Appendix 18. Model Archival Summary for Chloride Concentration at Station 06892350; Kansas River at De Soto, Kansas

This model archival summary summarizes the chloride concentration (Cl) model developed to compute 15-minute Cl 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 Cl 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.

Chloride 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 Cl concentration at the USGS National Water Quality Laboratory in Lakewood, Colorado.

Model Development

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

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

Cl concentration-based model:

$$\log_{10}(Cl) = 1.97 \times \log_{10}(SC) - 3.78$$

where

Cl = chloride in milligrams per liter (mg/L); and,

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

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

The log-transformed model may be retransformed to the original units so that Cl 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 1.01. The retransformed model, accounting for BCF is:

$$Chl = SC^{1.97} + 0.000168$$

Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(Cl) = 1.66 \times \log_{10}(SC) - 2.93$

Chloride Concentration Record

The Cl 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 Chloride; 06892350; Kansas River at De Soto, KS

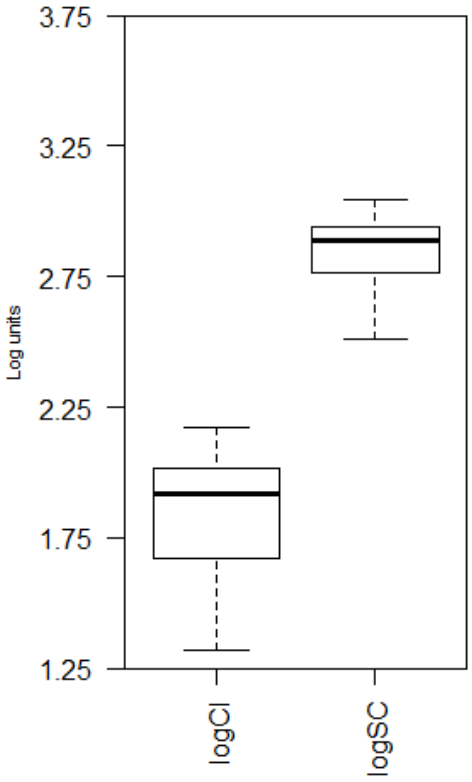
Model Form

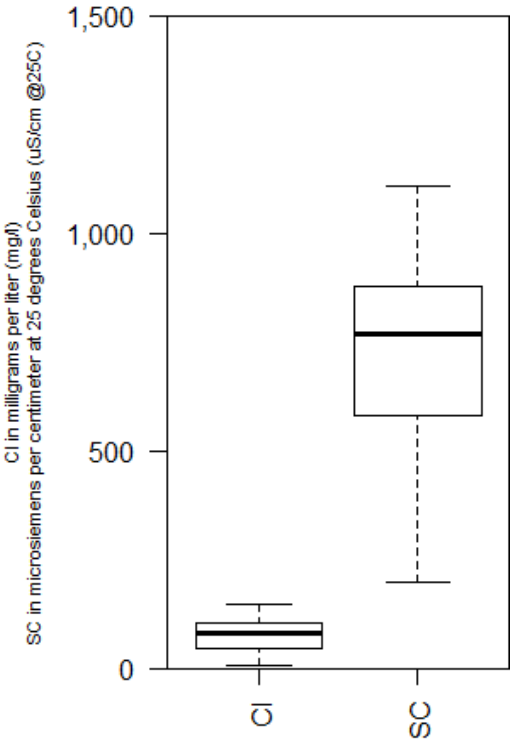
logCl = + 1.97 * logSC + -3.78

Variable Summary Statistics

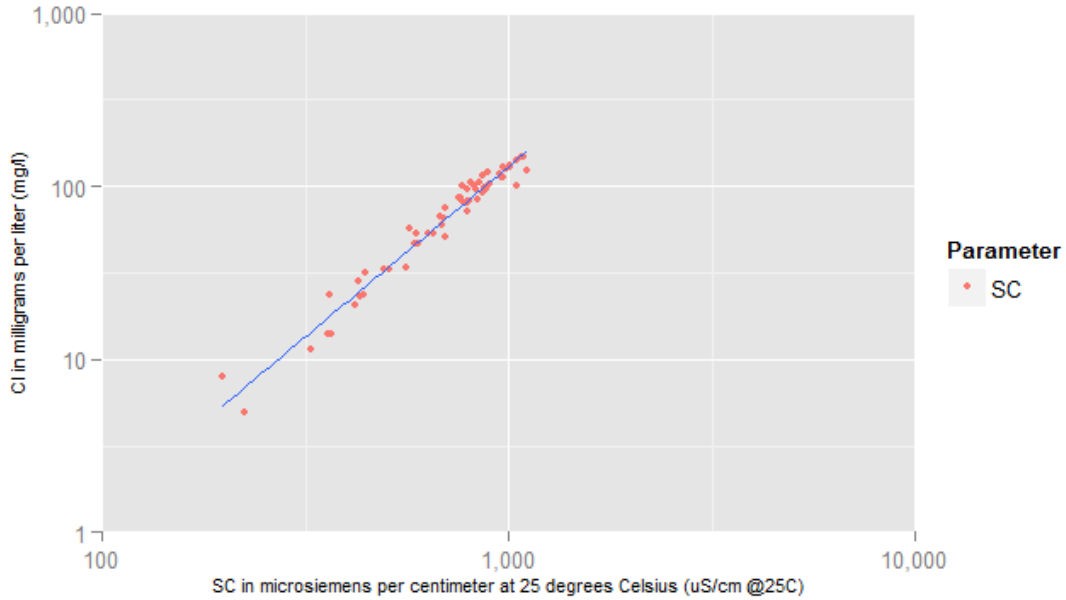
	logCl	logSC	Cl	SC
Minimum	0.693	2.29	4.93	197
1st Quartile	1.670	2.76	46.90	569
Median	1.920	2.89	82.30	770
Mean	1.790	2.83	76.30	726
3rd Quartile	2.020	2.95	105.00	882
Maximum	2.170	3.05	149.00	1110

Box Plot(s) of sample data





Exploratory Plot



Model Calibration

Basic Data

Number of Observations	59
Standard error (RMSE)	0.0667
Upper Model standard percentage error (MSPE)	16.6
Lower Model standard percentage error (MSPE)	14.2
Coefficient of determination (R^2)	0.961
Adjusted Coefficient of Determination (Adj. R^2)	0.96
Bias Correction Factor (BCF)	1.01

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-3.78	0.1490	-25.4	8.43e-33
logSC	1.97	0.0524	37.5	6.88e-42

Correlation Matrix

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

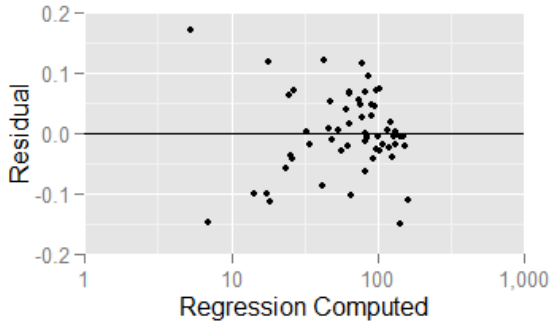
Test Criteria

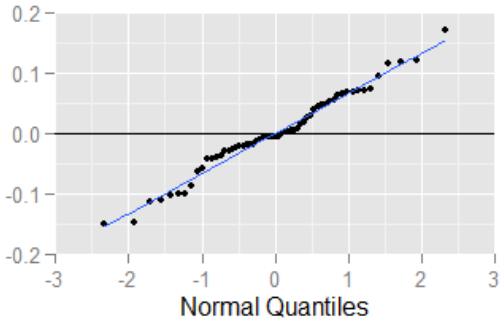
Leverage	Cook's D	DFFITS
0.05084746	0.10555214	0.26037782

