

Appendix 6.9. Model Archive Summary for Atrazine Concentration at U.S. Geological Survey station 375350097262800; Little Arkansas River upstream of ASR Facility near Sedgwick, Kansas, during May 2016 through August 2021

This model archive summary summarizes the atrazine model developed to compute hourly or daily atrazine. 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).

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: 375350097262800

Site Name: Little Arkansas River upstream of ASR Facility near Sedgwick, Kansas

Location: Latitude 37°53'49.7", longitude 97°26'28.0" referenced to North American Datum of 1983, in NE 1/4 NW 1/4 NW 1/4 sec.9, T.25 S., R.1 W., Sedgwick County, Kansas; hydrologic unit 11030012.

Equipment: A Sutron Satlink II High Data Rate Collection Platform (DCP) collected and transmitted stage data measured by a Sutron submersible pressure transducer and water-quality data measured by a water-quality and nitrate monitors. The DCP transmitted real-time stage and water-quality data via satellite. The primary reference gage is the top of the PVC well casing at 34.74 feet. The transducer is enclosed in a vertical two-inch PVC pipe on the west side of the building between gates 1 and 2 of the ASR intake building. Gage height was measured during April 2011 through December 2021. A YSI 6600 water-quality monitor equipped with water temperature, specific conductance, pH, dissolved oxygen, and turbidity (YSI Model 6136) sensors collected data during April 2011 through October 2015. 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 October 2015 through December 2021. A Hach Nitratax monitor collected nitrate data during March 2016 through December 2021.

Date model was developed: June 1, 2022

Model calibration data period: May 3, 2016 through August 25, 2021

Model Data

All data were collected using USGS protocols (U.S. Geological Survey, variously dated; Wagner and others, 2006; Sauer and Turnipseed, 2010) and are stored in the National Water Information System (NWIS) database (U.S. Geological Survey, 2022). Explanatory variables were evaluated individually and in combination. Potential explanatory variables included gage height, 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 19 concomitant values of discretely collected atrazine and continuously measured turbidity during May 2016 through August 2021. Discrete samples were collected over a range of gage height and turbidity conditions. No samples in the final model dataset 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, DFFITS, Cook's D (Cook, 1977), and leverage. Two samples (collected on February 24, 2020, and May 12, 2021) were not representative of the dataset and exceeded DFFITS and Cook's D outlier criteria and were removed from the model dataset. Removing data points based only on outlier criteria may overestimate the certainty of the model.

Atrazine

Discrete samples were collected near the northeast corner of the ASR intake building using 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 6 samples per year with a weighted basket sampler with a 1-Liter Teflon bottle or a DH-81 with a 1-Liter Teflon bottle, cap, and nozzle. Samples were analyzed for atrazine by or the USGS National Water Quality Laboratory according to standard methods (American Public Health Association and others, 1995).

Continuous Data

Concomitant turbidity values were time interpolated. If no concomitant continuous data were available within two 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 total organic carbon to turbidity 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 atrazine were examined for homoscedasticity (departures from zero did not change substantially over the range of model-calculated values).

Turbidity and seasonal components were selected as the best predictors of atrazine based on residual plots, high coefficient of determination (R^2), and low model standard percentage error (MSPE). Turbidity was positively correlated with atrazine.

Model Summary

Summary of atrazine regression analysis at station 375350097262800:

Atrazine-based model:

$$\log_{10}(ATR) = 0.646 \times \log_{10}(TBY) + 0.556 \times \sin(2\pi D) - 0.243 \times \cos(2\pi D) - 1.22$$

where,

\log_{10} = logarithm base 10;

ATR = atrazine, in micrograms per liter ($\mu\text{g/L}$);

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

D = date in decimal years

The log-transformed model may be retransformed to original units so that atrazine 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 1.23.

