Appendix 9. Model Archive Summary for Nitrate at U.S. Geological Survey Site 07144780, North Fork Ninnescah River above Cheney Reservoir, Kansas, during January 1, 1999, through December 31, 2019

This model archive summary summarizes the nitrate (NO_3) model developed to compute hourly or daily nitrate concentrations during January 1, 1999, onward. This model supersedes all prior models used during this period. The methods used follow U.S. Geological Survey (USGS) guidance as referenced in relevant Office of Surface Water/Office of Water Quality Technical Memoranda and USGS Techniques and Methods, book 3, chapter C4 (Rasmussen and others, 2009).

Site and Model Information

Site number: 07144780
Site name: North Fork Ninnescah River above Cheney Reservoir, Kansas
Location: Lat 37°51'45", long 98°00'49" referenced to North American Datum of 1927, in NE 1/4 SE 1/4 NE 1/4 sec.19, T.25 S., R.6 W., Reno County, Kans., Hydrologic Unit 11030014, on right bank at upstream side of county highway bridge, 10 miles south of Hutchinson, 18.1 miles upstream from Cheney Dam.

Equipment: A YSI 6600 Extended Deployment System water-quality monitor equipped with sensors for water temperature, specific conductance, pH, dissolved oxygen, and turbidity (a YSI Model 6026 turbidity sensor [November 9, 1998, to December 1, 2010] and a YSI Model 6136 turbidity sensor [October 17, 2009, to November 12, 2015; March 31, 2017, to June 7, 2017]) (YSI Incorporated, 2007, 2012a). The YSI 6600 water-quality monitor was in operation during November 9, 1998, through November 12, 2015.

A Xylem YSI EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, pH, and turbidity (YSI Incorporated, 2012b). The YSI EXO2 water-quality monitor began operation on November 13, 2015. Monitors were housed in a 4-inch diameter polyvinyl chloride (PVC) pipe and placed in a location representative of the stream cross section. Monitor readings were recorded and satellite transmitted hourly.

Date model was developed: April 26, 2019

Model calibration data period: January 26, 1999, to September 28, 2017

Model Data

All data were collected using USGS protocols (U.S. Geological Survey, 2006; Wagner and others, 2006; Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010) and are stored in the National Water Information System (NWIS) database (https://doi.org/10.5066/F7P55KJN; U.S. Geological Survey, 2020). Explanatory variables were evaluated individually and in combination. Potential explanatory variables included streamflow, water temperature, specific conductance, pH, dissolved oxygen, and turbidity. Seasonal components (sine and cosine variables) were also evaluated as explanatory variables.

The regression model is based on 109 concomitant values of discretely collected nitrate samples and continuously measured specific conductance during January 26, 1999, through September 28, 2017. Discrete samples were collected over a range of streamflows. One sample was less than the minimum reporting level (less than [<] 0.01 milligram per liter); therefore, a Tobit regression model was developed to compute estimates of nitrate using the absolute maximum likelihood estimation and approach (Hald, 1949; Cohen, 1950; Tobin, 1958; Helsel and others, 2020). Summary statistics and the complete model-calibration data are provided below. Potential outliers were identified using the methods described in Rasmussen and others (2009). Additionally, outlier test criteria, including leverage and Cook's distance (Cook's D), were used to estimate potential outlier influence on the final Tobit regression model (Cook, 1977). The sample collected on March 22, 2006, had a large specific conductance value likely from previous road salt application and was removed from the model calibration dataset. Previous

outliers (Stone and others, 2013) 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.

Nitrate

Discrete samples were collected from the downstream side of the bridge or instream within 50 feet of the bridge using equal-width-increment, multiple vertical, single vertical, or grab methods following U.S. Geological Survey (2006) and Rasmussen and others (2014). Discrete samples were collected on a semifixed to event-based schedule ranging from 2 to 17 samples per year with a Federal Interagency Sedimentation Project U.S. 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 nitrate by the Wichita Municipal Water and Wastewater Laboratory in Wichita, Kans., according to standard methods (American Public Health Association and others, 1995).

