

Appendix 30. Model Archival Summary for Actinomycetes Concentration at Station 06887500; Kansas River at Wamego, Kansas

This model archival summary summarizes the actinomycetes concentration (Act) model developed to compute 15-minute Act 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 53 concurrent measurements of Act 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 limits. Three samples were estimated due to colony counts outside of the ideal range. 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. None of the Act samples were deemed outliers.

Actinomycetes 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 Act concentration at the Ohio Water Microbiology Laboratory in Columbus, Ohio.

Model Development

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

Turb was selected as the best predictor of Act 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 Mallows' 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 Act concentration at site number 06887500.

Act concentration-based model:

$$\log_{10}(Act) = 1.35 \times \log_{10}(Turb) + 0.198$$

where

Act = Actinomycetes in colonies per milliliter (col/mL); and,

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

Turbidity makes physical and statistical sense as explanatory variables for Act.

The log-transformed model may be retransformed to the original units so that Act 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 1.58. The retransformed model, accounting for BCF is:

$$Act = Turb^{1.35} + 2.49$$

Previous Models

No previous models have been published.

Actinomycetes Concentration Record

The Act 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 Actinomycetes; 06887500; Kansas River at Wamego, KS

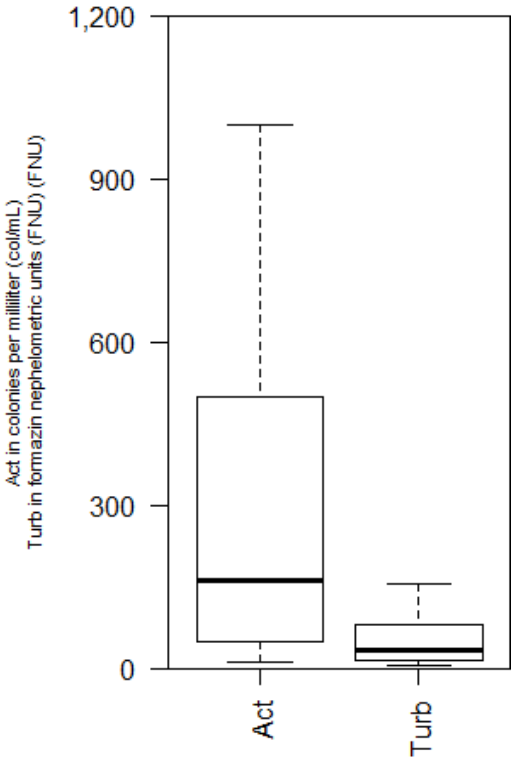
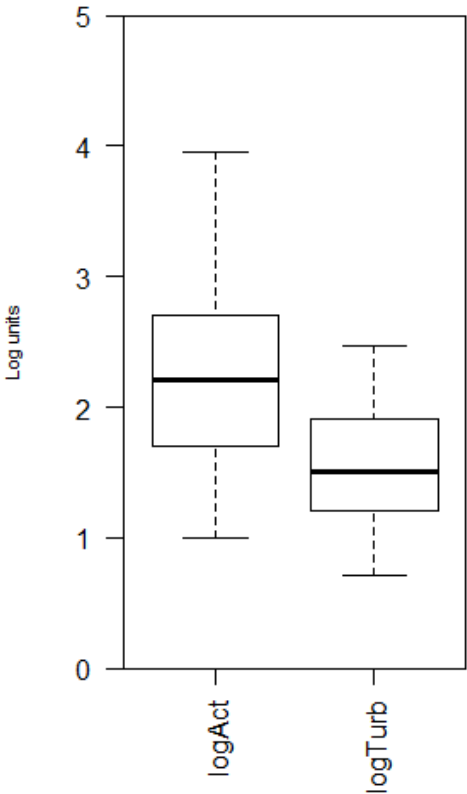
Model Form