Model Statistics, Data, and Plots

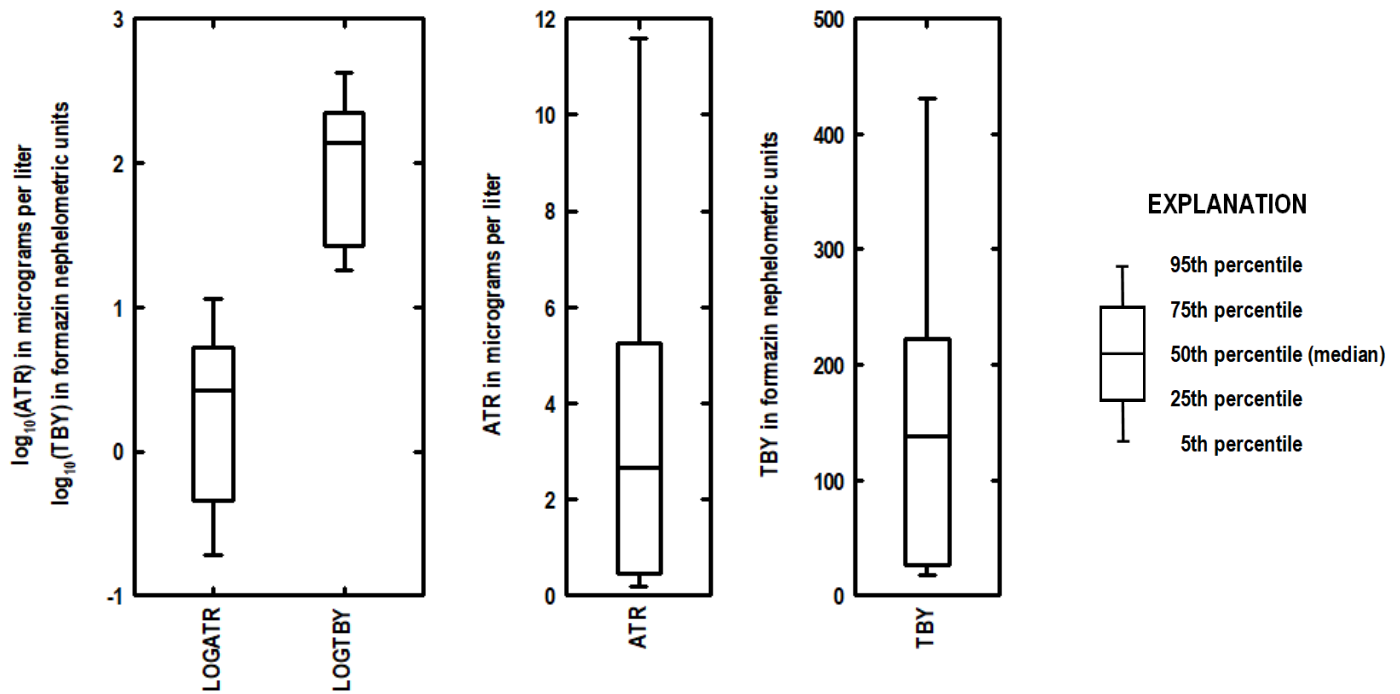
Model

$$\text{LOGATR} = + 0.646 * \text{LOGTBY} + 0.556 * \text{SIN2PID} - 0.243 * \text{COS2PID} - 1.22$$