Continuous Data

Specific conductance was measured with a YSI 6600 sensor during November 9, 1998, through November 12, 2015, and a YSI EXO2 sensor during November 13, 2015, through December 31, 2019. Concomitant specific conductance values were time interpolated. If continuous data were not available (2 or more hours of specific conductance values bracketing the sample collection time were missing) because of fouling, changes in equipment, or unsuitable site conditions, then the field monitor specific conductance value measured during sampling was substituted. If no concomitant continuous data were available, the sample was not included in the dataset.

Model Development

Stepwise regression analysis was done using R programming language (R Core Team, 2019) to relate discretely collected nitrate 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 calculated nitrate were examined for homoscedasticity (departures from zero did not change substantially over the range of model calculated values).

About 1 percent of the model-calibration dataset consisted of censored results (less than minimum reporting level). Tobit regression models were developed using absolute maximum likelihood estimation methods using the *smwrQW* (v.0.7.9) package in R programming language (R Core Team, 2019).

Specific conductance and seasonality were selected as best predictors of nitrate based on residual plots, a higher pseudocoefficient of determination (pseudo-R^2), and relatively low estimated standard residual error (RSE). Seasonality was included as an explanatory variable because nitrate seems to have a cyclical pattern potentially influenced by groundwater during low seasonal flow.

Model Summary

Summary of final nitrate regression analysis at USGS site 07144780:

Nitrate-based model:

$$NO3 = 0.00037 \times SPC + 0.0265 \times \sin(2\pi D) + 0.538 \times \cos(2\pi D) + 0.532 ,$$

where,

NO3 = nitrate, filtered, in milligrams per liter as nitrogen;
SPC = specific conductance, in microsiemens per centimeter at 25 degrees Celsius; and
D = date, in decimal years.

Previous Models

Version	Model Equation	Reference
1.0	$NO3 = -0.389 \times \log_{10}(TBY6026) + 0.0519 \times \sin(2\pi D) + 0.619 \times \cos(2\pi D) + 1.63$	Stone and others (2013)

Model Statistics, Plots and Data

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

Model

$NO3 = 0.00037 \times SPC + 0.0265 \times \sin(2\pi D) + 0.538 \times \cos(2\pi D) + 0.532$

Computation method: Absolute Maximum Likelihood Estimation (AMLE)

Explanatory Variables

```
Coefficients:
             Estimate Std. Error z-score p-value
(Intercept) 0.5324609  8.767e-02  6.0736  0.0000
SPC         0.0003698  8.572e-05  4.3141  0.0000
sin2piD     0.0264661  3.955e-02  0.6692  0.4943
cos2piD     0.5381386  5.703e-02  9.4362  0.0000
```

Basic Model Statistics

For a detailed definition and explanation of the terms used below, refer to Helsel and others (2020).

```
Estimated residual standard error (Unbiased) = 0.2947
Number of observations = 109, number censored = 1 (0.9 percent)

Log-likelihood (model) = -20.56
Log-likelihood (intercept only) = -68.16
  Chi-square = 95.19
  degrees of freedom = 3
  p-value = <0.0001

Computation method: AMLE

Pseudo-R-squared: 0.5832

  Akaike Information Criterion: 51.12
  Bayesian Information Criterion: 64.57


Variance inflation factors
    SPC 1.08
sin2piD 1.03
cos2piD 1.11
```