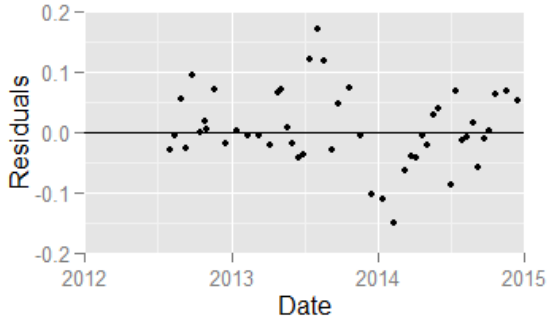
Flagged Observations

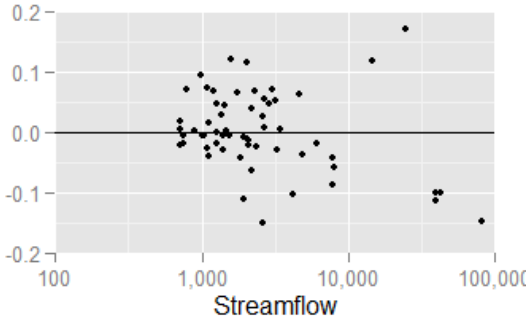
	logCl	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
7/15/2013 10:40	1.7550	1.6350	0.11980	1.815	1.853	0.02075	0.03490	0.2698
8/5/2013 11:30	0.8985	0.7290	0.16950	2.834	3.031	0.19630	0.98070	1.4980
8/19/2013 10:00	1.3700	1.2530	0.11720	1.815	1.853	0.06289	0.11060	0.4802
1/13/2014 8:00	2.0940	2.2050	-0.11130	-1.707	-1.737	0.04449	0.06781	-0.3747
2/10/2014 7:40	2.0070	2.1580	-0.15060	-2.302	-2.396	0.03856	0.10630	-0.4798
5/18/2015 15:30	1.0570	1.1570	-0.09951	-1.556	-1.576	0.08080	0.10640	-0.4672
6/6/2015 19:50	0.6929	0.8406	-0.14760	-2.415	-2.527	0.16050	0.55780	-1.1050
6/15/2015 14:50	1.1410	1.2410	-0.10060	-1.560	-1.580	0.06486	0.08439	-0.4162
6/29/2015 12:30	1.1460	1.2600	-0.11420	-1.767	-1.801	0.06167	0.10260	-0.4618