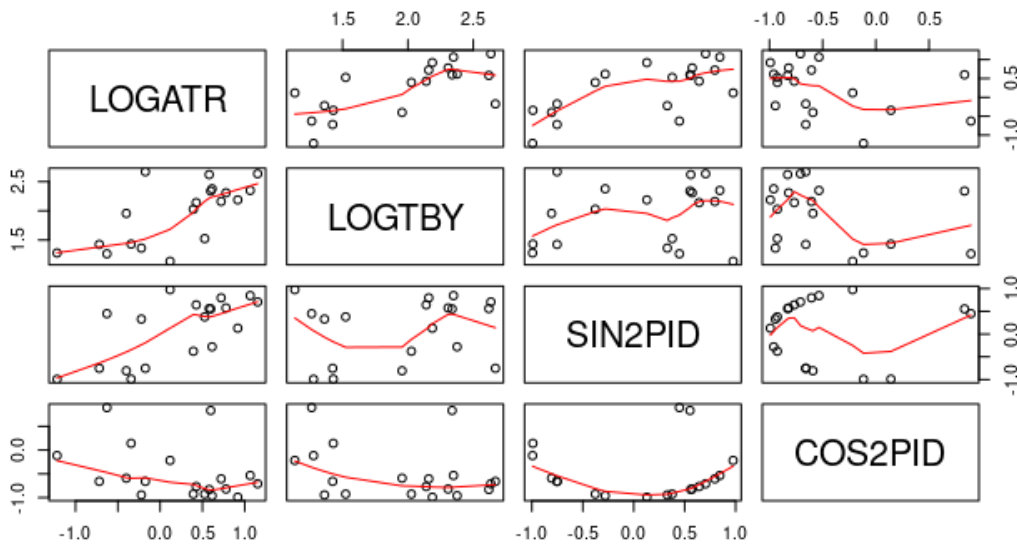
Variable Summary Statistics

	LOGATR	ATR	LOGTBY	SIN2PID	COS2PID	TBY
Minimum	-1.220	0.0605	1.13	-0.993	-0.992	13.6
1st Quartile	-0.343	0.4540	1.42	-0.753	-0.926	26.6
Median	0.427	2.6700	2.14	0.378	-0.658	138.0
Mean	0.220	3.6600	1.96	0.104	-0.494	157.0
3rd Quartile	0.721	5.2600	2.35	0.642	-0.218	223.0
Maximum	1.150	14.2000	2.67	0.976	0.894	467.0

Box Plots



Exploratory Plots



Basic Model Statistics

Number of Observations	19
Standard error (RMSE)	0.315
Average Model standard percentage error (MSPE)	79
Coefficient of determination (R^2)	0.813

Adjusted Coefficient of Determination (Adj. R²) 0.776
Bias Correction Factor (BCF) 1.23

Variance Inflation Factors (VIF)

LOGTBY SIN2PID COS2PID
1.18 1.05 1.13

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t)
(Intercept)	-1.220	0.295	-4.15	0.000863
LOGTBY	0.646	0.155	4.16	0.000846
SIN2PID	0.556	0.112	4.98	0.000164
COS2PID	-0.243	0.139	-1.74	0.102000

Correlation Matrix

	Intercept	LOGTBY	SIN2PID	COS2PID
Intercept	1.000	-0.943	0.1740	-0.1150
LOGTBY	-0.943	1.000	-0.2260	0.3410
SIN2PID	0.174	-0.226	1.0000	-0.0825
COS2PID	-0.115	0.341	-0.0825	1.0000

