

Model Archive Summary for Total Phosphorus Concentration at Station 05518000; Kankakee River at Shelby, Indiana

This model archive summarizes the total phosphorus (TP) concentration model developed to compute hourly TP from December 5, 2015 onward. The method used follows USGS guidance as referenced in relevant 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: 05518000

Site name: Kankakee River at Shelby, IN

LOCATION.-- Lat 41°10'58", long 87°20'25" referenced to North American Datum of 1927, in SW 1/4 NE 1/4 sec.33, T.32 N., R.8 W., Newton County, IN, Hydrologic Unit 07120001, on right bank at upstream side of Highway 55 bridge, 1.0 mi south of Shelby, 7.8 mi upstream from Beaver Lake Ditch, and at mile 68.0.

Equipment: A YSI 6600 water-quality monitor equipped with sensors for water temperature, specific conductance (SC), dissolved oxygen, and pH, and a YSI Model 6136 turbidity sensor. The monitor is housed in an 8-inch plastic pipe. The pipe is affixed to the downstream side of the pier approximately 70 feet from the left bank of the river. Readings from the YSI 6600 are recorded every 15 minutes and transmitted by way of satellite, hourly. A YSI Turbidity sensor started operation December 5, 2015.

Model number: 05518000.TP.WY18.1

Date model was created: September 26, 2018

Model calibration data period: December 5, 2015 to May 31, 2018

Model application date: December 5, 2015 onward

Computed by: Tim Lathrop September 26, 2018

Reviewed by:

Approved by:

Model Data

All data were collected using U.S. Geological Survey (USGS) protocols and are stored in the National Water Information System (NWIS) database (<http://dx.doi.org/10.5066/F7P55KJN>). The regression model is based on 32 concurrent measurements of continuous in-situ turbidity and discrete total phosphorus samples collected from January 5, 2016 through May 31, 2018. Samples were collected throughout the range of continuously observed hydrologic and seasonal conditions. Summary statistics and the complete model-calibration data are provided in the dataset. Studentized residuals from the final model were inspected for values greater than 3 or less than negative 3. Values outside of the 3 to -3 range are considered potential outliers and were investigated. For this model set, there was no samples with studentized residuals greater than 3 or less than -3.

Total phosphorus (TP) Data

Cross-section samples were collected from the downstream side of the bridge using the equal width-increment method. Cross-section samples were collected over the range of hydrologic and seasonal conditions. A FISP US D-

95 was used with an automatic bridge crane. Samples were analyzed for total phosphorous and orthophosphate at the USGS National Water Quality Laboratory in Denver, CO.

Surrogate Data

The continuous turbidity data used in this analysis were measured using a YSI 6920 V2 sonde from 2015 onward.

For more information on continuous turbidity records, see the Station Analysis (available upon request).

Model Development

Regression analysis was done using R by examining continuously measured data (water temperature, specific conductance, pH, dissolved oxygen, turbidity, streamflow, and nitrate plus nitrite) as explanatory variables for estimating total phosphorus concentration. A variety of models that predict TP and models that predict $\log_{10}(\text{TP})$ 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 TP were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of predicted values). This comparison of several models led to the conclusion that the most appropriate and reliable model would be one that estimated untransformed total phosphorus (TP) from turbidity ($\sqrt{\text{TURB}}$).

Continuous turbidity was selected as the best predictors of transformed total phosphorus based on residual plots, adjusted coefficient of determination (adjusted R^2), and model standard percentage error ($MSPE$). Values for all of 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 phosphorus at site number 05518000.

Total phosphorus-based model:

$$\text{TP} = (0.02255 \times \sqrt{\text{TURB}}) + 0.02047 \quad \text{where}$$

TP = total phosphorus, in milligrams per liter (mg/L), and
TURB = turbidity, in formazin nephelometric units (FNU).

The use of turbidity as an explanatory variable is appropriate physically and statistically. Total phosphorus at the Kankakee River at Shelby is largely composed of undissolved particles (79% on average). Phosphorus also binds to sediment in the river. Increased turbidity within the river reflects increased non-dissolved particles. Therefore, turbidity can be used as a surrogate of total phosphorus.

Un-transformed model:

$$\text{TP} = (0.02255 \times \sqrt{\text{TURB}}) + 0.02047$$

Previous models

None

Total Phosphorus Record

The total phosphorus record was computed using this regression model in the National Real-Time Water Quality (NRTWQ) Web site from 2015 onward. Data were computed at hourly intervals. The complete water quality record can be found at <http://nrtwq.usgs.gov/in>.