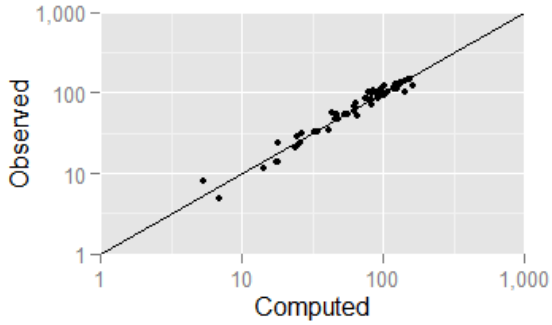
Statistical Plots

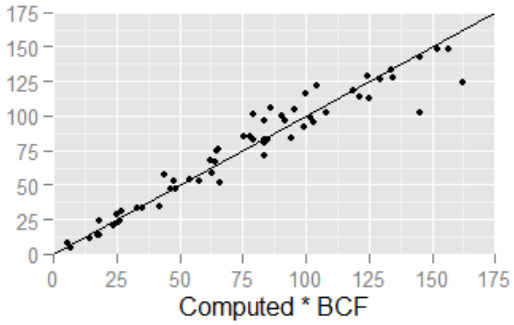


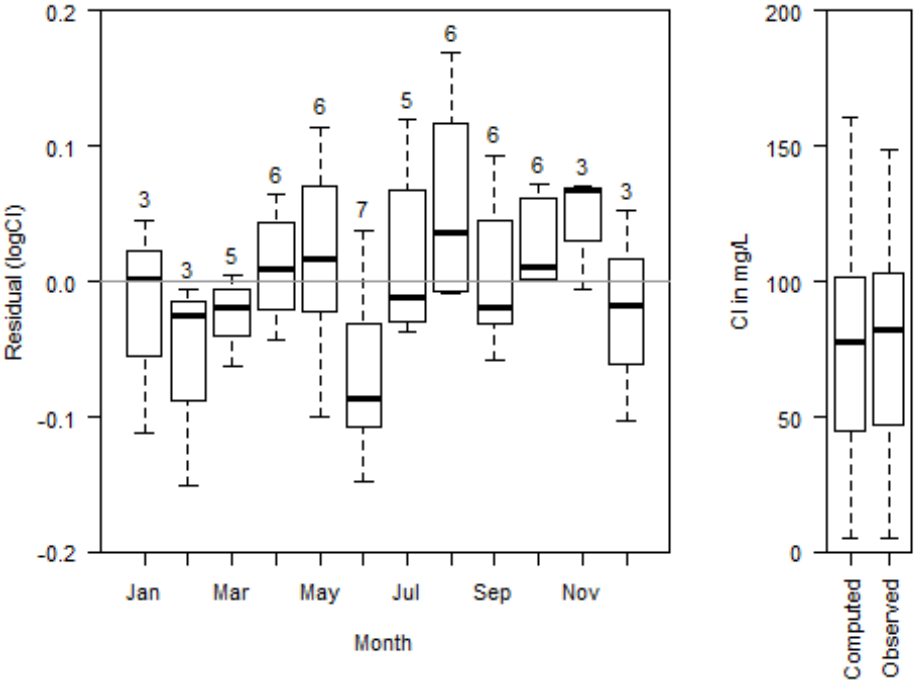


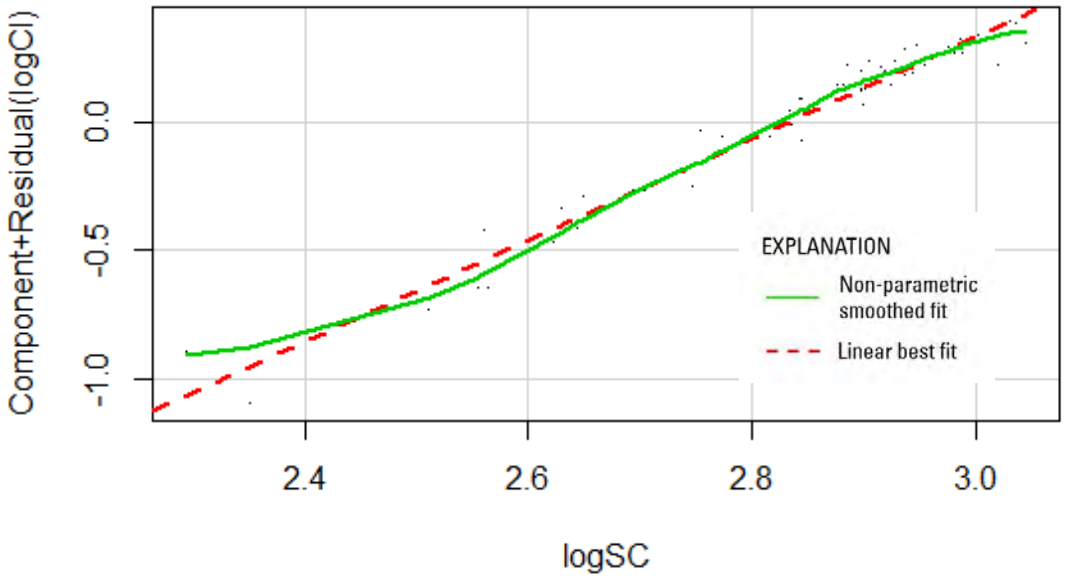









Residual (logCl)
0.2
0.1
0.0
-0.1
-0.2
3
3
5
6
6
7
5
6
6
6
3
3
Jan
Mar
May
Jul
Sep
Nov
Month
Cl in mg/L
200
150
100
50
0
Computed
Observed


