Appendix 9. Model Archival Summary for Calcium Concentration at Station 06887500; Kansas River at Wamego, Kansas

This model archival summary summarizes the Calcium concentration (Ca) model developed to compute 15-minute Ca 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 Ca 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 Ca samples were deemed outliers.

Calcium 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 or D-95, depth integrating sampler is used from the downstream side of the bridge. Samples are analyzed for Ca 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 Ca concentration. A variety of models that predict Ca, $(Ca)^2$, $\sqrt{Ca}$ and models that predict $\log_{10}(Ca)$ 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 Ca 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 $\log_{10}(Ca)$.

SC was selected as the best predictor of Ca 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 Ca concentration at site number 06887500.

Ca concentration-based model:

$$\log_{10}(Ca) = 0.646 \times \log_{10}(SC) - 0.0258$$

where

Ca = Calcium 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 Ca.

The log-transformed model may be retransformed to the original units so that Ca 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:

$Ca = SC^{0.646} + 0.942$

Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(Ca) = 0.787 \times \log_{10}(SC) - 0.465$

Calcium Concentration Record

The Ca 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 Calcium; 06887500; Kansas River at Wamego, KS
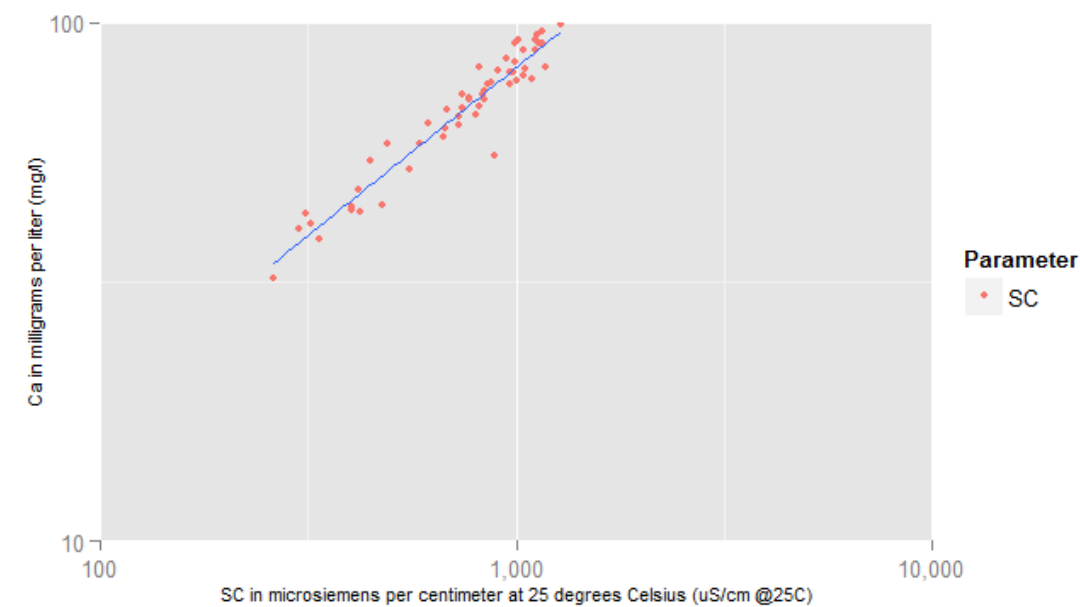
Model Form

logCa = + 0.646 * logSC + -0.0258

Variable Summary Statistics

	logCa	logSC	Ca	SC
Minimum	1.51	2.41	32.1	260
1st Quartile	1.77	2.77	58.3	585
Median	1.86	2.92	72.6	829
Mean	1.83	2.87	69.7	793
3rd Quartile	1.91	3.00	82.1	1000
Maximum	2.00	3.11	99.0	1280

