Appendix 12. Model Archive Summary for Orthophosphate 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 orthophosphate (OP) model developed to compute hourly or daily OP from January 1, 1999, through December 31, 2019. This model is used concomitantly with other models during this period to calculate concentrations when other explanatory variables are not available for the purposes of load and concentration model calculations. 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), and other standard USGS methods (Sauer and Turnipseed, 2010; Turnipseed and Sauer, 2010).

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 Sutron Satlink 2 High Data Rate Collection Platform and a Design Analysis Water Log H350/355 nonsubmersible pressure transducer transfers real-time stage, precipitation, 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 21.804 feet. The orifice is enclosed in 1 1/4-inch pipe, which runs from the gage house, under the bridge, and along an I-beam where it is attached to the concrete pier closest to the left edge of water.

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 U.S. Geological Survey (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 nine concomitant values of discretely collected OP samples and continuously measured streamflow during January 26, 1999, through September 28, 2017. Discrete samples were collected over a range of streamflows. A total of 34 samples were below the minimum reporting level (less than [<] 0.01 milligram per liter [mg/L], <0.02 mg/L, <0.04 mg/L); therefore, a Tobit regression model was developed to compute estimates of OP using the absolute maximum likelihood estimation 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). None of the samples in this dataset were deemed outliers or removed from the model calibration dataset.

Orthophosphate

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 the 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 1 to 16 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 sampling location. Samples are analyzed for OP by the Wichita Municipal Water and Wastewater Laboratory in Wichita, Kans., according to standard methods (American Public Health Association and others, 1995).

Continuous Data

The streamflow data used in this analysis were measured using a nonsubmersible pressure transducer from January 1, 1991, through December 31, 2019. The surrogate data used were time interpolated values from the continuous time series. If the continuous data were not 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 OP to streamflow 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 OP were examined for homoscedasticity (departures from zero did not change substantially over the range of model calculated values).

A total of 34.3 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).

Streamflow was selected as a good predictor of OP based on residual plots, a higher pseudocoefficient of determination (pseudo- R^2), and relatively low estimated standard residual error (RSE). This model was developed with the sole purpose to fill in gaps of missing data of the primary model for concentration and load estimations.

Model Summary

Summary of final OP regression analysis at USGS site 07144780.

OP-based model:

$$\log_{10}(OP) = 0.8036 \times \log_{10}(Q) - 3.346,$$

where,

OP = orthophosphate, in milligrams per liter as phosphorus, and Q = streamflow, in cubic feet per second.

The log-transformed model may be retransformed to original units so that OP 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 1.323. The retransformed model, accounting for BCF, is as follows:

$$OP = (Q^{0.8036} \times 10^{-3.346}) \times 1.323$$

Model Statistics and Data

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

Model

$$\log_{10}(OP) = 0.8036 \times \log_{10}(Q) - 3.346$$

Computation method: Absolute Maximum Likelihood Estimation (AMLE)

Explanatory Variables

Coefficients:

Estimate Std. Error z-score p-value

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) = 0.3488
Distribution: normal
Number of observations = 99, number censored = 34 (34.3 percent)

Log-likelihood (model) = -43.11
Log-likelihood (intercept only) = -105.1
Chi-square = 124
degrees of freedom = 1
p-value = <0.0001

