Appendix 14. Model Archival Summary for Sodium Concentration at Station 06892350; Kansas River at De Soto, Kansas

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

Sodium 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 Na 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 Na concentration. A variety of models that predict Na, $(Na)^2$, $\sqrt{Na}$ and models that predict $\log_{10}(Na)$ 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 Na 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 $\log_{10}$(Na).

Specific conductance was selected as the best predictor of Na 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 Na concentration at site number 06892350.

Na concentration-based model:

$$\log_{10}(Na) = 1.72 \times \log_{10}(SC) - 3.13$$

where

> *Na* = sodium 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 Na.

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

$$Na = SC^{1.72} + 0.000749$$

Previous Models

Start year	End year	Model
2000	2005	$\log_{10}(Na) = 1.49 \times \log_{10}(SC) - 2.51$

Sodium Concentration Record

The Na 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 Sodium; 06892350; Kansas River at De Soto, KS

Model Form
logNa = + 1.72 * logSC + -3.13

Variable Summary Statistics

	logNa	logSC	Na	SC
Minimum	0.723	2.29	5.29	197
1st Quartile	1.650	2.76	44.40	569
Median	1.830	2.89	68.30	770
Mean	1.730	2.83	63.60	726
3rd Quartile	1.930	2.95	86.00	882
Maximum	2.070	3.05	118.00	1110

Box Plot(s) of sample data
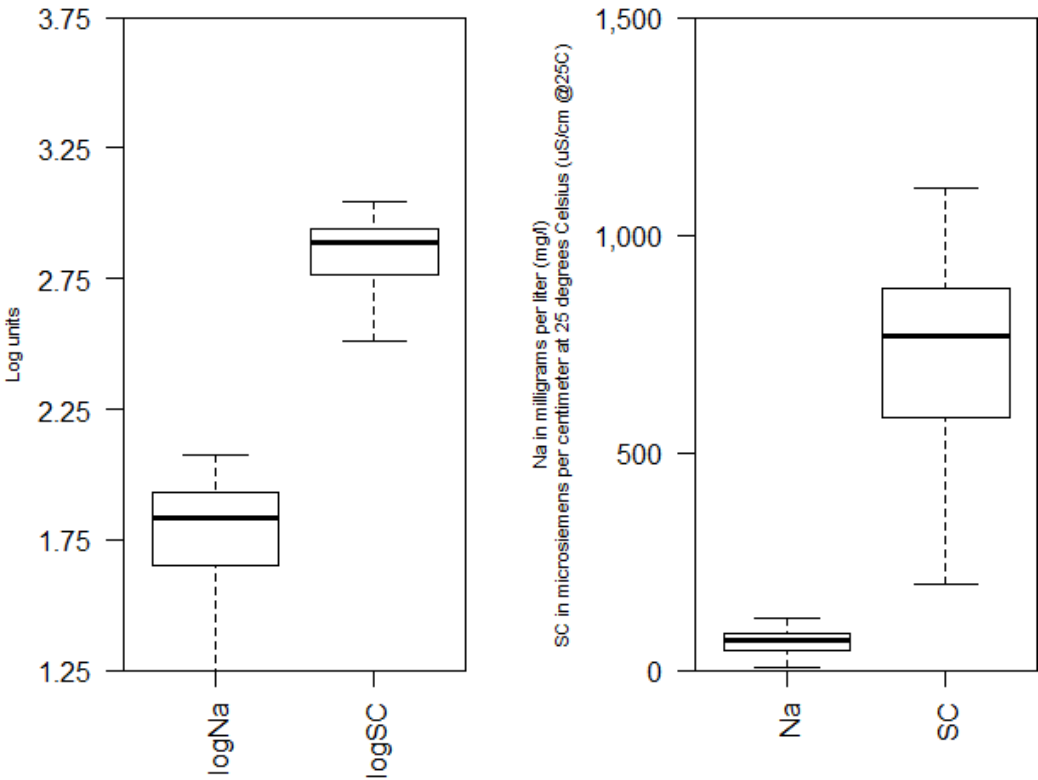
Exploratory Plot
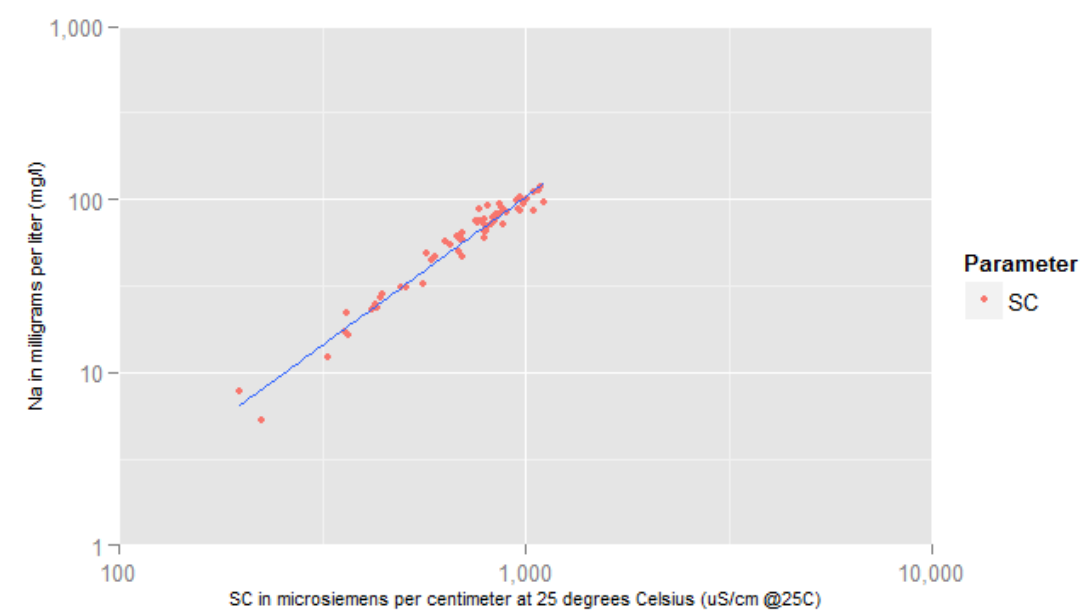
Model Calibration

