

# **Appendix 24. Model Archival Summary for Total Phosphorous Concentration at Station 06887500; Kansas River at Wamego, Kansas**

This model archival summary summarizes the total phosphorous concentration (TP) model developed to compute 15-minute TP from July 19, 2012 onward. This model supersedes all previous models.

## **Site and Model Information**

Site number: 06887500

Site name: Kansas River at Wamego, Kansas

Location: Lat 39°11'54", long 96°18'19" referenced to North American Datum of 1927, in SW 1/4 NW 1/4 SE 1/4 sec.9, T.10 S., R.10 E., Pottawatomie County, KS, Hydrologic Unit 10270102.

Equipment: An YSI 6600 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, pH, turbidity, and chlorophyll was installed from August 2012 through May 2014. From June 2014 to the present (2015) a Xylem YSI EXO2 water-quality monitor equipped with sensors for water temperature, specific conductance, dissolved oxygen, pH, turbidity, and chlorophyll. The monitor is housed in a 4-inch diameter galvanized steel pipe. Readings from the water-quality monitor are recorded every 15 minutes and transmits data by way of satellite, hourly.

Date model was created: October 15, 2015

Model calibration data period: July 19, 2012 – June 29, 2015

Model application date: July 19, 2012 onward

## **Model-Calibration Dataset**

All data were collected using U.S. Geological Survey (USGS) protocols and are stored in the National Water Information System (NWIS) database. Linear regression models were developed using the open-source software package “R.” Explanatory variables selected as inputs to linear regression were physicochemical properties: specific conductance, pH, water temperature, dissolved oxygen, turbidity, chlorophyll fluorescence, and streamflow. Seasonal components (sine and cosine variables) were also evaluated as explanatory variables in the models to determine if seasonal changes affected the model. All combinations of physicochemical properties and a seasonal component were evaluated to determine which combinations produced the best models.

The final selected regression model is based on 49 concurrent measurements of TP concentration and turbidity (Turb) collected from July 19, 2012 through June 29, 2015. Samples were collected throughout the range of continuously observed hydrologic conditions. No samples were below laboratory detection. Summary statistics and the complete model-calibration dataset are provided below. Studentized residuals from the final model were inspected for values greater than 3 or less than negative 3. Values outside of that range are considered potential outliers and are investigated. Six samples collected during icy conditions were found to have erroneously high sand fractions, and were removed from the dataset (November 19, 2012, December 17, 2012, January 14, 2013, December 16, 2013, January 13, 2014, and February 10, 2014). No other potential outliers were found to have errors in collection, processing, or analysis, therefore they were retained.

## **Total Phosphorous Sampling Details**

Cross-section samples are typically collected either from the downstream side of the bridge or instream within 100 feet of the bridge. The equal-width-increment (EWI) method is used, and samples typically are composited for analysis. Cross-section samples are collected every 2 weeks from March through October, once a month from November through February, and during selected runoff events. A FISP US DH-95, D-95, or D-96A1 depth integrating sampler is used from the bridge; and a DH-81 or DH-95 hand sampler is used for boat samples. Samples are analyzed for TP concentration at the USGS National Water Quality Laboratory in Lakewood, Colorado.

## **Model Development**

Regression analysis was done using R by examining Turb, day of year, streamflow, and other continuously measured data as explanatory variables for estimating TP concentration. A variety of models that predict TP,  $(TP)^2$ ,  $\sqrt{TP}$  and models that predict  $\log_{10}(TP)$  were evaluated. The distribution of residuals was examined for normality, and plots of residuals (the difference between the

measured and computed values) as compared to computed TP were examined for homoscedasticity (meaning that their departures from zero did not change substantially over the range of computed values). This comparison lead to the conclusion that the most appropriate and reliable model would be one that estimated  $\log_{10}(TP)$ .

