Appendix 8. Model Archival Summary for Total Dissolved Solids Concentration at Station 06892350; Kansas River at De Soto, Kansas

This model archival summary summarizes the total dissolved solids concentration (TDS) model developed to compute 15-minute TDS 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 59 concurrent measurements of TDS concentration and specific conductance (SC) 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. All other potential outliers were not found to have errors associated with collection, processing, or analysis, and were therefore considered valid.

Total Dissolved Solids 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 TDS concentration at the USGS National Water Quality Laboratory in Lakewood, Colorado.

Model Development

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

SC was selected as the best predictor of TDS 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 TDS concentration at site number 06892350.

TDS concentration-based model:

$$\log_{10}(TDS) = 0.938 \times \log_{10}(SC) - 0.0345$$

where

TDS = total dissolved solids in milligrams per liter (mg/L); and,

SC = specific conductance in microsiemens per centimeter at 25 degrees Celsius (µs/cm)

SC makes physical and statistical sense as explanatory variables for TDS.

The log-transformed model may be retransformed to the original units so that TDS 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 1.00. The retransformed model, accounting for BCF is:

$$TDS = SC^{0.938} + 0.924$$

Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(TDS) = 0.966 \times \log_{10}(SC) - 0.115$

Total Dissolved Solids Concentration Record

The DS 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 from Total Dissolved Solids; 06892350; Kansas River at De Soto, KS