$\log \text{Act} = + 1.35 * \log \text{Turb} + 0.198$

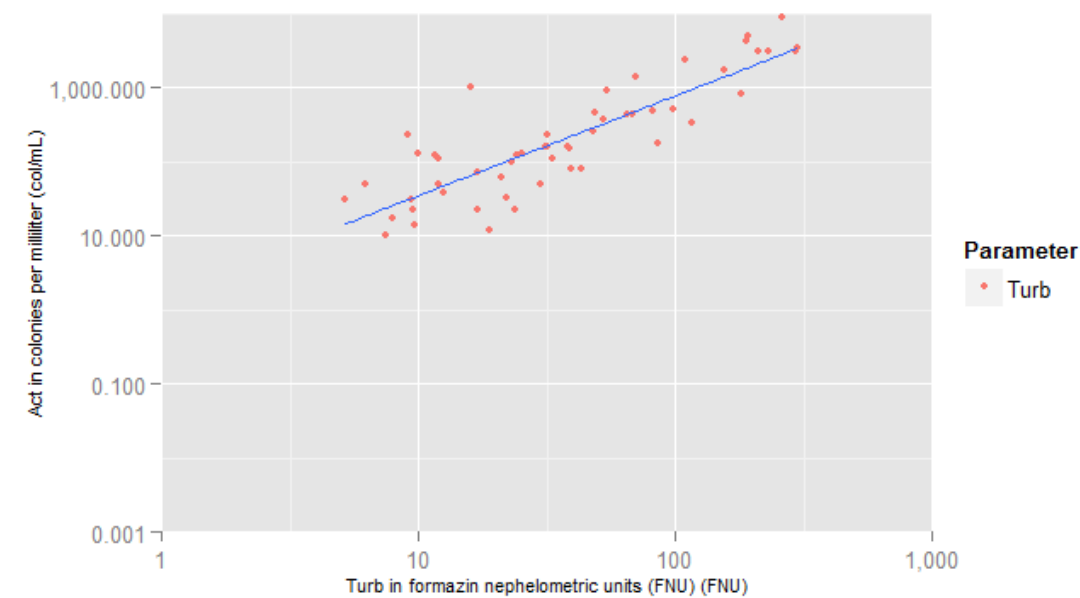
Variable Summary Statistics

	logAct	logTurb	Act	Turb
Minimum	1.00	0.716	10	5.2
1st Quartile	1.70	1.200	50	16.0
Median	2.20	1.510	160	32.0
Mean	2.30	1.560	847	67.1
3rd Quartile	2.70	1.910	500	81.2
Maximum	3.95	2.480	9000	299.0

Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

Number of Observations	53
Standard error (RMSE)	0.393
Upper Model standard percentage error (MSPE)	147
Lower Model standard percentage error (MSPE)	59.5
Coefficient of determination (R²)	0.739
Adjusted Coefficient of Determination (Adj. R²)	0.734
Bias Correction Factor (BCF)	1.58

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(> t)
(Intercept)	0.198	0.183	1.08	2.84e-01
logTurb	1.350	0.112	12.00	1.71e-16

