Appendix 1.11. Model Archive Summary for Alkalinity Concentration at U.S. Geological Survey site 07143672; Little Arkansas River at Highway 50 near Halstead, Kansas, during June 2013 through December 2019

This model archive summary summarizes the alkalinity model developed to compute hourly or daily alkalinity. Model development methods follow U.S. Geological Survey (USGS) guidance from Office of Surface Water/Office of Water Quality Technical Memoranda and USGS Techniques and Methods, book 3, chap. C4 (Rasmussen and others, 2009).

Site and Model Information

Site Number: 07143672

Site Name: Little Arkansas River at Highway 50 near Halstead, Kansas

Location: Latitude 38°01'43", longitude 97°32'25" referenced to North American Datum of 1927, in NW 1/4 NE 1/4 NE 1/4 sec.28, T.23 S., R.2 W., Harvey County, Kansas, hydrologic unit 11030012.

Equipment: A Sutron Satlink II High Data Rate Collection Platform and a Design Analysis Water Log H350/355 nonsubmersible pressure transducer transfers real-time stage and water-quality data via satellite. The primary reference gage is a Type-A wire-weight gage located on the downstream bridge guardrail. Check-bar elevation is 33.396 feet. The orifice tube is enclosed in 1.25-inch steel conduit trenched into the ground down to the edge of water, where the orifice emerges from the bank and culminates in a 2-inch open-end orifice tethered to a steel fencepost near the left edge of water. Gage height was measured during May 1998 through December 2019. A YSI 6600 water-quality monitor equipped with water temperature, specific conductance, pH, dissolved oxygen, and turbidity (a YSI Model 6026 [December 1998 through December 2006] and YSI Model 6136 [July 2004 through December 2017]) sensors collected data during May 1998 through December 2017. A YSI EXO2 water-quality monitor equipped with water temperature, specific conductance, pH, dissolved oxygen, turbidity, and fluorescent dissolved organic matter sensors collected data during January 2017 through December 2019. A Hach Nitratax monitor collected nitrate data during February 2017 through December 2019.

Date model was developed: June 1, 2020

Model calibration data period: June 3, 2013 through December 10, 2019

Model Data

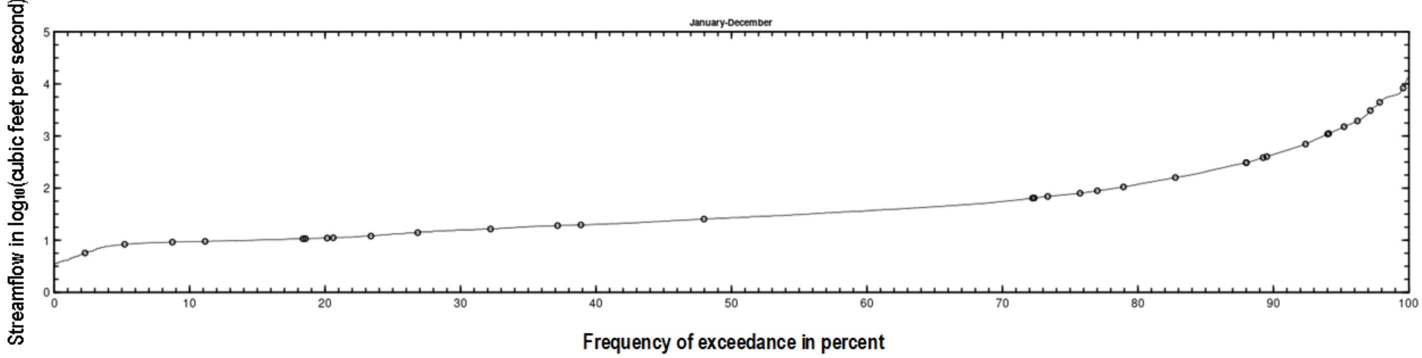
All data were collected using USGS protocols (U.S. Geological Survey, variously dated; Wagner and others, 2006; Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010) and are stored in the National Water Information System (NWIS) database (U.S. Geological Survey, 2021). Explanatory variables were evaluated individually and in combination. Potential explanatory variables included streamflow, water temperature, specific conductance, pH, dissolved oxygen, YSI EXO2 turbidity, nitrate, and fluorescent dissolved organic matter. Seasonal components (sine and cosine variables) also were evaluated as explanatory variables.