Turbidity and the day of the year were selected as the best predictors of TP based on residual plots, relatively high adjusted coefficient of determination (adjusted  $R^2$ ) and relatively low model standard percentage error ( $MSPE$ ), prediction error sum of squares (PRESS), and Mallow's  $C_p$ . Values for all of the afore mentioned statistics and metrics were computed for various models and are included below along with all relevant sample data and more in-depth statistical information.

## Model Summary

Summary of final regression analysis for TP concentration at site number 06887500.

TP concentration-based model:

$$\log_{10}(TP) = 0.27 \times \log_{10}(Turb) - 0.0763 \times \sin\left(\frac{2\pi DY}{365}\right) + 0.00703 \times \cos\left(\frac{2\pi DY}{365}\right) - 0.893$$

where

$TP$  = total phosphorous in milligrams per liter (mg/L) as P;

$Turb$  = turbidity in formazin nephelometric units (FNU);

$Sin \& Cos$  = seasonality component; and,

$DY$  = day of the year

Turbidity and the day of the year make physical and statistical sense as explanatory variables for TP.

The log-transformed model may be retransformed to the original units so that TP 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). For this model, the calculated BCF is 1.04. The retransformed model, accounting for BCF is:

$$TP = \frac{Turb^{0.27} \times 10^{0.00703 \times \cos\left(\frac{2\pi DY}{365}\right)} \times 0.133}{10^{0.0763 \sin\left(\frac{2\pi DY}{365}\right)}}$$

## Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(TP) = 0.368 \times \log_{10}(Turb) - 1.20$

## Total Phosphorous Concentration Record

The TP record is computed using this regression model and stored at the National Real-Time Water Quality (NRTWQ) Web site. Data are computed at 15-minute intervals. The complete water-quality record can be found at <http://nrtwq.usgs.gov/ks>.

