Appendix 22. Model Archival Summary for Total Nitrogen Concentration at Station 06887500; Kansas River at Wamego, Kansas

This model archival summary summarizes the total nitrogen (particulate plus dissolved) concentration (TN) model developed to compute 15-minute TN 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 was installed 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 36 concurrent measurements of TN concentration, turbidity (Turb), and chlorophyll (Chl) 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 limits. Three samples were estimated due to particulate N falling below the MDL of 0.03. Therefore, 0.015+Diss N was used to calculate TN in these cases. 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 Nitrogen 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 particulate and dissolved nitrogen concentrations at the USGS National Water Quality Laboratory in Lakewood, Colorado. Total nitrogen was calculated as the sum of particulate and dissolved nitrogen.

Model Development

Regression analysis was done using R by examining Turb, Chl, streamflow, and other continuously measured data as explanatory variables for estimating TN concentration. A variety of models that predict TN, $(TN)^2$, $\sqrt{TN}$ and models that predict $\log_{10}(TN)$ 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 TN 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 $\log_{10}$(TN).

Turb and Chl were selected as the best predictors of TN 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 TN concentration at site number 06887500.

TN concentration-based model:

$$\log_{10}(TN) = 0.314 \times \log_{10}(Turb) - 0.149 \times \log_{10}(Chl) - 0.131$$

where

TN = total nitrogen (particulate plus dissolved) in milligrams per liter (mg/L);

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

Chl = chlorophyll in micrograms per liter (µg/L)

Turbidity and Chl make physical and statistical sense as explanatory variables for TN.

The log-transformed model may be retransformed to the original units so that TN 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 1.02. The retransformed model, accounting for BCF is:

$$TN = Turb^{0314} + Chl^{-0.149} + 0.754$$

Previous Models

Start year	End year	Model
2000	2005	logTN = 0.237logTurb – 0.179

Total Nitrogen Concentration Record

The TN 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 Nitrogen; 06887500; Kansas River at Wamego, KS

