Appendix 17. Model Archival Summary for Chloride Concentration at Station 06887500; Kansas River at Wamego, 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: 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 55 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. None of the Cl samples were deemed outliers.

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 SC, 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 06887500.

Cl concentration-based model:

$$\log_{10}(Cl) = 1.82 \times \log_{10}(SC) - 3.33$$

where

Cl = chloride 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 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:

$$Chl = SC^{1.82} + 0.000472$$

Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(Cl) = 1.60 \times \log_{10}(SC) - 2.73$

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; 06887500; Kansas River at Wamego, KS

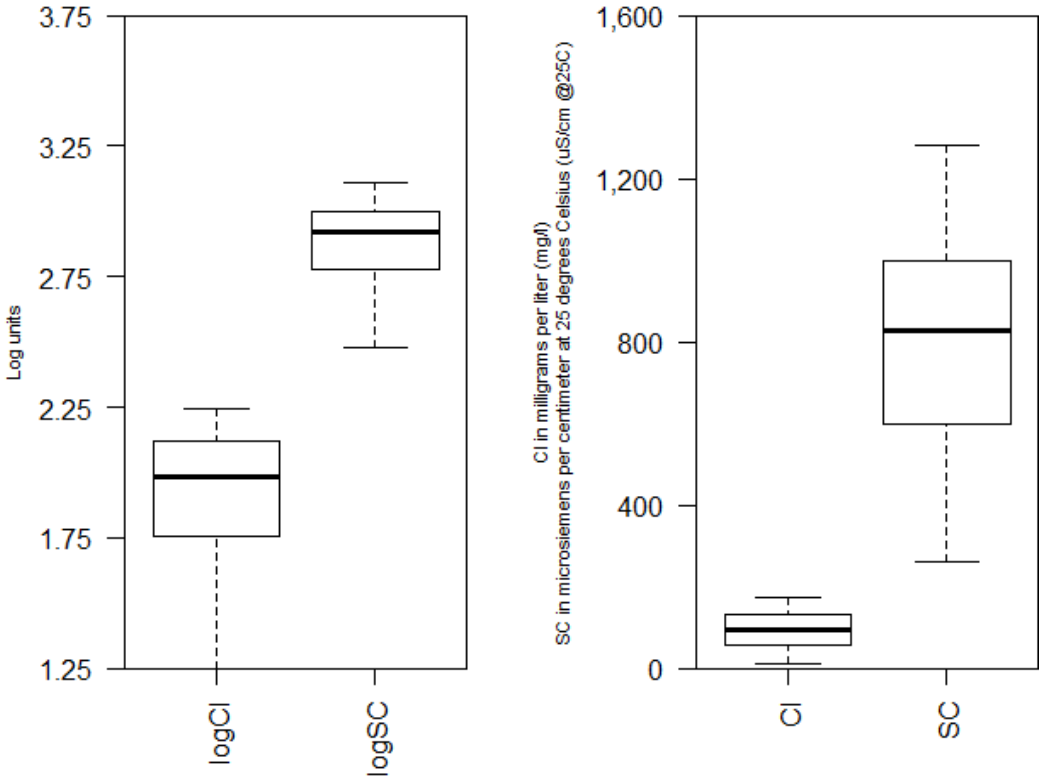
Model Form

logCl = + 1.82 * logSC + -3.33