The regression model is based on 33 concomitant values of discretely collected alkalinity and continuously measured specific conductance during June 2013 through December 2019. Discrete samples were collected over a range of streamflow and specific conductance conditions. No samples had concentrations that were below laboratory detection limits. Summary statistics and the complete model-calibration dataset are provided below. Outliers and influential points were identified using studentized residuals, DFITS, Cook's D (Cook, 1977), and leverage. Outliers in previously published versions of this model (Christensen and others, 2003; Rasmussen and others, 2016) were examined and retained in the dataset if there were no clear issues, explanations, or conditions that would cause a result to be invalid for model calibration. One sample (collection date March 12, 2013) was not representative of the dataset and exceeded Cook's D and DFITS outlier criteria and was removed from the model dataset to avoid erroneous inflation of model-computed values at the upper range of surrogate relations. Removing data points based only on outlier criteria may only overestimate the certainty of the model.

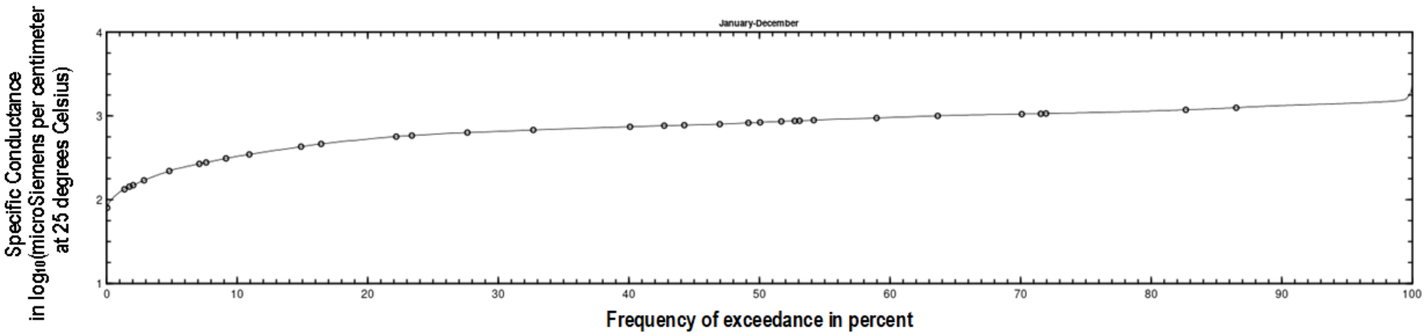
Alkalinity

Discrete samples were collected from the downstream side of the bridge or instream within 50 feet of the bridge using equal-width-increment, multi-vertical, single vertical or grab-dip methods following U.S. Geological Survey (variously dated) and Rasmussen and others (2014). Discrete samples were collected on a semifixed to event-based schedule ranging from 2 to 9 samples per year with a FISP US DH–95 or D–95 with a Teflon bottle, cap, and nozzle depth-integrating sampler, a DH–81 with a Teflon bottle, cap, and nozzle hand sampler or a grab sample with a Teflon bottle depending on sample location. Samples were analyzed for alkalinity by the U.S. Geological Survey Kansas Water Science Center according to standard methods (Rounds, 2012).

Alkalinity Samples Plotted on Streamflow Duration Curve



Alkalinity Samples Plotted on Specific Conductance Duration Curve



Continuous Data

Concomitant specific conductance values were time interpolated. If no concomitant continuous data were available within 2 hours of sample collection, the sample was not included in the dataset.

Model Development

Ordinary least squares regression analysis was done using R (version 4.0.0) programming language (R Core Team, 2020) to relate discretely collected alkalinity to specific conductance and other continuously measured data. The distribution of residuals was examined for normality and plots of residuals (the difference between the measured and model-calculated values) compared to model-computed alkalinity were examined for homoscedasticity (departures from zero did not change substantially over the range of model-calculated values). Previously published explanatory variables were also strongly considered for continuity; however, the best explanatory variable(s) were ultimately selected.

Specific conductance was selected as the best predictor of alkalinity based on residual plots, high coefficient of determination (R^2), and low model standard percentage error (MSPE). Specific conductance was positively related to total alkalinity because it measures water’s capacity to conduct an electrical current and is related to the concentration of ionized substances in water (Hem, 1992).

Model Summary

Summary of final alkalinity regression analysis at USGS site number 07143672:

Alkalinity-based model:

$$\log_{10}(ALK) = 0.974 \times \log_{10}(SC) - 0.531$$

where,

> $\log_{10}$ = logarithm base 10;
> ALK = alkalinity, in milligrams per liter as calcium carbonate (mg/L as $CaCO_3$); and
> SC = specific conductance, in microsiemens per centimeter at 25 degrees Celsius (μS/cm)