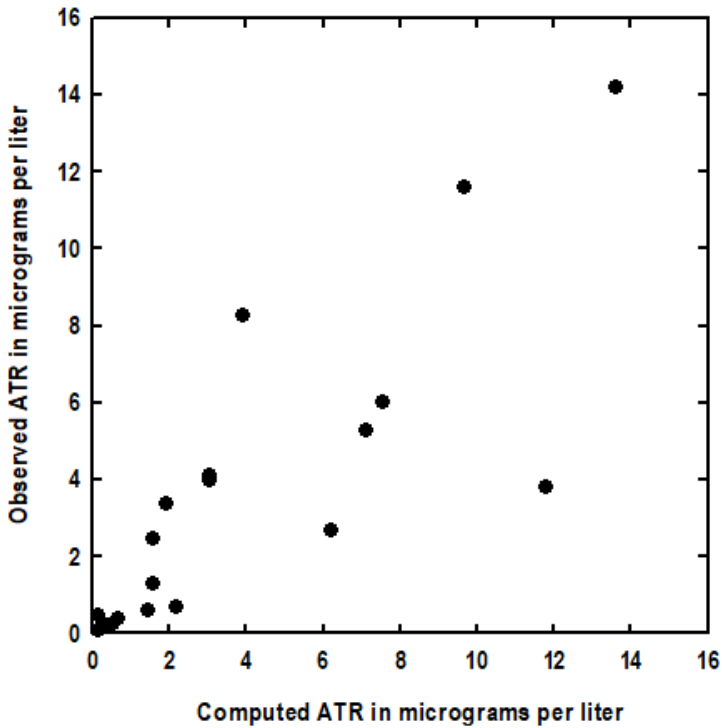
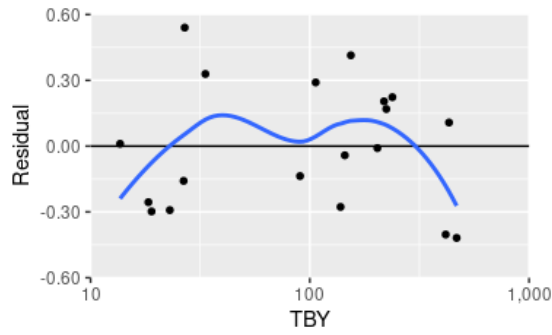
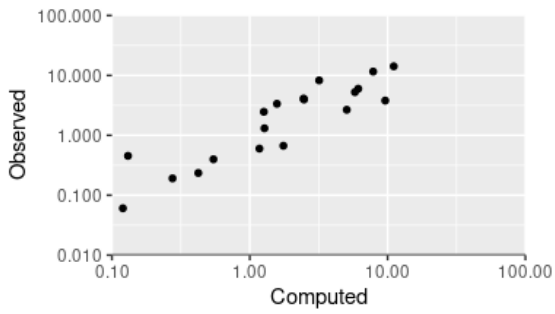
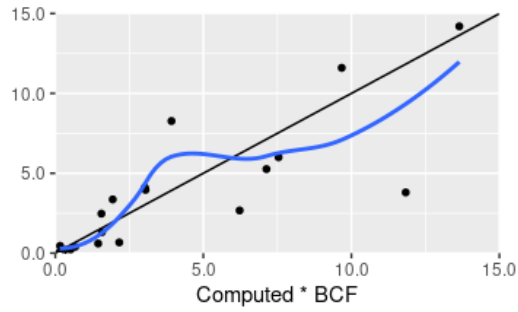
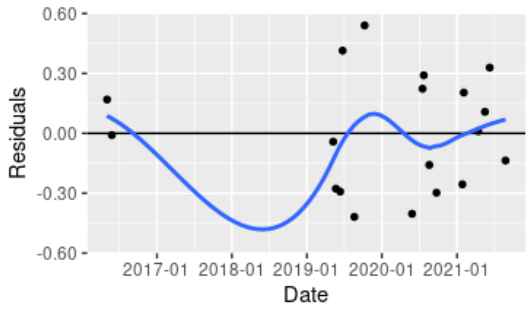
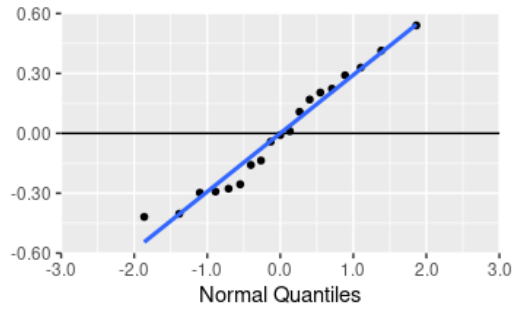
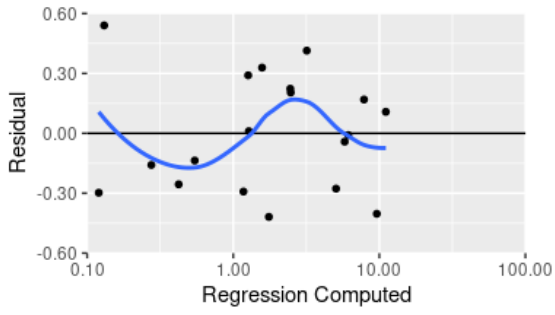
Outlier Test Criteria