Basic Data

```
Number of Observations                             59
Standard error (RMSE)                          0.0581
Upper Model standard percentage error (MSPE)     14.3
Lower Model standard percentage error (MSPE)     12.5
Coefficient of determination (R²)               0.961
Adjusted Coefficient of Determination (Adj. R²)  0.961
Bias Correction Factor (BCF)                     1.01
```

Explanatory Variables

```
            Coefficients Standard Error t value Pr(>|t|)
(Intercept)        -3.13         0.1290   -24.2 1.18e-31
logSC               1.72         0.0456    37.6 6.04e-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

|	logNa	Estimate	Residual	Standard Residual	Studentized Residual	Leverage	Cook's D	DFFITS
9/24/2012 13:30	1.9660	1.8550	0.11080	1.9270	1.9750	0.02015	0.038170	0.2832
8/5/2013 11:30	0.8880	0.8062	0.08174	1.5700	1.5910	0.19630	0.301000	0.7862
8/19/2013 10:00	1.3400	1.2630	0.07646	1.3600	1.3700	0.06289	0.062070	0.3550
1/13/2014 8:00	1.9790	2.0940	-0.11520	-2.0300	-2.0890	0.04449	0.095940	-0.4508
2/10/2014 7:40	1.9290	2.0530	-0.12360	-2.1710	-2.2470	0.03856	0.094520	-0.4500
5/4/2015 13:30	1.9390	1.8210	0.11770	2.0450	2.1060	0.01863	0.039700	0.2901