Outlier Test Criteria

```
leverage   cooksD
 0.08257  0.84457
```

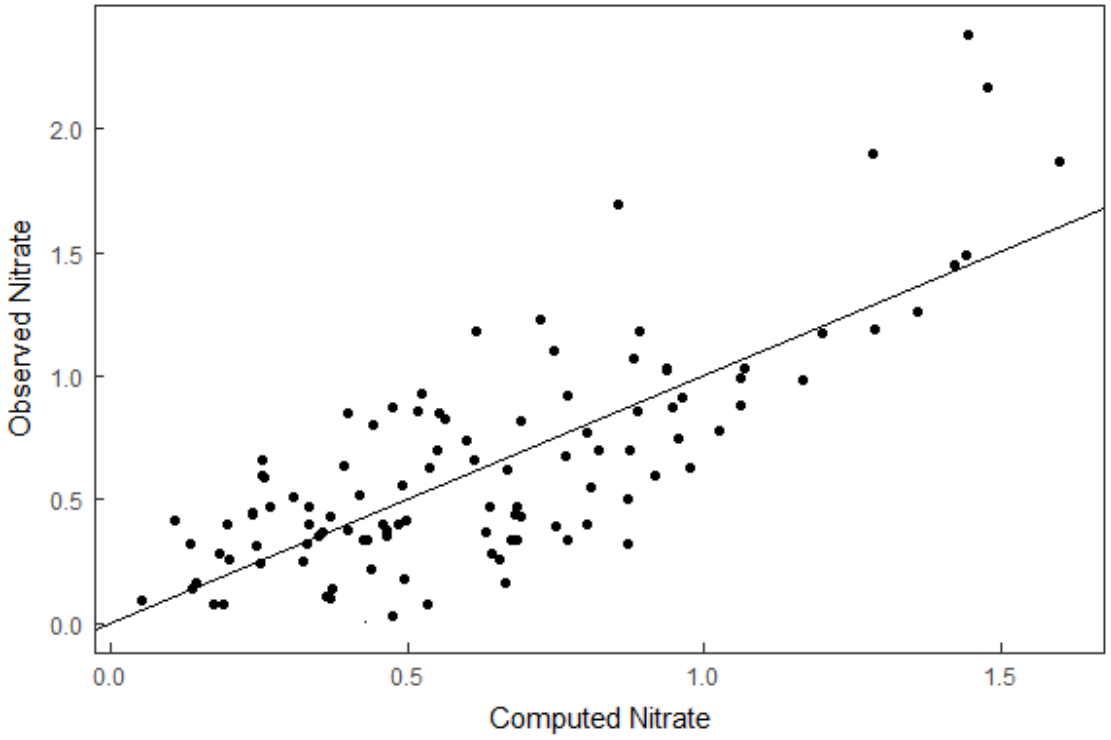
Flagged Observations

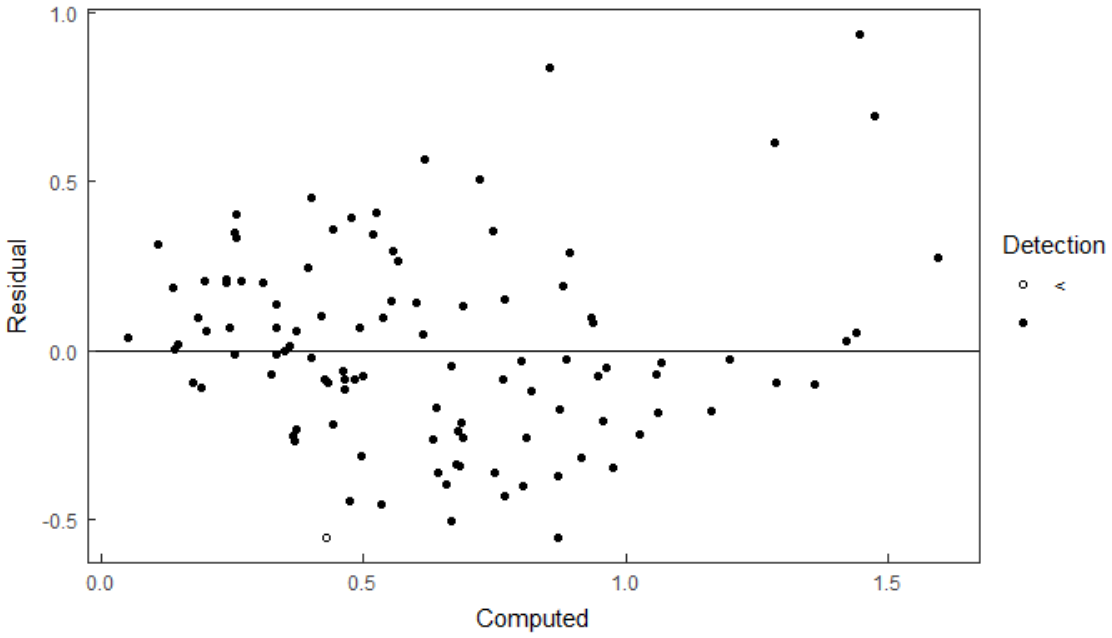
```
    Observations exceeding at least one test criterion
   Nitrate  ycen   yhat  resids leverage  cooksD
26    0.32 FALSE 0.8712 -0.5512  0.09838 0.10587
53    1.87 FALSE 1.5970  0.2730  0.08530 0.02187
```