Computation method: AMLE

Pseudo-R-squared: 0.6984

Akaike Information Criterion: 92.23
Bayesian Information Criterion: 100
```

Outlier Test Criteria

leverage cooksD 0.0303 0.6981

Flagged Observations

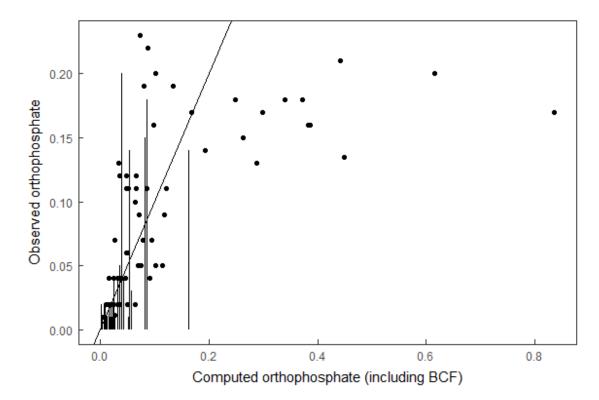
```
Observations exceeding at least one test criterion
   logOrthoP ycen
                             resids leverage
                      yhat
                                                 cooksD
18
     -0.8697 FALSE -0.4699 -0.399739 0.04679 3.382e-02
            TRUE -2.5672 -0.039113 0.05402 3.796e-04
23
     -2.0000
24
    -2.0000 TRUE -2.3088 -0.115788 0.03575 2.119e-03
25
    -0.7447 FALSE -0.5893 -0.155456 0.03857 4.144e-03
26
    -0.7447 FALSE -0.7248 -0.019892 0.03049 5.276e-05
27
    -2.0000 TRUE -2.3368 -0.104827 0.03749 1.828e-03
     -0.8239 FALSE -0.7013 -0.122653 0.03180 2.097e-03
33
41
     -0.7959 FALSE -0.5342 -0.261705  0.04224 1.296e-02
     -2.0000 TRUE -2.4183 -0.076615 0.04289 1.130e-03
47
49
     -2.0000 TRUE -2.2512 -0.140463 0.03234 2.801e-03
     -2.0000 TRUE -2.2512 -0.140463 0.03234 2.801e-03
50
57
     -0.6778 FALSE -0.4757 -0.202123  0.04637 8.562e-03
     -0.7959 FALSE -0.5386 -0.257280 0.04193 1.243e-02
61
66
    -0.8861 FALSE -0.6628 -0.223238  0.03402 7.467e-03
73
     -0.6990 FALSE -0.3318 -0.367180 0.05761 3.595e-02
74
     -0.7696 FALSE -0.6474 -0.122121 0.03494 2.299e-03
75
     -0.7696 FALSE -0.1996 -0.569995  0.06928 1.068e-01
79
     -1.6990
            TRUE -2.3054 -0.032005 0.03554 1.609e-04
80
     -1.6990
            TRUE -2.7641 -0.001315 0.07122 5.868e-07
     -0.7447 FALSE -0.5505 -0.194186 0.04112 6.932e-03
88
```

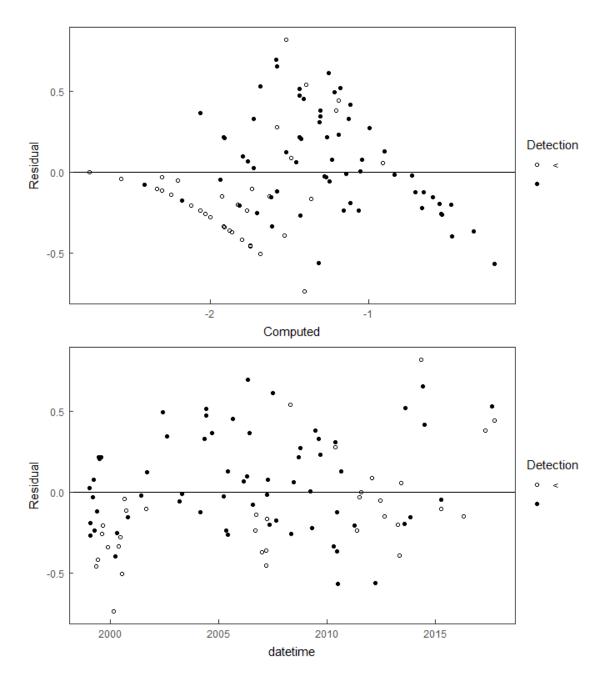
Bias correction factor

[1] 1.323394

2.5 % 97.5 % (Intercept) -3.6633322 -3.0277723 logQ 0.6785882 0.9286889

Plots





Variable Summary Statistics

Independent Variable (xvar) - Specific Conductance

Min. 1st Qu. Median Mean 3d Qu. Max. 5.291 78.989 223.226 654.116 512.975 8,216.570

Standard Deviation

[1] 1240.962

Dependent Variable (yvar) - Orthophosphate

Min. 1st Qu. Median Mean 3d Qu. Max. <0.01 <0.01 0.04 0.053 0.135 0.23

Standard Deviation

[1] 0.0939

Model-Calibration Data Set

			logOrthoP				Computed_logOP			
1	1999-01-26			2.017			-1.725			
2	1999-01-31			2.779			-1.112			
3	1999-02-03			2.382						
4	1999-03-17			2.584		384.00	-1.269		-0.03235	
5	1999-04-06			2.632			-1.231			
6	1999-04-16			2.843	0.05	696.50	-1.061			
7	1999-05-13			1.990		97.75	-1.746			
8	1999-05-24			2.199	0.02		-1.579			
9	1999-06-10				<0.01	84.00	-1.799			
	1999-06-25			2.375	0.06		-1.437			
	1999-07-02			2.386		243.25	-1.428			
	1999-07-14			1.788		61.33	-1.909			