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.956
E.vars	-0.956	1.000

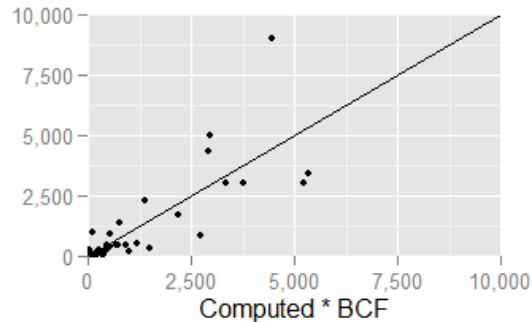
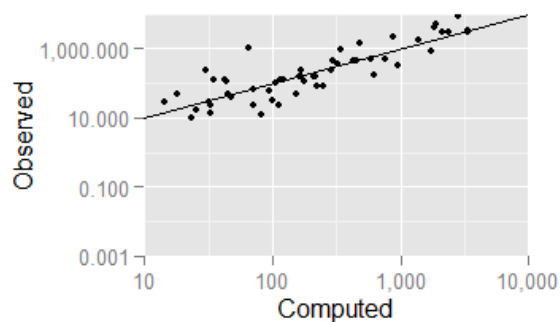
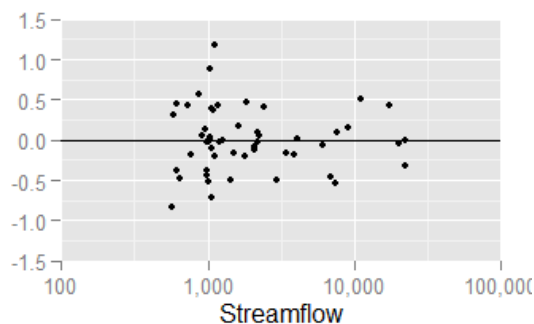
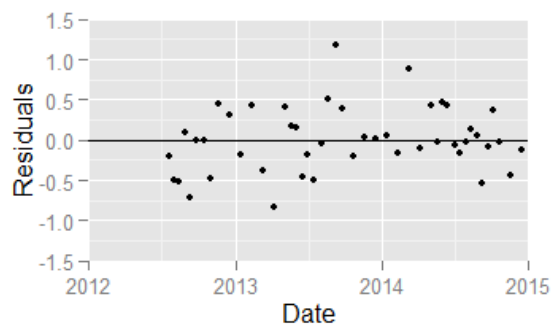
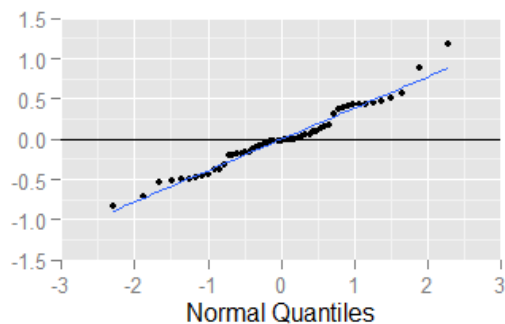
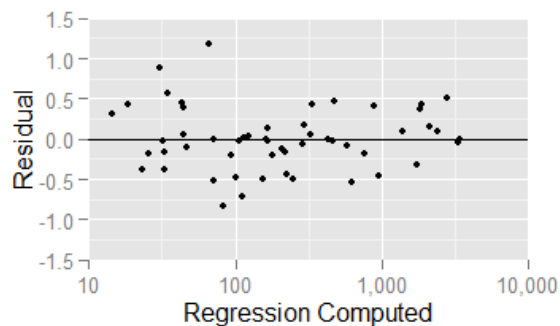
Test Criteria