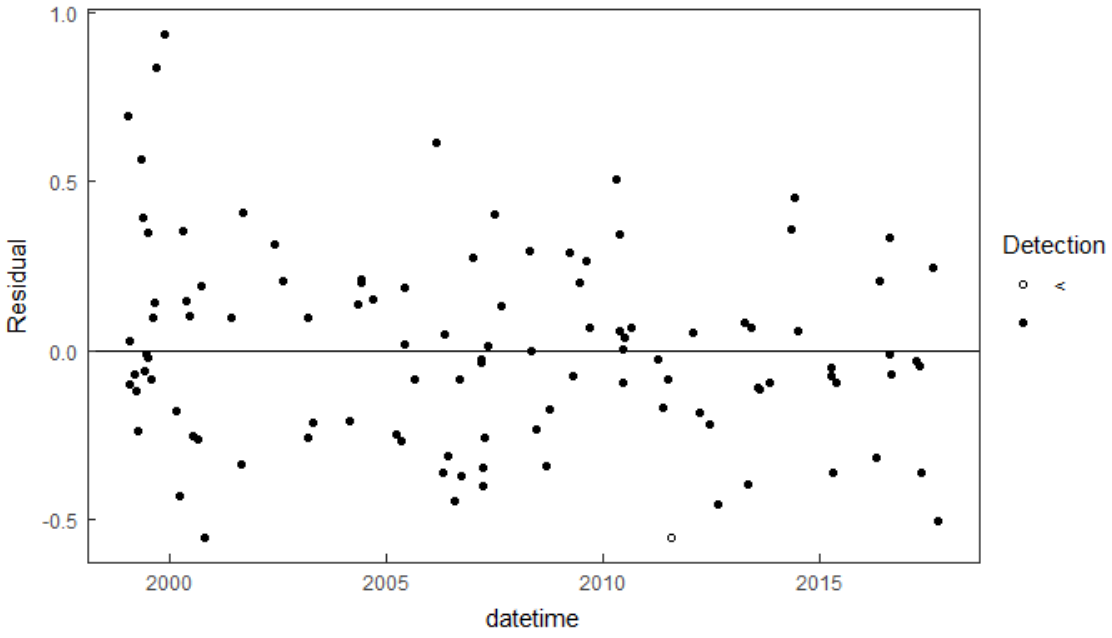
95% Confidence Intervals

```
                   2.5 %       97.5 %
(Intercept)  0.3606352627 0.7042865690
SPC          0.0002017926 0.0005378037
sin2piD     -0.0510453473 0.1039776464
cos2piD      0.4263633220 0.6499138899
```

Plots







Variable Summary Statistics

Independent Variable (xvar) - Specific Conductance

```
   Min. 1st Qu.  Median    Mean  3d Qu.     Max.
  156.5   559.0   897.0   849.1  1,153.5  1,432.5
```

Standard Deviation

```
[1] 347.8235
```

Dependent Variable (yvar) - Nitrate

```
    Min.  1st Qu.   Median     Mean    3d Qu.      Max.
   <0.01     0.34      0.5   0.6333       0.86     2.38
```