Model Statistics, Data, and Plots

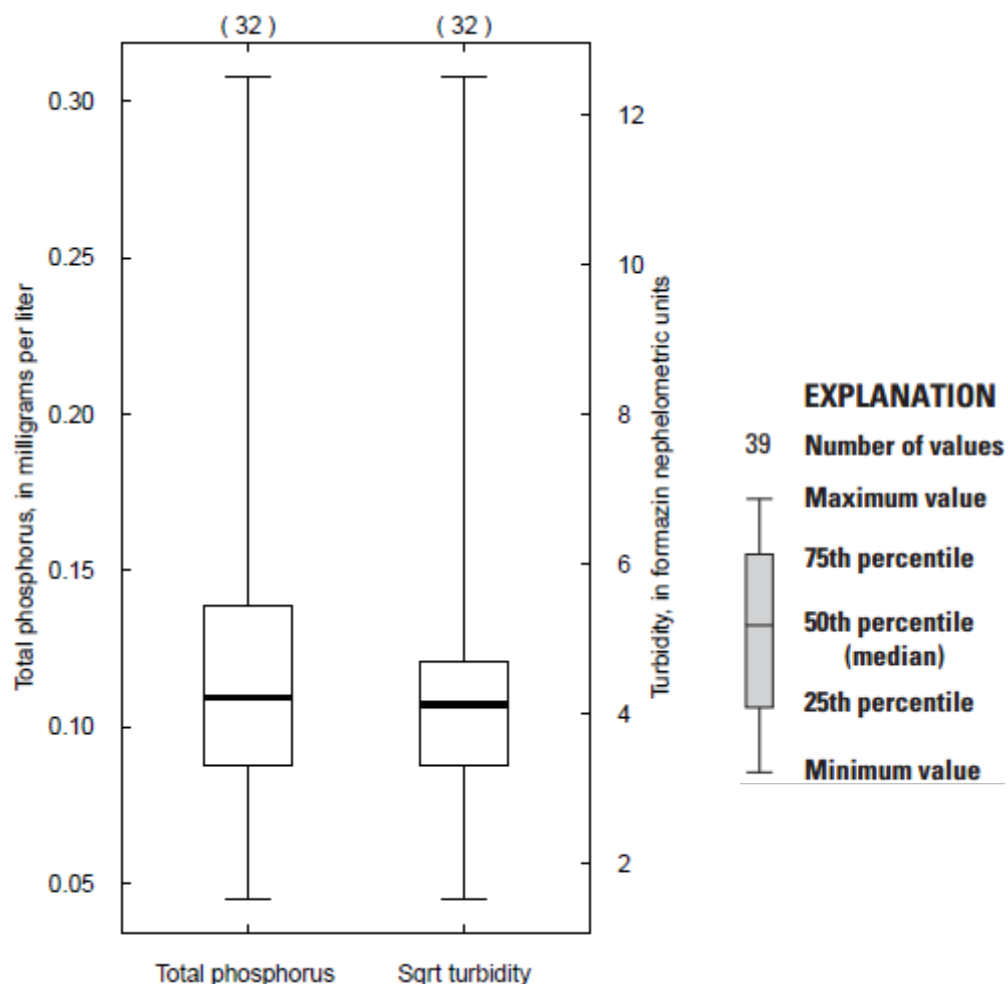
Model

$$TP = (0.02255 \times \text{sqrtTURB}) + 0.02047$$

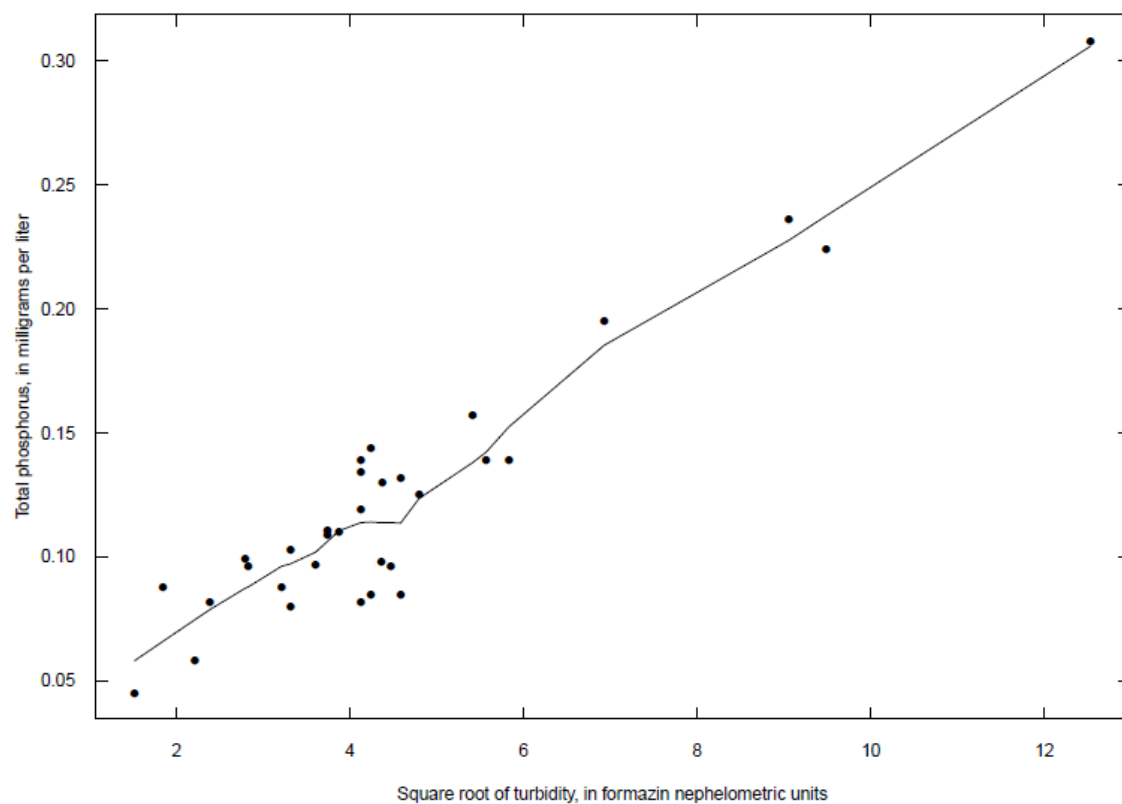
Variable Summary Statistics

	TP	logTP	Turb	sqrtTurb	Q	logQ
Q Minimum	0.045	-1.347	2.3	1.52	676	2.83
1st Quartile	0.088	-1.056	11.0	3.32	1,674	3.22
Median	0.110	-0.961	17.0	4.12	2,520	3.40
Mean	0.123	-0.944	25.6	4.54	2,492	3.34
3rd Quartile	0.139	-0.857	21.5	4.69	3,415	3.53
Maximum	0.308	-0.511	157.0	12.53	4,413	3.64