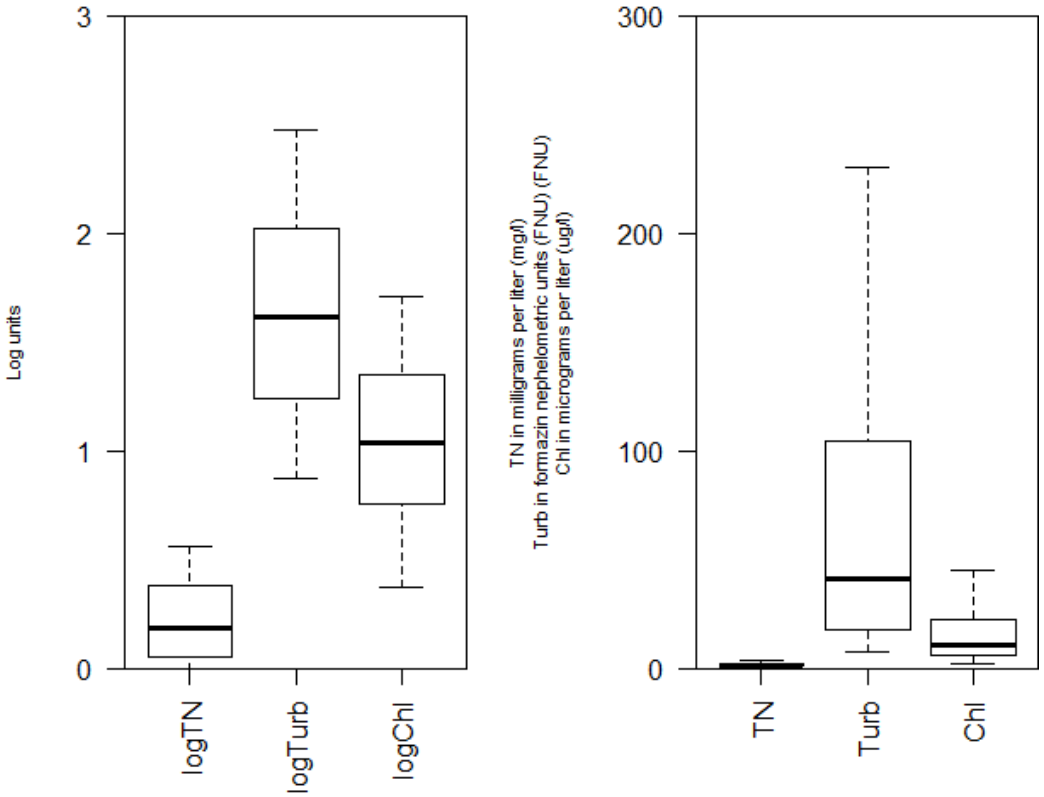
Model Form

logTN = + 0.314 * logTurb + -0.149 * logChl + -0.131

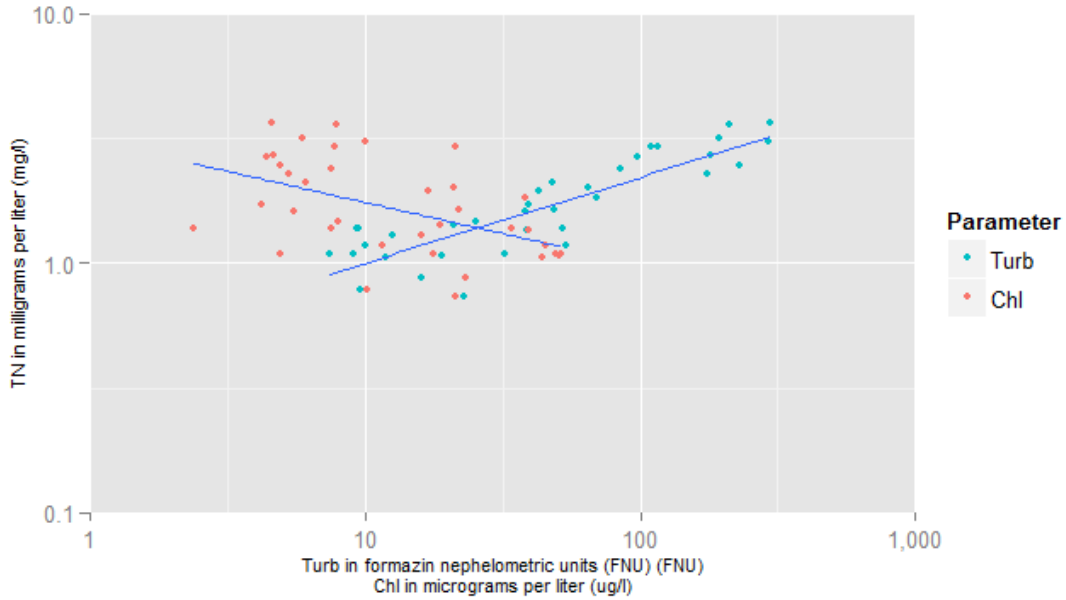
Variable Summary Statistics

	logTN	logTurb	logChl	TN	Turb	Chl
Minimum	-0.134	0.87	0.375	0.734	7.42	2.37
1st Quartile	0.053	1.24	0.756	1.130	17.50	5.70
Median	0.186	1.62	1.040	1.540	41.20	10.90
Mean	0.218	1.63	1.100	1.820	76.10	18.20
3rd Quartile	0.383	2.02	1.350	2.420	104.00	22.50
Maximum	0.559	2.48	1.710	3.620	299.00	51.20

Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

Number of Observations	36
Standard error (RMSE)	0.0817
Upper Model standard percentage error (MSPE)	20.7
Lower Model standard percentage error (MSPE)	17.1
Coefficient of determination (R^2)	0.829
Adjusted Coefficient of Determination (Adj. R^2)	0.819
Bias Correction Factor (BCF)	1.02

```
Variance Inflation Factors (VIF)
 logTurb   logChl
1.067732 1.067732
```

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.131	0.0721	-1.81	7.95e-02
logTurb	0.314	0.0295	10.60	3.47e-12
logChl	-0.149	0.0373	-3.99	3.42e-04

Correlation Matrix

	Intercept	logTurb	logChl
Intercept	1.000	-0.812	-0.738
logTurb	-0.812	1.000	0.252
logChl	-0.738	0.252	1.000