## Remarks

None

# R Output for Total Phosphorous; 06887500; Kansas River at Wamego, KS

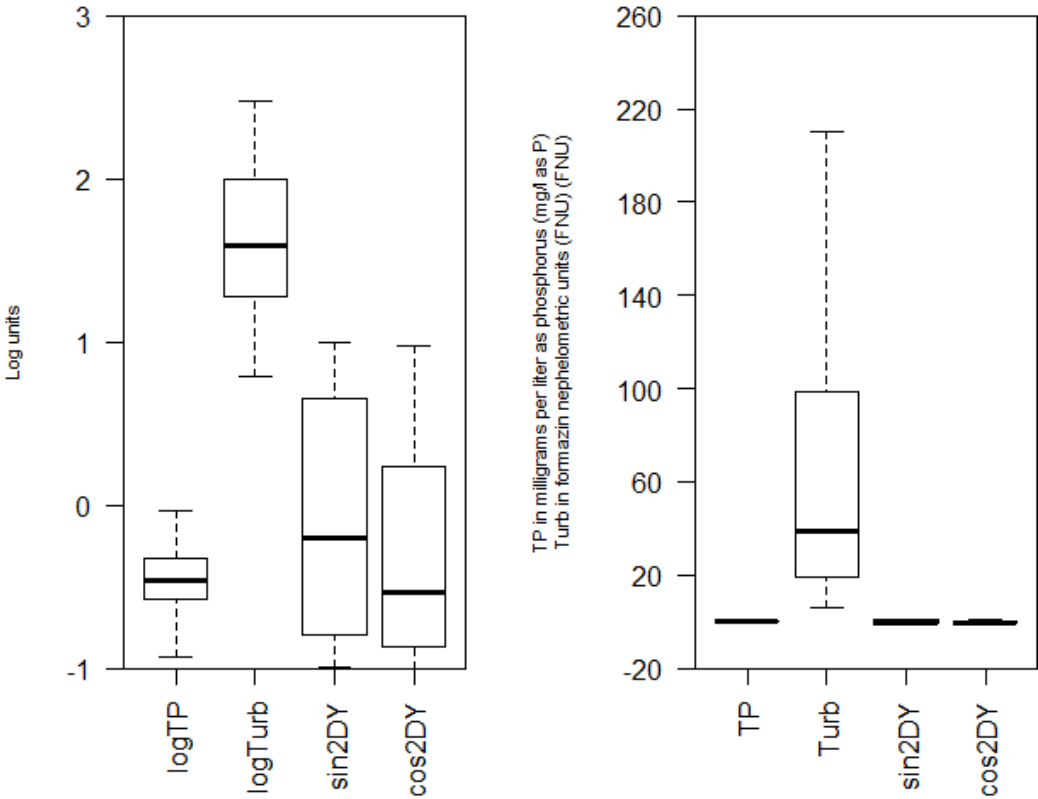
## Model Form

$$\log TP = +0.27 * \log Turb + -0.0763 * \sin 2DY + 0.00703 * \cos 2DY + -0.893$$

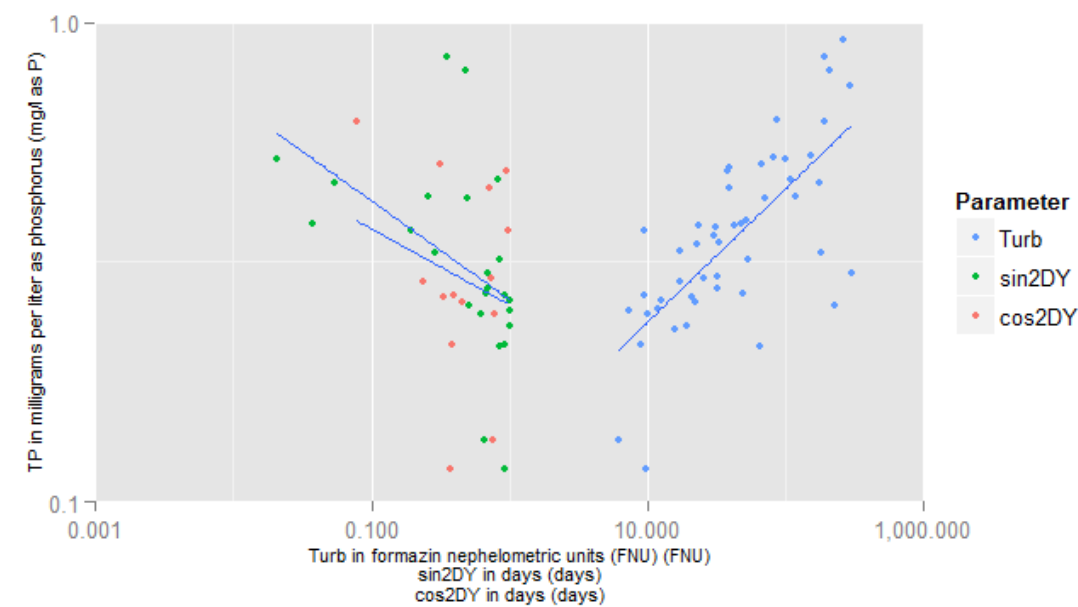
## Variable Summary Statistics

	logTP	logTurb	sin2DY	cos2DY	TP	Turb
Minimum	-0.9320	0.792	-0.9970	-1.000	0.117	6.2
1st Quartile	-0.5730	1.280	-0.7990	-0.868	0.267	19.0
Median	-0.4610	1.590	-0.2020	-0.539	0.346	39.0
Mean	-0.4490	1.630	-0.0741	-0.316	0.389	74.9
3rd Quartile	-0.3290	1.990	0.6530	0.235	0.469	98.5
Maximum	-0.0376	2.480	0.9980	0.981	0.917	299.0

## Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

Number of Observations	49
Standard error (RMSE)	0.121
Upper Model standard percentage error (MSPE)	32.1
Lower Model standard percentage error (MSPE)	24.3
Coefficient of determination (R <sup>2</sup> )	0.6
Adjusted Coefficient of Determination (Adj. R <sup>2</sup> )	0.574
Bias Correction Factor (BCF)	1.04
Variance Inflation Factors (VIF)	
logTurb	sin2DY
1.583840	1.036293
	1.565410

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t )
(Intercept)	-0.89300	0.0712	-12.50	2.72e-16
logTurb	0.27000	0.0462	5.84	5.36e-07
sin2DY	-0.07630	0.0245	-3.12	3.18e-03
cos2DY	0.00703	0.0352	0.20	8.43e-01

Correlation Matrix

	Intercept	logTurb	sin2DY	cos2DY
Intercept	1.000	-0.962	-0.144	-0.476
logTurb	-0.962	1.000	0.181	0.600
sin2DY	-0.144	0.181	1.000	0.146
cos2DY	-0.476	0.600	0.146	1.000

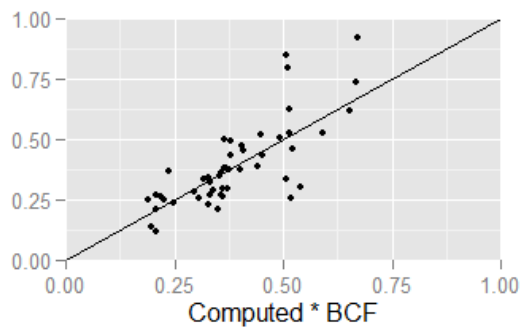
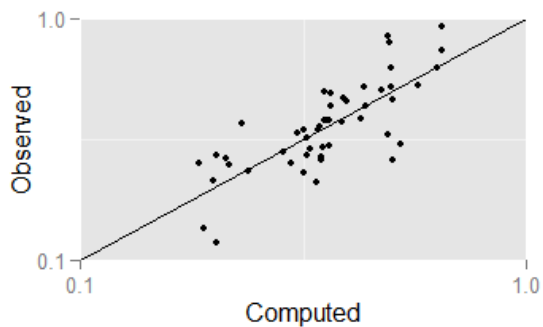
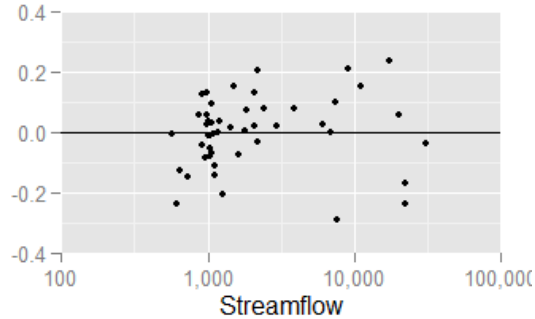
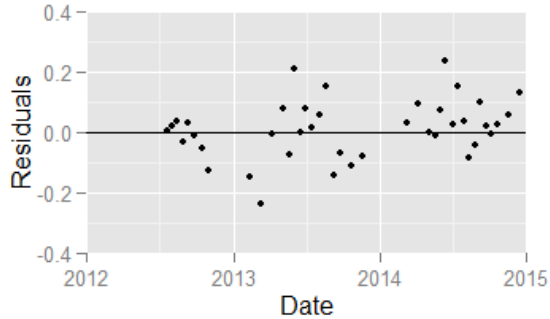
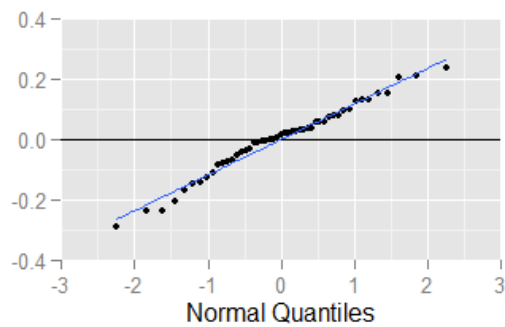
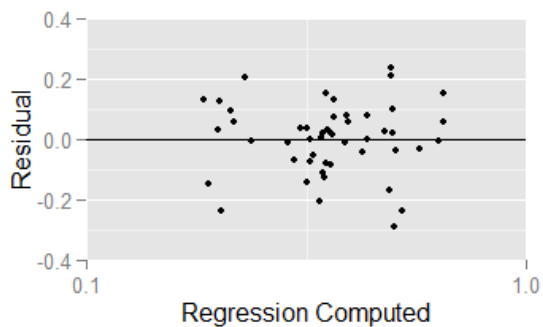
## Test Criteria