Box Plot(s) of sample data
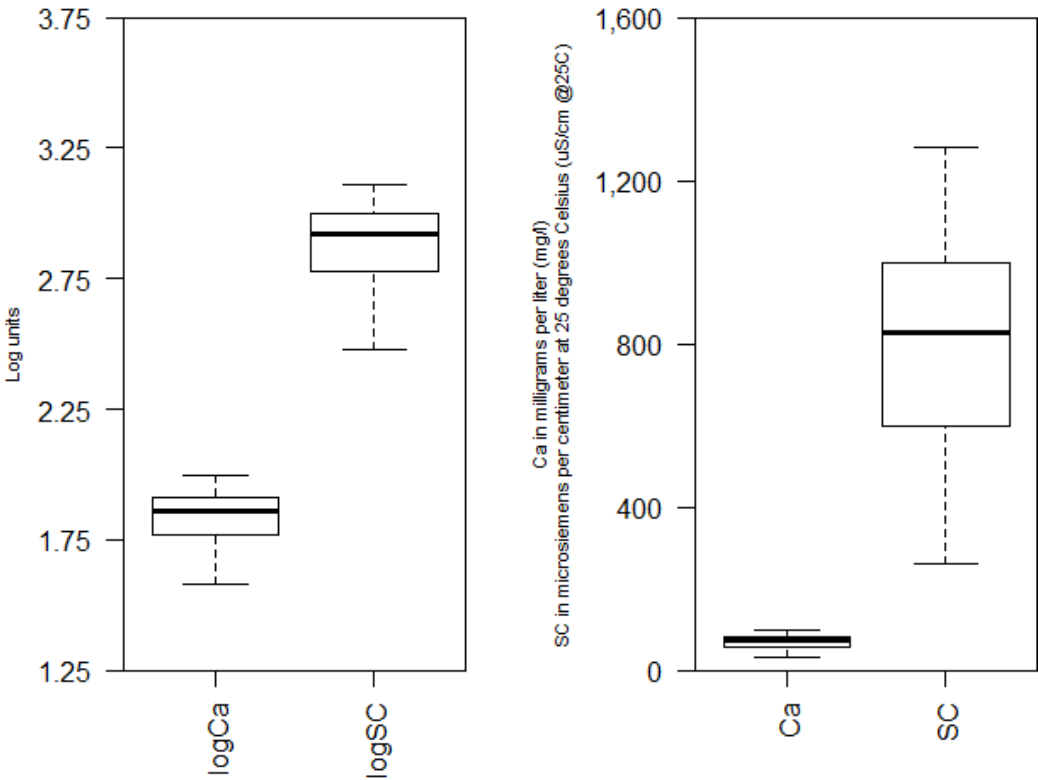
Exploratory Plot

Model Calibration

Basic Data

```
Number of Observations                          55
Standard error (RMSE)                       0.0334
Upper Model standard percentage error (MSPE)  7.98
Lower Model standard percentage error (MSPE)  7.39
Coefficient of determination (R²)            0.925
Adjusted Coefficient of Determination (Adj. R²)  0.923
Bias Correction Factor (BCF)                     1
```

Explanatory Variables

	Coefficients	Standard Error	t value	Pr(>|t|)
(Intercept)	-0.0258	0.0728	-0.354	7.24e-01
logSC	0.6460	0.0253	25.500	1.90e-31

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

|	logCa	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
8/27/2012 11:20	1.744	1.879	-0.13550	-4.1060	-4.9260	0.02177	0.187600	-0.7348
2/11/2013 8:20	1.914	1.960	-0.04560	-1.3970	-1.4100	0.04225	0.043060	-0.2962
3/11/2013 9:00	1.891	1.938	-0.04663	-1.4230	-1.4370	0.03482	0.036530	-0.2730
6/3/2013 8:50	1.646	1.706	-0.05932	-1.8140	-1.8550	0.03887	0.066550	-0.3731
6/17/2013 8:30	1.612	1.594	0.01849	0.5825	0.5788	0.09374	0.017540	0.1861
7/1/2013 8:10	1.638	1.659	-0.02084	-0.6435	-0.6399	0.05752	0.012640	-0.1581
8/5/2013 7:30	1.506	1.535	-0.02921	-0.9424	-0.9414	0.13650	0.070190	-0.3743
6/11/2014 9:00	1.646	1.659	-0.01321	-0.4080	-0.4048	0.05752	0.005081	-0.1000
6/30/2014 8:30	1.767	1.713	0.05361	1.6370	1.6640	0.03627	0.050450	0.3229
5/18/2015 10:30	1.581	1.607	-0.02600	-0.8151	-0.8125	0.08533	0.030990	-0.2482
6/1/2015 8:00	1.732	1.687	0.04509	1.3840	1.3960	0.04577	0.045940	0.3058
6/15/2015 7:50	1.631	1.588	0.04349	1.3730	1.3850	0.09781	0.102200	0.4559
6/29/2015 8:20	1.602	1.576	0.02693	0.8540	0.8518	0.10610	0.043280	0.2934