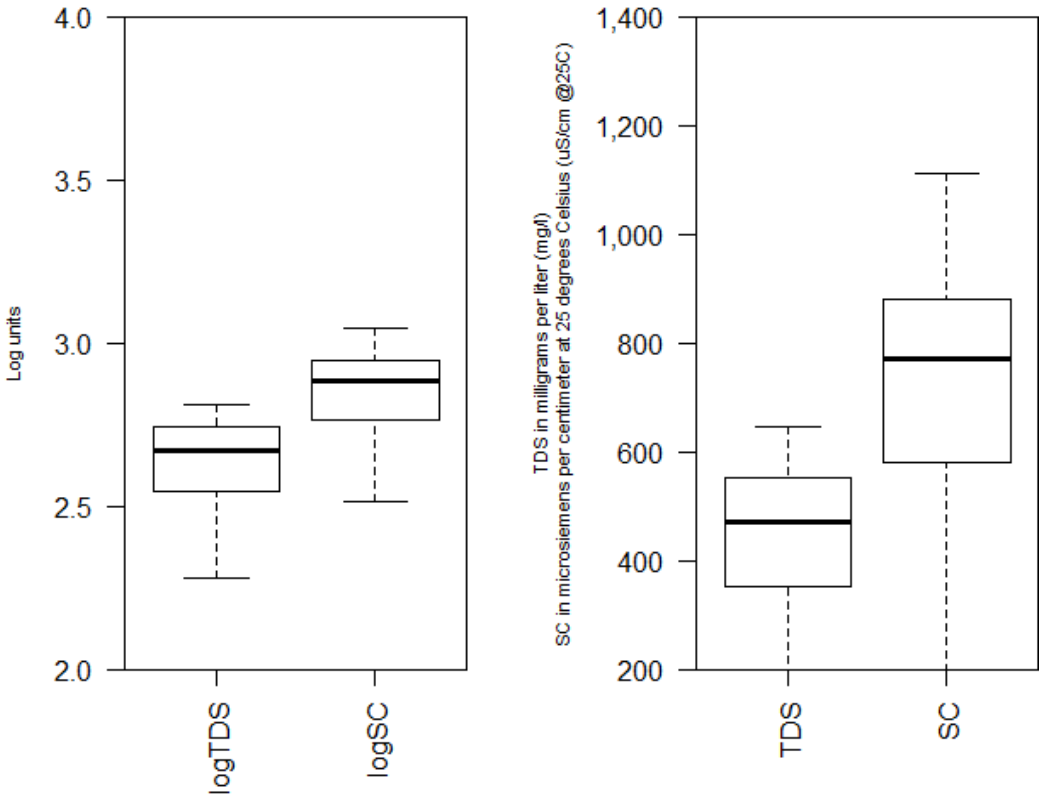
Model Form

logTDS = + 0.938 * logSC + -0.0345

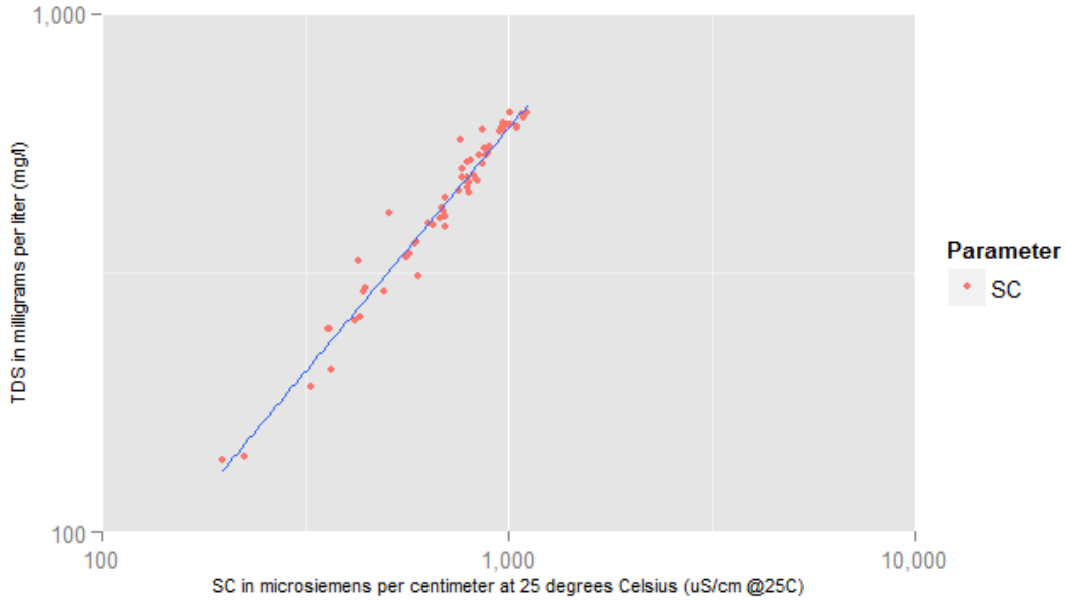
Variable Summary Statistics

	logTDS	logSC	TDS	SC
Minimum	2.14	2.29	138	197
1st Quartile	2.54	2.76	345	569
Median	2.67	2.89	470	770
Mean	2.62	2.83	446	726
3rd Quartile	2.74	2.95	552	882
Maximum	2.81	3.05	645	1110

Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

Number of Observations	59
Standard error (RMSE)	0.0325
Upper Model standard percentage error (MSPE)	7.77
Lower Model standard percentage error (MSPE)	7.21
Coefficient of determination (R^2)	0.96
Adjusted Coefficient of Determination (Adj. R^2)	0.959
Bias Correction Factor (BCF)	1

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.0345	0.0724	-0.477	6.35e-01
logSC	0.9380	0.0255	36.800	2.10e-41

Correlation Matrix

	Intercept	E.vars
Intercept	1.000	-0.998
E.vars	-0.998	1.000

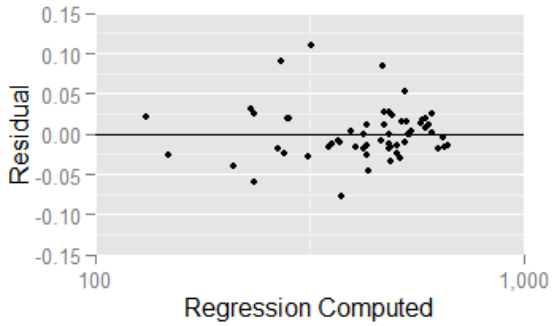
Test Criteria

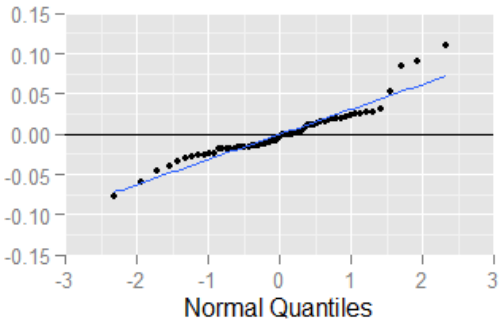
Leverage	Cook's D	DFFITS
0.05084746	0.10555214	0.26037782