5/18/2015 15:30	1.0860	1.1790	-0.09376	-1.6840	-1.7120	0.08080	0.124600	-0.5076
6/6/2015 19:50	0.7234	0.9036	-0.18020	-3.3860	-3.7560	0.16050	1.097000	-1.6420
6/15/2015 14:50	1.2300	1.2530	-0.02362	-0.4205	-0.4175	0.06486	0.006132	-0.1099
6/29/2015 12:30	1.2190	1.2690	-0.05096	-0.9058	-0.9044	0.06167	0.026960	-0.2319

Statistical Plots
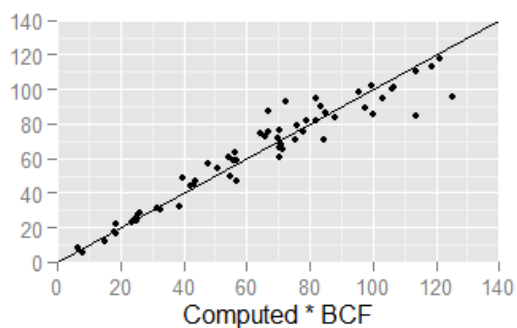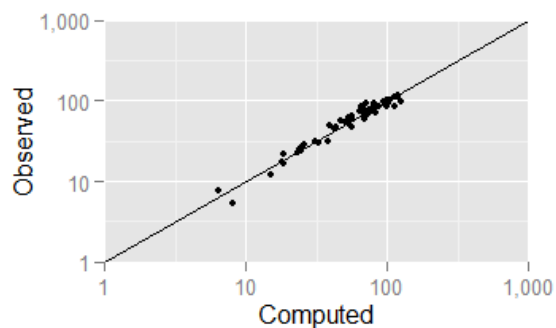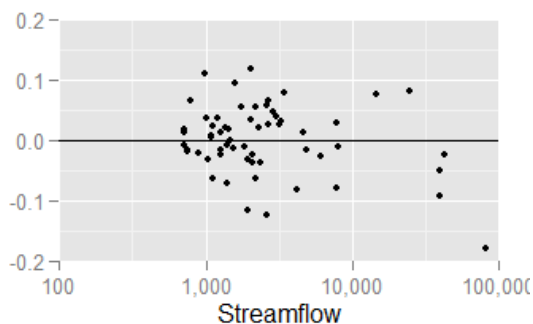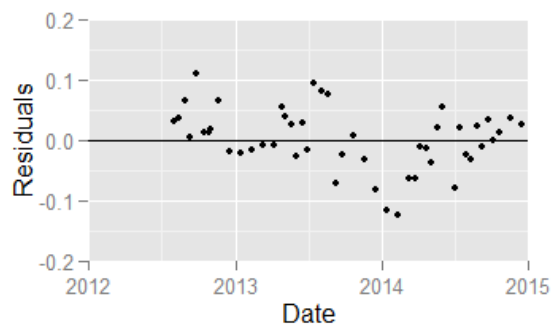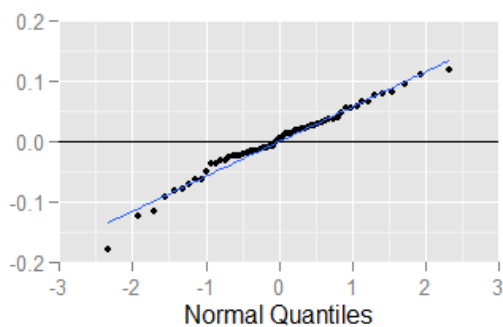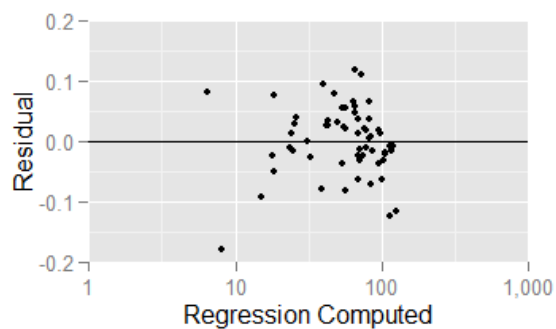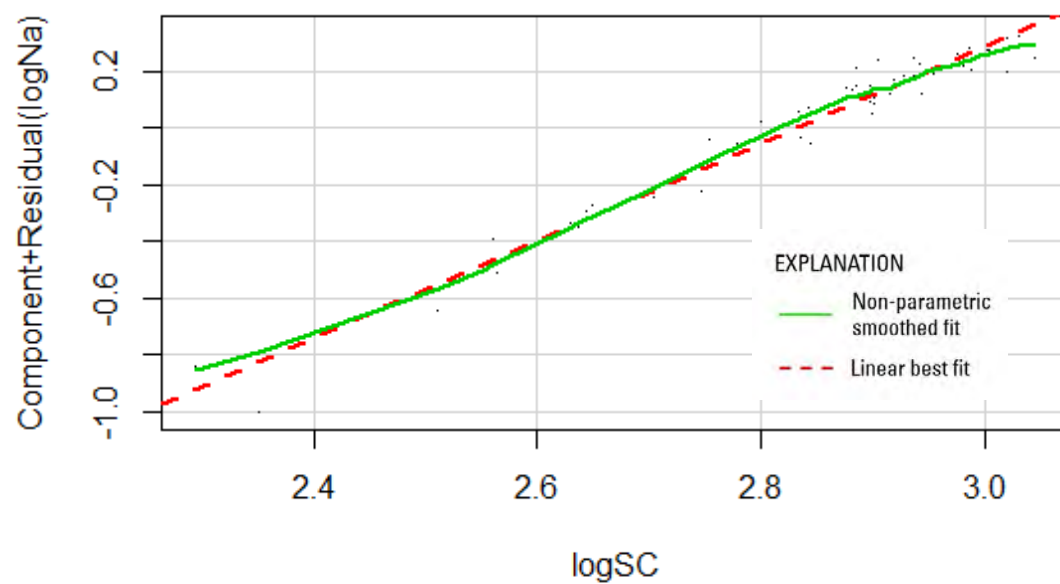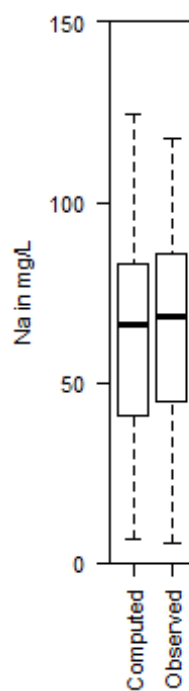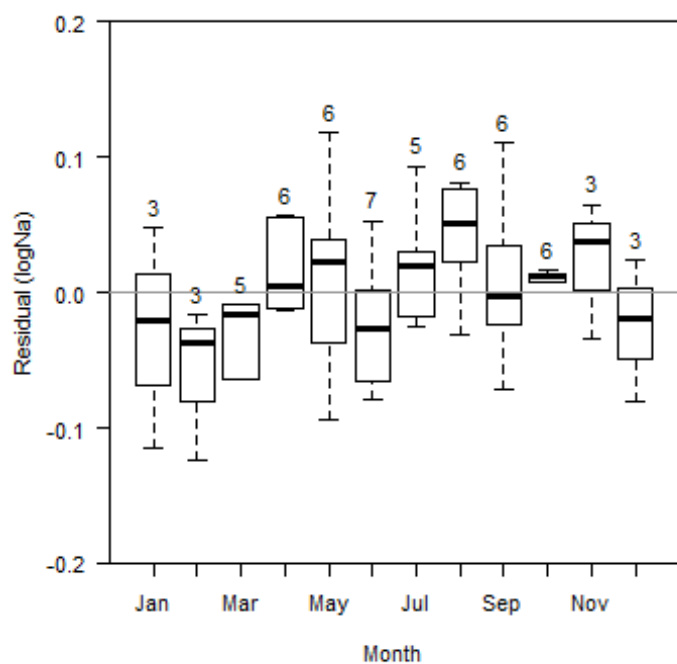
Models considered

Model Formula	Number of Variables	Standard Error	R2	Adjusted R2	Cp	PRESS	VIF	MSPE
logNa ~ logSC	1	0.05807	96.13	96.06	39.32	0.2196	1	± 13
logNa ~ SC	1	0.1015	88.17	87.96	233.2	0.6679	1	± 24
logNa ~ logQ	1	0.1192	83.7	83.41	342.2	0.8877	1	± 28
logNa ~ Q + logSC	2	0.04731	97.48	97.39	8.509	0.1364	2.089	± 11
logNa ~ logQ + logSC	2	0.04809	97.39	97.3	10.54	0.1521	4.375	± 11
logNa ~ SC + logSC	2	0.05435	96.67	96.55	28.17	0.2461	19.88	± 13
logNa ~ Q + logQ + logSC	3	0.04527	97.73	97.61	4.3	0.1476	3.472	± 10