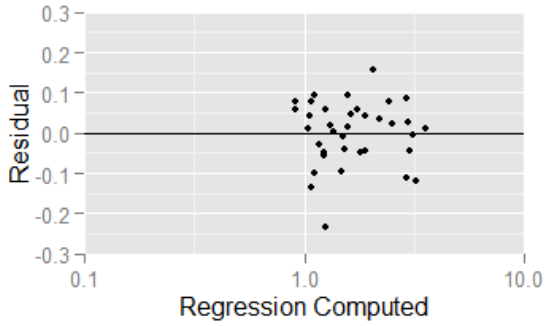
Test Criteria

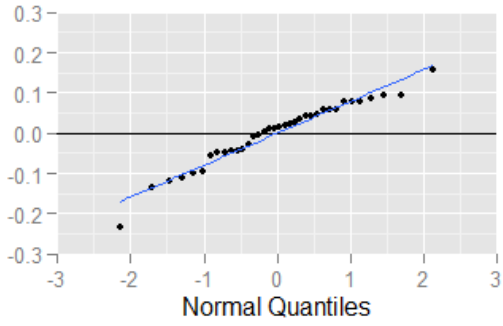
```
 Leverage  Cook's D    DFFITS
0.1666667 0.1936416 0.4714045
```

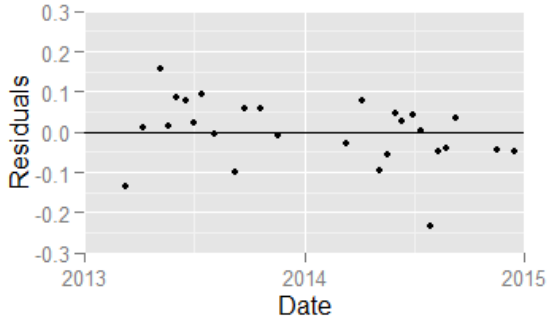
Flagged Observations

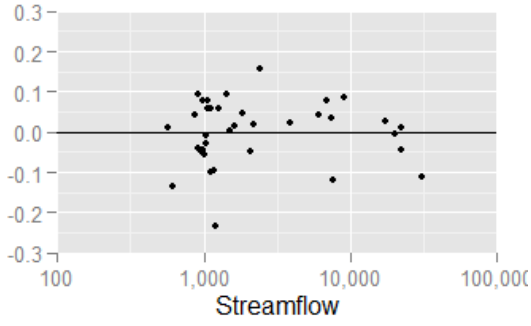
	logTN	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
3/11/2013 9:00	-0.1079	0.02792	-0.13580	-1.7430	-1.8010	0.08979	0.099860	-0.5656
5/6/2013 8:00	0.4669	0.31150	0.15540	1.9700	2.0650	0.06809	0.094510	0.5582
7/28/2014 10:00	-0.1343	0.09857	-0.23290	-2.9140	-3.3300	0.04309	0.127500	-0.7066
1/12/2015 8:40	0.1383	0.11830	0.02001	0.2802	0.2762	0.23560	0.008066	0.1534
6/1/2015 8:00	0.3881	0.50680	-0.11870	-1.5380	-1.5720	0.10680	0.094290	-0.5436