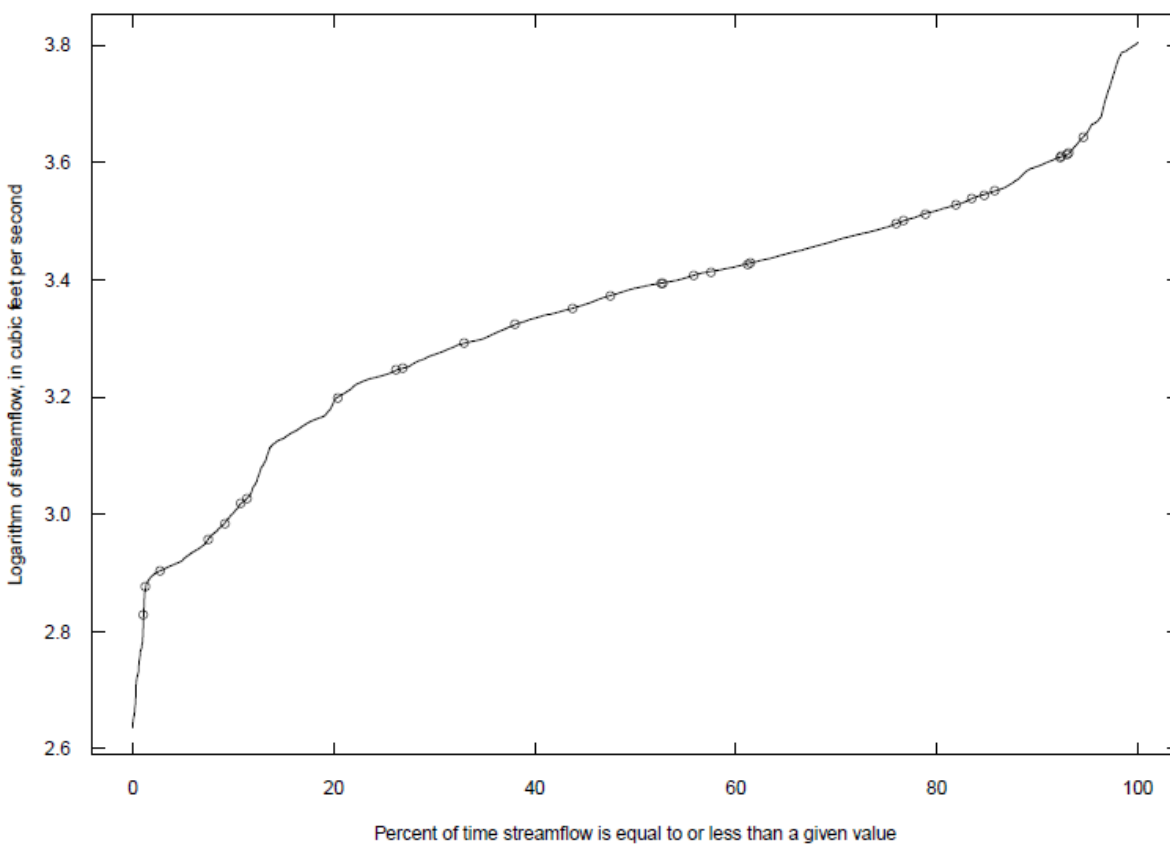
Box Plots



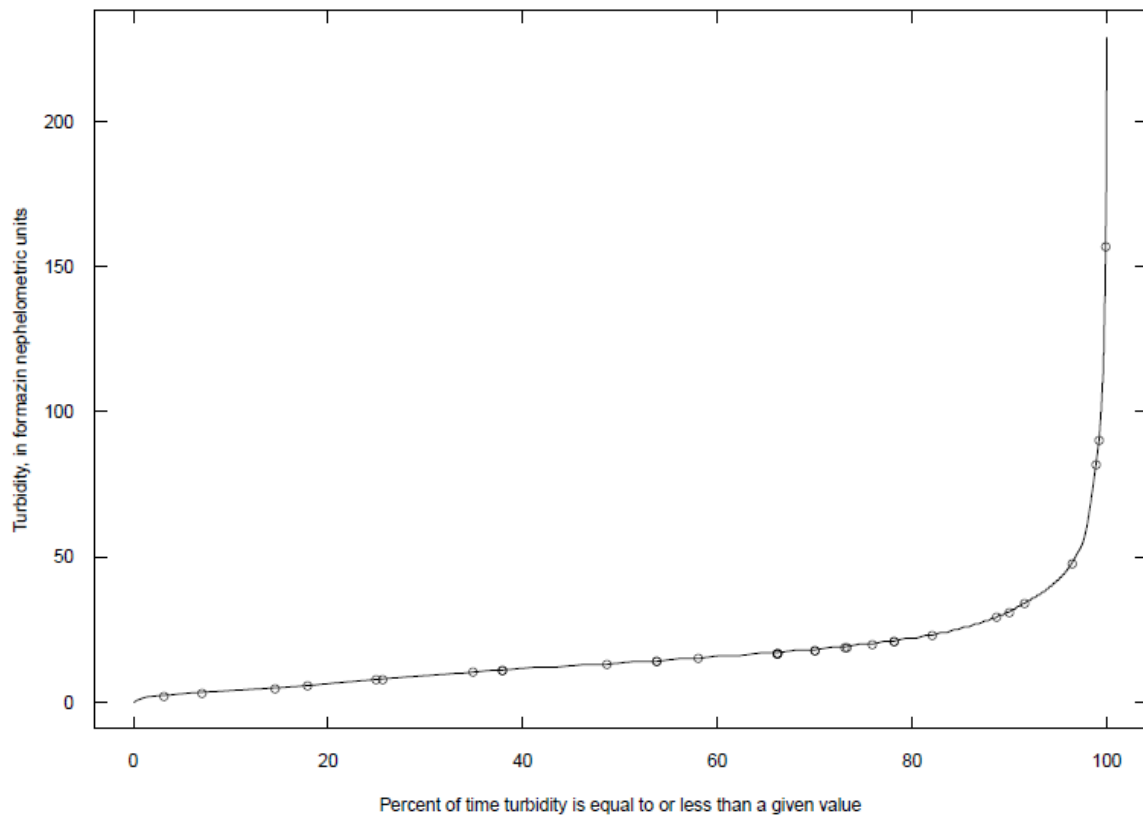
Exploratory Plots



December 5, 2015–May 31, 2018



December 5, 2015–May 31, 2018



Basic Model Statistics

Number of Observations	32
Standard error (RMSE)	0.0177
Average Model standard percentage error (MSPE)	
Upper MSPE	14.4
Lower MSPE	-14.4
Model average percent prediction interval (PI)	
Model Upper PI	6.68
Model Lower PI	-6.68
Coefficient of determination (R^2)	0.896
Adjusted coefficient of determination	0.892
PRESS statistic	0.0103