	1999-07-29			1.778	0.02		-1.917			
	1999-08-12						-2.033			
	1999-08-26				<0.01		-2.124			
	1999-12-02			1.787	<0.01	61.25	-1.909			
	2000-02-25					259.33	-1.406			
	2000-03-24			3.578		3786.67				
	2000-04-27			2.040		109.75	-1.706			
	2000-05-25			1.778	<0.01	60.00	-1.917			
	2000-06-21			1.672	<0.01		-2.002		-0.27716	
	2000-07-26			2.068	<0.01	117.00	-1.684	0.02742	-0.50629	
23	2000-08-29	11:00:00		0.968	<0.01		-2.567	0.00358	-0.03911	
24	2000-09-28	10:30:00		1.290		19.50	-2.309	0.00650	-0.11579	
25	2000-10-26	10:50:00	-0.745	3.430	0.18	2690.00	-0.589	0.34059	-0.15546	
	2001-06-06			3.261	0.18	1824.17	-0.725	0.24928	-0.01989	
	2001-09-04			1.255	<0.01	18.00	-2.337			
28	2001-09-19	10:25:00	-1.4	2.273	0.04	187.33	-1.519			
	2002-06-12					446.33	-1.216			
	2002-08-14		-0.959	2.539		346.08	-1.305	0.06555	0.34636	
	2003-03-18			2.615	0.05	412.00	-1.244	0.07541	-0.05691	
32	2003-04-21	11:30:00			0.07	551.00	-1.143	0.09525	-0.01224	
	2004-03-05					1951.67	-0.701			
	2004-05-14				0.16	575.83	-1.127	0.09869	0.33139	
35	2004-06-14	09:45:00			0.12		-1.436	0.04852	0.51482	
	2004-06-14		-0.959		0.11		-1.436		0.47704	
	2004-09-08			1.591	0.02	39.00	-2.067		0.36794	
	2005-03-24			2.576	0.05	376.75	-1.275	0.07018	-0.02569	
	2005-05-16			2.721	0.04		-1.159		-0.23908	
40	2005-06-10	10:55:00		3.048		1116.67	-0.896			
	2005-06-13		-0.796			3150.00	-0.534			
	2005-08-29		-0.959			255.17	-1.411		0.45273	
	2006-03-02			1.967	0.02	92.67	-1.765			
	2006-05-01			1.928	0.02	84.75	-1.796			
	2006-05-12		-0.886		0.13	155.50	-1.584		0.69814	
	2006-06-05			1.591	0.02	39.00	-2.067		0.36794	
	2006-07-31			1.154	<0.01	14.25	-2.418		-0.07662	
	2006-09-07			1.591	<0.01	39.00	-2.067			
	2006-09-21			1.362	<0.01	23.00	-2.251			
	2006-09-21			1.362	<0.01	23.00	-2.251			
	2007-01-09			1.845	<0.01	70.00	-1.863		-0.37115	
	2007-03-14			1.824	<0.01	66.67	-1.880		-0.35907	
	2007-03-22			1.987	<0.01	97.00	-1.749		-0.45541	
	2007-03-26			2.471	0.03		-1.360		-0.16296	
	2007-03-31		-0.854			1325.00	-0.837		-0.01745	
56	2007-04-16	12:15:00	-0.959	2.871	0.11	743.50	-1.038	0.12119	0.07947	

E 7	2007-05-07 10:30:0	0 -0.678	2 E71	Q 21	3725.00	-0.476	0.44243	-0.20212	
	2007-05-07 10:30:0						0.44243	0.61529	
	2007-08-29 10:25:0			0.23	401.00	-1.254			
			1.447	<0.01	28.00	-2.183	0.00869	-0.17342	
	2008-04-24 11:40:0			0.14		-1.396	0.05315	0.54216	
	2008-05-09 11:35:0				3110.31	-0.539	0.38274	-0.25728	
	2008-06-19 09:45:0		2.349	0.04	223.23	-1.458	0.04608	0.06007	
	2008-09-15 10:55:0		2.591	0.09	389.55	-1.264	0.07209	0.21792	
	2008-10-16 10:10:0			0.19	848.02	-0.992	0.13470	0.27093	
	2009-03-31 11:20:0		2.853	0.09	713.40	-1.053	0.11723	0.00675	