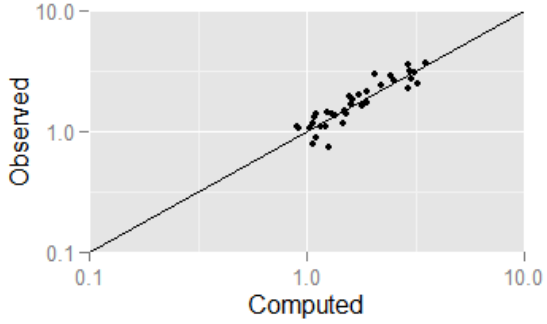
Statistical Plots

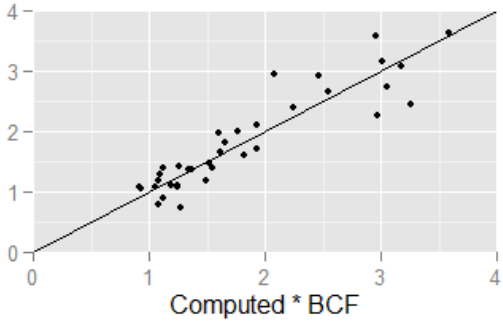


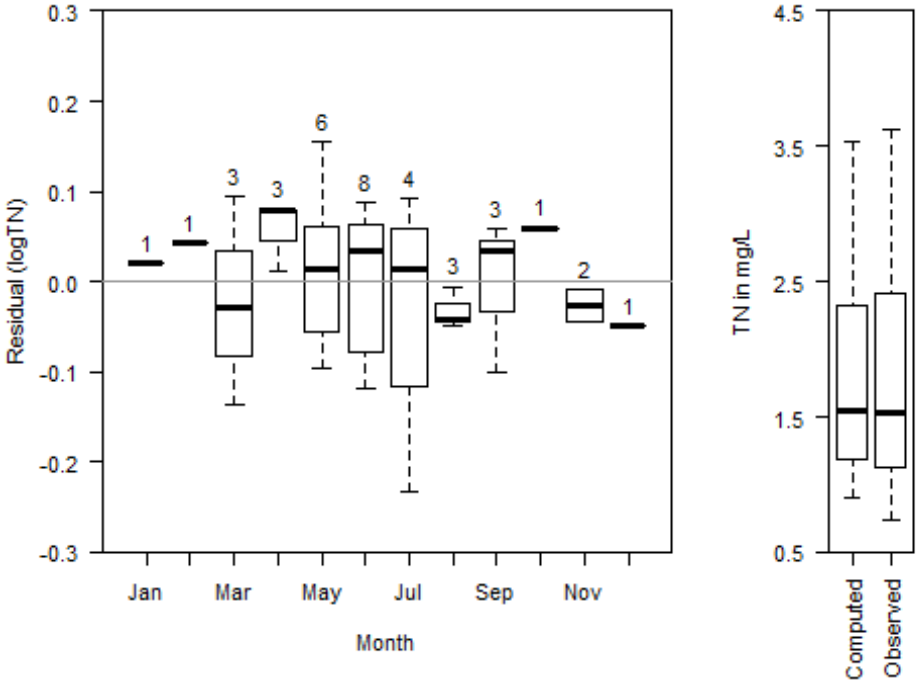




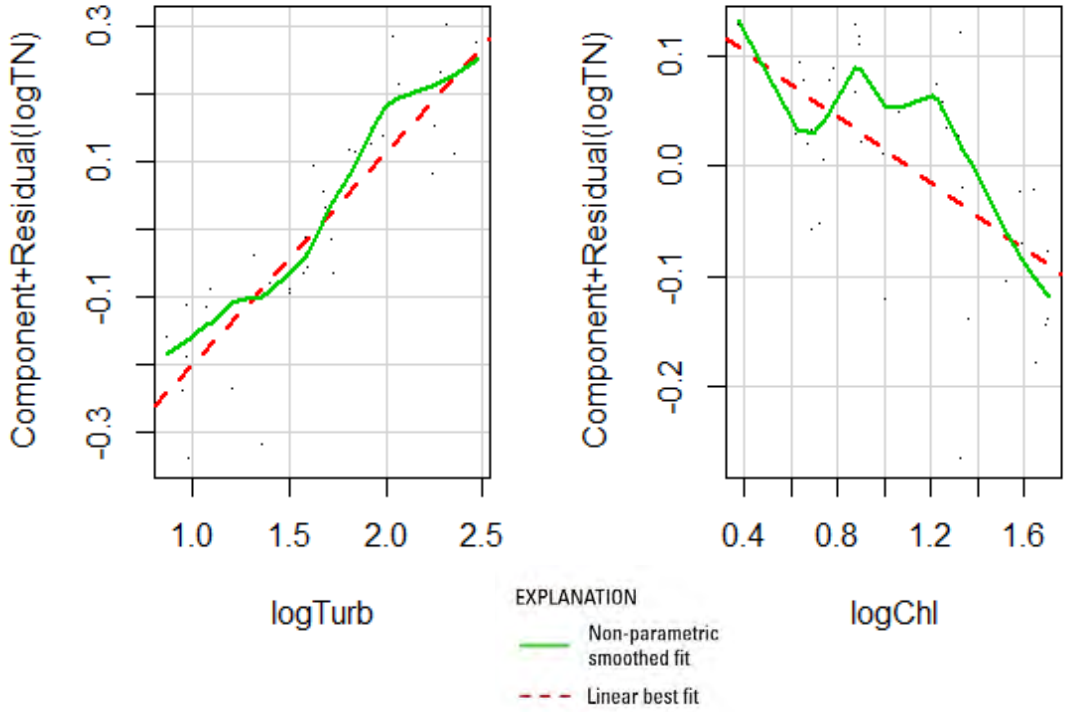








Component + Residual Plots



Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logTN ~ logTurb	1	0.09804	74.69	73.94	21.77	0.363	1	± 23
logTN ~ logQ	1	0.1091	68.63	67.71	34.63	0.4584	1	± 25
logTN ~ Turb	1	0.1113	67.37	66.41	37.31	0.4758	1	± 26
logTN ~ logTurb + Chl	2	0.0798	83.72	82.74	4.574	0.2437	1.055	± 18
logTN ~ logTurb + logChl	2	0.0817	82.94	81.9	6.241	0.2557	1.068	± 19
logTN ~ Temp + logTurb	2	0.08869	79.89	78.67	12.71	0.309	1.361	± 21
logTN ~ logTurb + Chl + sin2DY	3	0.07927	84.42	82.96	5.088	0.251	1.06	± 18
logTN ~ Temp + logTurb + Chl	3	0.07939	84.38	82.91	5.181	0.2628	1.894	± 18