The log-transformed model may be retransformed to original units so that ALK 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; Duan, 1983). For this model, the calculated BCF is 1.03. The retransformed model, accounting for BCF is:

$$ALK = 0.3033 \times SC^{0.974}$$

Model Statistics, Data, and Plots

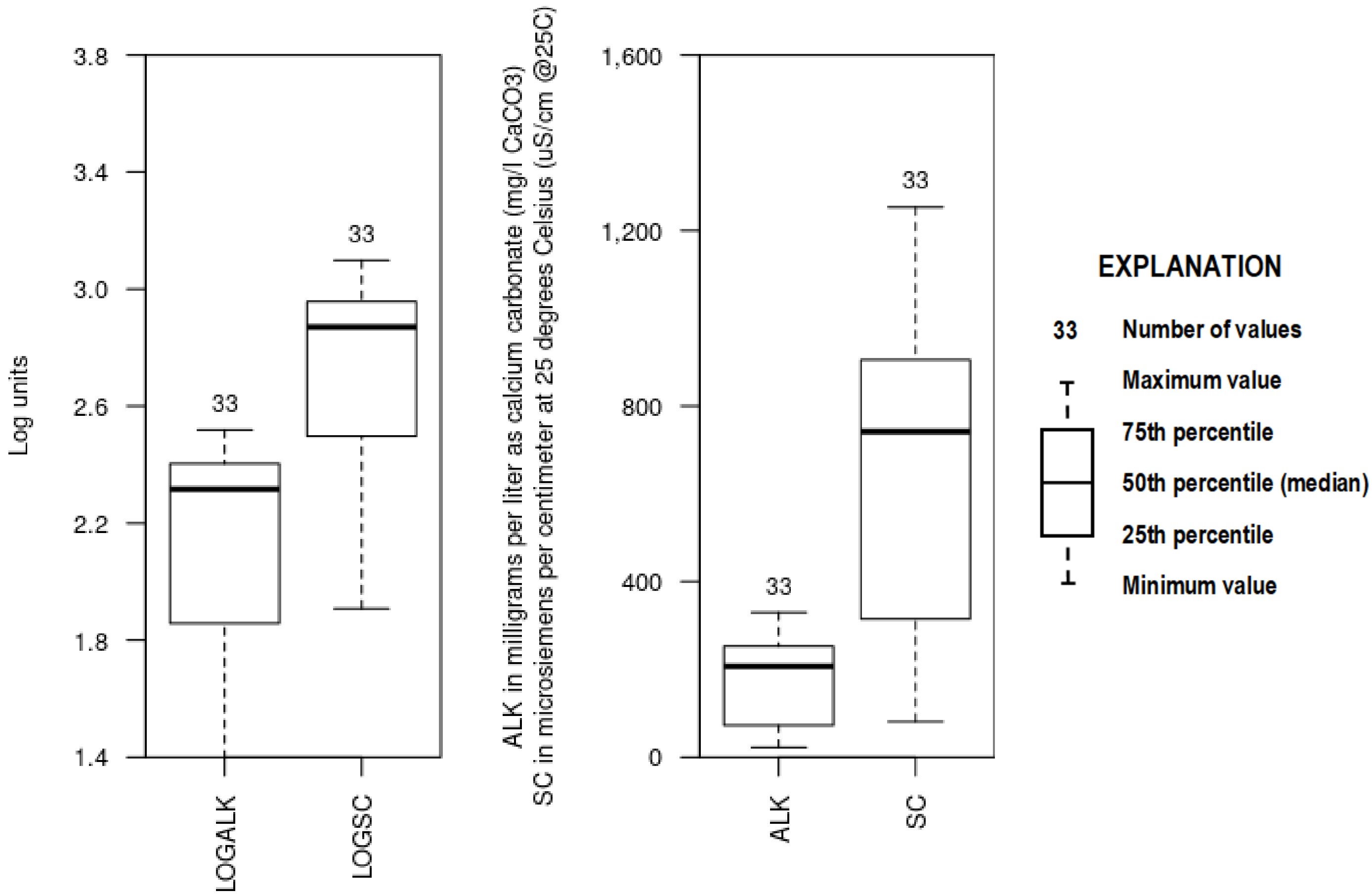
Model

LOGALK = + 0.974 * LOGSC - 0.531

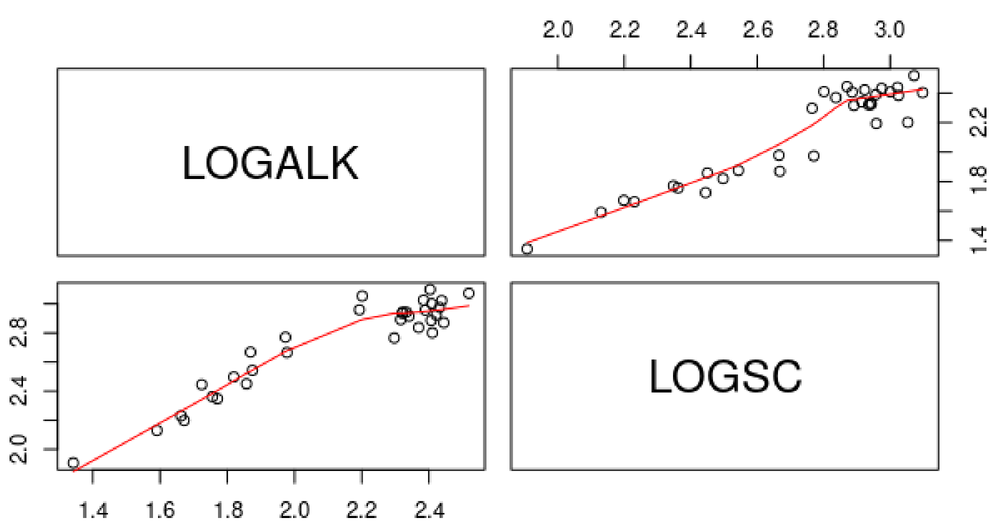
Variable Summary Statistics

```
               LOGALK ALK LOGSC     SC
Minimum          1.34  22  1.91   80.8
1st Quartile     1.86  72  2.50  314.0
Median           2.32 207  2.87  742.0
Mean             2.13 168  2.73  656.0
3rd Quartile     2.40 253  2.96  906.0
Maximum          2.52 330  3.10 1250.0
```

Box Plots



Exploratory Plots



Basic Model Statistics

Number of Observations	33
Standard error (RMSE)	0.108