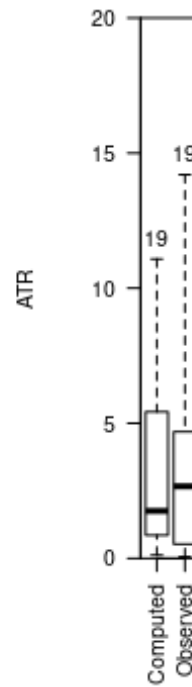
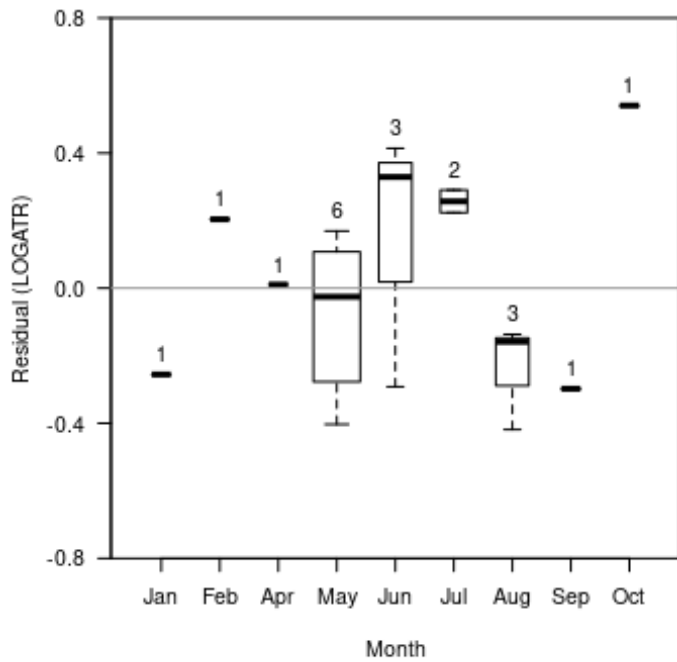
Leverage	Cook's D	DFFITS
0.632	0.309	0.918

Flagged Observations

	LOGATR	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
8/20/2019 10:00	-0.175	0.244	-0.419	-1.59	-1.68	0.301	0.272	-1.11
10/9/2019 10:20	-0.343	-0.883	0.541	2.01	2.28	0.272	0.379	1.39
1/27/2021 9:30	-0.629	-0.373	-0.256	-1.07	-1.08	0.425	0.212	-0.926
2/3/2021 9:30	0.598	0.394	0.204	0.922	0.917	0.506	0.218	0.928