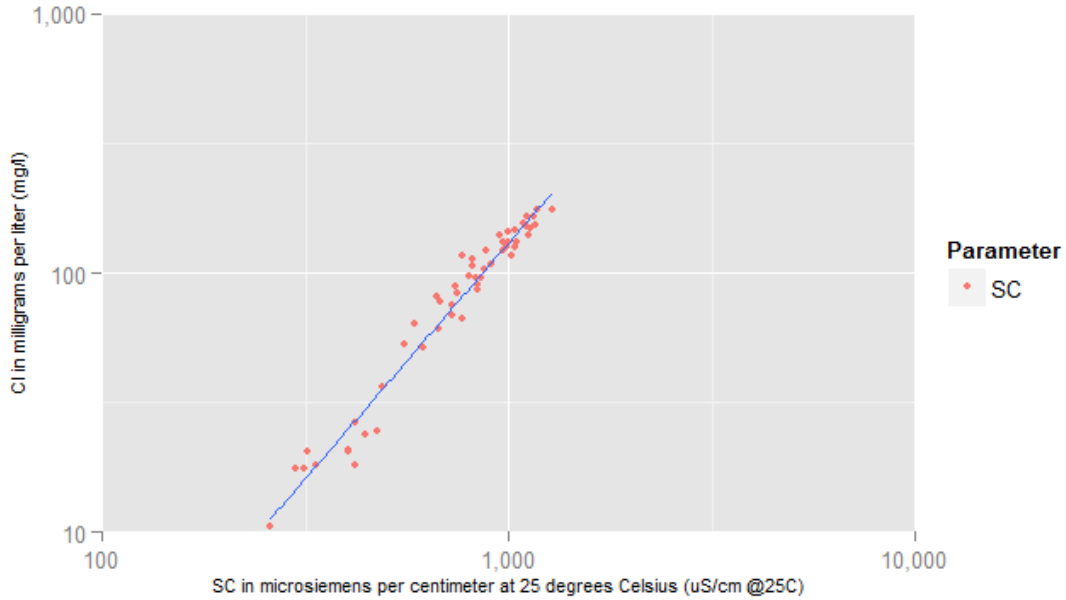
Variable Summary Statistics

	logCl	logSC	Cl	SC
Minimum	1.02	2.41	10.5	260
1st Quartile	1.73	2.77	53.2	585
Median	1.98	2.92	96.3	829
Mean	1.88	2.87	93.4	793
3rd Quartile	2.12	3.00	131.0	1000
Maximum	2.24	3.11	175.0	1280

Box Plot(s) of sample data



Exploratory Plot



Model Calibration

Basic Data

Number of Observations	55
Standard error (RMSE)	0.0646
Upper Model standard percentage error (MSPE)	16
Lower Model standard percentage error (MSPE)	13.8
Coefficient of determination (R^2)	0.963
Adjusted Coefficient of Determination (Adj. R^2)	0.962
Bias Correction Factor (BCF)	1.01

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-3.33	0.1410	-23.6	8.03e-30
logSC	1.82	0.0491	37.0	1.50e-39

Correlation Matrix

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

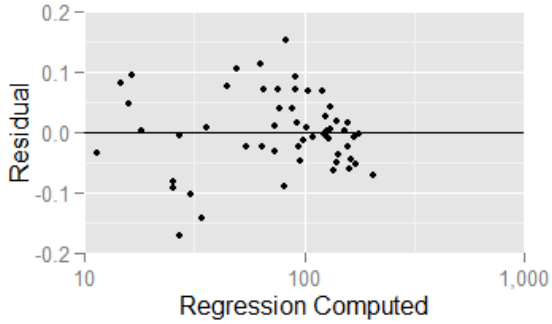
Test Criteria

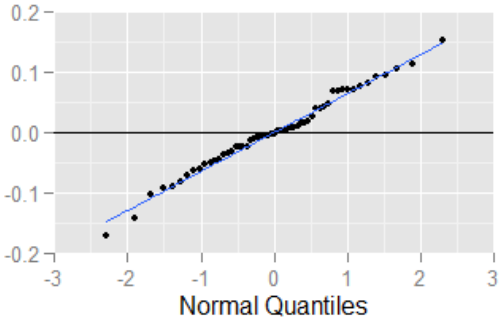
Leverage	Cook's D	DFFITS
0.05454545	0.10556635	0.26967994