Average Model standard percentage error (MSPE)	25.2
Coefficient of determination (R^2)	0.892
Adjusted Coefficient of Determination (Adj. R^2)	0.888
Bias Correction Factor (BCF)	1.03

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.531	0.1670	-3.17	3.38e-03
LOGSC	0.974	0.0609	16.00	1.60e-16

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.994
E.vars	-0.994	1.000

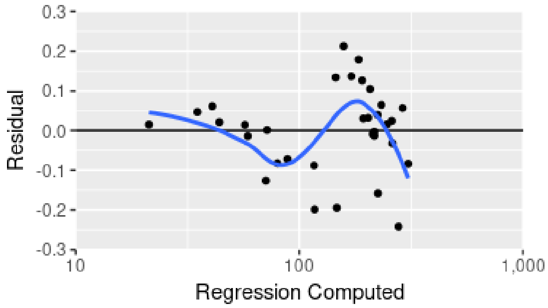
Outlier Test Criteria

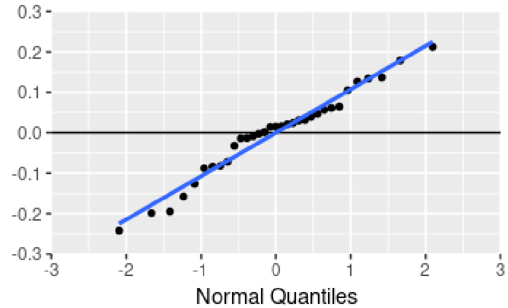
Leverage	Cook's D	DFFITS
0.182	0.194	0.492

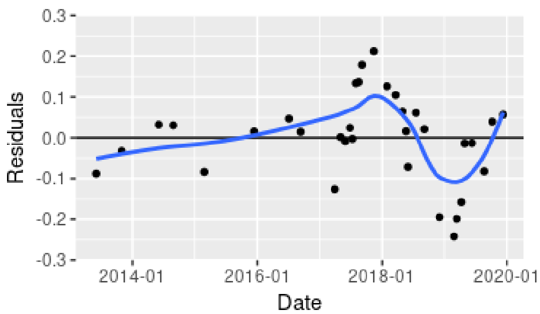
Flagged Observations

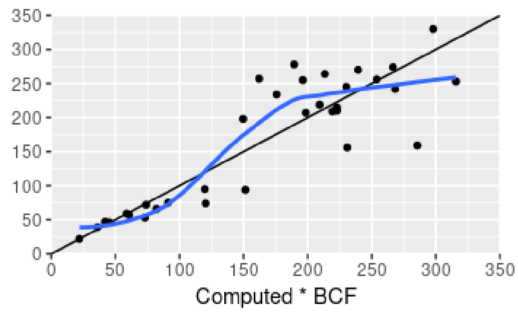
	LOGALK	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
12/14/2015 10:35	1.34	1.33	0.0153	0.162	0.159	0.2440	0.00423	0.0905
2/26/2019 11:40	2.20	2.44	-0.2420	-2.310	-2.490	0.0633	0.18000	-0.6490

