

Chloride Exploration

By Laura De Cicco

August 20, 2014

Contents

1	Load Data	1
2	Convert Units	2
3	Summarize Data	2

Figures

Figure 1	Boxplots of chloride	6
Figure 2	Map of all chloride values	7
Figure 3	Map of chloride values above maxVal	8

Tables

Table 1	Sites with reported values higher than the defined maxVal, sorted by maximum reported value	4
----------------	---	---

1 Load Data

The following code is used to pull raw data from the Water Quality Portal <http://www.waterqualitydata.us/>.

```
library(dataRetrieval)
library(USGSHydroTools)
```

```

parameter <- "Chloride"
state <- "US:55"

minVal <- 0
maxVal <- 1000

sitesReturned <- getWQPSites(statecode=state,
                             characteristicName=parameter)
dataReturned <- getGeneralWQPData(statecode=state,
                                   characteristicName=parameter)

```

2 Convert Units

To compare chloride values across different agencies, we need to convert reported units to a standard choice. In this report, we are using mg/L.

```

#Reported units:
unique(dataReturned$ResultMeasure.MeasureUnitCode)

[1] "mg/l"      " "      "mg/l"      "ppm"
[5] "ueq/L"     " "ueq/L" "mg/kg"     "ug/l"

if(state == "US:55"){
  # WI:
  cFactor <- c(1, 1, 1, 1, 0.001, 0.001, 1, 0.001)
  names(cFactor) <- unique(dataReturned$ResultMeasure.MeasureUnitCode)
} else {
  cFactor <- c(1, 1, 1, 1, 0.001, 0.001)
  names(cFactor) <- unique(dataReturned$ResultMeasure.MeasureUnitCode)
}

dataReturned$ResultMeasureValue <- dataReturned$ResultMeasureValue*
  cFactor[dataReturned$ResultMeasure.MeasureUnitCode]

```

3 Summarize Data

R provides many options for statistical tests.

```

Site: WIDNR_WQX-573077
One Sample t-test

```

```
data: ResultMeasureValue
t = 1, df = 16, p-value = 0.332
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
-24431 68094
sample estimates:
mean of x
21832
```

Site: USGS-040871475
One Sample t-test

```
data: ResultMeasureValue
t = 3.946, df = 93, p-value = 0.0001542
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
1107 3350
sample estimates:
mean of x
2229
```

Site: USGS-040871476
One Sample t-test

```
data: ResultMeasureValue
t = 3.378, df = 69, p-value = 0.001202
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
978.2 3799.3
sample estimates:
mean of x
2389
```

Site: USGS-05429131
One Sample t-test

```
data: ResultMeasureValue
t = 2.317, df = 40, p-value = 0.0257
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
132.5 1941.8
sample estimates:
mean of x
1037
```

Site: USGS-04087211
One Sample t-test

```

data: ResultMeasureValue
t = 1.829, df = 23, p-value = 0.08037
alternative hypothesis: true mean is not equal to 0
95 percent confidence interval:
 -136.9 2228.2
sample estimates:
mean of x
 1046

```

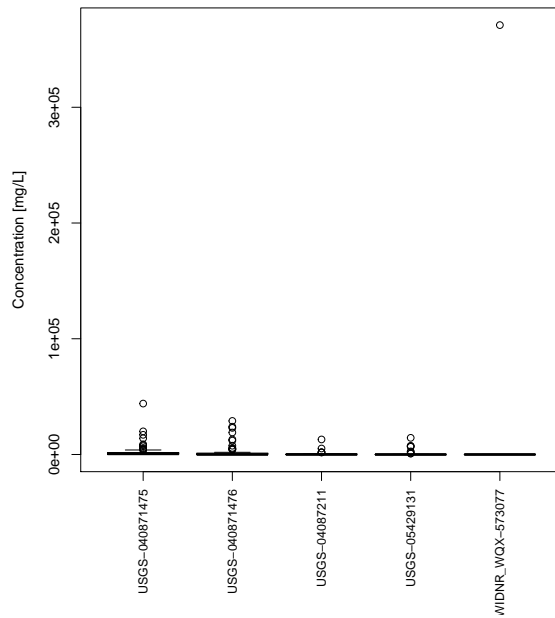
Some basic information about the data. There are 2142 NA values, and 312 values above the defined maxVal. Table 1 shows the sites that report values over maxVal.