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t)
(Intercept)	0.02047	0.00711	2.88	0.0073
sqr tturb	0.0225	0.00141	16.05	2.9e-16

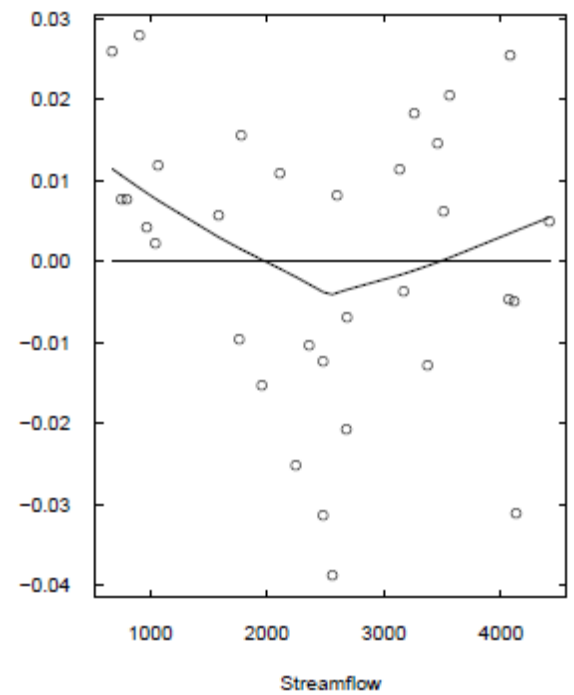
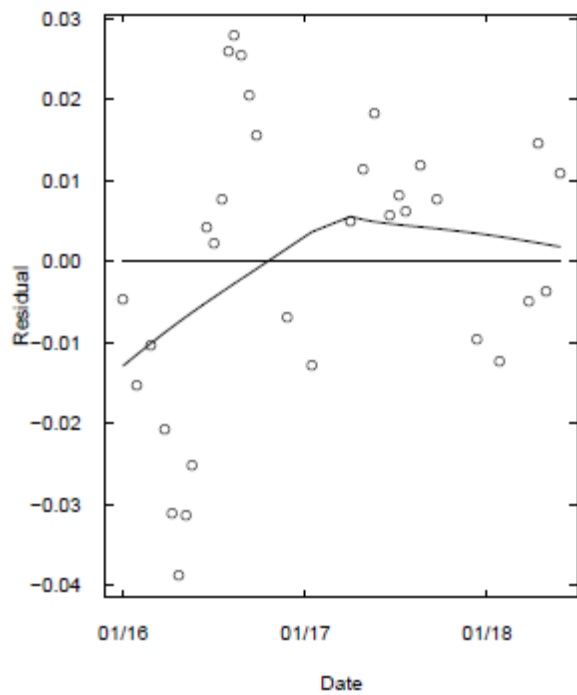
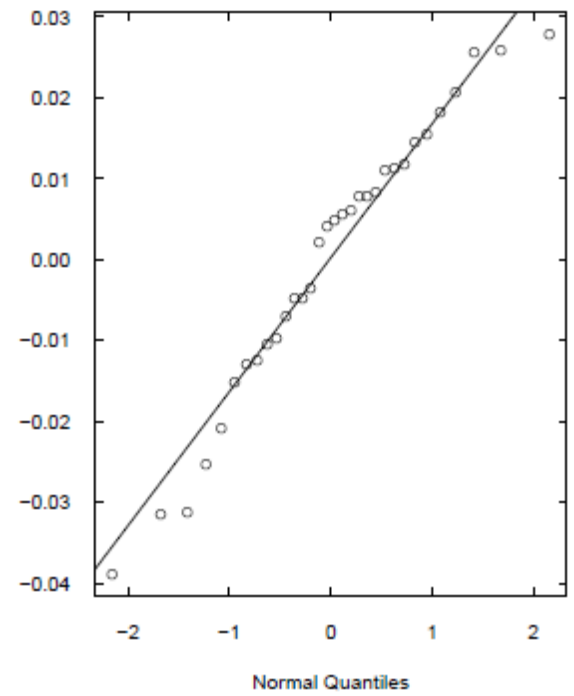
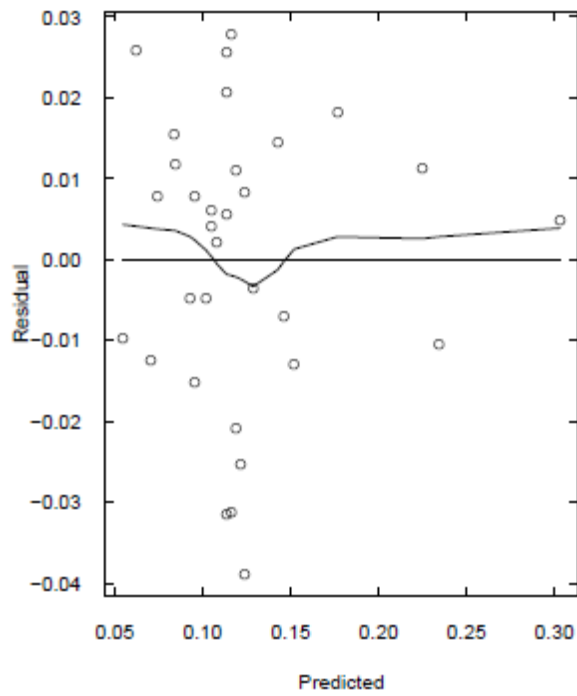
Outlier Test Criteria