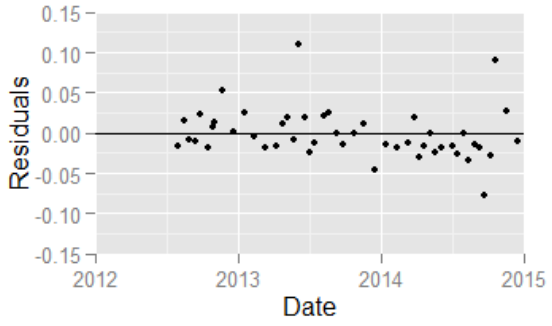
Flagged Observations

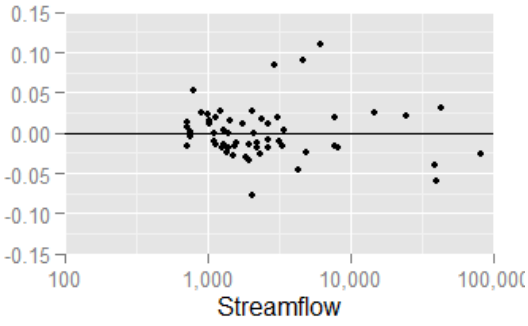
	logTDS	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
6/3/2013 13:10	2.614	2.504	0.10980	3.4260	3.8110	0.02705	0.16320	0.6355
8/5/2013 11:30	2.139	2.118	0.02110	0.7244	0.7214	0.19630	0.06408	0.3565
8/19/2013 10:00	2.393	2.368	0.02495	0.7933	0.7907	0.06289	0.02112	0.2048
9/22/2014 10:00	2.495	2.573	-0.07772	-2.4150	-2.5260	0.01879	0.05583	-0.3496
10/20/2014 14:00	2.524	2.434	0.09051	2.8460	3.0460	0.04226	0.17880	0.6399
1/12/2015 11:30	2.755	2.670	0.08500	2.6400	2.7940	0.01841	0.06536	0.3825
5/18/2015 15:30	2.282	2.322	-0.04047	-1.2990	-1.3070	0.08080	0.07418	-0.3876
6/6/2015 19:50	2.146	2.171	-0.02547	-0.8557	-0.8537	0.16050	0.07002	-0.3733
6/15/2015 14:50	2.393	2.363	0.03047	0.9697	0.9692	0.06486	0.03261	0.2552
6/29/2015 12:30	2.312	2.371	-0.05893	-1.8720	-1.9160	0.06167	0.11520	-0.4911