Statistical Plots
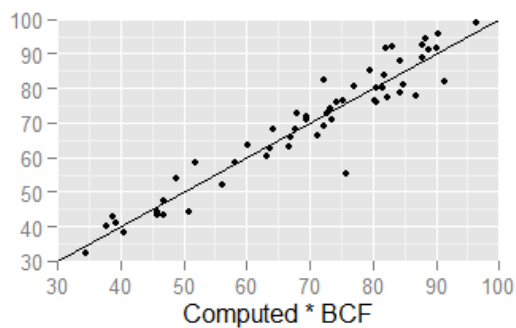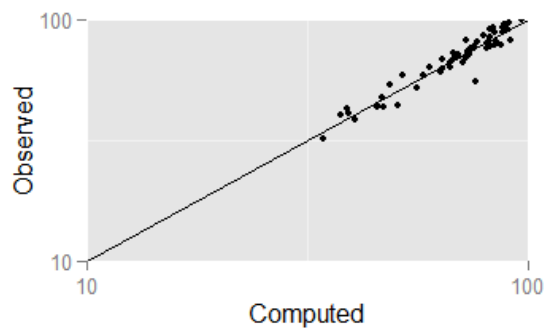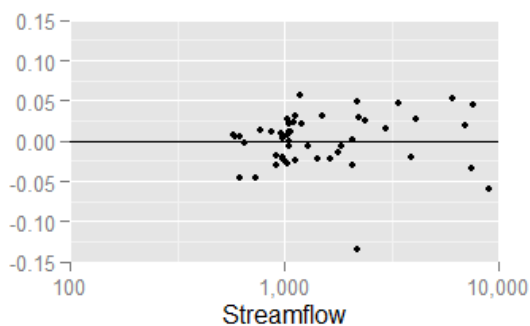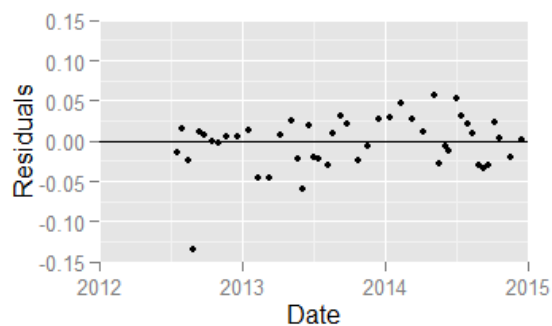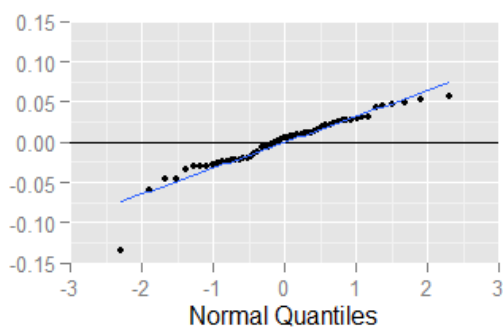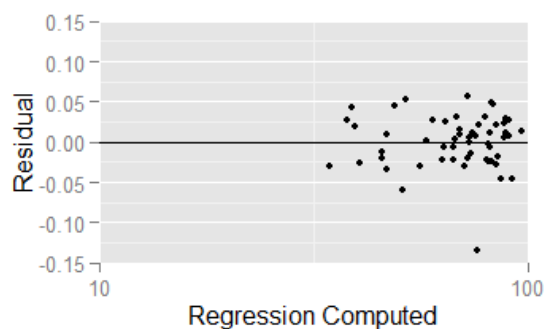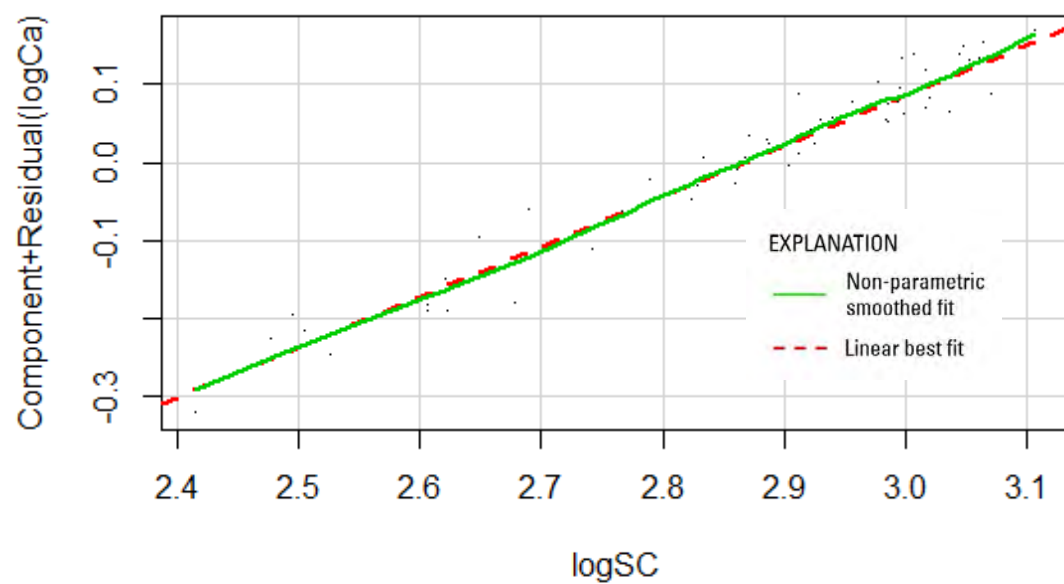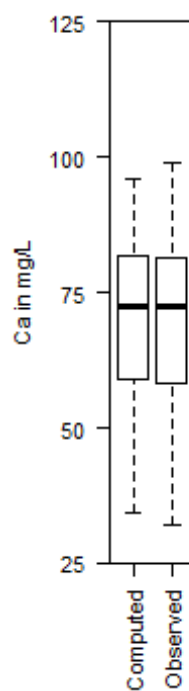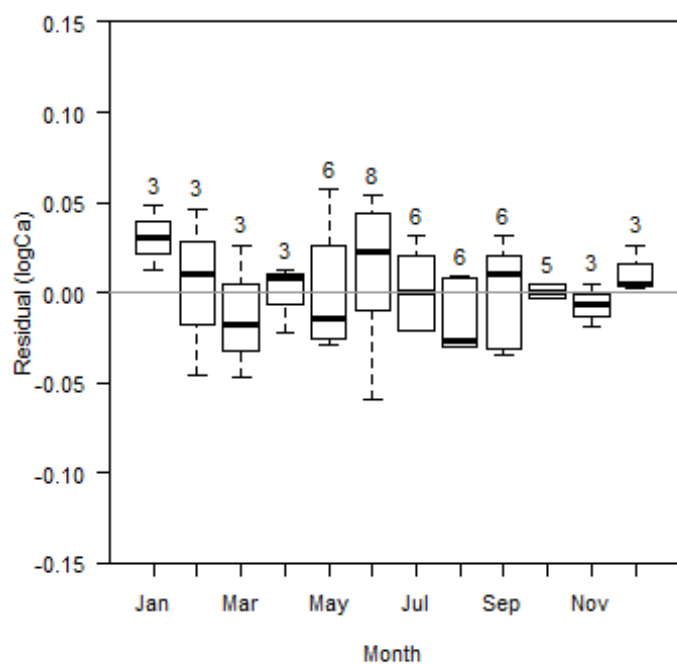
Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logCa ~ logSC	1	0.03335	92.48	92.34	-0.5297	0.06333	1	± 7.7
logCa ~ SC	1	0.0394	89.5	89.31	19.42	0.08941	1	± 9.1
logCa ~ logQ	1	0.06563	70.87	70.32	144.4	0.243	1	± 15
logCa ~ logQ + logSC	2	0.03322	92.68	92.4	0.1308	0.06408	5.074	± 7.7
logCa ~ logSC + sin2DY	2	0.03343	92.58	92.3	0.7594	0.06491	1.001	± 7.7
logCa ~ logSC + cos2DY	2	0.03344	92.58	92.29	0.7851	0.06451	1.527	± 7.7