logNa ~ logQ + SC + logSC	3	0.04613	97.64	97.52	6.412	0.1699	4.594	± 11
logNa ~ Q + SC + logSC	3	0.04757	97.49	97.36	10.06	0.1697	2.925	± 11
logNa ~ Q + logQ + SC + logSC	4	0.04515	97.78	97.62	5	0.1821	4.787	± 10

Data

	Date	logNa	logSC	Na	SC	Computed logNa	Computed Na	Residual	Normal Quantiles
1	2012-07-30	1.732	2.816	53.95	655	1.701	50.69	0.0307	0.582
2	2012-08-13	1.955	2.942	90.13	875.5	1.917	83.39	0.0374	0.74
3	2012-08-27	1.871	2.877	74.22	753.2	1.805	64.42	0.0652	1.13
4	2012-09-10	1.913	2.938	81.92	866	1.909	81.84	0.00408	0
5	2012-09-24	1.966	2.906	92.46	805.3	1.855	72.25	0.111	1.93
6	2012-10-15	1.856	2.898	71.7	791	1.842	70.06	0.0137	0.17
7	2012-10-26	2.008	2.987	101.9	971	1.995	99.59	0.0136	0.127
8	2012-10-29	1.993	2.977	98.45	948	1.977	95.58	0.0165	0.213
9	2012-11-19	1.975	2.938	94.49	867	1.91	82	0.0652	1.22
10	2012-12-17	2.005	3.004	101.2	1010	2.024	106.6	-0.0187	-0.437
11	2013-01-14	2.002	3.003	100.4	1008	2.022	106.2	-0.0205	-0.484
12	2013-02-11	2.053	3.031	113.1	1075	2.07	118.6	-0.0169	-0.345
13	2013-03-11	2.044	3.021	110.7	1050	2.053	113.9	-0.00869	-0.0848
14	2013-04-08	2.072	3.037	118	1090	2.081	121.4	-0.00881	-0.127
15	2013-04-25	1.803	2.843	63.47	697	1.748	56.4	0.055	0.986