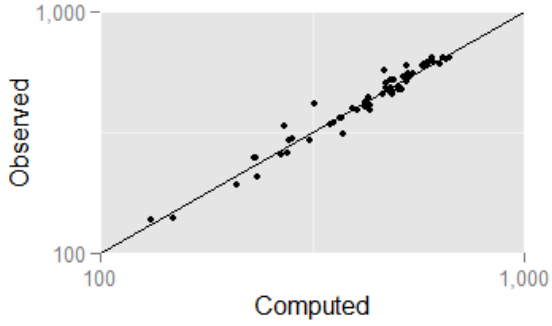
Statistical Plots

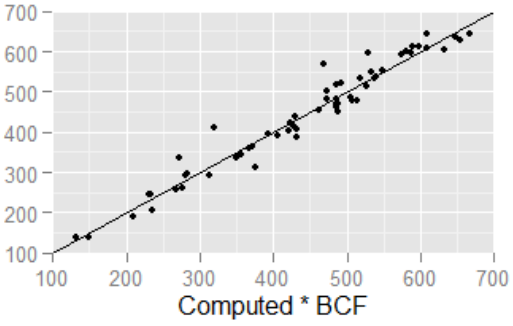


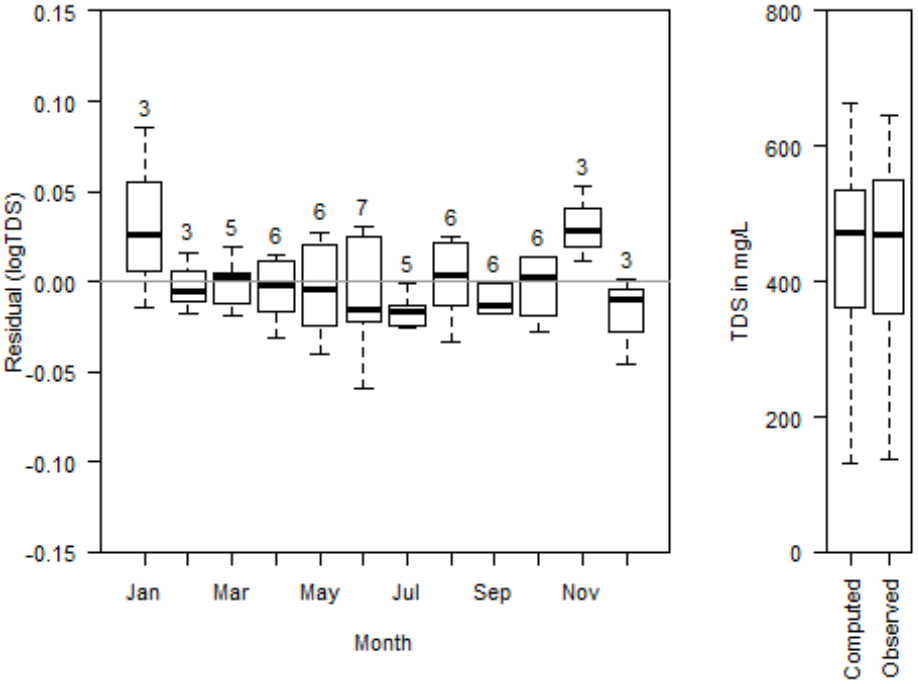


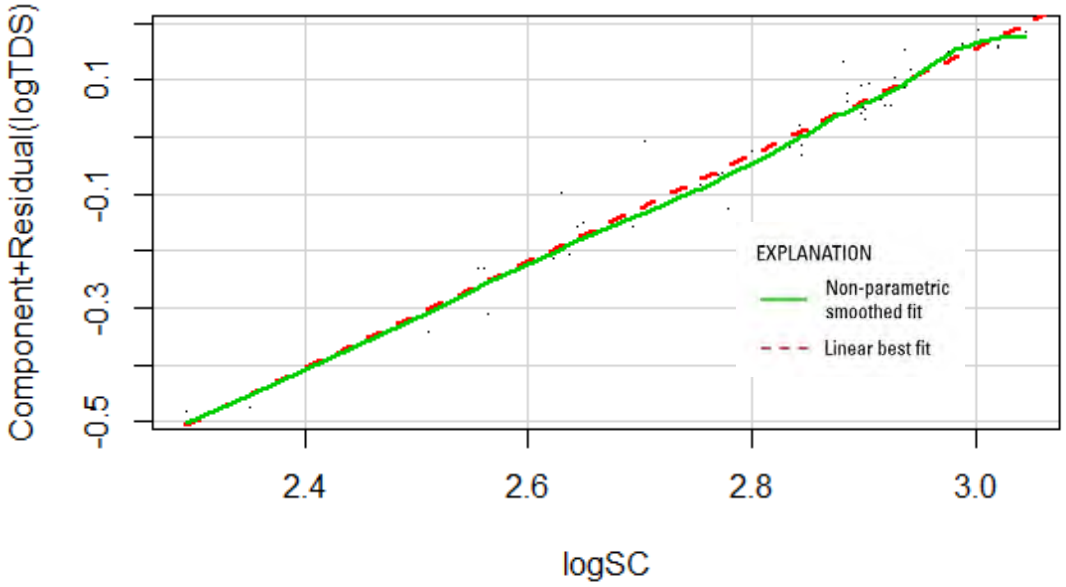









Residual (logTDS)
0.15
0.10
0.05
0.00
-0.05
-0.10
-0.15
3
3
5
6
6
7
5
6
6
6
3
3
Jan
Mar
May
Jul
Sep
Nov
Month
TDS in mg/L
800
600
400
200
0
Computed
Observed