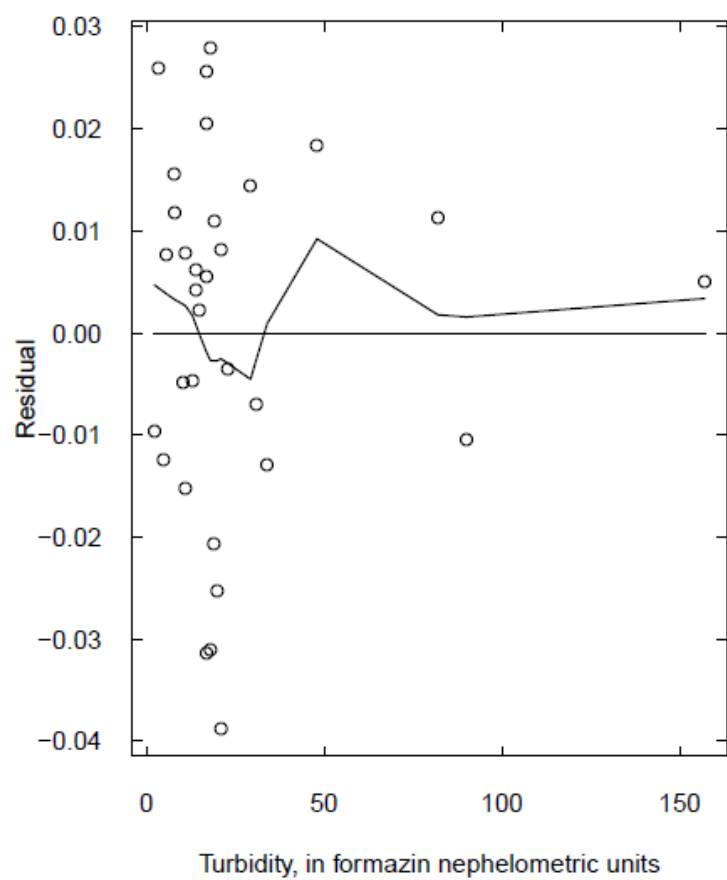
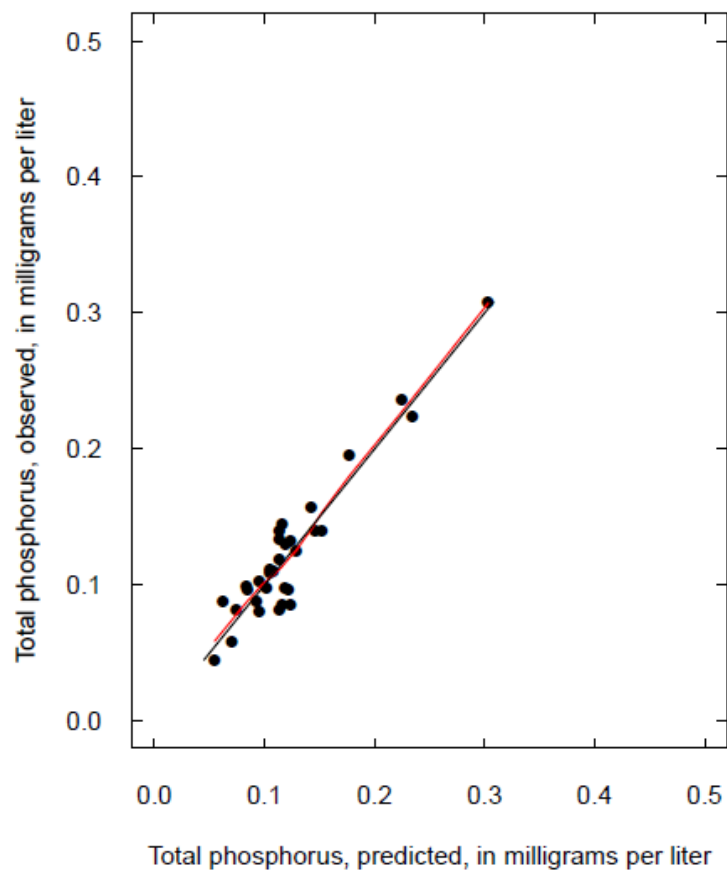
Leverage	Cook's D	DFFITS
0.188	0.807	0.500