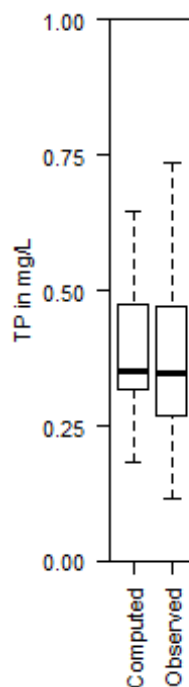
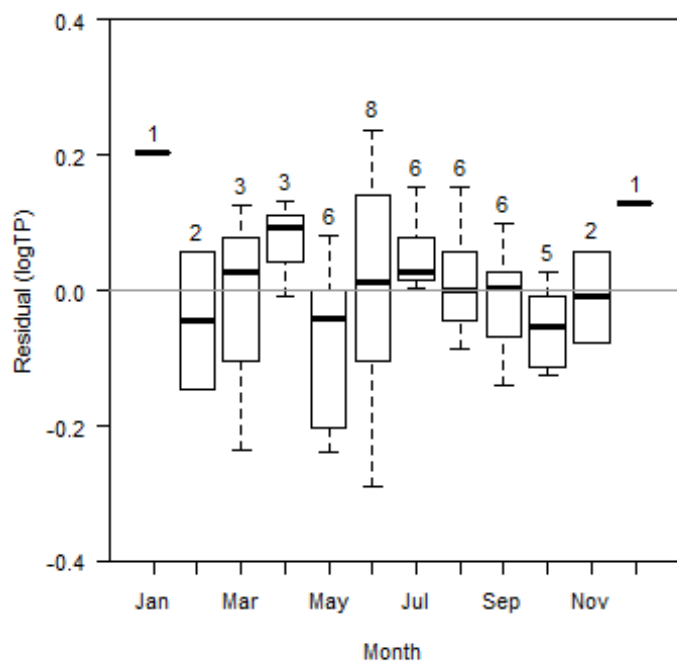
Leverage	Cook's D	DFFITS
0.1836735	0.2632863	0.4948717