Component+Residual(logTDS)
0.1
-0.1
-0.3
-0.5
2.4
2.6
2.8
3.0
logSC
EXPLANATION
Non-parametric smoothed fit
Linear best fit

Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logTDS ~ logSC	1	0.03249	95.96	95.89	1.895	0.06458	1	± 7.5
logTDS ~ SC	1	0.04813	91.13	90.97	69.86	0.1492	1	± 11
logTDS ~ logQ	1	0.07893	76.14	75.72	280.7	0.3825	1	± 18
logTDS ~ Q + logSC	2	0.03218	96.1	95.96	1.831	0.06412	2.089	± 7.4
logTDS ~ logQ + logSC	2	0.03251	96.02	95.88	2.977	0.06626	4.375	± 7.5
logTDS ~ SC + logSC	2	0.03278	95.96	95.81	3.895	0.06686	19.88	± 7.6
logTDS ~ Q + SC + logSC	3	0.03223	96.16	95.95	3.01	0.0649	2.925	± 7.4
logTDS ~ Q + logQ + logSC	3	0.03247	96.1	95.89	3.826	0.06667	3.472	± 7.5
logTDS ~ logQ + SC + logSC	3	0.0328	96.03	95.81	4.932	0.06807	4.594	± 7.6
logTDS ~ Q + logQ + SC + logSC	4	0.03252	96.16	95.88	5	0.06749	4.787	± 7.5

Data

	Date	logTDS	logSC	TDS	SC	Computed logTDS	Computed TDS	Residual	Normal Quantiles
1	2012-07-30	2.591	2.816	390.1	655	2.608	406.2	-0.0164	-0.532
2	2012-08-13	2.74	2.942	550.1	875.5	2.726	533.3	0.0147	0.633
3	2012-08-27	2.657	2.877	453.8	753.2	2.664	463.1	-0.00753	-0.0424
4	2012-09-10	2.71	2.938	512.9	866	2.721	527.8	-0.0112	-0.17
5	2012-09-24	2.716	2.906	519.9	805.3	2.692	493.1	0.0242	0.986
6	2012-10-15	2.666	2.898	463.1	791	2.684	484.8	-0.0187	-0.797
7	2012-10-26	2.775	2.987	595	971	2.768	587.7	0.00662	0.345
8	2012-10-29	2.772	2.977	591.4	948	2.758	574.6	0.0137	0.532
9	2012-11-19	2.775	2.938	595	867	2.722	528.4	0.0528	1.55
10	2012-12-17	2.785	3.004	609.9	1010	2.784	609.8	0.00128	0.213
11	2013-01-14	2.809	3.003	644.5	1008	2.783	608.5	0.0262	1.13
12	2013-02-11	2.804	3.031	637.5	1075	2.809	646.5	-0.00492	0
13	2013-03-11	2.781	3.021	603.5	1050	2.8	632.4	-0.0191	-0.856
14	2013-04-08	2.798	3.037	628.3	1090	2.815	655	-0.0169	-0.582
15	2013-04-25	2.644	2.843	440.4	697	2.633	430.6	0.011	0.484
16	2013-05-06	2.471	2.649	295.9	446	2.451	283.2	0.0202	0.856
17	2013-05-20	2.556	2.77	360	589.2	2.564	367.8	-0.00808	-0.0848
18	2013-06-03	2.614	2.706	410.9	508	2.504	320	0.11	2.32
19	2013-06-17	2.465	2.644	292.1	441	2.446	280.3	0.0191	0.797
20	2013-07-01	2.415	2.637	260.2	433.5	2.439	275.8	-0.024	-0.919
21	2013-07-15	2.537	2.755	344.6	569.4	2.55	356.2	-0.0131	-0.257
22	2013-08-05	2.139	2.294	137.8	197	2.118	131.6	0.0211	0.919
23	2013-08-19	2.393	2.561	247.1	363.8	2.368	234	0.025	1.06
24	2013-09-09	2.728	2.945	534	882	2.729	537	-0.0012	0.0424
25	2013-09-23	2.687	2.917	486.3	826	2.702	504.9	-0.0152	-0.391
26	2013-10-21	2.73	2.947	536.5	886	2.731	539.3	-0.00105	0.127
27	2013-11-18	2.787	2.996	612.2	990.6	2.776	598.8	0.0109	0.391
28	2013-12-16	2.589	2.846	388.6	701	2.635	432.9	-0.0457	-1.71
29	2014-01-13	2.808	3.045	643.4	1110	2.822	666.2	-0.014	-0.345
30	2014-02-10	2.782	3.021	605.4	1050	2.8	632.4	-0.0177	-0.633
31	2014-03-10	2.673	2.899	470.9	793.3	2.686	486.1	-0.0127	-0.213
32	2014-03-24	2.788	2.988	613.5	973.2	2.769	588.9	0.019	0.74
33	2014-04-07	2.679	2.925	477.2	841.2	2.709	513.7	-0.0308	-1.31
34	2014-04-21	2.673	2.902	470.5	797.8	2.688	488.8	-0.0153	-0.484
35	2014-05-05	2.625	2.836	422.2	684.9	2.626	423.5	-0.000216	0.17
36	2014-05-19	2.679	2.919	477.7	830.2	2.704	507.4	-0.025	-0.986