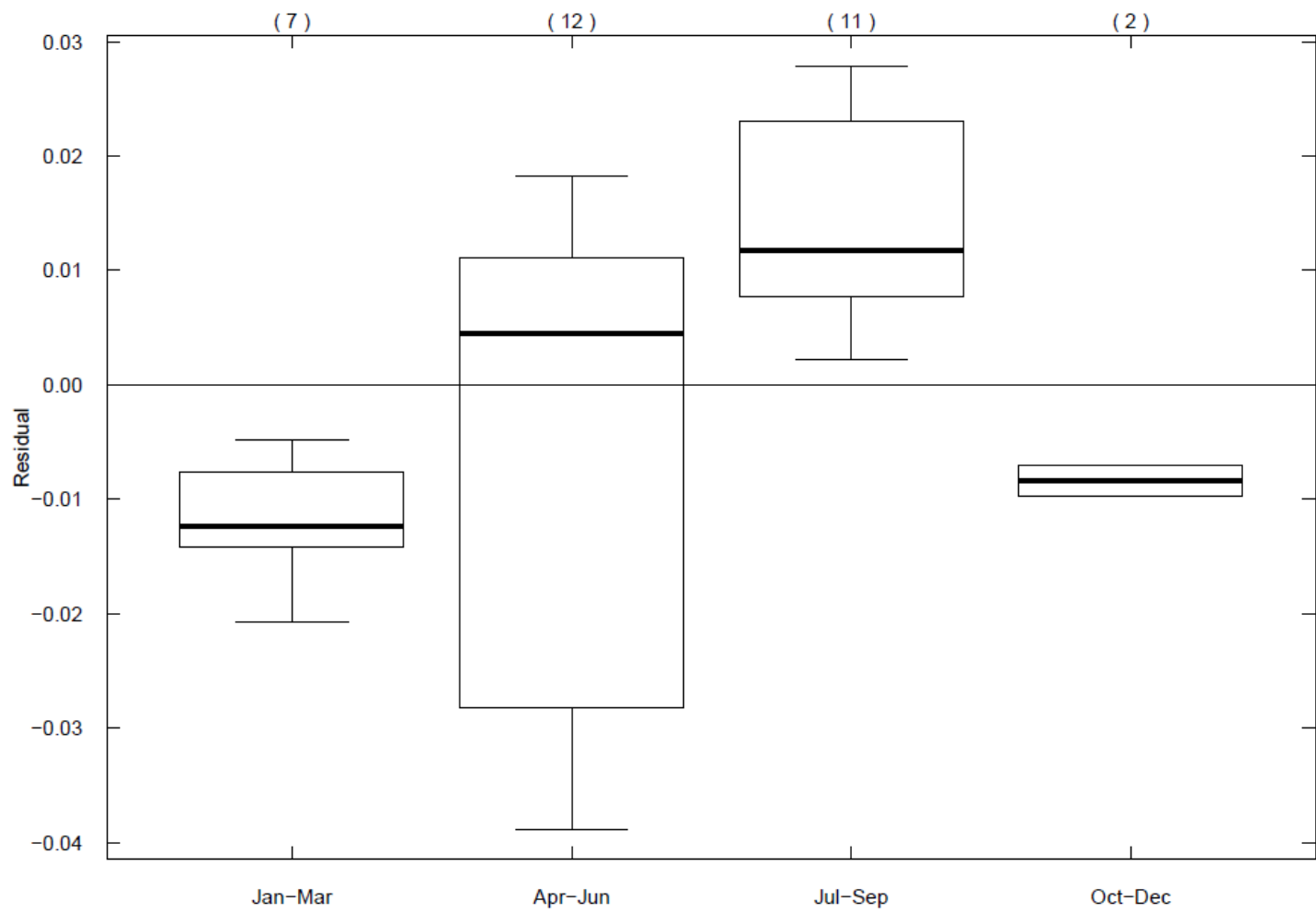
Flagged Observations

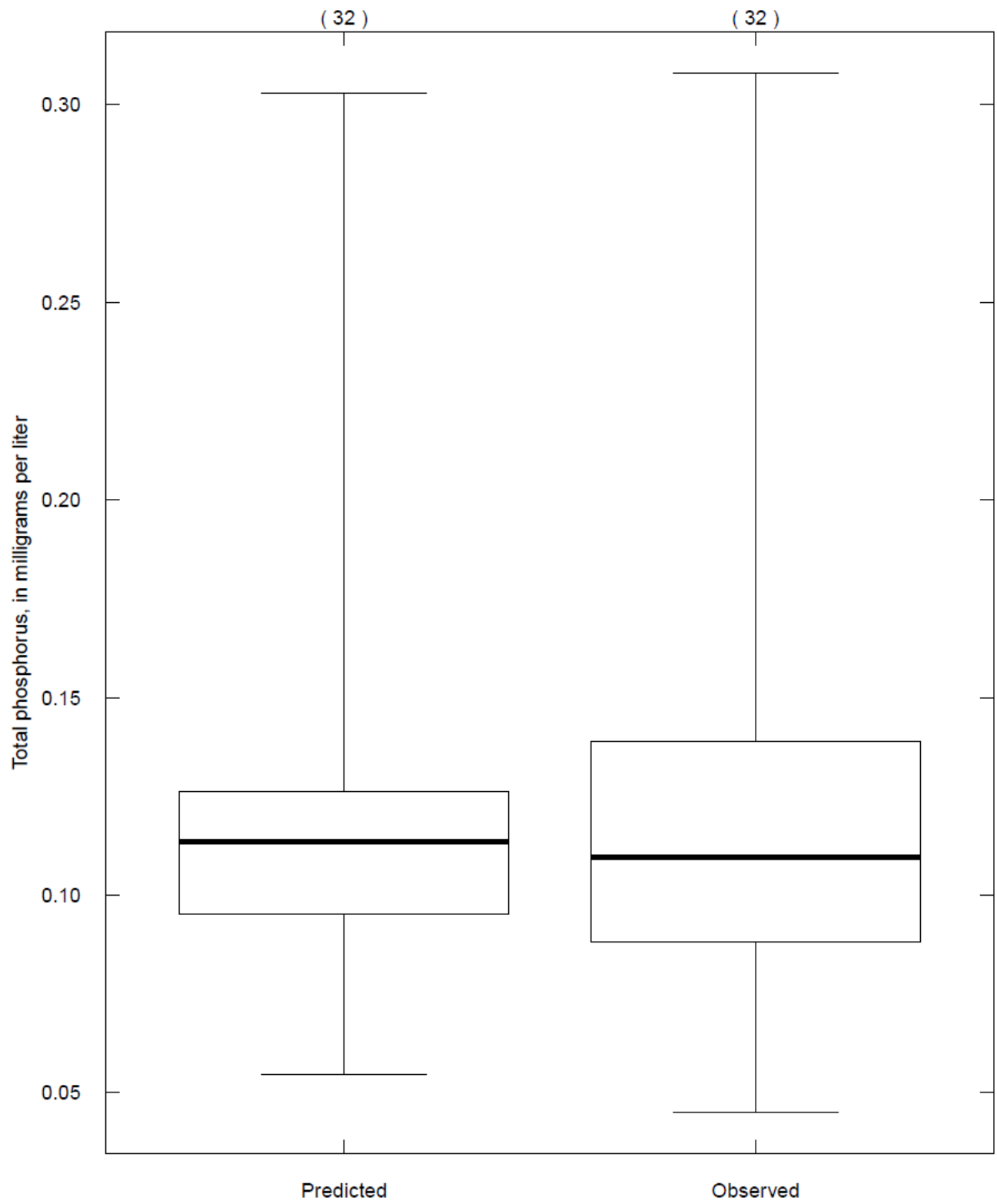
	tp	yhat	resids	stnd.res	stud.res	leverage	cooksD	dfits
4/6/2017	0.308	0.303	0.00494	0.371	0.365	0.433*	0.0525	0.319