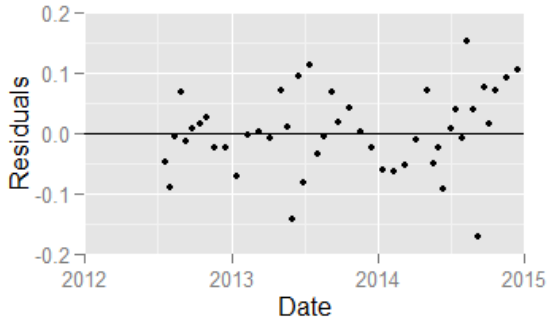
Flagged Observations

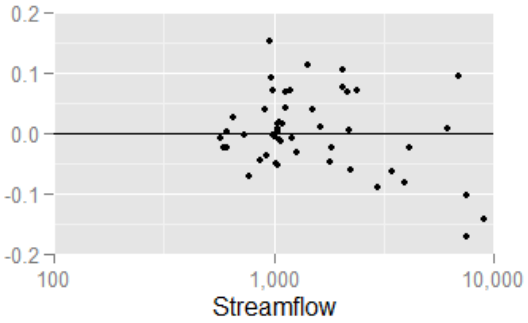
	logCl	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
06/03/2013 8:50	1.391	1.532	-0.14130	-2.23000	-2.32000	0.03887	1.006e-01	-0.46660
06/17/2013 8:30	1.312	1.219	0.09279	1.50800	1.52700	0.09374	1.177e-01	0.49120
07/01/2013 8:10	1.320	1.402	-0.08167	-1.30200	-1.31100	0.05752	5.172e-02	-0.32380
08/05/2013 7:30	1.020	1.054	-0.03360	-0.55950	-0.55590	0.13650	2.474e-02	-0.22100
06/11/2014 9:00	1.309	1.402	-0.09291	-1.48100	-1.49800	0.05752	6.694e-02	-0.37020
08/11/2014 9:20	2.064	1.914	0.15080	2.35500	2.46500	0.01843	5.204e-02	0.33770
09/08/2014 8:00	1.261	1.433	-0.17200	-2.73400	-2.92200	0.05253	2.072e-01	-0.68810
05/18/2015 10:30	1.259	1.256	0.00264	0.04272	0.04232	0.08533	8.513e-05	0.01292
06/01/2015 8:00	1.375	1.479	-0.10420	-1.65100	-1.67900	0.04577	6.538e-02	-0.36780
06/15/2015 7:50	1.248	1.201	0.04685	0.76330	0.76020	0.09781	3.158e-02	0.25030
06/29/2015 8:20	1.248	1.167	0.08100	1.32600	1.33500	0.10610	1.043e-01	0.46010