Standard Deviation

```
[1] 0.4497
```

Model Calibration Dataset

	datetime	Nitrate	SPC	Computed_Nitrate	residuals
1	1999-01-26 11:50:00	2.17	1210	1.4767	0.693497
2	1999-01-31 14:25:00	1.26	949	1.3605	-0.100302
3	1999-02-03 10:45:00	1.45	1154	1.4224	0.027787
4	1999-03-17 11:40:00	0.99	979	1.0599	-0.069706
5	1999-04-06 13:55:00	0.7	829	0.8216	-0.121386
6	1999-04-16 12:55:00	0.44	697	0.6807	-0.240591
7	1999-05-13 10:25:00	1.18	1130	0.6160	0.564243
8	1999-05-24 10:45:00	0.87	950	0.4761	0.394164
9	1999-06-10 12:00:00	0.4	1133	0.4596	-0.059315
10	1999-06-25 11:15:00	0.32	896	0.3323	-0.012074
11	1999-07-02 10:15:00	0.6	705	0.2549	0.345270
12	1999-07-14 11:20:00	0.38	1080	0.4007	-0.020464
13	1999-07-29 09:55:00	0.38	1147	0.4657	-0.085472
14	1999-08-12 10:35:00	0.63	1155	0.5361	0.094168
15	1999-08-26 10:50:00	0.74	1081	0.6002	0.140061
16	1999-09-22 11:20:00	1.69	1163	0.8560	0.834258
17	1999-12-02 10:35:00	2.38	1227	1.4461	0.934064
18	2000-02-25 10:40:00	0.986	822	1.1653	-0.179169
19	2000-03-24 13:50:00	0.34	390	0.7699	-0.429806
20	2000-04-27 10:45:00	1.1	1165	0.7481	0.354152
21	2000-05-25 10:20:00	0.7	1188	0.5522	0.148026
22	2000-06-21 12:00:00	0.52	1117	0.4191	0.101124
23	2000-07-26 11:50:00	0.11	896	0.3656	-0.255371
24	2000-08-29 11:00:00	0.37	1087	0.6324	-0.262127
25	2000-09-28 10:30:00	1.07	1054	0.8798	0.190435
26	2000-10-26 10:50:00	0.32	345	0.8713	-0.551210
27	2001-06-06 11:35:00	0.28	342	0.1830	0.097093
28	2001-09-04 11:05:00	0.34	1100	0.6766	-0.336392
29	2001-09-19 10:25:00	0.93	343	0.5253	0.404732
30	2002-06-12 11:10:00	0.42	204	0.1085	0.311547
31	2002-08-14 11:35:00	0.47	398	0.2676	0.202529
32	2003-03-18 12:00:00	1.03	667	0.9356	0.094523
33	2003-03-19 12:20:00	0.55	351	0.8098	-0.259722
34	2003-04-21 11:30:00	0.47	832	0.6860	-0.215817
35	2004-03-05 12:10:00	0.75	448	0.9569	-0.206858
36	2004-05-14 10:35:00	0.47	408	0.3349	0.135211
37	2004-06-14 09:45:00	0.44	582	0.2387	0.201477
38	2004-06-14 09:50:00	0.45	584	0.2394	0.210737
39	2004-09-08 10:25:00	0.92	1240	0.7698	0.150431
40	2005-03-24 10:15:00	0.78	1060	1.0270	-0.246815
41	2005-05-16 11:40:00	0.1	520	0.3693	-0.269126
42	2005-06-10 10:55:00	0.32	258	0.1359	0.184213
43	2005-06-13 09:25:00	0.16	310	0.1444	0.015726
44	2005-08-29 09:35:00	0.4	707	0.4841	-0.083971
45	2006-03-02 09:50:00	1.9	1250	1.2856	0.614602
46	2006-05-01 11:15:00	0.39	1242	0.7516	-0.361303
47	2006-05-12 10:30:00	0.66	1100	0.6124	0.047831
48	2006-06-05 10:15:00	0.18	1171	0.4944	-0.314136
49	2006-07-31 10:30:00	0.03	1150	0.4749	-0.444630
50	2006-09-07 10:50:00	0.68	1280	0.7679	-0.087585
51	2006-09-21 10:00:00	0.5	1230	0.8716	-0.371322
52	2006-09-21 10:05:00	0.5	1230	0.8716	-0.371322
53	2007-01-09 10:30:00	1.87	1430	1.5973	0.272951