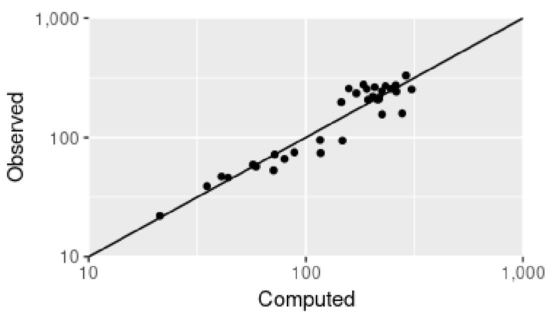
Statistical Plots

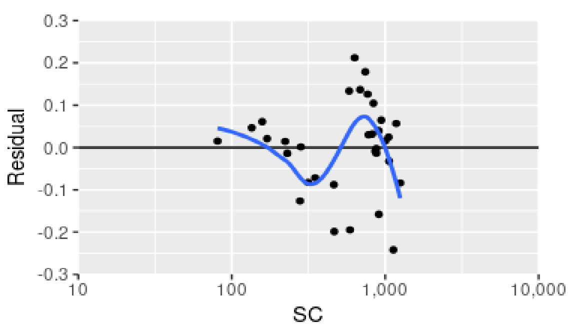


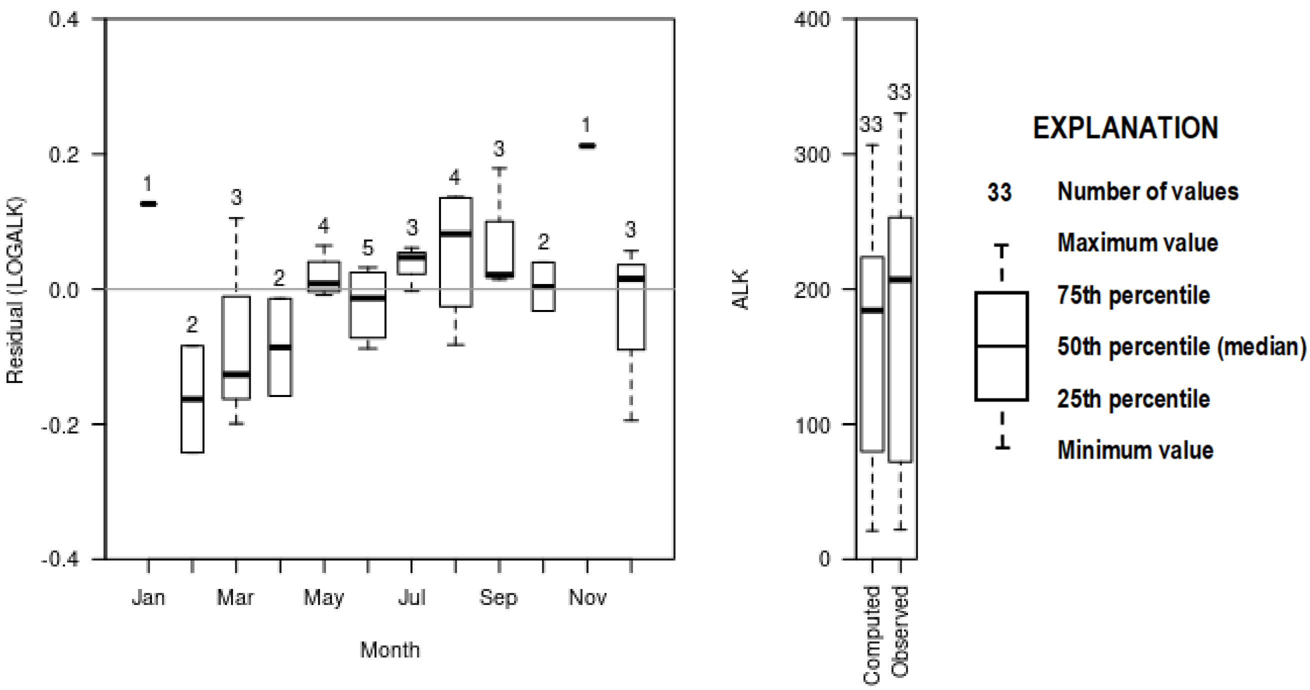


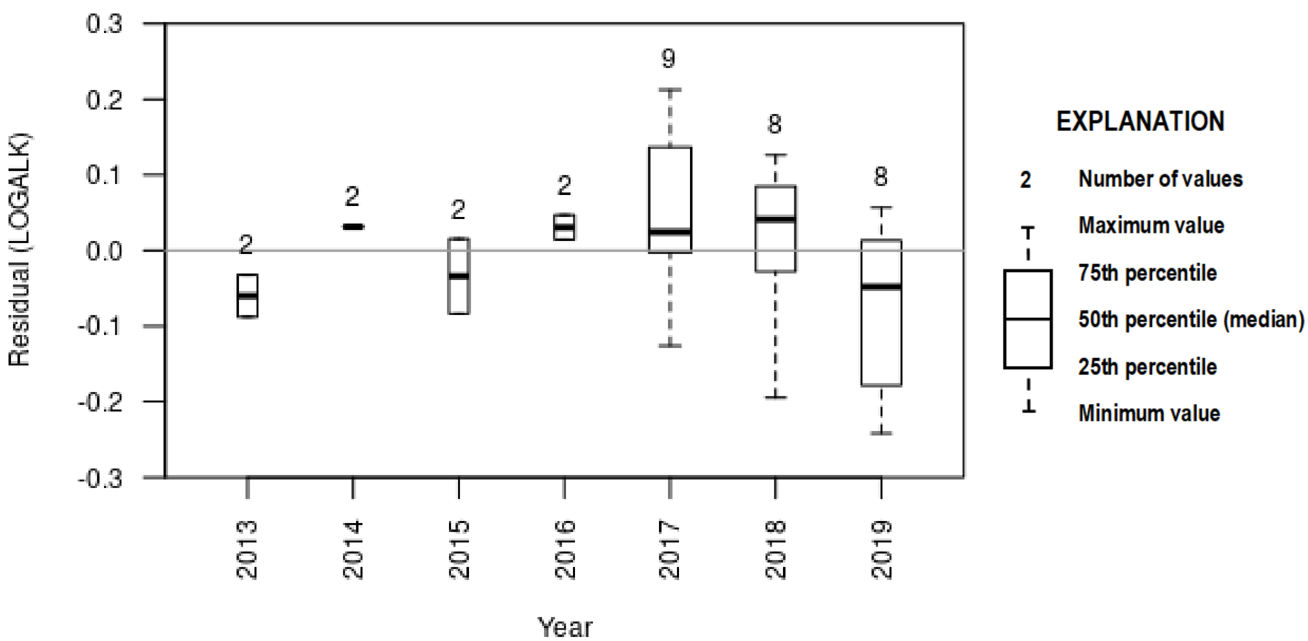




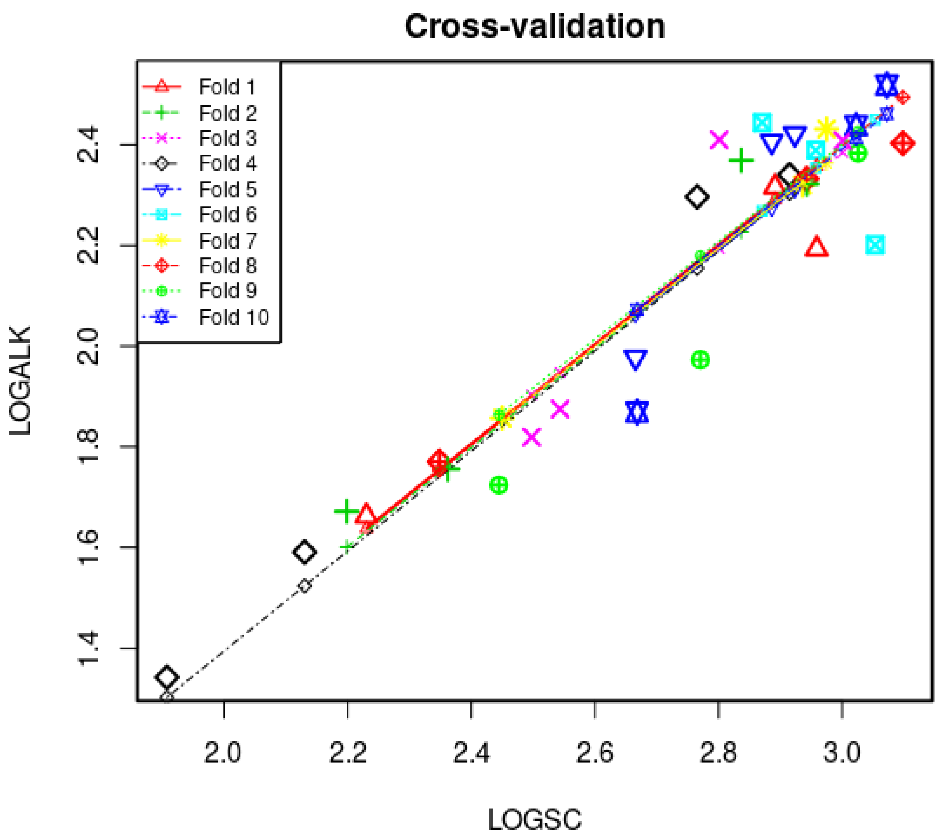




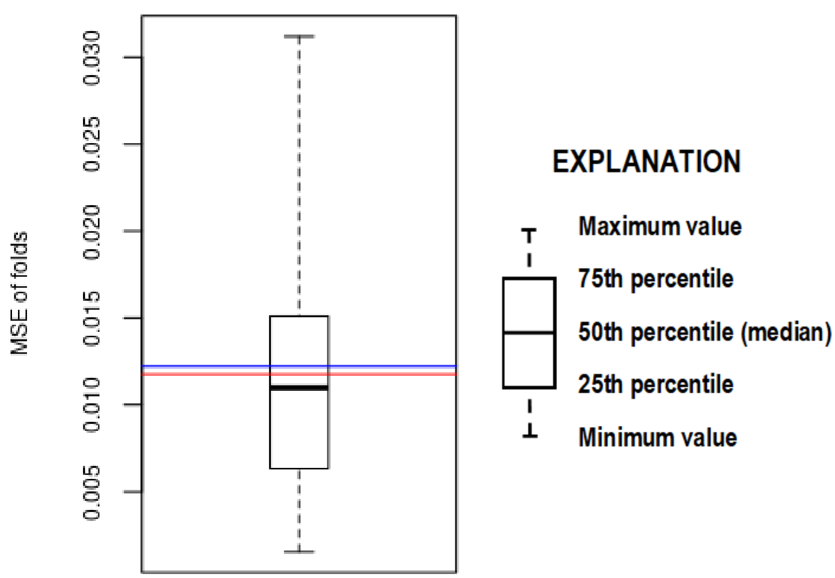

Residual (LOGALK)
Month
Jan
Mar
May
Jul
Sep
Nov
ALK
Computed
Observed
EXPLANATION
33 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value


Residual (LOGALK)
Year
2013
2014
2015
2016
2017
2018
2019
EXPLANATION
2 Number of values
Maximum value
75th percentile
50th percentile (median)
25th percentile
Minimum value

Cross Validation




```
              Minimum MSE of folds:  0.00154
                 Mean MSE of folds:  0.01220
               Median MSE of folds:  0.01100
              Maximum MSE of folds:  0.03120
 (Mean MSE of folds) / (Model MSE):  1.04000
```




```
Red line - Model MSE

Blue line - Mean MSE of folds
```