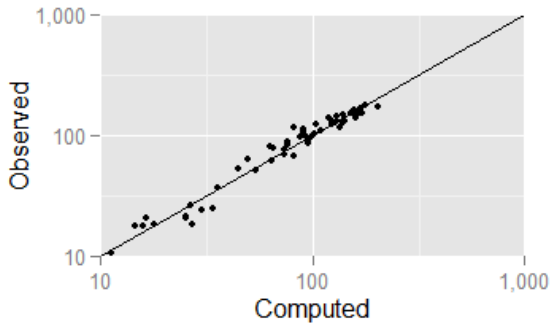
Statistical Plots

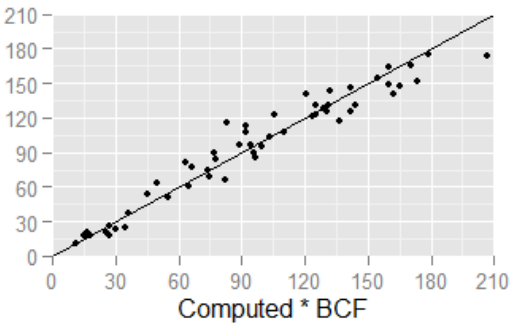


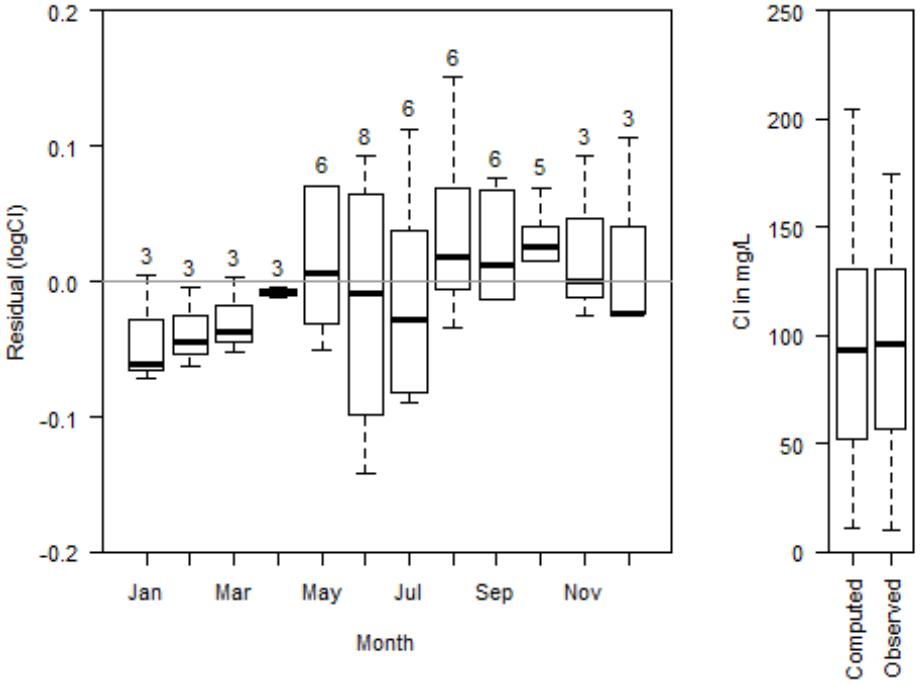


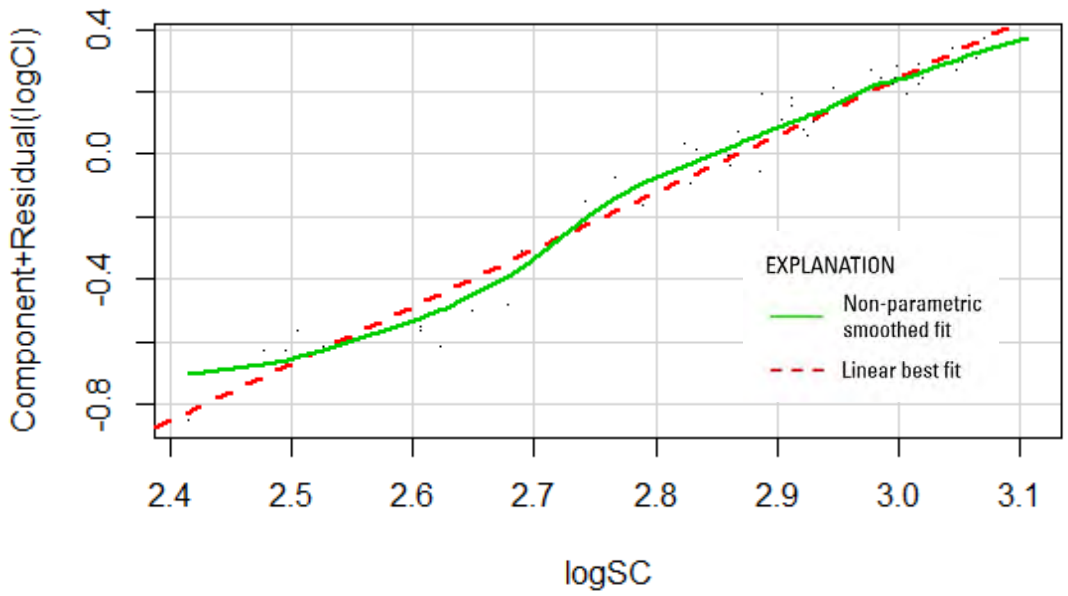









Residual (logCl)
0.2
0.1
0.0
-0.1
-0.2
3
3
3
3
6
8
6
6
6
5
3
3
Jan
Mar
May
Jul
Sep
Nov
Month
Cl in mg/L
250
200
150
100
50
0
Computed
Observed


Component+Residual(logCl)
0.4
0.0
-0.4
-0.8
2.4
2.5
2.6
2.7
2.8
2.9
3.0
3.1
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	0.06462	96.28	96.21	17.5	0.24	1	± 15
logCl ~ SC	1	0.09895	91.27	91.1	109.6	0.5674	1	± 23
logCl ~ logQ	1	0.1356	83.61	83.3	250.4	1.046	1	± 32
logCl ~ logQ + logSC	2	0.05948	96.9	96.79	7.935	0.211	5.074	± 14
logCl ~ SC + logSC	2	0.06383	96.44	96.3	16.58	0.2395	27.19	± 15
logCl ~ Q + logSC	2	0.0649	96.31	96.17	18.79	0.2663	2.882	± 15
logCl ~ Q + logQ + logSC	3	0.0582	97.09	96.92	6.458	0.2107	5.104	± 13
logCl ~ logQ + SC + logSC	3	0.05994	96.92	96.74	9.707	0.2182	5.89	± 14
logCl ~ Q + SC + logSC	3	0.0644	96.44	96.23	18.47	0.2786	4.885	± 15
logCl ~ Q + logQ + SC + logSC	4	0.05684	97.28	97.06	5	0.218	7.509	± 13

Data

	Date	logCl	logSC	Cl	SC	Computed logCl	Computed Cl	Residual	Normal Quantiles
1	2012-07-19	1.933	2.925	85.68	842	1.981	96.77	-0.0482	-0.806
2	2012-07-30	1.821	2.886	66.28	770	1.911	82.26	-0.0892	-1.38
3	2012-08-13	2.088	2.987	122.4	970	2.093	125.1	-0.00513	-0.137
4	2012-08-27	2.088	2.946	122.4	884	2.02	105.7	0.0683	0.87
5	2012-09-10	1.979	2.931	95.21	854	1.992	99.29	-0.0136	-0.323
6	2012-09-24	2.016	2.941	103.7	872.3	2.009	103.2	0.00693	0.229
7	2012-10-15	1.984	2.918	96.28	828.7	1.969	94	0.015	0.421
8	2012-10-29	2.119	2.987	131.5	970.5	2.093	125.2	0.0259	0.522