37	2014-06-02	2.606	2.834	403.3	682.3	2.624	422.1	-0.0185	-0.74
38	2014-06-30	2.528	2.748	337.5	559.7	2.543	350.5	-0.0153	-0.437
39	2014-07-14	2.608	2.845	405.6	699.3	2.634	431.9	-0.0261	-1.13
40	2014-07-28	2.684	2.899	483.3	792.8	2.685	485.8	-0.00108	0.0848
41	2014-08-11	2.654	2.902	451.2	798.2	2.688	489	-0.0337	-1.42
42	2014-08-25	2.617	2.841	414	693.4	2.631	428.5	-0.0137	-0.301
43	2014-09-08	2.409	2.623	256.3	420.2	2.427	267.8	-0.018	-0.685
44	2014-09-22	2.495	2.779	312.7	601.5	2.573	375	-0.0777	-2.32
45	2014-10-06	2.466	2.695	292.1	495.5	2.494	312.6	-0.0283	-1.22
46	2014-10-20	2.524	2.631	334.6	427.8	2.434	272.4	0.0905	1.93
47	2014-11-17	2.713	2.899	516.4	792.8	2.685	485.9	0.0277	1.31
48	2014-12-15	2.559	2.775	362	595.1	2.568	371.2	-0.0097	-0.127
49	2015-01-12	2.755	2.883	568.5	763	2.67	468.7	0.085	1.71
50	2015-02-09	2.779	2.981	600.9	958	2.762	580.3	0.0164	0.685
51	2015-03-09	2.742	2.955	552.4	902	2.738	548.4	0.0044	0.301
52	2015-03-23	2.597	2.801	394.9	633	2.594	393.4	0.00294	0.257
53	2015-04-06	2.727	2.928	533.6	848	2.713	517.5	0.0145	0.582
54	2015-04-20	2.684	2.886	483.1	769.5	2.673	472.5	0.0109	0.437
55	2015-05-04	2.701	2.886	501.8	769.3	2.673	472.3	0.0275	1.22
56	2015-05-18	2.282	2.512	191.3	325.1	2.322	210.5	-0.0405	-1.55
57	2015-06-07	2.146	2.351	139.9	224.5	2.171	148.8	-0.0255	-1.06
58	2015-06-15	2.393	2.555	247.2	359	2.363	231.1	0.0305	1.42
59	2015-06-29	2.312	2.565	205.3	366.9	2.371	235.8	-0.0589	-1.93

Definitions and National Water Information System (parameter code)

TDS: Total dissolved solids in mg/L (70300)
SC: Specific conductance in uS/cm @25C (00095)