## Flagged Observations

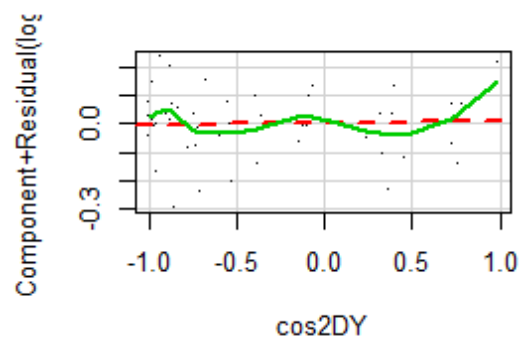
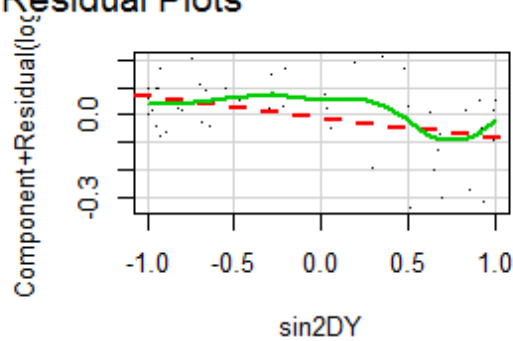
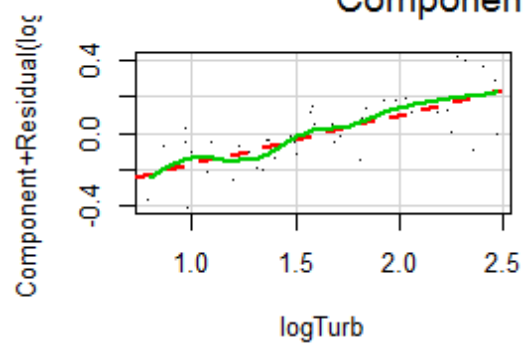
	logTP	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
3/11/2013 9:00	-0.93180	-0.6952	-0.2367	-2.057	-2.137	0.09691	0.11350	-0.7001
6/3/2013 8:50	-0.10130	-0.3085	0.2072	1.787	1.834	0.08300	0.07229	0.5517
6/11/2014 9:00	-0.07212	-0.3095	0.2374	2.036	2.113	0.07202	0.08043	0.5886
1/12/2015 8:40	-0.43420	-0.6383	0.2042	1.794	1.841	0.11610	0.10560	0.6670
5/18/2015 10:30	-0.52290	-0.2827	-0.2402	-2.128	-2.219	0.13050	0.16990	-0.8595
6/1/2015 8:00	-0.59010	-0.3000	-0.2901	-2.513	-2.680	0.09084	0.15780	-0.8473

## Statistical Plots





### Component + Residual Plots



EXPLANATION

- Non-parametric smoothed fit
- Linear best fit

## Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logTP ~ logTurb	1	0.1311	51.01	49.97	19.59	0.9063	1 ± 31	
logTP ~ Turb	1	0.1514	34.71	33.32	41.09	1.239	1 ± 36	
logTP ~ logQ	1	0.1544	32.06	30.61	44.59	1.259	1 ± 36	
logTP ~ logTurb + sin2DY	2	0.1198	60	58.26	9.745	0.775	1.014 ± 28	
logTP ~ logTurb + sin4DY	2	0.1274	54.77	52.8	16.64	0.8648	1.107 ± 30	
logTP ~ logTurb + logChl	2	0.1302	52.73	50.68	19.33	0.9034	1.157 ± 30	
logTP ~ logTurb + logChl + sin2DY	3	0.1171	62.57	60.08	8.352	0.7498	1.182 ± 27	
logTP ~ logTurb + sin2DY + cos4DY	3	0.1185	61.71	59.15	9.492	0.7704	1.16 ± 28	
logTP ~ logTurb + sin2DY + sin4DY	3	0.1197	60.95	58.34	10.5	0.7789	1.168 ± 28	
logTP ~ Q + logQ + logTurb + sin2DY	4	0.1131	65.86	62.76	6.013	0.7347	6.058 ± 26	
logTP ~ Turb + logTurb + logChl + sin2DY	4	0.1151	64.67	61.46	7.585	0.7483	7.464 ± 27	
logTP ~ logTurb + Chl + logChl + sin2DY	4	0.1155	64.4	61.16	7.941	0.736	1.261 ± 27	
logTP ~ Turb + logTurb + Chl + logChl + sin2DY	5	0.1124	67.09	63.27	6.391	0.7211	7.596 ± 26	
logTP ~ Q + logQ + logTurb + logChl + sin2DY	5	0.1133	66.55	62.66	7.108	0.7455	6.349 ± 26	
logTP ~ Turb + logTurb + logChl + sin2DY + sin4DY	5	0.1137	66.29	62.37	7.455	0.7341	8.322 ± 26	
logTP ~ Turb + logTurb + Chl + logChl + sin2DY + sin4DY	6	0.1114	68.39	63.87	6.684	0.7159	8.402 ± 26	
logTP ~ Q + logQ + logTurb + Chl + logChl + sin2DY	6	0.1115	68.38	63.86	6.698	0.7344	6.349 ± 26	
logTP ~ Q + Turb + logTurb + Chl + logChl + sin2DY	6	0.1133	67.34	62.68	8.06	0.7535	2.619 ± 26	