Table 1. Sites with reported values higher than the defined maxVal, sorted by maximum reported value

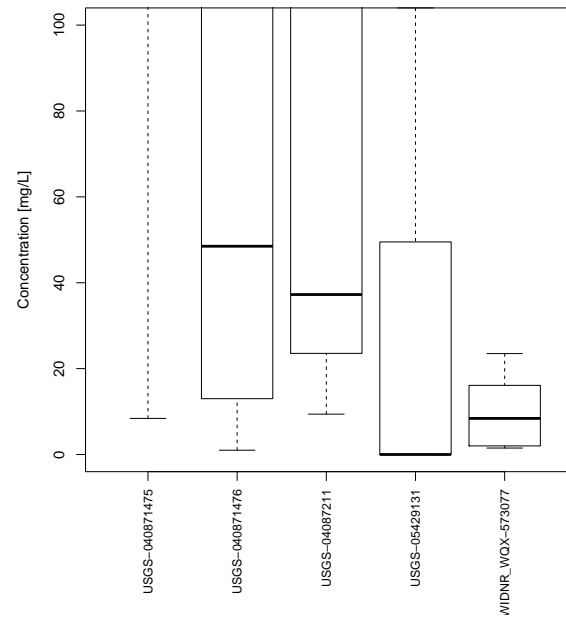
MonitoringLocationIdentifier	count	mean	max
WIDNR_WQX-573077	1	371000	371000
USGS-040871475	36	5306	44000
USGS-040871476	18	8972	29000
USGS-05429131	6	6758	14400
USGS-04087211	5	4756	13000
WIDNR_WQX-413919	5	4756	13000
USGS-05429130	5	7140	11300
USGS-430356089183501	4	3900	10000
USGS-452232092400601	6	7580	9940
USGS-04087114	5	3876	9840
WIDNR_WQX-413918	5	3876	9840
USGS-040871478	4	3730	8630
USGS-040871488	52	2810	7700
USGS-430532089315601	2	6950	7100
USGS-054064785	4	6460	7060
WIDNR_WQX-133432	4	6460	7060
USGS-04087119	13	2790	6470
USGS-04087088	10	2807	6400
USGS-434528087433001	2	4570	5100
USGS-040871482	6	2062	4930
USGS-040869416	1	4770	4770
USGS-434521087424701	10	3566	4380
USGS-040869415	17	2344	4330
USGS-04087159	8	2454	4210
USGS-462957090541001	2	4100	4150

USGS-040871602	1	4110	4110
USGS-04086946	3	2560	3560
WIDNR_WQX-413780	3	2560	3560
WIDNR_WQX-10010491	1	3473	3473
USGS-430532089315602	4	2120	3250
USGS-434333087483502	1	3140	3140
USGS-430309089260701	2	2250	3100
USGS-04087120	8	2005	2940
USGS-463631091533301	1	2940	2940
USGS-040871473	7	1729	2900
USGS-430208087543200	1	2590	2590
WIDNR_WQX-413749	6	1968	2530
USGS-430356089183502	2	2190	2510
USGS-040871472	1	2290	2290
USGS-04087070	1	1930	1930
USGS-431039087552401	1	1890	1890
USGS-040869462	5	1486	1820
USGS-430402089183501	1	1820	1820
WIDNR_WQX-413781	5	1486	1820
WIDNR_WQX-053530	2	1755	1760
USGS-04087142	3	1400	1700
USGS-441043088084101	1	1700	1700
USGS-04087214	1	1610	1610
USGS-04087204	4	1228	1480
USGS-434425090462401	1	1380	1380
USGS-040872015	1	1300	1300
USGS-462956090541101	1	1300	1300
USGS-441007088114801	2	1170	1200
WIDNR_WQX-413913	2	1135	1150
USGS-430247091085101	5	1084	1100
USGS-442955088071503	1	1100	1100
USGS-430748089181800	1	1070	1070
WIDNR_WQX-10010400	1	1016	1016
WIDNR_WQX-10034960	1	1003	1003

Figure 1 shows a box plot concentration at the sites with the top 5 reported maximum values.