9	2012-11-19	1.952	2.923	89.45	837	1.976	95.73	-0.0249	-0.471
10	2012-12-17	2.175	3.045	149.5	1110	2.199	159.9	-0.0244	-0.421
11	2013-01-14	2.24	3.107	173.7	1280	2.312	207.1	-0.0717	-1.18
12	2013-02-11	2.243	3.072	175	1180	2.247	178.6	-0.00423	-0.0454
13	2013-03-11	2.187	3.037	153.9	1090	2.185	154.7	0.00252	0.0909
14	2013-04-08	2.219	3.061	165.6	1151	2.228	170.7	-0.00877	-0.229
15	2013-05-06	1.886	2.834	76.92	682	1.815	65.99	0.0711	1.09
16	2013-05-20	1.873	2.861	74.64	726	1.864	73.93	0.00877	0.323
17	2013-06-03	1.391	2.678	24.6	476.6	1.532	34.42	-0.141	-1.9
18	2013-06-17	1.312	2.506	20.49	320.3	1.219	16.72	0.0928	1.51
19	2013-07-01	1.32	2.606	20.9	404	1.402	25.49	-0.0817	-1.27
20	2013-07-15	1.911	2.824	81.43	667.5	1.798	63.46	0.113	1.9
21	2013-08-05	1.02	2.415	10.47	259.8	1.054	11.43	-0.0336	-0.63
22	2013-08-19	1.424	2.621	26.55	418	1.429	27.12	-0.00469	-0.0909
23	2013-09-09	2.146	2.979	140	952	2.078	120.9	0.0683	0.806
24	2013-09-23	2.165	3.017	146.2	1040	2.148	142	0.0173	0.471
25	2013-10-21	2.157	3	143.7	1000	2.117	132.3	0.0407	0.686
26	2013-11-18	2.107	2.994	128.1	987.2	2.107	129.2	0.000837	0.0454
27	2013-12-16	1.712	2.79	51.5	616.3	1.735	54.9	-0.0232	-0.372
28	2014-01-13	2.146	3.049	139.9	1120	2.206	162.5	-0.0603	-1.01
29	2014-02-10	2.067	3.007	116.8	1017	2.13	136.4	-0.0627	-1.09
30	2014-03-10	2.182	3.064	152	1160	2.234	173.2	-0.052	-0.939
31	2014-04-07	2.099	2.996	125.5	991.4	2.11	130.2	-0.0112	-0.276
32	2014-05-05	2.029	2.913	106.9	818.7	1.959	91.96	0.0701	1.01
33	2014-05-19	2.097	3.017	125.1	1040	2.148	142	-0.0505	-0.87
34	2014-06-02	1.781	2.829	60.46	675.1	1.807	64.79	-0.0254	-0.522

