Appendix 23. Model Archival Summary for Total Nitrogen Concentration at Station 06892350; Kansas River at De Soto, 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: 06892350
Site name: Kansas River at De Soto, Kansas
Location: Lat 38°59'00", long 94°57'52" referenced to North American Datum of 1927, in NE 1/4 SE 1/4 SE 1/4 sec.28, T.12 S., R.22 E., Leavenworth County, KS, Hydrologic Unit 10270104.

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 40 concurrent measurements of TN concentration, turbidity (Turb), and temperature (Temp) 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. 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. One sample, June 11, 2014, was found to have potential errors in collection and processing, and has been removed from the dataset. 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, Temp, 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 the 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 06892350.

TN concentration-based model:

$$TN = -0.0362(Temp) + 1.93 \times \log_{10}(Turb) - 0.735$$

where

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

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

Temp = water temperature in degrees Celsius.

Turb and Temp make physical and statistical sense as explanatory variables for TN.

Previous Models

Start year	End year	Model
2000	2005	logTN =0.239logTurb - 0.263

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; 06892350; Kansas River at De Soto, KS

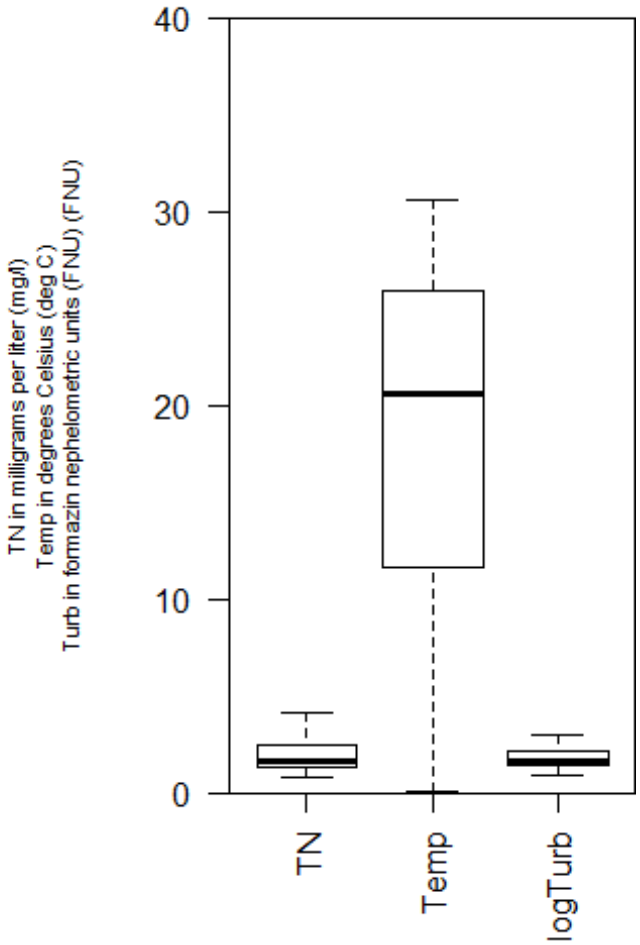
Model Form

TN = + -0.0362 * Temp + 1.93 * logTurb + -0.735

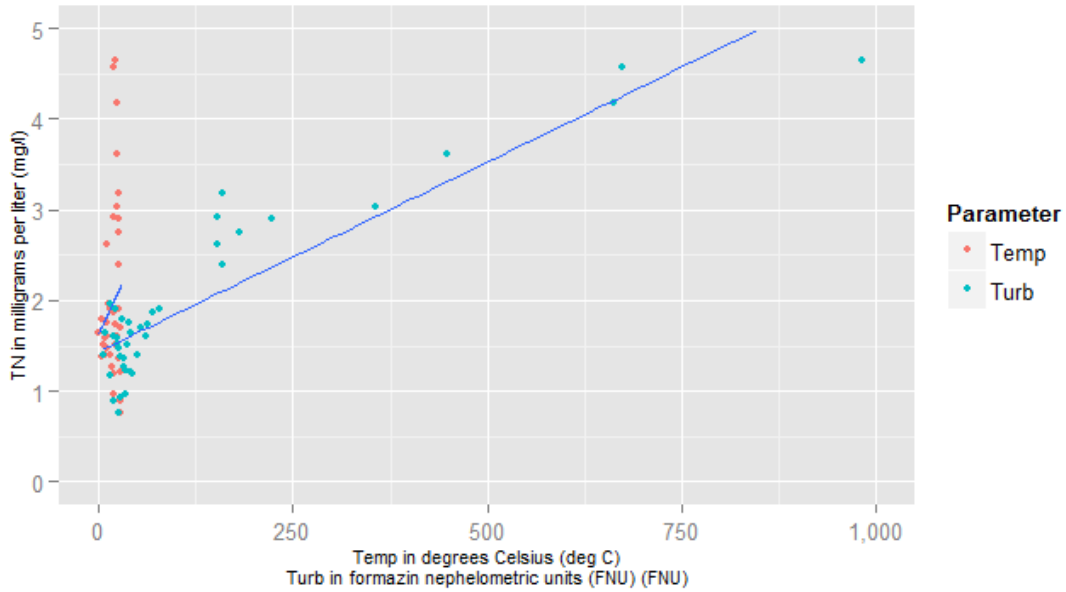
Variable Summary Statistics

	TN	Temp	logTurb	Turb
Minimum	0.758	0.04	0.869	7.4
1st Quartile	1.370	11.60	1.420	26.1
Median	1.640	20.60	1.610	40.4
Mean	1.980	18.70	1.760	129.0
3rd Quartile	2.500	25.90	2.190	153.0
Maximum	4.650	30.60	2.990	982.0

Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

```
Number of Observations                            40
Standard error (RMSE)                          0.385
Upper Model standard percentage error (MSPE)    19.5
Lower Model standard percentage error (MSPE)    19.5
Coefficient of determination (R²)              0.854
Adjusted Coefficient of Determination (Adj. R²) 0.846
```

```
Variance Inflation Factors (VIF)
   Temp logTurb
1.24847 1.24847
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)      -0.7350        0.22500   -3.27 2.31e-03
Temp             -0.0362        0.00844   -4.29 1.25e-04
logTurb           1.9300        0.13300   14.50 7.91e-17
```

Correlation Matrix

```
          Intercept   Temp logTurb
Intercept     1.000 -0.239  -0.728
Temp         -0.239  1.000  -0.446
logTurb      -0.728 -0.446   1.000
```

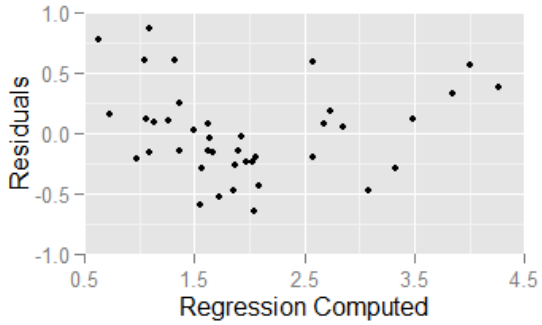
Test Criteria

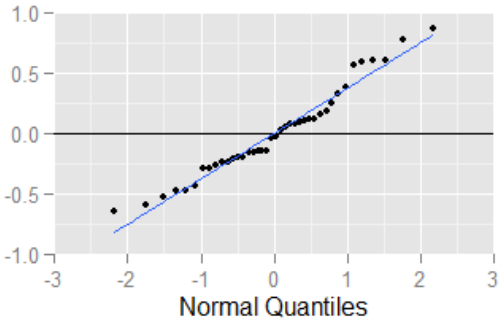
Leverage	Cook's D	DFFITS
0.1500000	0.1937589	0.4472136