16	2013-05-06	1.454	2.649	28.45	446	1.415	26.22	0.0391	0.797
17	2013-05-20	1.647	2.77	44.36	589.2	1.622	42.27	0.0246	0.484
18	2013-06-03	1.485	2.706	30.58	508	1.512	32.78	-0.0265	-0.685
19	2013-06-17	1.435	2.644	27.22	441	1.407	25.72	0.0283	0.532
20	2013-07-01	1.377	2.637	23.8	433.5	1.394	24.97	-0.0172	-0.391
21	2013-07-15	1.69	2.755	49	569.4	1.597	39.87	0.0933	1.71
22	2013-08-05	0.888	2.294	7.726	197	0.8062	6.455	0.0817	1.55
23	2013-08-19	1.34	2.561	21.86	363.8	1.263	18.49	0.0765	1.31
24	2013-09-09	1.851	2.945	71.01	882	1.923	84.45	-0.0716	-1.22
25	2013-09-23	1.851	2.917	70.93	826	1.874	75.47	-0.0232	-0.532
26	2013-10-21	1.934	2.947	85.98	886	1.926	85.11	0.00808	0.0424
27	2013-11-18	1.976	2.996	94.63	990.6	2.009	103.1	-0.0334	-0.797
28	2013-12-16	1.671	2.846	46.88	701	1.752	56.95	-0.0808	-1.42
29	2014-01-13	1.979	3.045	95.28	1110	2.094	125.3	-0.115	-1.71
30	2014-02-10	1.929	3.021	84.96	1050	2.053	113.9	-0.124	-1.93
31	2014-03-10	1.779	2.899	60.14	793.3	1.844	70.41	-0.0648	-1.13
32	2014-03-24	1.932	2.988	85.59	973.2	1.996	99.99	-0.0639	-1.06
33	2014-04-07	1.875	2.925	75.06	841.2	1.888	77.87	-0.0123	-0.213
34	2014-04-21	1.835	2.902	68.34	797.8	1.848	71.1	-0.0136	-0.257