35	2014-06-11	1.309	2.606	20.36	404	1.402	25.49	-0.0929	-1.51
36	2014-06-30	1.562	2.69	36.44	490.3	1.555	36.23	0.00705	0.276
37	2014-07-14	1.922	2.872	83.54	744.4	1.884	77.36	0.0379	0.575
38	2014-07-28	2.031	2.957	107.3	905	2.038	110.3	-0.00733	-0.183
39	2014-08-11	2.064	2.888	116	773	1.914	82.85	0.151	2.29
40	2014-08-25	1.984	2.905	96.33	803	1.944	88.78	0.0401	0.63
41	2014-09-08	1.261	2.624	18.24	420.2	1.433	27.38	-0.172	-2.29
42	2014-09-22	1.726	2.743	53.21	553.3	1.65	45.14	0.0761	1.18
43	2014-10-06	2.214	3.045	163.6	1110	2.199	159.9	0.0148	0.372
44	2014-10-20	1.949	2.869	88.9	740	1.879	76.54	0.0696	0.939
45	2014-11-17	2.051	2.913	112.6	818.5	1.959	91.92	0.0927	1.38
46	2014-12-15	1.8	2.767	63.1	585.2	1.694	49.98	0.106	1.67
47	2015-01-12	2.119	2.999	131.5	996.7	2.114	131.5	0.00467	0.183
48	2015-02-09	2.168	3.053	147.2	1130	2.213	165.1	-0.0453	-0.745
49	2015-03-09	2.116	3.02	130.7	1048	2.154	144	-0.0373	-0.686
50	2015-04-06	2.085	2.985	121.6	965.5	2.089	124.1	-0.00406	0
51	2015-05-04	1.835	2.862	68.4	728	1.866	74.3	-0.0313	-0.575
52	2015-05-18	1.259	2.526	18.16	336	1.256	18.24	0.00264	0.137
53	2015-06-01	1.375	2.649	23.7	445.5	1.479	30.45	-0.104	-1.67
54	2015-06-15	1.248	2.496	17.7	313.3	1.201	16.06	0.0468	0.745
55	2015-06-29	1.248	2.477	17.7	300	1.167	14.85	0.081	1.27

Definitions and National Water Information System (parameter code)

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