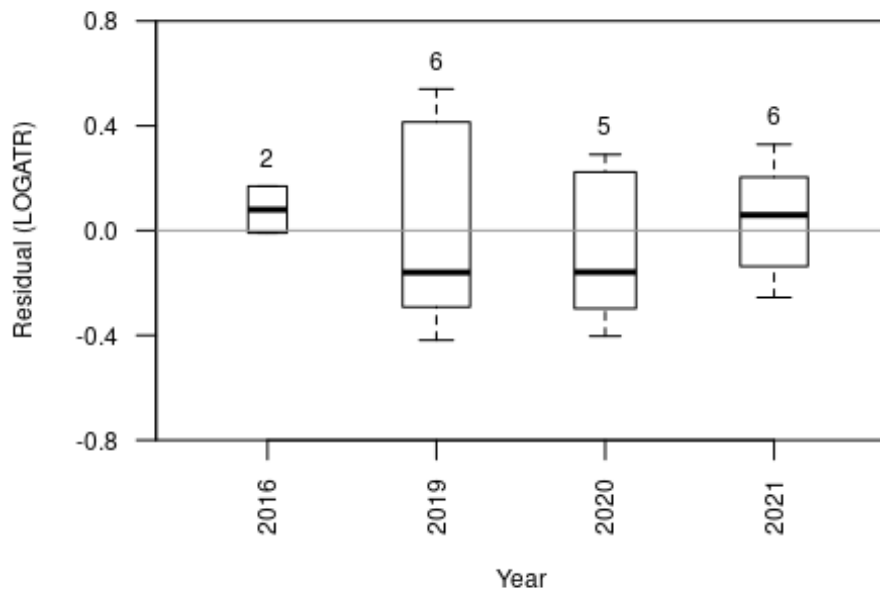
Statistical Plots





EXPLANATION

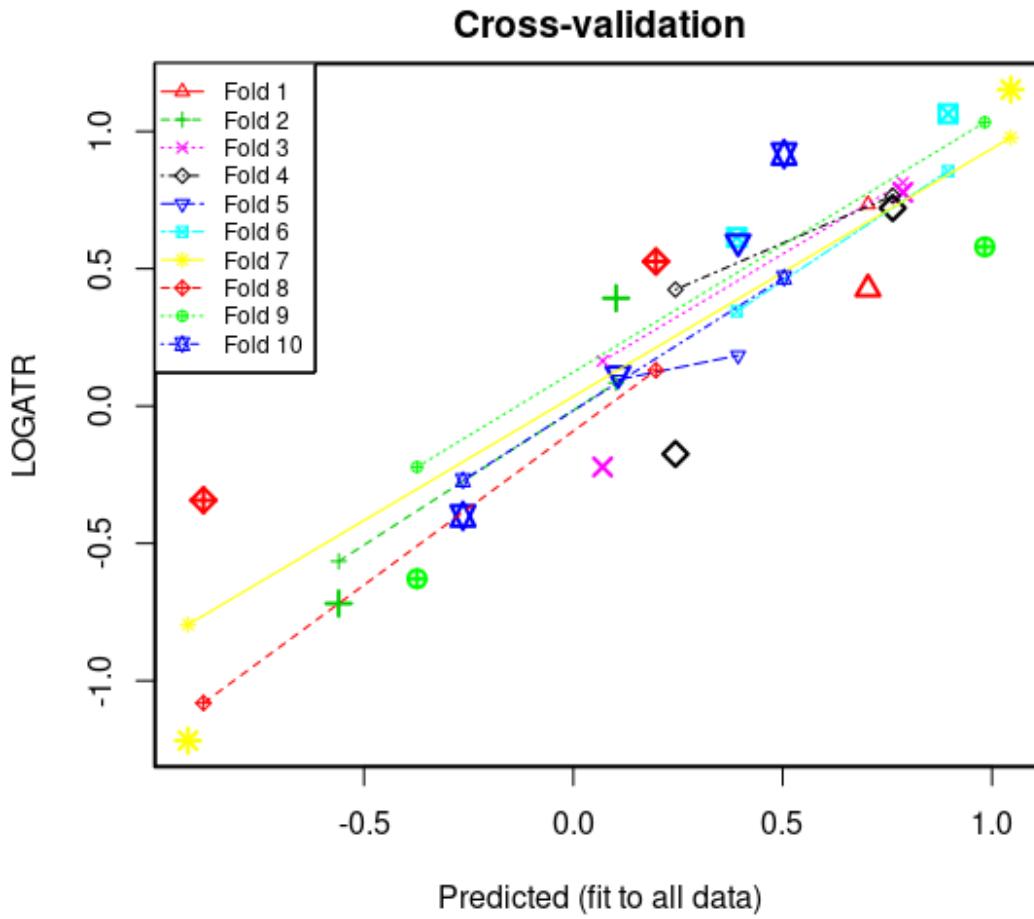
- 19 Number of values
- T Maximum value
- 75th percentile
- 50th percentile (median)
- 25th percentile
- Minimum value