## Data

	Date	logTP	logTurb	sin2DY	cos2DY	TP	Turb	Computed logTP	Computed TP	Residual	Normal Quantiles
0											
1	2012-07-19	-0.4609	1.524	-0.2924	-0.9563	0.346	33.4	-0.4656	0.3543	0.00467	-0.051
2	2012-07-30	-0.4461	1.479	-0.4681	-0.8837	0.358	30.1	-0.4639	0.3557	0.0178	0.051
3	2012-08-13	-0.4763	1.233	-0.6639	-0.7478	0.334	17.1	-0.5143	0.3167	0.038	0.42
4	2012-08-27	-0.2782	2.188	-0.8242	-0.5663	0.527	154	-0.243	0.5915	-0.0352	-0.42
5	2012-09-10	-0.4237	1.375	-0.9347	-0.3554	0.377	23.71	-0.4525	0.3651	0.0289	0.311
6	2012-09-24	-0.4271	1.496	-0.9924	-0.1227	0.374	31.33	-0.4138	0.3992	-0.0133	-0.365
7	2012-10-15	-0.5391	1.23	-0.9719	0.2354	0.289	17	-0.4846	0.3392	-0.0545	-0.594
8	2012-10-29	-0.585	1.342	-0.8882	0.4595	0.26	22	-0.4592	0.3596	-0.126	-1.02
9	2013-02-11	-0.8697	0.7924	0.6532	0.7572	0.135	6.2	-0.7232	0.1958	-0.146	-1.21
10	2013-03-11	-0.9318	0.9845	0.9297	0.3684	0.117	9.65	-0.6952	0.2089	-0.237	-1.62
11	2013-04-08	-0.6326	1.279	0.9946	-0.1036	0.233	19	-0.624	0.246	-0.00867	-0.258
12	2013-05-06	-0.3288	2.041	0.8332	-0.553	0.469	110	-0.4089	0.4038	0.08	0.788
13	2013-05-20	-0.567	1.69	0.6766	-0.7364	0.271	49	-0.493	0.3326	-0.074	-0.72
14	2013-06-03	-0.1013	2.322	0.4814	-0.8765	0.792	210	-0.3085	0.5088	0.207	1.84
15	2013-06-17	-0.3625	2.067	0.2585	-0.966	0.434	116.7	-0.361	0.4508	-0.0015	-0.153
16	2013-07-01	-0.2848	1.993	0.02068	-0.9998	0.519	98.5	-0.363	0.4488	0.0781	0.72
17	2013-07-15	-0.4248	1.633	-0.2184	-0.9759	0.376	43	-0.4417	0.3743	0.0169	0
18	2013-08-05	-0.1325	2.468	-0.5491	-0.8357	0.737	294	-0.1901	0.6682	0.0575	0.476
19	2013-08-19	-0.03763	2.415	-0.7337	-0.6795	0.917	260	-0.1893	0.6694	0.152	1.32
20	2013-09-09	-0.6402	1.204	-0.9262	-0.377	0.229	16	-0.4995	0.3277	-0.141	-1.11
21	2013-09-23	-0.5969	1.076	-0.9897	-0.1435	0.253	11.9	-0.5277	0.3071	-0.0692	-0.655
22	2013-10-21	-0.5735	1.322	-0.9434	0.3316	0.267	21	-0.4613	0.3579	-0.112	-0.935
23	2013-11-18	-0.5346	1.404	-0.6821	0.7313	0.292	25.33	-0.4564	0.3619	-0.0782	-0.788
24	2014-03-10	-0.6737	0.9576	0.923	0.3848	0.212	9.07	-0.7018	0.2057	0.0281	0.258
25	2014-04-07	-0.58	1.097	0.9961	-0.08811	0.263	12.5	-0.6731	0.2197	0.093	0.859
26	2014-05-05	-0.4935	1.732	0.8422	-0.5392	0.321	53.99	-0.4929	0.3327	-0.000573	-0.102
27	2014-05-19	-0.5544	1.505	0.6897	-0.7241	0.279	32	-0.5439	0.2958	-0.0105	-0.311
28	2014-06-02	-0.3655	1.845	0.4966	-0.868	0.431	70	-0.4384	0.3772	0.0729	0.655
29	2014-06-11	-0.07212	2.284	0.3564	-0.9343	0.847	192.5	-0.3095	0.5075	0.237	2.25
30	2014-06-30	-0.4214	1.679	0.03765	-0.9993	0.379	47.79	-0.4491	0.3681	0.0277	0.205
31	2014-07-14	-0.3019	1.591	-0.2018	-0.9794	0.499	39	-0.4545	0.3635	0.153	1.45
32	2014-07-28	-0.4634	1.362	-0.4308	-0.9024	0.344	23	-0.4984	0.3285	0.035	0.365
33	2014-08-11	-0.5287	1.505	-0.6334	-0.7738	0.296	32	-0.4433	0.373	-0.0854	-0.859
34	2014-08-25	-0.4145	1.72	-0.7995	-0.6007	0.385	52.5	-0.3714	0.4402	-0.0432	-0.534
35	2014-09-08	-0.2041	1.932	-0.9196	-0.3929	0.625	85.42	-0.3037	0.5144	0.0995	0.935