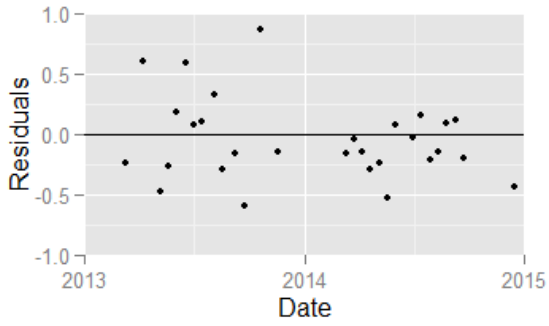
Flagged Observations

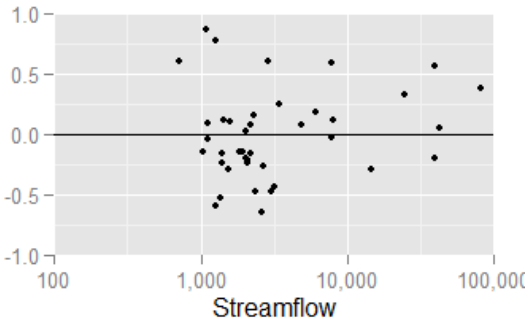
	TN	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
10/21/2013 10:30	1.953	1.0830	0.8701	2.326	2.483	0.05474	0.10440	0.5974
1/12/2015 11:30	1.649	1.0430	0.6058	1.729	1.779	0.17090	0.20530	0.8075
3/9/2015 12:20	1.404	0.6341	0.7699	2.119	2.229	0.10830	0.18170	0.7769
5/18/2015 15:30	4.567	3.9990	0.5684	1.606	1.642	0.15410	0.15670	0.7012
6/6/2015 19:50	4.647	4.2660	0.3811	1.100	1.103	0.18920	0.09412	0.5329