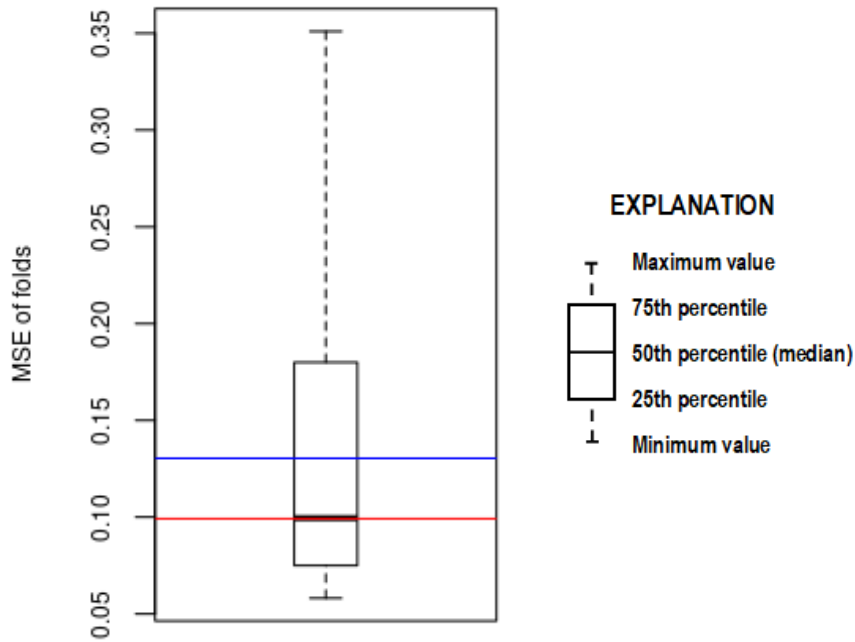
EXPLANATION

- 2 Number of values
- T Maximum value
- 75th percentile
- 50th percentile (median)
- 25th percentile
- Minimum value

Cross Validation



Minimum MSE of folds: 0.0581
Mean MSE of folds: 0.1300
Median MSE of folds: 0.0994
Maximum MSE of folds: 0.3510
(Mean MSE of folds) / (Model MSE): 1.3100



Red line - Model MSE

Blue line - Mean MSE of folds

Model-Calibration Dataset

	Date	LOGATR	LOGTBY	ATR	TBY	Computed LOGATR	Computed ATR	Residual	Normal Quantiles
1	5/3/2016	1.06	2.35	11.6	223	0.895	9.67	0.169	0.402
2	5/26/2016	0.778	2.31	6	203	0.787	7.54	-0.00887	0
3	5/9/2019	0.721	2.16	5.26	144	0.763	7.13	-0.0418	-0.131
4	5/22/2019	0.427	2.14	2.67	138	0.704	6.22	-0.277	-0.708
5	6/12/2019	-0.222	1.36	0.6	23	0.0699	1.45	-0.292	-0.887
6	6/24/2019	0.918	2.19	8.27	154	0.503	3.92	0.414	1.38
7	8/20/2019	-0.175	2.67	0.669	467	0.244	2.16	-0.419	-1.86
8	10/9/2019	-0.343	1.43	0.454	26.9	-0.883	0.161	0.541	1.86
9	5/27/2020	0.58	2.62	3.8	417	0.983	11.8	-0.403	-1.38
10	7/17/2020	0.614	2.38	4.11	238	0.39	3.03	0.223	0.708
11	7/23/2020	0.393	2.03	2.47	106	0.102	1.56	0.291	0.887
12	8/19/2020	-0.719	1.42	0.191	26.6	-0.56	0.339	-0.158	-0.402
13	9/23/2020	-1.22	1.28	0.0605	19	-0.92	0.148	-0.298	-1.1
14	1/27/2021	-0.629	1.26	0.235	18.4	-0.373	0.521	-0.256	-0.549
15	2/3/2021	0.598	2.34	3.96	218	0.394	3.05	0.204	0.549
16	4/14/2021	0.117	1.13	1.31	13.6	0.107	1.57	0.0106	0.131
17	5/17/2021	1.15	2.63	14.2	431	1.04	13.6	0.108	0.264
18	6/9/2021	0.526	1.52	3.36	33.4	0.197	1.94	0.329	1.1
19	8/25/2021	-0.4	1.96	0.398	90.3	-0.264	0.671	-0.136	-0.264

Definitions

ATR: Atrazine in ug/l (39632)

TBY: Turbidity in FNU (63680)

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.
- 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>.]
- 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., 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>.]
- 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, 2022, 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>.]