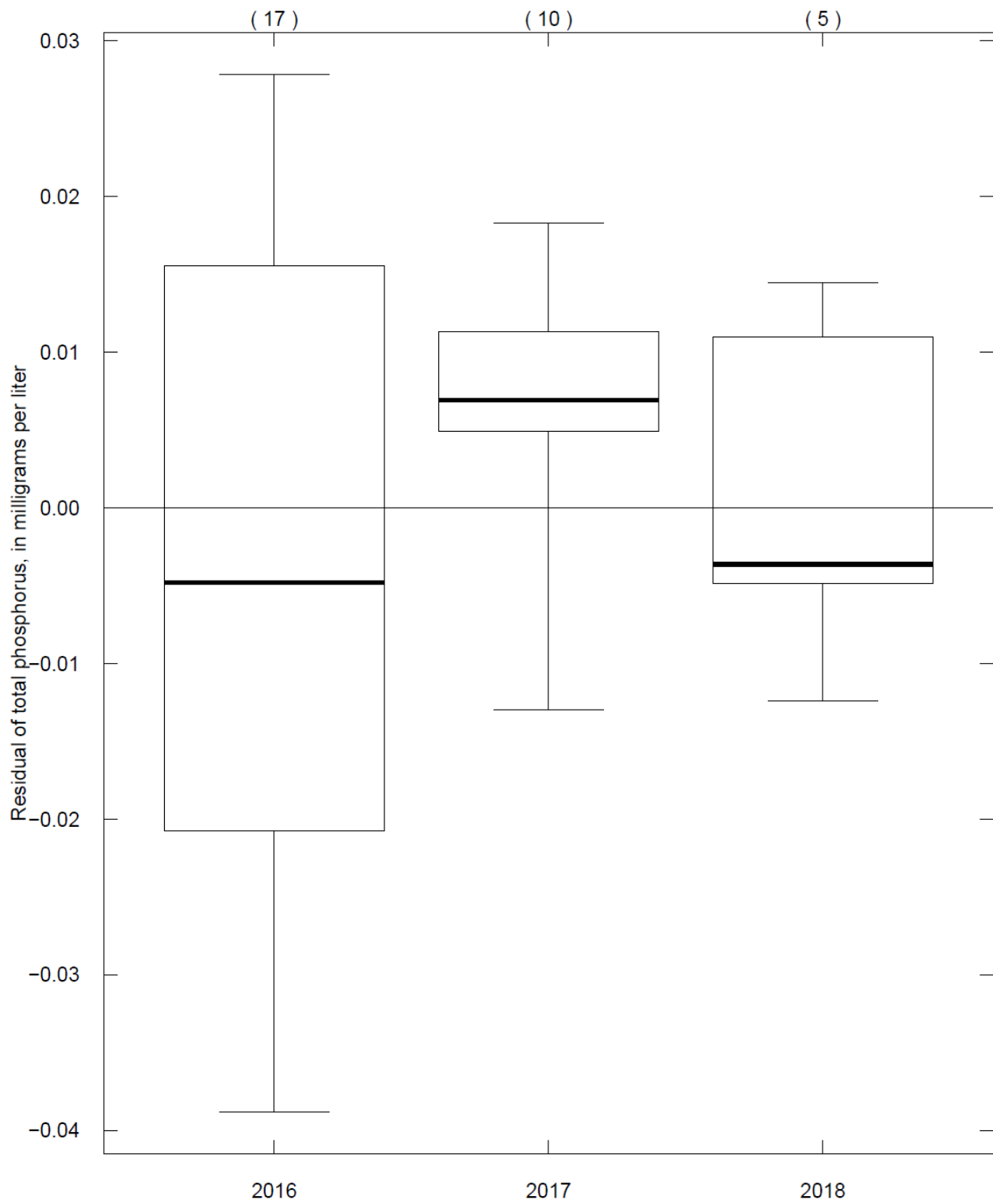
Statistical Plots



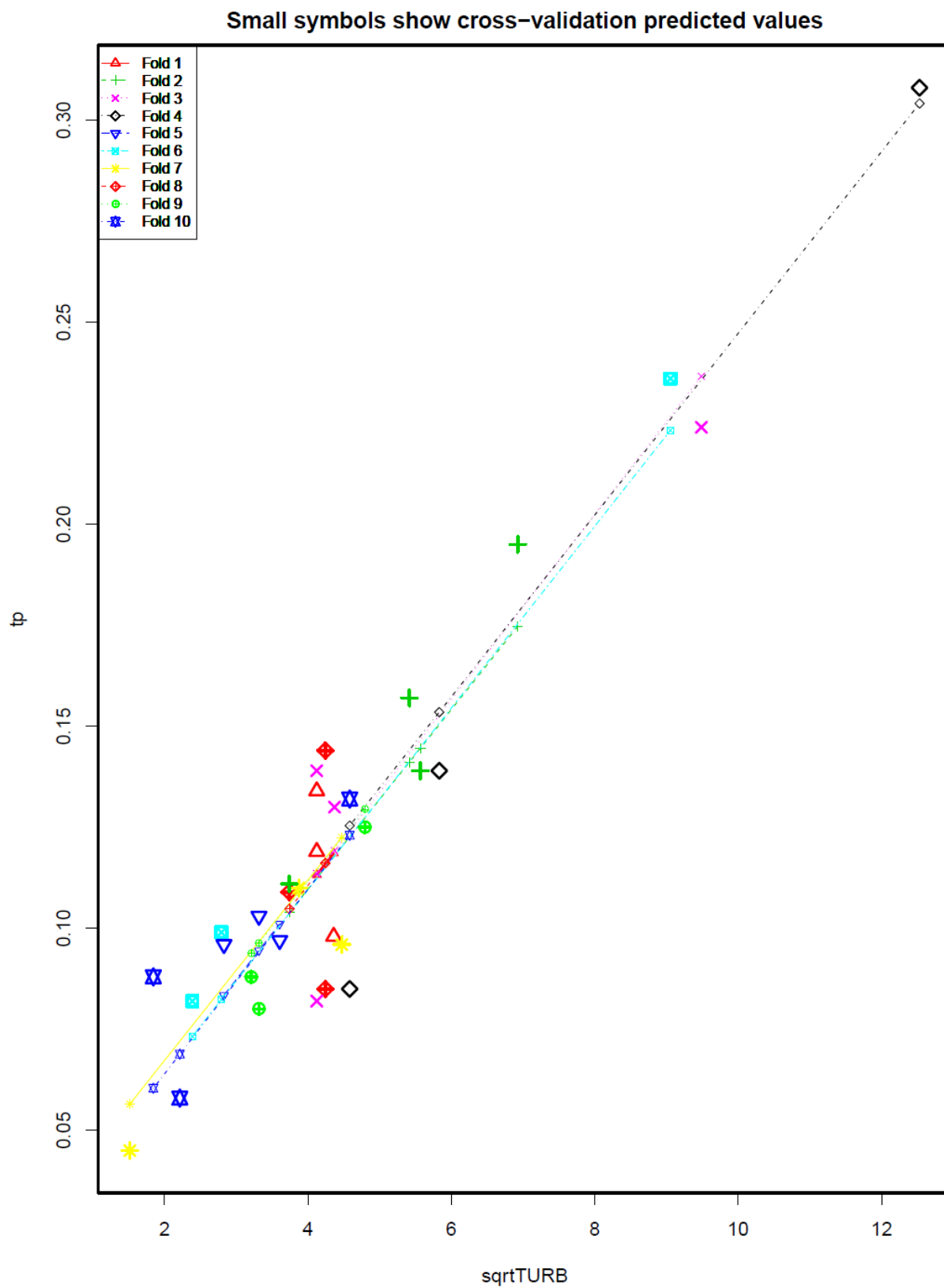








Cross Validation



Mean MSE of folds: 0.000314

Model MSE: 0.000313

(Mean MSE of folds) / (Model MSE): 1.003

Model-Calibration Data Set

Definitions

TP: Total phosphorus in milligrams per liter (62855)

TURB: Turbidity in formazin nephelometric units (63680)

datetime	q	do	ph	sc	wtemp	turb	no3	tp	Computed tp	Residuals	Normal Quantiles
1/5/2016 10:45	4,070	11.2	7.5	582	2.4	13	3	0.097	0.102	-0.005	-0.278
2/2/2016 9:45	1,957	10.8	7.8	632	5.6	11	1.72	0.08	0.095	-0.015	-0.947
2/29/2016 15:45	2,360	11.1	7.8	577	7	90	2.08	0.224	0.234	-0.010	-0.626
3/29/2016 9:00	2,678	9.9	8	598	9.9	19	2.33	0.098	0.119	-0.021	-1.078
4/12/2016 9:30	4,132	10.13	7.8	574	8.5	18	2.64	0.085	0.116	-0.031	-1.410
4/25/2016 17:00	2,558	9.1	8	611	17.3	21	1.68	0.085	0.124	-0.039	-2.150
5/10/2016 9:00	2,481	9.37	7.9	605	13.5	17	2.08	0.082	0.113	-0.031	-1.670
5/23/2016 15:20	2,247	8.7	7.7	601	18.8	20	1.74	0.096	0.121	-0.025	-1.220
6/21/2016 18:30	968	7.6	7.9	656	26.1	14	1.26	0.109	0.105	0.004	-0.039
7/6/2016 17:50	1,044	8.2	7.9	646	23.8	15	1.48	0.11	0.108	0.002	-0.118
7/21/2016 9:00	752	7.1	7.7	663	25.1	5.7	1.01	0.082	0.074	0.008	0.278
8/4/2016 11:00	676	7.2	7.9	656	25.9	3.4	1.07	0.088	0.062	0.026	1.670
8/15/2016 14:10	909	6.8	7.8	632	23.9	18	1	0.144	0.116	0.028	2.150