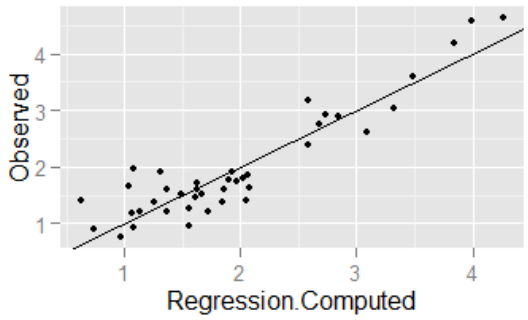
Statistical Plots

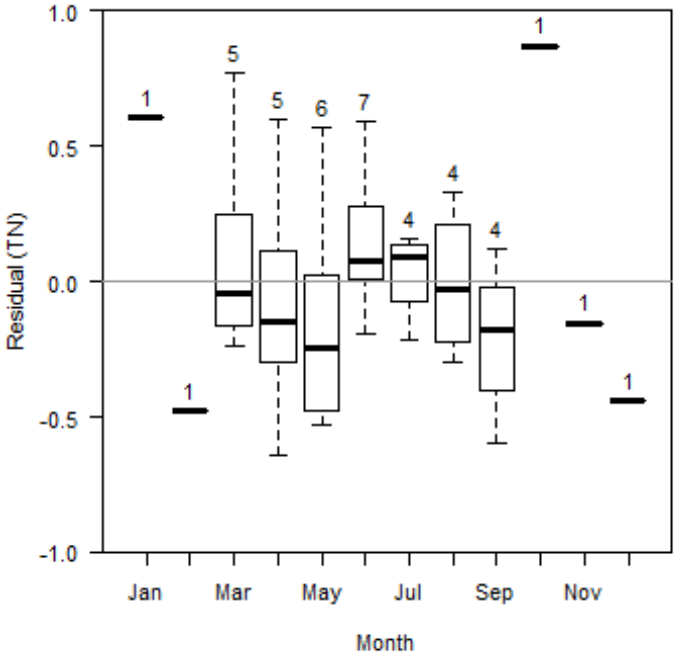


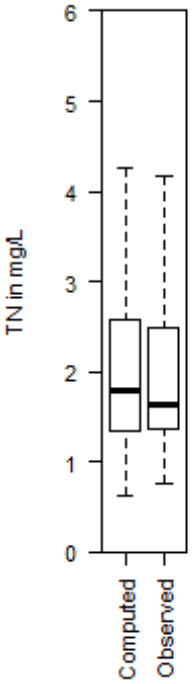




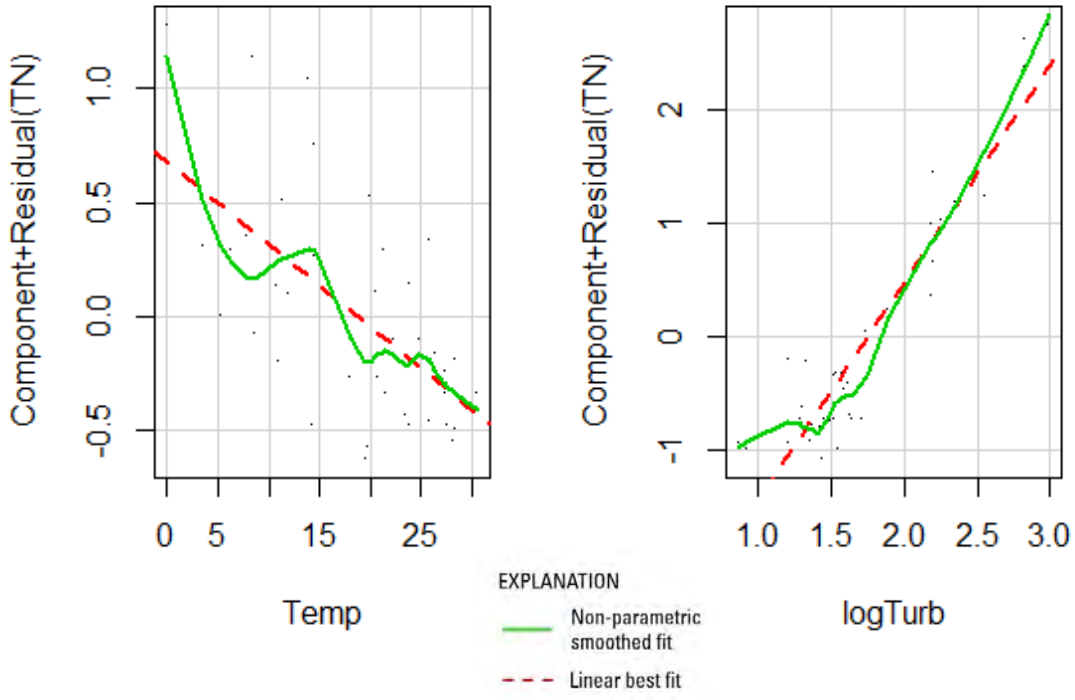








Component + Residual Plots



Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
TN ~ Turb	1	0.4268	81.51	81.03	21.91	9.13	1	± 22
TN ~ logTurb	1	0.4645	78.1	77.52	32.6	9.482	1	± 24
TN ~ logSC	1	0.5163	72.94	72.23	48.75	11.45	1	± 26
TN ~ Turb + logTurb	2	0.376	86.03	85.27	9.759	6.898	3.787	± 19
TN ~ Temp + logTurb	2	0.3848	85.37	84.57	11.84	6.515	1.248	± 19
TN ~ logTemp + logTurb	2	0.3921	84.8	83.98	13.59	7.977	1.21	± 20
TN ~ Temp + Turb + logTurb	3	0.3276	89.68	88.82	0.3261	5.214	1.384	± 17
TN ~ logTemp + Turb + logTurb	3	0.3371	89.07	88.16	2.233	6.398	1.364	± 17
TN ~ Turb + logTurb + cos2DY	3	0.3519	88.09	87.1	5.296	5.978	4.224	± 18
TN ~ Temp + logSC + Turb + logTurb	4	0.3193	90.47	89.38	-0.146	4.956	1.433	± 16
TN ~ Temp + SC + Turb + logTurb	4	0.3213	90.35	89.25	0.2305	5.107	1.466	± 16
TN ~ Temp + logTemp + Turb + logTurb	4	0.3258	90.08	88.94	1.078	7.69	2.49	± 16
TN ~ Temp + logSC + Turb + logTurb + sin4DY	5	0.3202	90.69	89.32	1.16	5.369	1.461	± 16
TN ~ Q + logQ + Temp + Turb + logTurb	5	0.3205	90.67	89.3	1.213	4.959	6.924	± 16
TN ~ Temp + logTemp + logSC + Turb + logTurb	5	0.3219	90.59	89.2	1.484	6.489	2.883	± 16
TN ~ Q + logQ + Temp + Turb + logTurb + sin4DY	6	0.3148	91.27	89.68	1.352	4.927	8.187	± 16
TN ~ Q + logQ + Temp + Turb + logTurb + Chl	6	0.3198	90.98	89.34	2.243	4.957	7.828	± 16
TN ~ Q + Temp + logSC + Turb + logTurb + sin4DY	6	0.321	90.92	89.27	2.445	5.585	3.814	± 16