logTN ~ Q + logTurb + Chl	3	0.07996	84.15	82.67	5.66	0.2529	2.365	± 19
logTN ~ Q + logQ + Temp + logTurb	4	0.07325	87.12	85.45	1.367	0.2131	6.557	± 17
logTN ~ Q + logQ + logTurb + Chl	4	0.07683	85.82	84	4.109	0.2269	7.109	± 18
logTN ~ Q + logQ + logTurb + cos2DY	4	0.07759	85.54	83.68	4.706	0.2419	6.559	± 18
logTN ~ Q + logQ + Temp + logTurb + Chl	5	0.07338	87.49	85.4	2.578	0.2273	7.499	± 17
logTN ~ Q + logQ + Temp + logTemp + logTurb	5	0.07351	87.44	85.35	2.674	0.3558	6.94	± 17
logTN ~ Q + logQ + Temp + logTurb + sin2DY	5	0.0741	87.24	85.11	3.101	0.2189	6.588	± 17
logTN ~ Q + logQ + Temp + logTurb + Chl + logChl	6	0.07335	87.91	85.41	3.675	0.2295	7.711	± 17
logTN ~ Q + logQ + Temp + logTurb + Chl + sin2DY	6	0.07391	87.73	85.19	4.068	0.233	7.5	± 17
logTN ~ Q + logQ + Temp + logTemp + logTurb + Chl	6	0.07399	87.7	85.16	4.119	0.579	8.164	± 17

