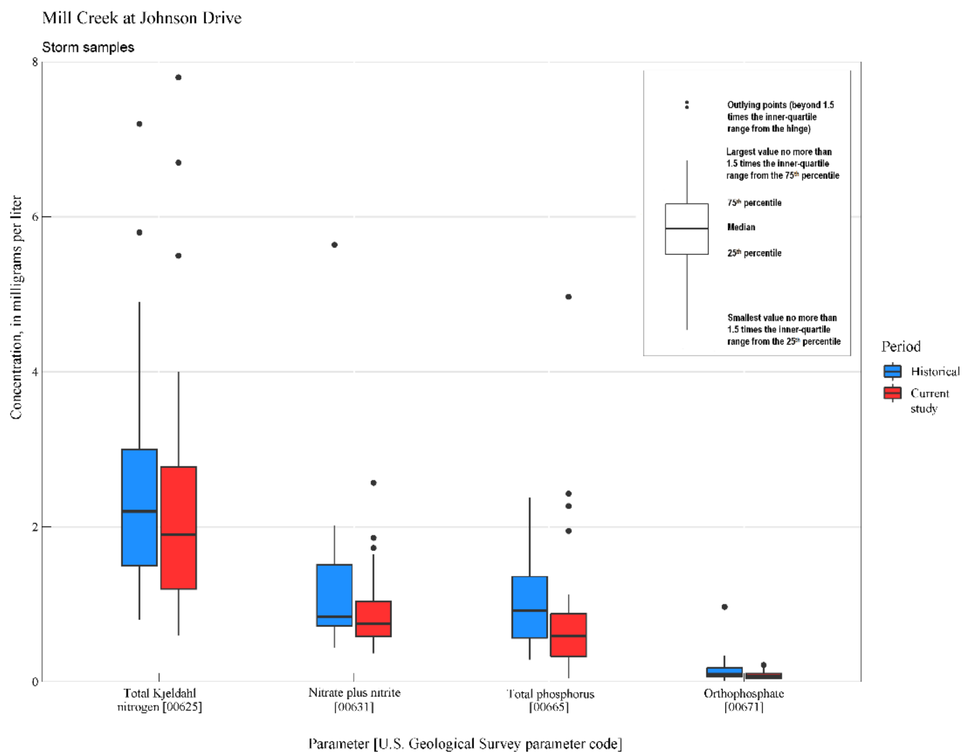


Figure 6.8. Comparison of historical and current-study nutrient data at Mill Creek at Johnson Drive. *A*. Storm-event samples. *B*. Routine samples.