Data

	Date	TN	Temp	logTurb	Turb	Computed TN	Residual	Normal Quantiles
1	2013-03-11	1.785	3.51	1.495	31.28	2.022	-0.237	-0.71
2	2013-04-08	1.911	14.4	1.332	21.5	1.314	0.597	1.34
3	2013-05-06	2.612	10.94	2.188	154	3.088	-0.476	-1.2
4	2013-05-20	1.6	23.54	1.788	61.4	1.862	-0.262	-0.793
5	2013-06-03	2.911	20.57	2.184	152.9	2.734	0.177	0.71
6	2013-06-17	3.171	25.83	2.204	160	2.582	0.589	1.2
7	2013-07-01	2.753	25.97	2.258	181	2.68	0.0729	0.22
8	2013-07-15	1.365	26.33	1.526	33.6	1.257	0.108	0.417
9	2013-08-05	4.177	23.77	2.822	663.3	3.848	0.329	0.881
10	2013-08-19	3.028	23.84	2.551	356	3.324	-0.296	-0.881
11	2013-09-09	0.925	27.5	1.459	28.75	1.084	-0.159	-0.284
12	2013-09-23	0.957	19.52	1.554	35.84	1.557	-0.6	-1.75
13	2013-10-21	1.953	13.95	1.204	16	1.083	0.87	2.17
14	2013-11-18	1.464	10.76	1.422	26.4	1.618	-0.154	-0.22
15	2014-03-10	1.502	6.35	1.362	23	1.662	-0.16	-0.35
16	2014-03-24	1.591	7.88	1.376	23.75	1.633	-0.0423	-0.0312
17	2014-04-07	1.756	11.9	1.591	39	1.903	-0.147	-0.0937
18	2014-04-21	1.262	18.02	1.527	33.67	1.559	-0.297	-0.977
19	2014-05-05	1.731	21.5	1.803	63.57	1.965	-0.234	-0.632
20	2014-05-19	1.197	19.83	1.648	44.5	1.727	-0.53	-1.52
21	2014-06-02	1.703	27.4	1.739	54.83	1.628	0.0749	0.284
22	2014-06-30	1.9	27.3	1.892	78	1.927	-0.0269	0.0312
23	2014-07-14	0.894	28.35	1.294	19.67	0.7351	0.159	0.632
24	2014-07-28	0.758	28.05	1.411	25.75	0.9715	-0.214	-0.558
25	2014-08-11	1.217	28.34	1.621	41.8	1.367	-0.15	-0.157
26	2014-08-25	1.22	30.56	1.539	34.6	1.128	0.0918	0.35
27	2014-09-08	3.605	24.77	2.652	449	3.484	0.121	0.558
28	2014-09-22	1.863	20.68	1.839	69	2.064	-0.201	-0.486
29	2014-12-15	1.639	8.7	1.623	41.99	2.081	-0.442	-1.08

30	2015-01-12	1.649	0.04	0.9227	8.37	1.043	0.606	1.52
31	2015-02-09	1.371	5.3	1.439	27.5	1.849	-0.478	-1.34
32	2015-03-09	1.404	8.5	0.8692	7.4	0.6341	0.77	1.75
33	2015-03-23	1.61	11.34	1.301	20	1.364	0.246	0.793
34	2015-04-06	1.18	14.53	1.206	16.07	1.066	0.114	0.486
35	2015-04-20	1.404	14.2	1.708	51	2.045	-0.641	-2.17
36	2015-05-04	1.518	22.28	1.574	37.49	1.495	0.0232	0.0937
37	2015-05-18	4.567	19.98	2.829	674.5	3.999	0.568	1.08
38	2015-06-07	4.647	21.3	2.992	982.5	4.266	0.381	0.977
39	2015-06-15	2.891	26.2	2.348	223	2.846	0.0446	0.157
40	2015-06-29	2.383	25.89	2.204	160	2.58	-0.197	-0.417

Definitions and National Water Information System (parameter code)

TN: Nitrogen, mixed forms (NH3), (NH4), organic, (NO2) and (NO3) in mg/L (00600)
Temp: Temperature, water in deg C (00010)
Turb: Turbidity in FNU (63680)