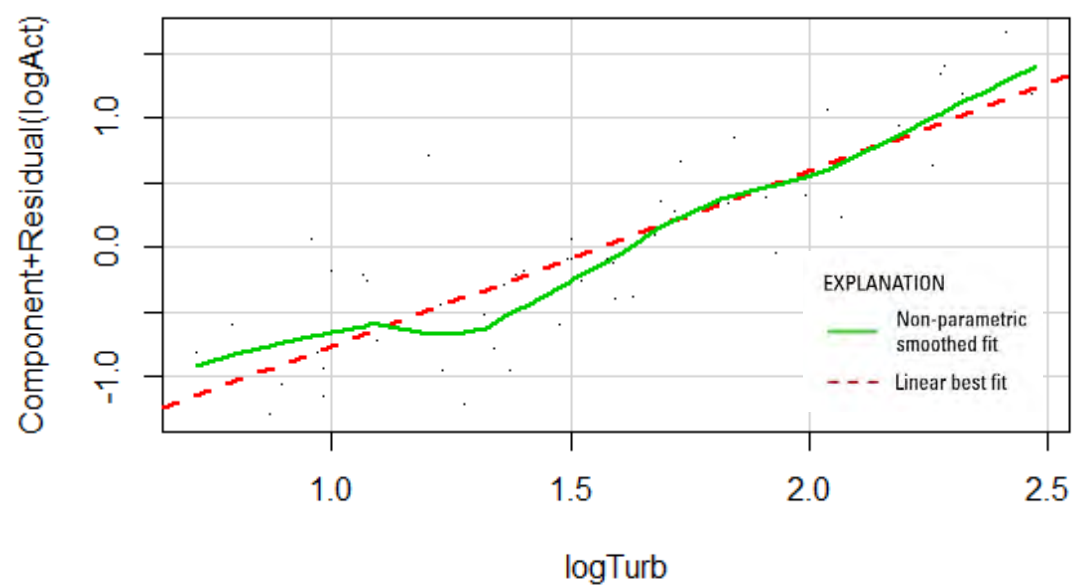
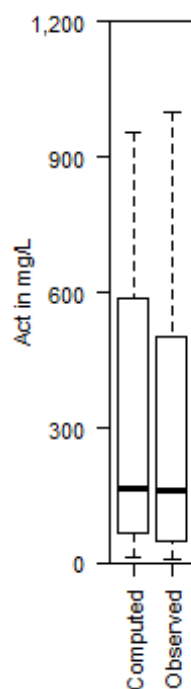
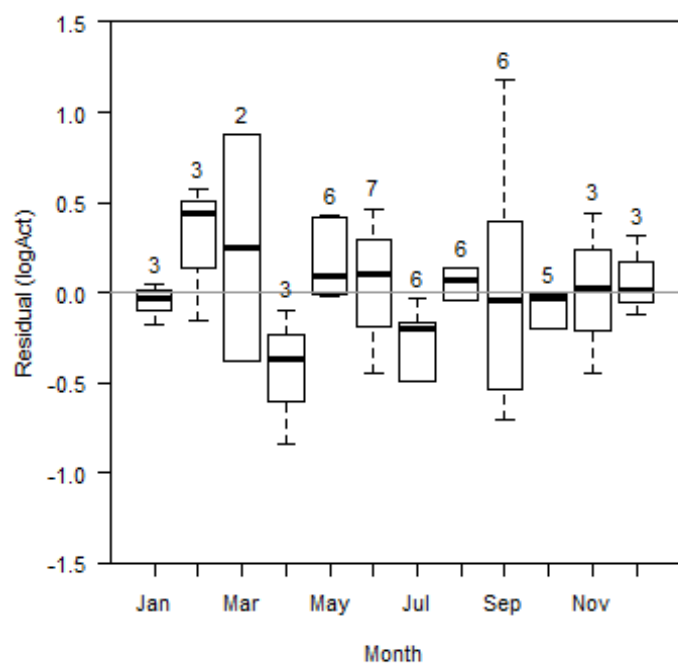
Leverage	Cook's D	DFFITS
0.05660377	0.10557428	0.27472113