logCa ~ logQ + logSC + cos2DY	3	0.03337	92.75	92.33	1.613	0.0654	5.122	± 7.7
logCa ~ logQ + logSC + sin2DY	3	0.03344	92.72	92.3	1.812	0.06694	5.432	± 7.7
logCa ~ logSC + sin2DY + cos2DY	3	0.0335	92.7	92.27	1.993	0.06616	1.532	± 7.7
logCa ~ logQ + logSC + sin2DY + cos2DY	4	0.03356	92.81	92.24	3.221	0.06854	5.505	± 7.7
logCa ~ logQ + SC + logSC + cos2DY	4	0.03366	92.77	92.19	3.507	0.06765	5.89	± 7.8
logCa ~ Q + logQ + logSC + cos2DY	4	0.03368	92.76	92.18	3.555	0.06875	5.142	± 7.8
logCa ~ logQ + SC + logSC + sin2DY + cos2DY	5	0.03383	92.85	92.12	5.001	0.07038	6.087	± 7.8
logCa ~ Q + logQ + logSC + sin2DY + cos2DY	5	0.03388	92.82	92.09	5.138	0.07194	5.168	± 7.8
logCa ~ Q + logQ + SC + logSC + cos2DY	5	0.034	92.77	92.03	5.502	0.07092	7.546	± 7.8
logCa ~ Q + logQ + SC + logSC + sin2DY + cos2DY	6	0.03418	92.85	91.95	7	0.07379	7.574	± 7.9