	2009-04-27 12:15:0				2178.89	-0.663	0.28754	-0.22324	
	2009-06-17 10:40:0			0.12	349.14	-1.302	0.06602	0.38108	
	2009-08-20 10:50:0		2.014	0.04	103.19	-1.727	0.02479	0.32936	
69	2009-09-10 11:30:0	0 -0.959	2.682	0.11	481.29	-1.190	0.08544	0.23126	
70	2010-04-23 10:00:0	0 <-1.7	2.159	<0.02	144.33	-1.610	0.03246	-0.33715	
71	2010-05-17 16:40:0	0 -1.3	2.197	0.05	157.26	-1.580	0.03478	0.27925	
72	2010-05-27 10:00:0	0 -1	2.533	0.1	341.23	-1.310	0.06481	0.30990	
73	2010-06-14 11:30:0	0 -0.699	3.750	0.2	5625.32	-0.332	0.61618	-0.36718	
74	2010-06-16 10:15:0	0 -0.77	3.357	0.17	2277.10	-0.648	0.29791	-0.12212	
75	2010-07-06 10:30:0	0 -0.77	3.915	0.17	8216.57	-0.200	0.83548	-0.56999	
76	2010-08-25 11:00:0	0 -0.77	3.047	0.17	1114.21	-0.897	0.16774	0.12734	
77	2011-04-13 10:00:0	0 <-1.7	1.904	<0.02	80.15	-1.816	0.02023	-0.20861	
78	2011-05-23 10:20:0	0 <-1.7	1.965	<0.02	92.25	-1.767	0.02265	-0.23681	
79	2011-06-28 10:00:0	0 <-1.7	1.294	<0.02	19.69	-2.306	0.00655	-0.03200	
80	2011-07-27 11:20:0	0 <-1.7	0.724	<0.02	5.29	-2.764	0.00228	-0.00132	
81	2012-02-06 09:45:0	0 -1.4	2.313	0.04	205.38	-1.487	0.04310	0.08916	
82	2012-03-23 10:15:0	0 <-1.7	2.529	<0.02	337.76	-1.314	0.06428	-0.56158	
83	2012-06-20 09:15:0	0 <-1.7	1.414	<0.02	25.93	-2.209	0.00817	-0.05139	
84	2012-08-27 09:30:0	0 <-1.7	1.762	<0.02	57.83	-1.930	0.01557	-0.15004	
85	2013-04-11 10:10:0	0 <-1.7	1.891	<0.02	77.82	-1.826	0.01976	-0.20291	
86	2013-05-10 10:00:0	0 <-1.7	2.258	<0.02	180.98	-1.531	0.03893	-0.39314	
87	2013-05-31 10:00:0	0 -0.854	3.033	0.14	1078.66	-0.908	0.16342	0.05434	
88	2013-08-05 10:05:0			0.18	3005.69	-0.551	0.37236	-0.19419	
89	2013-08-16 08:30:0	0 -0.658	2.699	0.22	499.95	-1.177	0.08809	0.51901	
	2013-10-31 10:00:0		2.153	<0.04	142.12	-1.616	0.03206	-0.15610	
	2014-05-13 10:00:0			0.2	186.89	-1.520	0.03995	0.82105	
92	2014-06-10 10:30:0	0 -0.921	2.199	0.12	157.96	-1.579	0.03490	0.65791	
	2014-07-02 09:10:0			0.2	597.26	-1.115	0.10163	0.41555	
	2015-04-08 09:45:0		1.754	<0.04	56.75	-1.936	0.01533	-0.04512	
	2015-04-14 09:55:0		1.999	<0.04	99.67	-1.740	0.02411	-0.10304	
	2016-04-19 10:25:0		2.138	<0.04	137.55	-1.627	0.03123	-0.15068	
	2017-04-20 12:00:0			0.15	463.31	-1.203	0.08287	0.37925	
	2017-08-11 11:00:0		2.066	0.07	116.31	-1.686	0.02729	0.53065	
	2017-09-28 10:30:0			0.18	484.79	-1.187	0.08594	0.44261	
		0.7.43		0.10	10 1.75	1.10,	0.0005T	J ZUI	

Definitions

OP: orthophosphate, in milligrams per liter as phosphorus (00671)

Q: streamflow, instantaneous, in cubic feet per second (00061)

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.]
- 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.]
- 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.]