54	2007-03-14 10:20:00	1.17	1283		1.1987	-0.028500
55	2007-03-22 10:00:00	1.03	1120		1.0674	-0.037150
56	2007-03-26 10:40:00	0.63	974		0.9771	-0.346867
57	2007-03-31 12:30:00	0.4	629		0.8032	-0.403103
58	2007-04-16 12:15:00	0.43	722		0.6899	-0.259713
59	2007-05-07 10:30:00	0.37	311		0.3585	0.011584
60	2007-06-29 10:25:00	0.66	704		0.2563	0.403888
61	2007-09-04 11:25:00	0.82	1140		0.6914	0.128816
62	2008-04-24 11:40:00	0.85	574		0.5551	0.295046
63	2008-05-09 11:35:00	0.35	353		0.3509	-0.000831
64	2008-06-19 09:45:00	0.14	979		0.3722	-0.231978
65	2008-09-15 10:55:00	0.34	842		0.6834	-0.343223
66	2008-10-16 10:10:00	0.7	591		0.8741	-0.174029
67	2009-03-31 11:20:00	1.18	871		0.8926	0.287529
68	2009-04-27 12:15:00	0.42	466		0.4979	-0.077799
69	2009-06-17 10:40:00	0.51	787		0.3086	0.201555
70	2009-08-20 10:50:00	0.83	1098		0.5646	0.265609
71	2009-09-10 11:30:00	0.56	464		0.4912	0.068907
72	2010-04-23 10:00:00	1.23	980		0.7229	0.507314
73	2010-05-17 16:40:00	0.86	945		0.5194	0.340868
74	2010-05-27 10:00:00	0.43	716		0.3719	0.058322
75	2010-06-14 11:30:00	0.14	304		0.1387	0.001357
76	2010-06-16 10:15:00	0.08	417		0.1748	-0.094641
77	2010-07-06 10:30:00	0.09	156		0.0518	0.038239
78	2010-08-25 11:00:00	0.4	378		0.3328	0.067280
79	2011-04-13 10:00:00	0.86	1180		0.8869	-0.026642
80	2011-05-23 10:20:00	0.47	1373		0.6388	-0.168503
81	2011-06-28 10:00:00	0.34	1160		0.4261	-0.085858
82	2011-07-27 11:20:00	<0.01	1070		0.4300	-0.552442
83	2012-02-06 09:45:00	1.49	1242		1.4405	0.049700
84	2012-03-23 10:15:00	0.88	1158		1.0631	-0.182870
85	2012-06-20 09:15:00	0.22	1170		0.4408	-0.220511
86	2012-08-27 09:30:00	0.08	860		0.5332	-0.453010
87	2013-04-11 10:10:00	1.02	1270		0.9386	0.081698
88	2013-05-10 10:00:00	0.26	1180		0.6569	-0.396685
89	2013-05-31 10:00:00	0.31	431		0.2447	0.065441
90	2013-08-05 10:05:00	0.08	322		0.1906	-0.110548
91	2013-08-16 08:30:00	0.35	897		0.4644	-0.114156
92	2013-10-31 10:00:00	1.19	1380		1.2878	-0.097555
93	2014-05-13 10:00:00	0.8	659		0.4419	0.358233
94	2014-06-10 10:30:00	0.85	972		0.3998	0.450408
95	2014-07-02 09:10:00	0.26	559		0.2010	0.059168
96	2015-04-08 09:45:00	0.91	1260		0.9626	-0.052305
97	2015-04-14 09:55:00	0.87	1369		0.9477	-0.077442
98	2015-04-21 10:15:00	0.39	1010		0.7517	-0.361456
99	2015-05-26 10:45:00	0.34	868		0.4338	-0.093639
100	2016-04-19 10:25:00	0.6	1432		0.9169	-0.316603
101	2016-05-25 10:10:00	0.4	225		0.1961	0.203993
102	2016-08-07 12:30:00	0.59	464		0.2584	0.331729
103	2016-08-13 11:30:00	0.24	361		0.2539	-0.013757
104	2016-08-27 09:20:00	0.25	292		0.3230	-0.072965
105	2017-03-29 10:45:00	0.77	576		0.8021	-0.031981
106	2017-04-20 12:00:00	0.62	762		0.6688	-0.048654
107	2017-05-02 09:50:00	0.28	972		0.6430	-0.362825

```
108 2017-08-11 11:00:00    0.64  789           0.3948  0.245352
109 2017-09-28 10:30:00    0.16  502           0.6664 -0.506328
```

Definitions

```
NO3: nitrate, filtered, in milligrams per liter as nitrogen (00618)
SPC: specific conductance, in microsiemens per centimeter at 25 degrees Celsius (00095)
D: date, in decimal years
Leverage: an outlier's measure in the x-direction (Helsel and others, 2020).
p-value: the probability that the independent variable has no effect on the dependent
      variable (Helsel and others, 2020).
Pseudo-R-squared: pseudocoefficient of determination. An estimation of the proportion of
      variance in the response variable explained by the model (McKelvey and Zavoina,
      1975).
z-score: the estimated coefficient divided by its associated standard error (Helsel and
      others, 2020).
```

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

References Cited

American Public Health Association, American Water Works Association, and Water Environment Federation, 1995, Standard methods for the examination of water and wastewater (19th ed.): Washington, D.C., American Public Health Association, 905 p.