Data

	Date	logTN	logTurb	logChl	TN	Turb	Chl	Computed logTN	Computed TN	Residual	Normal Quantiles	Censored Values
1	2013-03-11	-0.1079	0.9845	1.009	0.78	9.65	10.2	0.02792	1.083	-0.136	-1.7	--
2	2013-04-08	0.02816	1.279	1.706	1.067	19	50.77	0.0163	1.055	0.0119	-0.0346	--
3	2013-05-06	0.4669	2.041	1.33	2.93	110	21.4	0.3115	2.081	0.155	2.13	--
4	2013-05-20	0.2154	1.69	1.338	1.642	49	21.8	0.2001	1.61	0.0153	0.0346	--
5	2013-06-03	0.5516	2.322	0.8943	3.561	210	7.84	0.4646	2.961	0.0869	1.28	--
6	2013-06-17	0.463	2.067	0.8921	2.904	116.7	7.8	0.3849	2.464	0.0781	0.907	--
7	2013-07-01	0.4234	1.993	0.6435	2.651	98.5	4.4	0.3989	2.545	0.0245	0.174	--
8	2013-07-15	0.2911	1.633	1.229	1.955	43	16.95	0.1986	1.605	0.0925	1.46	--
9	2013-08-05	0.4891	2.468	0.9983	3.084	294	9.96	0.495	3.175	-0.00587	-0.245	--
10	2013-09-09	-0.05651	1.204	1.365	0.878	16	23.2	0.04359	1.123	-0.1	-1.14	--
11	2013-09-23	0.02036	1.076	1.645	1.048	11.9	44.2	-0.03849	0.9295	0.0588	0.627	--
12	2013-10-21	0.1538	1.322	1.274	1.425	21	18.8	0.09427	1.262	0.0595	0.714	--
13	2013-11-18	0.1661	1.404	0.9031	1.466	25.33	8	0.1751	1.52	-0.00901	-0.317	--
14	2014-03-10	0.03862	0.9576	0.6931	1.093	9.07	4.933	0.06652	1.184	-0.0279	-0.391	--
15	2014-04-07	0.1119	1.097	1.207	1.294	12.5	16.11	0.0336	1.097	0.0783	1.02	--
16	2014-05-05	0.0693	1.732	1.654	1.173	53.99	45.09	0.1663	1.489	-0.097	-1.02	--
17	2014-05-19	0.03302	1.505	1.696	1.079	32	49.7	0.0887	1.246	-0.0557	-0.907	--
18	2014-06-02	0.2598	1.845	1.584	1.819	70	38.4	0.2121	1.655	0.0478	0.545	--
19	2014-06-11	0.4989	2.284	0.7709	3.154	192.5	5.9	0.4712	3.006	0.0277	0.245	--
20	2014-06-30	0.3197	1.679	0.7874	2.088	47.79	6.129	0.2789	1.93	0.0408	0.391	--
21	2014-07-14	0.1329	1.591	1.592	1.358	39	39.08	0.1312	1.374	0.00168	-0.174	--
22	2014-07-28	-0.1343	1.362	1.328	0.734	23	21.3	0.09857	1.274	-0.233	-2.13	--
23	2014-08-11	0.03703	1.505	1.709	1.089	32	51.22	0.08674	1.24	-0.0497	-0.807	--
24	2014-08-25	0.1396	1.72	1.53	1.379	52.5	33.88	0.181	1.541	-0.0414	-0.466	--
25	2014-09-08	0.3777	1.932	0.8778	2.386	85.42	7.548	0.3445	2.245	0.0331	0.317	--
26	2014-11-17	0.2338	1.597	0.6232	1.713	39.5	4.2	0.2774	1.924	-0.0437	-0.545	--
27	2014-12-15	0.2057	1.58	0.7404	1.606	38	5.5	0.2547	1.826	-0.0489	-0.714	--
28	2015-01-12	0.1383	0.9713	0.3747	1.375	9.36	2.37	0.1183	1.334	0.02	0.104	--
29	2015-02-09	0.06744	1	1.063	1.168	10	11.55	0.02473	1.075	0.0427	0.466	--
30	2015-03-09	0.1399	0.9759	0.8762	1.38	9.46	7.52	0.04496	1.126	0.0949	1.7	--
31	2015-04-06	0.03543	0.8704	1.25	1.085	7.42	17.8	-0.04394	0.9179	0.0794	1.14	--
32	2015-05-04	0.301	1.813	1.321	2	65	20.96	0.2412	1.77	0.0599	0.807	--
33	2015-05-18	0.5591	2.475	0.6604	3.623	298.8	4.575	0.5476	3.583	0.0115	-0.104	--

34	2015-06-01	0.3881	2.362	0.6944	2.444	230	4.948	0.5068	3.263	-0.119	-1.46	--
35	2015-06-15	0.4342	2.26	0.6698	2.718	181.8	4.675	0.4785	3.056	-0.0442	-0.627	--
36	2015-06-29	0.3555	2.243	0.7229	2.267	175	5.283	0.4654	2.966	-0.11	-1.28	--

Definitions and National Water Information System (parameter code)

TN: Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3) in mg/L (00600)
Turb: Turbidity in FNU (63680)
Chl: in ug/L (32318)