Model-Calibration Dataset

	Date	LOGALK	LOGSC	ALK	SC	Computed LOGALK	Computed ALK	Residual	Normal Quantiles
1	6/3/2013	1.98	2.67	95	463	2.07	120	-0.088	-0.959
2	10/30/2013	2.38	3.03	242	1060	2.42	268	-0.0324	-0.556
3	6/4/2014	2.34	2.91	219	822	2.31	209	0.032	0.387
4	8/28/2014	2.32	2.89	207	778	2.29	199	0.0305	0.307
5	2/25/2015	2.4	3.1	253	1250	2.49	316	-0.0838	-0.846
6	12/14/2015	1.34	1.91	22	80.8	1.33	21.9	0.0153	0
7	7/5/2016	1.59	2.13	39	135	1.54	36	0.0469	0.556
8	9/12/2016	1.77	2.35	59	223	1.76	58.7	0.0145	-0.0756
9	3/30/2017	1.72	2.44	53	279	1.85	72.9	-0.126	-1.09
10	5/3/2017	1.86	2.45	72	282	1.86	73.8	0.00145	-0.152
11	5/30/2017	2.32	2.94	209	862	2.33	219	-0.00862	-0.307
12	6/27/2017	2.44	3.02	274	1050	2.41	267	0.0244	0.228
13	7/12/2017	2.33	2.94	215	876	2.34	223	-0.00286	-0.228
14	8/1/2017	2.3	2.77	198	583	2.16	150	0.134	1.23
15	8/17/2017	2.37	2.84	234	687	2.23	176	0.137	1.42