Flagged Observations

	logAct	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
09/10/2012 9:00	1.342	2.048	-0.706000	-1.818000	-1.862000	0.02175	3.675e-02	-0.277600
12/17/2012 8:15	1.477	1.162	0.315400	0.836300	0.833800	0.07725	2.927e-02	0.241200
02/11/2013 8:20	1.699	1.265	0.434500	1.146000	1.149000	0.06719	4.728e-02	0.308500
04/08/2013 7:40	1.079	1.919	-0.839800	-2.167000	-2.252000	0.02545	6.130e-02	-0.363800
06/03/2013 8:50	3.477	3.323	0.153900	0.405700	0.402300	0.06575	5.791e-03	0.106700
08/05/2013 7:30	3.477	3.520	-0.042710	-0.113700	-0.112600	0.08554	6.051e-04	-0.034450
08/19/2013 9:30	3.954	3.448	0.506200	1.343000	1.354000	0.07791	7.617e-02	0.393500
09/09/2013 8:00	3.000	1.819	1.181000	3.054000	3.346000	0.02935	1.411e-01	0.581800
03/10/2014 8:20	2.362	1.487	0.874900	2.285000	2.388000	0.04870	1.336e-01	0.540200
06/11/2014 9:00	3.699	3.272	0.426600	1.122000	1.124000	0.06120	4.100e-02	0.287100
10/06/2014 9:00	3.633	3.265	0.368800	0.969100	0.968500	0.06054	3.026e-02	0.245800
02/09/2015 10:20	2.114	1.544	0.570100	1.486000	1.504000	0.04467	5.159e-02	0.325200
04/06/2015 8:50	1.000	1.369	-0.369500	-0.969600	-0.969000	0.05791	2.889e-02	-0.240200
05/18/2015 10:30	3.531	3.529	0.002286	0.006092	0.006032	0.08657	1.758e-06	0.001857
06/01/2015 8:00	3.477	3.376	0.100800	0.266300	0.263800	0.07076	2.699e-03	0.072800
06/15/2015 7:50	2.924	3.239	-0.314600	-0.825700	-0.823100	0.05833	2.112e-02	-0.204900

Statistical Plots





Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logAct ~ logTurb	1	0.3926	73.9	73.39	5.58	8.47	1 ± 100	
logAct ~ Turb	1	0.4345	68.03	67.41	17.85	10.52	1 ± 120	
logAct ~ logQ	1	0.5993	39.19	38	78.17	19.56	1 ± 190	
logAct ~ Turb + logTurb	2	0.3837	75.55	74.58	4.119	8.174	4.884 ± 100	
logAct ~ logQ + logTurb	2	0.3964	73.91	72.87	7.554	8.695	2.189 ± 100	
logAct ~ Q + logTurb	2	0.3965	73.9	72.86	7.574	8.673	1.774 ± 100	
logAct ~ Q + Turb + logTurb	3	0.3759	77.01	75.6	3.073	7.755	3.154 ± 98	
logAct ~ logQ + Turb + logTurb	3	0.3824	76.21	74.75	4.746	8.115	2.808 ± 100	
logAct ~ Q + logQ + logTurb	3	0.4004	73.92	72.32	9.546	8.954	5.427 ± 110	
logAct ~ Q + logQ + Turb + logTurb	4	0.3795	77.05	75.13	5	7.972	7.746 ± 99	