36	2014-09-22	-0.2815	1.909	-0.9869	-0.1612	0.523	81.17	-0.3029	0.5154	0.0214	0.102
37	2014-10-06	-0.2083	2.279	-0.9969	0.07882	0.619	190	-0.2007	0.6521	-0.00763	-0.205
38	2014-10-20	-0.2967	1.831	-0.9493	0.3145	0.505	67.8	-0.3235	0.4915	0.0268	0.153
39	2014-11-17	-0.3449	1.597	-0.693	0.7209	0.452	39.5	-0.4036	0.4087	0.0587	0.594
40	2014-12-15	-0.3089	1.58	-0.2783	0.9605	0.491	38	-0.4381	0.3775	0.129	1.11
41	2015-01-12	-0.4342	0.9713	0.1943	0.9809	0.368	9.36	-0.6383	0.2381	0.204	1.62
42	2015-02-09	-0.6073	1	0.6278	0.7783	0.247	10	-0.6651	0.2238	0.0578	0.534
43	2015-03-09	-0.5702	0.9759	0.9165	0.4	0.269	9.46	-0.6963	0.2083	0.126	1.02
44	2015-04-06	-0.6021	0.8704	0.9975	-0.07011	0.25	7.42	-0.7342	0.1909	0.132	1.21
45	2015-05-04	-0.6757	1.813	0.8518	-0.5239	0.211	65	-0.4718	0.3493	-0.204	-1.45
46	2015-05-18	-0.5229	2.475	0.7007	-0.7134	0.3	298.8	-0.2827	0.5399	-0.24	-1.84
47	2015-06-01	-0.5901	2.362	0.5118	-0.8591	0.257	230	-0.3	0.5188	-0.29	-2.25
48	2015-06-15	-0.4789	2.26	0.292	-0.9564	0.332	181.8	-0.3115	0.5052	-0.167	-1.32
49	2015-06-29	-0.3363	2.243	0.05496	-0.9985	0.461	175	-0.2982	0.521	-0.0381	-0.476

Definitions and National Water Information System (parameter code)

TP: Phosphorus in mg/L as P (00665)

Turb: Turbidity in FNU (63680)

DY: Date in decimal years