16	9/5/2017	2.44	2.87	278	742	2.27	189	0.179	1.66
17	11/14/2017	2.41	2.8	257	632	2.2	162	0.212	2.1
18	1/30/2018	2.41	2.89	255	769	2.28	196	0.126	1.09
19	3/21/2018	2.42	2.92	264	838	2.32	213	0.105	0.959
20	5/1/2018	2.43	2.97	270	943	2.37	239	0.0647	0.846
21	5/22/2018	2.41	3	256	1000	2.39	254	0.0159	0.0756
22	6/2/2018	1.88	2.54	75	350	1.95	91	-0.0718	-0.646
23	7/18/2018	1.67	2.2	47	158	1.61	42	0.0613	0.742
24	9/6/2018	1.66	2.23	46	170	1.64	45.1	0.021	0.152
25	12/3/2018	1.97	2.77	94	589	2.17	151	-0.195	-1.42
26	2/26/2019	2.2	3.05	159	1130	2.44	286	-0.242	-2.1
27	3/14/2019	1.87	2.67	74	466	2.07	120	-0.199	-1.66
28	4/10/2019	2.19	2.96	156	909	2.35	231	-0.158	-1.23
29	4/29/2019	1.76	2.36	57	230	1.77	60.6	-0.0139	-0.47
30	6/11/2019	2.32	2.94	210	876	2.34	223	-0.0132	-0.387
31	8/21/2019	1.82	2.5	66	314	1.9	82.1	-0.0823	-0.742
32	10/8/2019	2.39	2.96	245	906	2.35	230	0.0395	0.47
33	12/10/2019	2.52	3.07	330	1180	2.46	298	0.0568	0.646