8/29/2016 15:15	4,077	5.84	7.3	552	22	17	1.3	0.139	0.113	0.026	1.410
9/15/2016 7:30	3,561	6.92	7.5	495	19.7	17	1.75	0.134	0.113	0.021	1.229
9/29/2016 7:50	1,776	8.55	7.7	621	15.3	7.8	1.22	0.099	0.083	0.016	0.946
11/29/2016 18:50	2,684	10.13	7.9	592	8.2	31	2.26	0.139	0.146	-0.007	-0.445
1/18/2017 12:50	3,372	10.79	7.6	543	5	34	2.23	0.139	0.152	-0.013	-0.830
4/6/2017 16:00	4,413	9.98	7.6	472	8.7	157	1.88	0.308	0.303	0.005	0.039
5/1/2017 14:20	3,134	8.65	7.8	513	12.7	82	1.86	0.236	0.225	0.011	0.626
5/23/2017 17:00	3,261	7.71	7.7	519	17.3	48	3.18	0.195	0.177	0.018	1.077
6/23/2017 13:00	1,585	7.02	7.7	633	22.5	17	1.45	0.119	0.113	0.006	0.117
7/12/2017 11:10	2,597	6.88	7.5	543	22.6	21	2.66	0.132	0.124	0.008	0.445
7/26/2017 10:50	3,510	6.84	7.5	559	22.1	14	1.96	0.111	0.105	0.006	0.197
8/24/2017 10:20	1,067	7.49	7.8	606	20.8	8	1.04	0.096	0.084	0.012	0.724
9/27/2017 9:30	801	6.83	7.8	633	22	11	1.03	0.103	0.095	0.008	0.360
12/15/2017 11:20	1,762	12.54	7.7	645	1.2	2.3	1.48	0.045	0.055	-0.010	-0.533
1/30/2018 9:40	2,480			610	2.8	4.9	2.12	0.058	0.070	-0.012	-0.724
3/30/2018 12:20	4,111	10.2	7.8	535	8.6	10	1.09	0.088	0.093	-0.005	-0.360
4/17/2018 18:50	3,458	11.01	8.1	529	6.4	29	2.81	0.157	0.143	0.014	0.830
5/4/2018 10:50	3,168	7.4	7.98	544	17.8	23	1.56	0.125	0.129	-0.004	-0.197
5/31/2018 10:10	2,108	6.35	7.9	619	23	19	1.25	0.13	0.119	0.011	0.533