(a) Maximum y scale



(b) Zoomed in to 100 mg/L

Figure 1. Boxplots of chloride

Concentration of Chloride

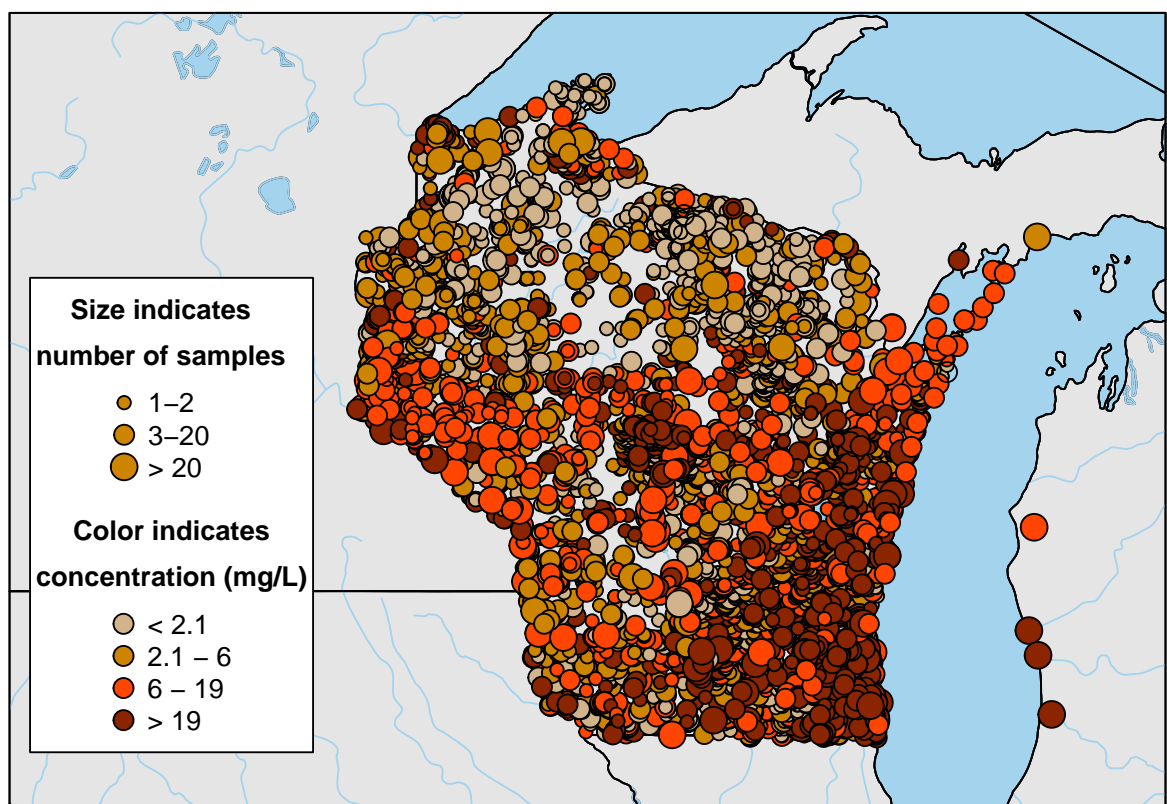


Figure 2. Map of all chloride values

Concentration outside user-defined range Chloride

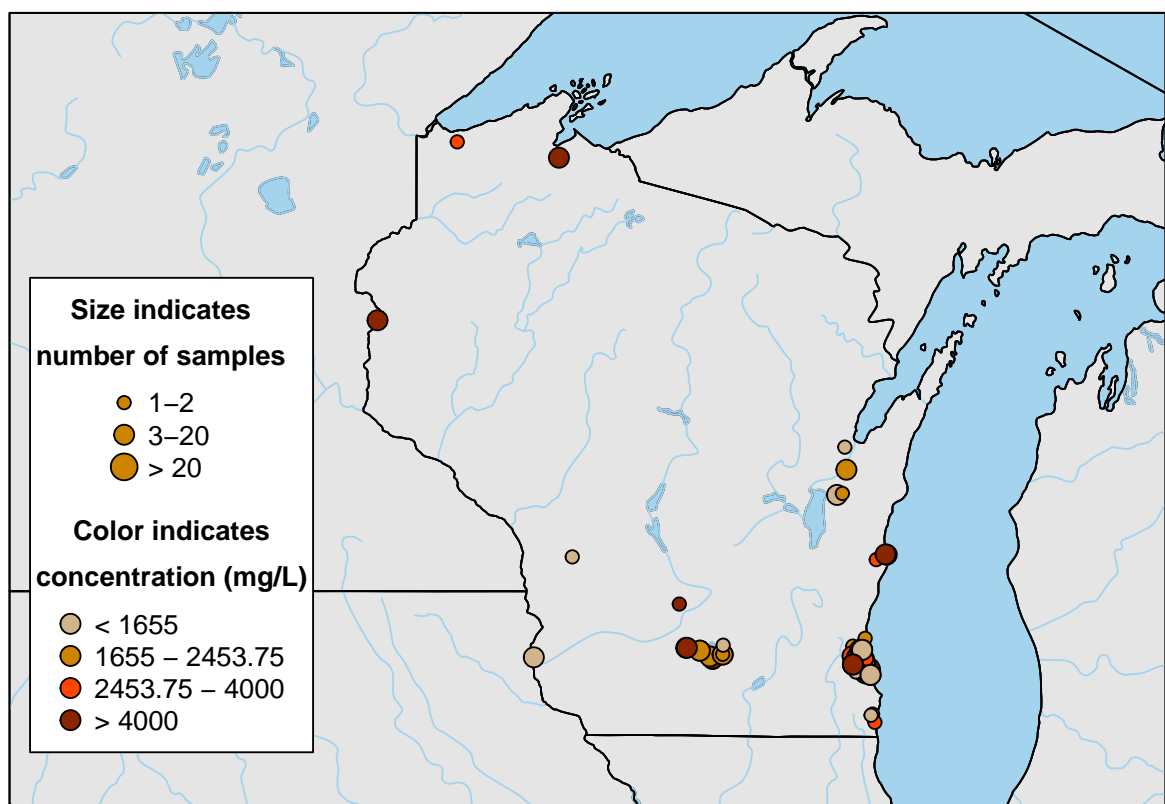


Figure 3. Map of chloride values above maxVal