Definitions

ALK: Alkalinity in mg/L $CaCO_3$ (39086)
SC: Specific conductance in µS/cm @25C (00095)

References Cited

Christensen, V.G., Ziegler, A.C., Rasmussen P.P., and Jian X., 2003, Continuous real-time water-quality monitoring of Kansas streams, *in* Proceedings of 2003 Spring Specialty Conference on Agricultural Hydrology and Water Quality, Kansas City, Mo., May 12–14, 2003: Middleburg, Va., American Water Resources Association Technical Publication Series No. TPS–03–1, compact disc. [Also available at https://nrtwq.usgs.gov/ks/methods/christensen2003.]

Cook, D.R., 1977, Detection of influential observation in linear regression: Technometrics, v. 19, no. 1, p. 15–18. [Also available at https://www.jstor.org/stable/1268249?seq=4#metadata_info_tab_contents.]

Duan, N., 1983, Smearing estimate—A nonparametric retransformation method: Journal of the American Statistical Association, v. 78, no. 383, p. 605–610. [Also available at https://doi.org/10.1080/01621459.1983.10478017.]

Hem, J.D., 1992, Study and interpretation of chemical characteristics of natural water: U.S. Geological Survey Water-Supply Paper 2254, 3rd ed., 263 p. [Also available at https://pubs.usgs.gov/wsp/wsp2254/.]

R Core Team, 2020, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, version 4.0.0. [Also available at https://www.r-project.org.]

Rasmussen, P.P., Eslick, P.J., and Ziegler, A.C., 2016, Relations between continuous real-time physical properties and discrete water-quality constituents in the Little Arkansas River, south-central Kansas, 1998–2014: U.S. Geological Survey Open-File Report 2016–1057, 16 p. [Also available at https://doi.org/10.3133/ofr20161057.]

Rasmussen, P.P., Gray, J.R., Glysson, G.D., and Ziegler, A.C, 2009, Guidelines and procedures for computing time-series suspended-sediment concentrations and loads from in-stream turbidity sensor and streamflow data: U.S. Geological Survey Techniques and Methods, book 3, chap. C4, 53 p. [Also available at https://doi.org/10.3133/tm3C4.]

Rasmussen, T.J., Bennett, T.J., Stone, M.L., Foster, G.M., Graham, J.L., and Putnam, J.E., 2014, Quality-assurance and data-management plan for water-quality activities in the Kansas Water Science Center, 2014: U.S. Geological Survey Open-File Report 2014–1233, 41 p. [Also available at https://doi.org/10.3133/ofr20141233.]

Rounds, S.A., 2012, Alkalinity and acid neutralizing capacity (ver. 4.0, September 2012): U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chap. 6.6, 45 p. [Also available at https://doi.org/10.3133/twri09A6.6.]

Sauer, V.B., and Turnipseed, D.P., 2010, Stage measurement at gaging stations: U.S. Geological Survey Techniques and Methods, book 3, chap. A7, 45 p. [Also available at https://doi.org/10.3133/tm3A7.]

Turnipseed, D.P., and Sauer, V.B., 2010, Discharge measurements at gaging stations: U.S. Geological Survey Techniques and Methods, book 3, chap. A8, 87 p. [Also available at https://doi.org/10.3133/tm3A8.]

U.S. Geological Survey, 2021, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed December 8, 2021, 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 [variously paged]. [Also available at https://water.usgs.gov/owq/FieldManual/.]

Wagner, R.J., Boulger, R.W., Jr., Oblinger, C.J., and Smith, B.A., 2006, Guidelines and standard procedures for continuous water-quality monitors—Station operation, record computation, and data reporting: U.S. Geological Survey Techniques and Methods, book 1, chap. D3, 96 p. [Also available at https://doi.org/10.3133/tm1D3.]