```
35 2014-05-05  1.697 2.836 49.79 684.9    1.734    54.72  -0.0373    -0.919
36 2014-05-19  1.899 2.919 79.22 830.2    1.878    76.13   0.0209     0.345
37 2014-06-02  1.785 2.834 60.96 682.3    1.732    54.37   0.0533     0.919
38 2014-06-30  1.505 2.748 32.02 559.7    1.584    38.71  -0.0788     -1.31
39 2014-07-14   1.77 2.845 58.88 699.3     1.75    56.72   0.0199     0.301
40 2014-07-28  1.819 2.899 65.87 792.8    1.843    70.33  -0.0248    -0.633
41 2014-08-11  1.817 2.902 65.61 798.2    1.849    71.16  -0.0316     -0.74
42 2014-08-25  1.766 2.841 58.32 693.4    1.744     55.9   0.0221     0.391
43 2014-09-08   1.36 2.623 22.91 420.2    1.371    23.67  -0.0106     -0.17
44 2014-09-22  1.672 2.779 46.96 601.5    1.638     43.8   0.0339     0.633
45 2014-10-06  1.493 2.695 31.09 495.5    1.493    31.41 -0.000747  -0.0424
46 2014-10-20  1.396 2.631 24.88 427.8    1.384    24.41   0.0119    0.0848
47 2014-11-17  1.881 2.899 75.98 792.8    1.844    70.34   0.0372     0.685
48 2014-12-15  1.654 2.775  45.1 595.1     1.63       43   0.0244     0.437
49 2015-01-12  1.862 2.883 72.86   763    1.815    65.86   0.0475     0.856
50 2015-02-09  1.948 2.981 88.69   958    1.985    97.32  -0.0366    -0.856
51 2015-03-09  1.923 2.955  83.8   902     1.94    87.76  -0.0164    -0.301
52 2015-03-23  1.754 2.801 56.72   633    1.676    47.81   0.0779     1.42
53 2015-04-06  1.912 2.928 81.57   848    1.894    78.95   0.0179     0.257
54 2015-04-20  1.878 2.886 75.52 769.5    1.821    66.83   0.0568     1.06
55 2015-05-04  1.939 2.886 86.84 769.3    1.821    66.79    0.118     2.32
56 2015-05-18  1.086 2.512 12.18 325.1    1.179    15.24  -0.0938    -1.55
57 2015-06-07 0.7234 2.351 5.289 224.5   0.9036    8.077    -0.18    -2.32
58 2015-06-15   1.23 2.555 16.97   359    1.253    18.07  -0.0236    -0.582
59 2015-06-29  1.219 2.565 16.54 366.9    1.269    18.76   -0.051    -0.986
```

Definitions and National Water Information System (parameter code)

Na: Sodium in mg/L (00930)
SC: Specific conductance in uS/cm @25C (00095)