Christensen, V.G., Graham, J.L., Milligan, C.R., Pope, L.M., and Ziegler, A.C., 2006, Water quality and relation to taste-and-odor compounds in the North Fork Ninnescah River and Cheney Reservoir, south-central Kansas, 1997–2003: U.S. Geological Survey Scientific Investigations Report 2006–5095, 43 p. [Also available at https://doi.org/10.3133/sir20065095.]

Cohen, A.C., Jr., 1950, Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples: Annals of Mathematical Statistics, v. 21, no. 4, p. 557–569, accessed October 2019 at https://doi.org/10.1214/aoms/1177729751.

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

Hald, A., 1949, Maximum likelihood estimation of the parameters of a normal distribution which is truncated at a known point: Scandinavian Actuarial Journal, v. 1949, no. 1, p. 119–134. [Also available at https://doi.org/10.1080/03461238.1949.10419767.]

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., 2020, Statistical methods in water resources—Supporting materials: U.S. Geological Survey data release, accessed August 2020 at https://doi.org/10.5066/P9JWL6XR.

McKelvey, R.D., and Zavoina, W., 1975, A statistical model for the analysis of ordinal level dependent variables: The Journal of Mathematical Sociology, v.4, no. 1, p. 103–120. [Also available at https://doi.org/10.1080/0022250X.1975.9989847.]

R Core Team, 2019, R—A language and environment for statistical computing: Vienna, Austria, R Foundation for Statistical Computing, accessed August 2019 at https://www.R-project.org/.

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., accessed April 2020 at https://doi.org/10.3133/ofr20141233.

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, 52 p. [Also available at https://doi.org/10.3133/tm3C4.]

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., accessed April 2020 at https://doi.org/10.3133/tm3A7.

Stone, M.L., Graham, J.L., and Gatotho, J.W., 2013, Continuous real-time water-quality monitoring and regression analysis to compute constituent concentrations and loads in the North Fork Ninnescah River upstream from Cheney Reservoir, south-central Kansas, 1999–2012: U.S. Geological Survey Scientific Investigations Report 2013–5071, 44 p., accessed July 2020 at https://doi.org/10.3133/sir20135071.

Tobin, J., 1958, Estimation of relationships for limited dependent variables: Econometrica, v. 26, no. 1, p. 24–36. [Also available at https://doi.org/10.2307/1907382.]

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., accessed April 2020 at https://doi.org/10.3133/tm3A8.

U.S. Geological Survey, 2006, Collection of water samples (ver. 2.0, September 2006): U.S. Geological Survey Techniques of Water Resources Investigations, book 9, chap. A4 [variously paged]. [Also available at https://doi.org/10.3133/twri09A4.]

U.S. Geological Survey, 2020, USGS water data for the Nation: U.S. Geological Survey National Water Information System database, accessed April 20, 2020, at https://doi.org/10.5066/F7P55KJN.

Wagner, R.J., Boulger, R.W., Jr., Oblinger, C.J., and Smith, B.A., 2006, Guidelines and standard procedures for continuous water-quality monitors—Station operation, record computation, and data reporting: U.S. Geological Survey Techniques and Methods, book 1, chap D3, 51 p. plus 8 attachments. [Also available at https://doi.org/10.3133/tm1D3.]

YSI Incorporated, 2007, YSI 6136 turbidity sensor: YSI Incorporated, 2 p., accessed November 2019 at https://www.ysi.com/File%20Library/Documents/Specification%20Sheets/E56-6136-Turbidity-Sensor.pdf.

YSI Incorporated, 2012a, 6-series multiparameter water quality sondes—User manual, revision J: YSI Incorporated, 379 p., accessed November 2019 at https://www.ysi.com/File%20Library/Documents/Manuals/069300-YSI-6-Series-Manual-RevJ.pdf.

YSI Incorporated, 2012b, EXO water quality field sensors—Features, specifications, and comparability to YSI 6-series sensors, revision B: YSI Incorporated, 14 p., accessed November 2019 at https://www.exowater.com/media/pdfs/EXO-6Series-Sensor-Comparison.pdf.