B. Routine samples by nutrient

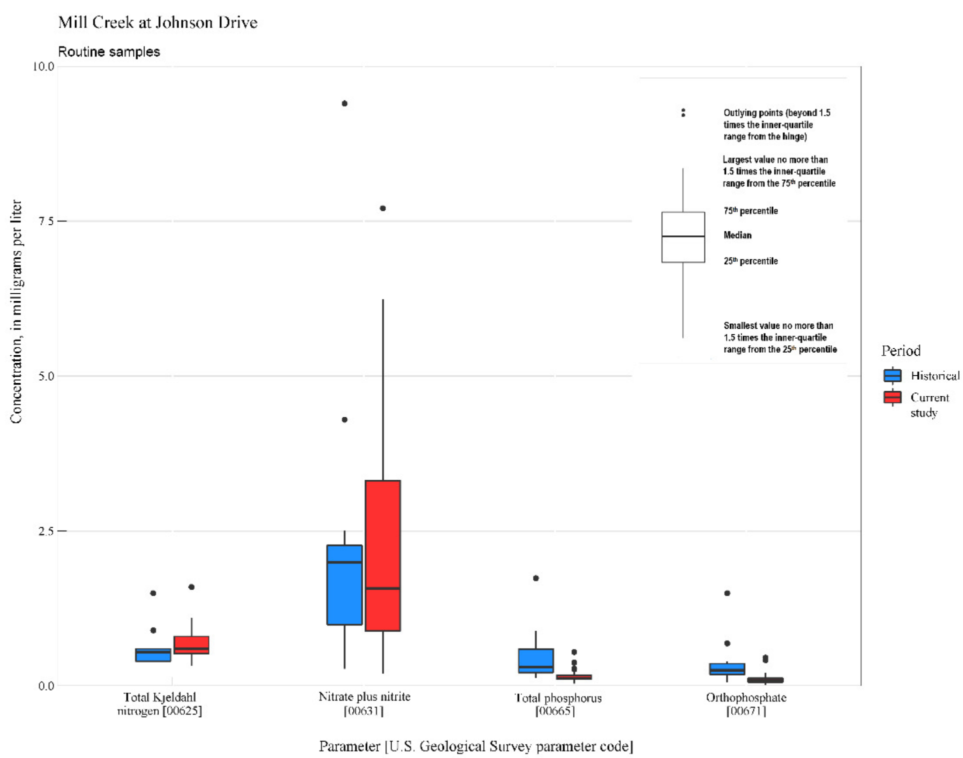


Figure 6.8. Comparison of historical and current-study nutrient data at Mill Creek at Johnson Drive. *A*. Storm-event samples. *B*. Routine samples.

Turkey Creek Watershed

There were three sites in the Turkey Creek watershed with historical nutrient data (table 6.6). Thirteen historical samples were collected between 2001 and 2010, four of which were storm-event samples and nine of which were routine samples. In the current study, there were 40 storm-event and 8 routine (low-flow) samples collected in this watershed (sites 17–18, fig. 1 and table 2 in main body of report). Current-study storm samples displayed wider variability in Kjeldahl nitrogen (fig. 6.9*A*) and historical routine samples displayed wider variability in all four nutrients (fig. 6.9*B*) but no other differences between historical and current-study data were apparent in either storm-event samples (fig. 6.9*A*) or in routine samples (fig. 6.9*B*), and statistically significant determinations were not possible due to the compositions of the datasets.

Table 6.6. Water-quality data sites with historical data and (or) data collected in the current study, Turkey Creek watershed, Johnson County, Kansas.

Site name	Site number	Number of historical storm-event samples	Number of historical routine samples	Samples collected in current study?
Turkey C at 67 St, Johnson Co, KS	390027094415600	4	4	No
Turkey C at Hwy 635, Johnson Co, KS	390219094402000	0	2	No
Turkey C at Kansas City nr I-35, KS	390424094365400	0	3	No