Component+Residual(logCl)
0.0
-0.5
-1.0
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
logCl ~ logSC	1	0.06672	96.11	96.04	48.86	0.2944	1	± 15
logCl ~ SC	1	0.1078	89.86	89.68	216	0.7472	1	± 25
logCl ~ logQ	1	0.1298	85.29	85.03	338	1.031	1	± 30
logCl ~ logQ + logSC	2	0.05043	97.82	97.74	5.294	0.1681	4.375	± 12
logCl ~ Q + logSC	2	0.0557	97.34	97.24	18.13	0.2041	2.089	± 13
logCl ~ SC + logSC	2	0.06635	96.22	96.09	47.91	0.3732	19.88	± 15
logCl ~ Q + logQ + logSC	3	0.04931	97.95	97.84	3.744	0.1631	3.472	± 11
logCl ~ logQ + SC + logSC	3	0.05086	97.82	97.7	7.23	0.2076	4.594	± 12
logCl ~ Q + SC + logSC	3	0.05521	97.43	97.29	17.62	0.2351	2.925	± 13
logCl ~ Q + logQ + SC + logSC	4	0.04943	97.98	97.83	5	0.2028	4.787	± 11

Data

	Date	logCl	logSC	Cl	SC	Computed logCl	Computed Cl	Residual	Normal Quantiles
1	2012-07-30	1.726	2.816	53.18	655	1.755	57.51	-0.0291	-0.633
2	2012-08-13	1.996	2.942	99.06	875.5	2.003	101.7	-0.0067	-0.127
3	2012-08-27	1.929	2.877	84.92	753.2	1.874	75.69	0.0549	0.797
4	2012-09-10	1.966	2.938	92.39	866	1.993	99.59	-0.0276	-0.582
5	2012-09-24	2.025	2.906	105.8	805.3	1.931	86.34	0.0934	1.42
6	2012-10-15	1.915	2.898	82.31	791	1.916	83.34	-0.000491	0.0848
7	2012-10-26	2.108	2.987	128.2	971	2.091	124.7	0.0167	0.391
8	2012-10-29	2.075	2.977	118.8	948	2.071	119	0.00419	0.213
9	2012-11-19	2.065	2.938	116	867	1.994	99.81	0.0703	1.22
10	2012-12-17	2.107	3.004	127.9	1010	2.125	134.8	-0.0178	-0.301
11	2013-01-14	2.124	3.003	133	1008	2.123	134.2	0.00115	0.127
12	2013-02-11	2.172	3.031	148.7	1075	2.178	152.3	-0.0056	-0.0424
13	2013-03-11	2.152	3.021	142.1	1050	2.158	145.4	-0.00529	0.0424
14	2013-04-08	2.169	3.037	147.7	1090	2.19	156.5	-0.0203	-0.437
15	2013-04-25	1.872	2.843	74.44	697	1.808	64.99	0.0639	0.919
16	2013-05-06	1.497	2.649	31.39	446	1.427	27.02	0.0701	1.13
17	2013-05-20	1.671	2.77	46.85	589.2	1.664	46.71	0.00628	0.301
18	2013-06-03	1.519	2.706	33.06	508	1.538	34.9	-0.0185	-0.345
19	2013-06-17	1.373	2.644	23.62	441	1.417	26.42	-0.0437	-0.919
20	2013-07-01	1.366	2.637	23.23	433.5	1.402	25.55	-0.0364	-0.74
21	2013-07-15	1.755	2.755	56.9	569.4	1.635	43.67	0.12	1.93
22	2013-08-05	0.8985	2.294	7.916	197	0.729	5.419	0.17	2.32
23	2013-08-19	1.37	2.561	23.44	363.8	1.253	18.1	0.117	1.71
24	2013-09-09	1.978	2.945	95.11	882	2.009	103.2	-0.0307	-0.685
25	2013-09-23	1.998	2.917	99.62	826	1.953	90.75	0.0454	0.633
26	2013-10-21	2.085	2.947	121.6	886	2.013	104.2	0.0721	1.31
27	2013-11-18	2.103	2.996	126.6	990.6	2.108	129.7	-0.00551	0
28	2013-12-16	1.71	2.846	51.23	701	1.813	65.72	-0.103	-1.42
29	2014-01-13	2.094	3.045	124.1	1110	2.205	162.2	-0.111	-1.55
30	2014-02-10	2.007	3.021	101.7	1050	2.158	145.4	-0.151	-2.32
31	2014-03-10	1.855	2.899	71.63	793.3	1.918	83.81	-0.0632	-1.06
32	2014-03-24	2.053	2.988	112.9	973.2	2.093	125.3	-0.0401	-0.797
33	2014-04-07	1.926	2.925	84.35	841.2	1.969	94.07	-0.0424	-0.856
34	2014-04-21	1.917	2.902	82.66	797.8	1.923	84.76	-0.00597	-0.0848