Data

	Date	logCa	logSC	Ca	SC	Computed logCa	Computed Ca	Residual	Normal Quantiles
1	2012-07-19	1.851	2.925	70.91	842	1.865	73.54	-0.0146	-0.421
2	2012-07-30	1.855	2.886	71.65	770	1.84	69.41	0.015	0.471
3	2012-08-13	1.88	2.987	75.87	970	1.905	80.58	-0.025	-0.806
4	2012-08-27	1.744	2.946	55.4	884	1.879	75.89	-0.135	-2.29
5	2012-09-10	1.881	2.931	76.04	854	1.869	74.21	0.0118	0.323
6	2012-09-24	1.884	2.941	76.48	872.3	1.875	75.24	0.00831	0.137
7	2012-10-15	1.861	2.918	72.57	828.7	1.861	72.78	-6.3e-05	-0.137
8	2012-10-29	1.902	2.987	79.85	970.5	1.905	80.61	-0.0029	-0.183
9	2012-11-19	1.869	2.923	73.92	837	1.864	73.25	0.00513	0.0454
10	2012-12-17	1.948	3.045	88.7	1110	1.943	87.92	0.00504	0
11	2013-01-14	1.996	3.107	98.98	1280	1.983	96.4	0.0127	0.421
12	2013-02-11	1.914	3.072	82.12	1180	1.96	91.46	-0.0456	-1.51
13	2013-03-11	1.891	3.037	77.83	1090	1.938	86.89	-0.0466	-1.67
14	2013-04-08	1.961	3.061	91.46	1151	1.953	90	0.00817	0.0909
15	2013-05-06	1.832	2.834	67.96	682	1.806	64.17	0.0261	0.745
16	2013-05-20	1.801	2.861	63.25	726	1.824	66.82	-0.0226	-0.745
17	2013-06-03	1.646	2.678	44.28	476.6	1.706	50.9	-0.0593	-1.9
18	2013-06-17	1.612	2.506	40.97	320.3	1.594	39.37	0.0185	0.522
19	2013-07-01	1.638	2.606	43.48	404	1.659	45.74	-0.0208	-0.575
20	2013-07-15	1.779	2.824	60.11	667.5	1.8	63.28	-0.0212	-0.63
21	2013-08-05	1.506	2.415	32.06	259.8	1.535	34.39	-0.0292	-1.09
22	2013-08-19	1.677	2.621	47.56	418	1.669	46.76	0.00855	0.183
23	2013-09-09	1.931	2.979	85.34	952	1.9	79.61	0.0314	1.18
24	2013-09-23	1.945	3.017	88.1	1040	1.925	84.29	0.0204	0.575
25	2013-10-21	1.889	3	77.36	1000	1.914	82.18	-0.0251	-0.87
26	2013-11-18	1.904	2.994	80.08	987.2	1.91	81.5	-0.00643	-0.276
27	2013-12-16	1.804	2.79	63.69	616.3	1.778	60.1	0.0264	0.806
28	2014-01-13	1.975	3.049	94.49	1120	1.945	88.43	0.03	1.01
29	2014-02-10	1.965	3.007	92.23	1017	1.918	83.08	0.0466	1.51
30	2014-03-10	1.982	3.064	95.88	1160	1.955	90.46	0.0265	0.87
31	2014-04-07	1.923	2.996	83.81	991.4	1.911	81.72	0.0121	0.372