Data

0	Date	logAct	logTurb	Act	Turb	Computed logAct	Computed Act	Residual	Normal Quantiles
1	2012-07-19	2.041	1.524	110	33.4	2.249	280.4	-0.207	-0.716
2	2012-07-30	1.699	1.479	50	30.1	2.188	243.7	-0.489	-1.36
3	2012-08-13	1.342	1.233	22	17.1	1.857	113.9	-0.515	-1.49
4	2012-08-27	3.23	2.188	1700	154	3.142	2193	0.0885	0.438
5	2012-09-10	1.342	1.375	22	23.71	2.048	176.8	-0.706	-1.88
6	2012-09-24	2.204	1.496	160	31.33	2.211	257.2	-0.00718	0.0944
7	2012-10-15	1.845	1.23	70	17	1.854	113	-0.00891	0.0471
8	2012-10-29	1.519	1.342	33	22	2.005	159.8	-0.486	-1.15
9	2012-11-19	2.079	1.068	120	11.69	1.635	68.26	0.444	1.25
10	2012-12-17	1.477	0.716	30	5.2	1.162	22.95	0.315	0.716
11	2013-01-14	1.23	0.8965	17	7.88	1.405	40.15	-0.174	-0.544
12	2013-02-11	1.699	0.7924	50	6.2	1.265	29.08	0.434	1.15
13	2013-03-11	1.146	0.9845	14	9.65	1.523	52.73	-0.377	-0.914
14	2013-04-08	1.079	1.279	12	19	1.919	131.2	-0.84	-2.28
15	2013-05-06	3.362	2.041	2300	110	2.945	1394	0.416	0.914
16	2013-05-20	2.653	1.69	450	49	2.473	469.6	0.181	0.657
17	2013-06-03	3.477	2.322	3000	210	3.323	3328	0.154	0.599
18	2013-06-17	2.531	2.067	340	116.7	2.98	1509	-0.448	-1.07
19	2013-07-01	2.699	1.993	500	98.5	2.881	1202	-0.182	-0.599
20	2013-07-15	1.908	1.633	81	43	2.396	393.9	-0.488	-1.25
21	2013-08-05	3.477	2.468	3000	294	3.52	5234	-0.0427	-0.19
22	2013-08-19	3.954	2.415	9000	260	3.448	4436	0.506	1.49
23	2013-09-09	3	1.204	1000	16	1.819	104.1	1.18	2.28
24	2013-09-23	2.041	1.076	110	11.9	1.646	69.92	0.396	0.844
25	2013-10-21	1.778	1.322	60	21	1.977	150.1	-0.199	-0.657
26	2013-11-18	2.114	1.404	130	25.33	2.087	193.2	0.0269	0.287
27	2013-12-16	2.079	1.386	120	24.33	2.064	183	0.0157	0.238
28	2014-01-13	1.699	1.079	50	12	1.65	70.71	0.0485	0.336
29	2014-02-10	1.362	0.9823	23	9.6	1.52	52.37	-0.158	-0.438
30	2014-03-10	2.362	0.9576	230	9.07	1.487	48.51	0.875	1.88
31	2014-04-07	1.58	1.097	38	12.5	1.674	74.7	-0.0945	-0.336
32	2014-05-05	2.954	1.732	900	53.99	2.529	535	0.425	0.988
33	2014-05-19	2.204	1.505	160	32	2.224	264.7	-0.0195	0
34	2014-06-02	3.146	1.845	1400	70	2.681	758.8	0.465	1.36
35	2014-06-11	3.699	2.284	5000	192.5	3.272	2960	0.427	1.07
36	2014-06-30	2.398	1.679	250	47.79	2.458	454	-0.0601	-0.238

37	2014-07-14	2.176	1.591	150	39	2.339	345.4	-0.163	-0.49
38	2014-07-28	2	1.362	100	23	2.031	169.7	-0.0307	-0.142
39	2014-08-11	2.362	1.505	230	32	2.224	264.7	0.138	0.544
40	2014-08-25	2.568	1.72	370	52.5	2.513	515.3	0.0552	0.386
41	2014-09-08	2.255	1.932	180	85.42	2.797	992	-0.542	-1.66
42	2014-09-22	2.681	1.909	480	81.17	2.768	926.1	-0.0864	-0.287
43	2014-10-06	3.633	2.279	4300	190	3.265	2909	0.369	0.779
44	2014-10-20	2.633	1.831	430	67.8	2.662	726.9	-0.029	-0.0944
45	2014-11-17	1.903	1.597	80	39.5	2.347	351.4	-0.444	-0.988
46	2014-12-15	2.204	1.58	160	38	2.324	333.5	-0.12	-0.386
47	2015-01-12	1.477	0.9713	30	9.36	1.505	50.61	-0.0281	-0.0471
48	2015-02-09	2.114	1	130	10	1.544	55.32	0.57	1.66
49	2015-04-06	1	0.8704	10	7.42	1.369	37.03	-0.369	-0.844
50	2015-05-04	2.633	1.813	430	65	2.638	686.8	-0.00435	0.142
51	2015-05-18	3.531	2.475	3400	298.8	3.529	5348	0.00229	0.19
52	2015-06-01	3.477	2.362	3000	230	3.376	3762	0.101	0.49
53	2015-06-15	2.924	2.26	840	181.8	3.239	2741	-0.315	-0.779

Definitions and National Water Information System (parameter code)

Act: Actinomycetes in col/mL (63688)

Turb: Turbidity in FNU (63680)