35	2014-05-05	1.771	2.836	59.05	684.9	1.793	62.78	-0.0217	-0.484
36	2014-05-19	1.985	2.919	96.63	830.2	1.957	91.67	0.0278	0.484
37	2014-06-02	1.828	2.834	67.27	682.3	1.79	62.33	0.0381	0.532
38	2014-06-30	1.533	2.748	34.16	559.7	1.621	42.23	-0.0872	-1.13
39	2014-07-14	1.878	2.845	75.51	699.3	1.811	65.42	0.0673	1.06
40	2014-07-28	1.905	2.899	80.42	792.8	1.918	83.7	-0.0124	-0.257
41	2014-08-11	1.914	2.902	82.08	798.2	1.924	84.84	-0.0094	-0.17
42	2014-08-25	1.819	2.841	65.98	693.4	1.803	64.33	0.0159	0.345
43	2014-09-08	1.317	2.623	20.76	420.2	1.376	24.03	-0.0587	-0.986
44	2014-09-22	1.672	2.779	46.96	601.5	1.682	48.64	-0.0104	-0.213
45	2014-10-06	1.519	2.695	33.02	495.5	1.517	33.23	0.00222	0.17
46	2014-10-20	1.452	2.631	28.34	427.8	1.391	24.89	0.0613	0.856
47	2014-11-17	1.985	2.899	96.58	792.8	1.918	83.72	0.067	0.986
48	2014-12-15	1.725	2.775	53.1	595.1	1.673	47.62	0.0522	0.74
49	2015-01-12	1.931	2.883	85.25	763	1.885	77.64	0.0455	0.685
50	2015-02-09	2.055	2.981	113.5	958	2.079	121.5	-0.0246	-0.532
51	2015-03-09	2.008	2.955	102	902	2.028	107.9	-0.0196	-0.391
52	2015-03-23	1.73	2.801	53.74	633	1.726	53.78	0.00468	0.257
53	2015-04-06	2.019	2.928	104.5	848	1.975	95.56	0.0438	0.582
54	2015-04-20	1.917	2.886	82.68	769.5	1.892	78.95	0.025	0.437
55	2015-05-04	2.007	2.886	101.5	769.3	1.892	78.9	0.114	1.55
56	2015-05-18	1.057	2.512	11.41	325.1	1.157	14.51	-0.0995	-1.22
57	2015-06-07	0.6929	2.351	4.931	224.5	0.8406	7.007	-0.148	-1.93
58	2015-06-15	1.141	2.555	13.83	359	1.241	17.63	-0.101	-1.31
59	2015-06-29	1.146	2.565	13.99	366.9	1.26	18.4	-0.114	-1.71

Definitions and National Water Information System (parameter code)

Cl: Chloride in mg/L (00940)
SC: Specific conductance in uS/cm @25C (00095)