***A.* Storm-event samples by nutrient**

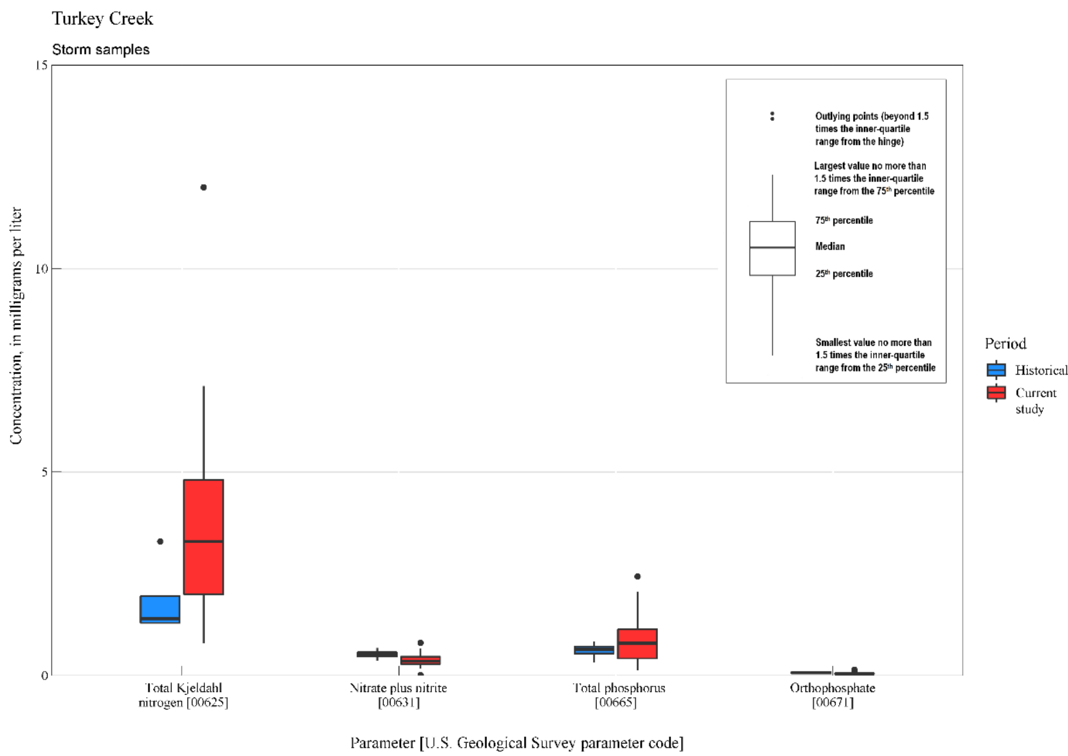


Figure 6.9. Comparison of historical and current-study nutrient data in the Turkey Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

B. **Routine samples by nutrient**

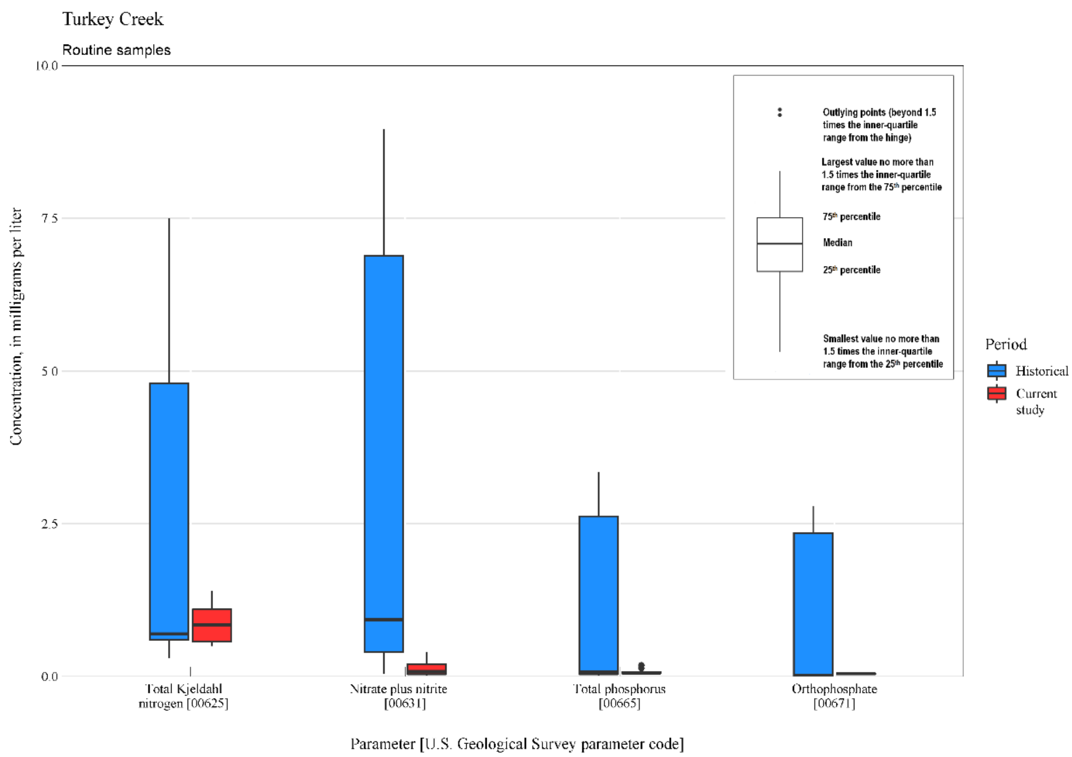


Figure 6.9. Comparison of historical and current-study nutrient data in the Turkey Creek watershed. *A*. Storm-event samples. *B*. Routine samples.

References Cited

U.S. Geological Survey, 2019, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed February 1, 2019, at https://doi.org/10.5066/F7P55KJN.