```
32 2014-05-05 1.915 2.913 82.19 818.7     1.857    72.21   0.0574      2.29
33 2014-05-19 1.896 3.017 78.68  1040     1.925    84.29  -0.0287     -1.01
34 2014-06-02 1.797 2.829 62.62 675.1     1.803    63.75 -0.00658    -0.323
35 2014-06-11 1.646 2.606 44.25   404     1.659    45.74  -0.0132    -0.372
36 2014-06-30 1.767  2.69 58.49 490.3     1.713    51.84   0.0536       1.9
37 2014-07-14 1.862 2.872 72.74 744.4     1.831    67.91   0.0311      1.09
38 2014-07-28 1.906 2.957 80.54   905     1.886    77.05   0.0205      0.63
39 2014-08-11  1.85 2.888 70.79   773     1.841    69.58  0.00868     0.229
40 2014-08-25 1.822 2.905 66.38   803     1.852    71.32  -0.0299     -1.18
41 2014-09-08 1.636 2.624 43.24 420.2      1.67    46.92  -0.0343     -1.38
42 2014-09-22 1.717 2.743 52.08 553.3     1.747    56.06  -0.0308     -1.27
43 2014-10-06 1.967 3.045 92.61  1110     1.943    87.92   0.0238     0.686
44 2014-10-20 1.833 2.869 68.14   740     1.829    67.65  0.00436   -0.0454
45 2014-11-17 1.838 2.913 68.89 818.5     1.857     72.2  -0.0192    -0.522
46 2014-12-15 1.766 2.767  58.3 585.2     1.763    58.13  0.00249   -0.0909
47 2015-01-12 1.961 2.999 91.41 996.7     1.913    82.01   0.0483      1.67
48 2015-02-09 1.959 3.053 90.93  1130     1.948    88.94   0.0108     0.276
49 2015-03-09 1.909  3.02 81.08  1048     1.927    84.71  -0.0178    -0.471
50 2015-04-06 1.882 2.985  76.2 965.5     1.904    80.34  -0.0218    -0.686
51 2015-05-04 1.819 2.862 65.86   728     1.824    66.94 -0.00583    -0.229
52 2015-05-18 1.581 2.526 38.14   336     1.607    40.61   -0.026    -0.939
53 2015-06-01 1.732 2.649 53.91 445.5     1.687    48.73   0.0451      1.38
54 2015-06-15 1.631 2.496 42.78 313.3     1.588    38.81   0.0435      1.27
55 2015-06-29 1.602 2.477 40.04   300     1.576    37.74   0.0269     0.939
```

Definitions and National Water Information System (parameter code)

Ca: Calcium in mg/L (00915)
SC: Specific conductance in uS/cm @25C (00095)