2012 NYC Water Quality Measures Analysis

1. Data

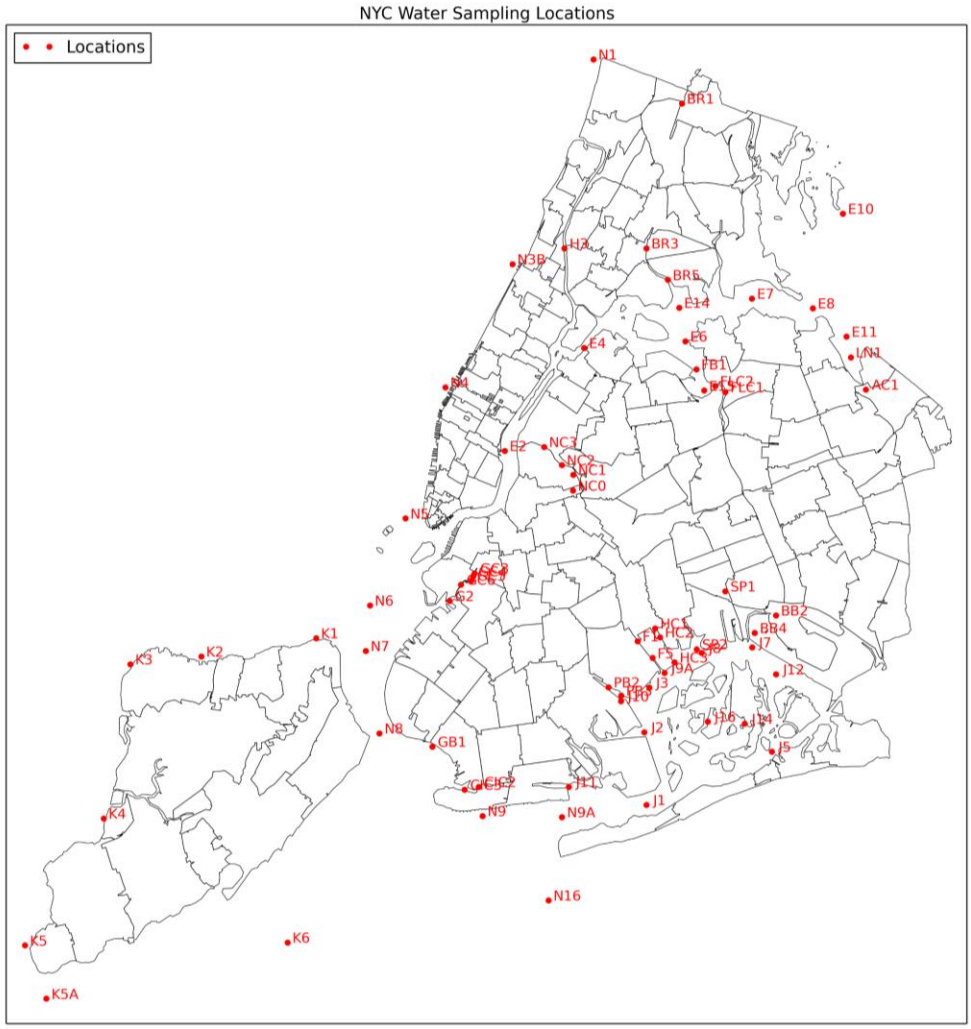

Figure 1, 70 NYC water sampling locations

The sampling locations are not distributed evenly, most of them are around the coasts of Brooklyn and Queens. After looking into the sampling data, we can find that: one record has a typo, the NCO should be NC0; J13, J15 and J17 are not included in the location information file, but they should be around other points that names start with J. There are total 1769 records left after dropping the ones without DO Top entries, then we can plot the number of records at each sampling point in NYC in 2012:

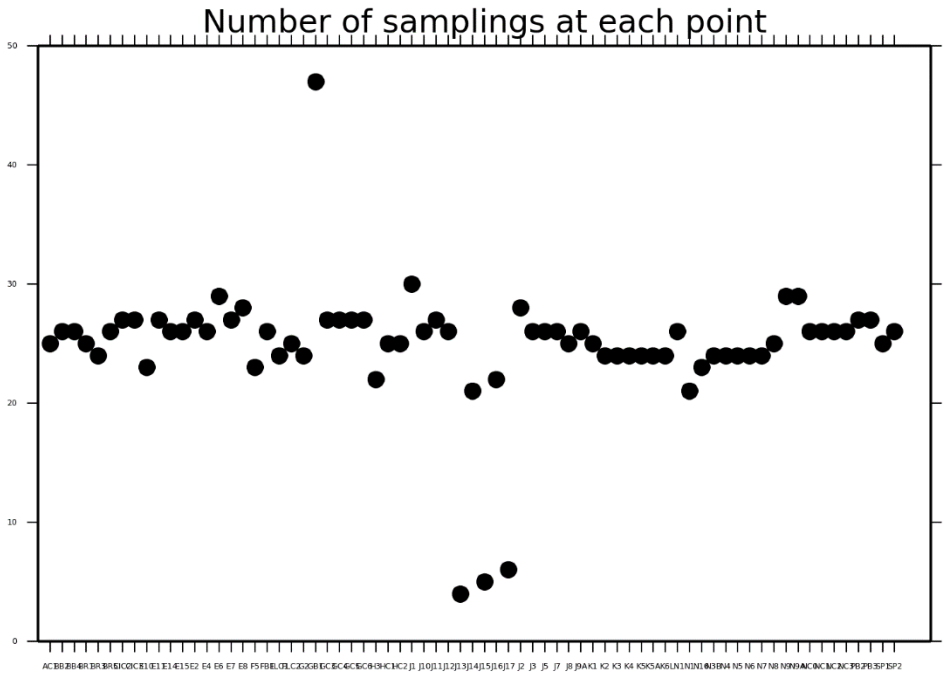

Figure 2, Number of samplings at each point

The average number is a little more than 25, which means there're averaged two records for each point every month in 2012.

2. Analysis

We can pick up the monthly averaged values of DO, FC, Enterococci and Transparency, and plot the following charts:

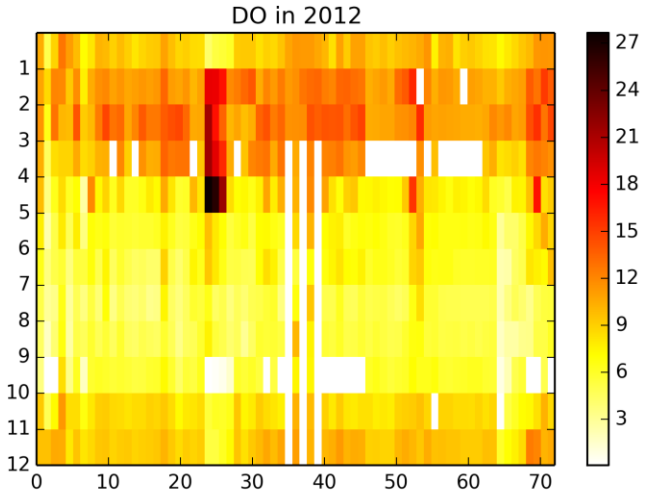

Figure 3, DO monthly averaged values at each sampling point

The x axis is the name of locations in the alphabetic order: 'AC1', 'BB2', 'BB4', 'BR1', 'BR3', 'BR5', 'CIC2', 'CIC3', 'E10', 'E11', 'E14', 'E15', 'E2', 'E4', 'E6', 'E7', 'E8', 'F5', 'FB1', 'FLC1', 'FLC2', 'G2', 'GB1', 'GC3', 'GC4', 'GC5', 'GC6', 'H3', 'HC1', 'HC2', 'J1', 'J10', 'J11', 'J12', 'J13', 'J14', 'J15', 'J16', 'J17', 'J2', 'J3', 'J5', 'J7', 'J8', 'J9A', 'K1', 'K2', 'K3', 'K4', 'K5', 'K5A', 'K6', 'LN1', 'N1', 'N16', 'N3B', 'N4', 'N5', 'N6', 'N7', 'N8', 'N9', 'N9A', 'NC0', 'NC1', 'NC2', 'NC3', 'PB2', 'PB3', 'SP1', 'SP2'.

It's very clear that the DO is higher from November to May, lower from June to October. Almost all the location follow this rule.

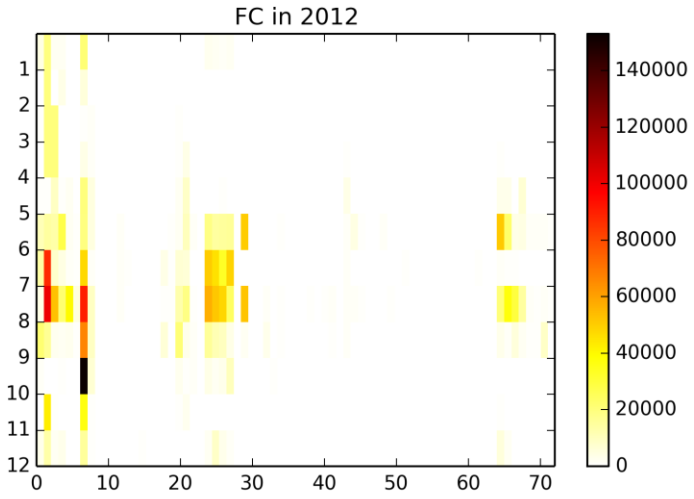

Figure 4, DO monthly averaged values at each sampling point

Areas around BB2, BB4 and CIC2 have high FC value through the year, and some other areas have relative high values from May to July.

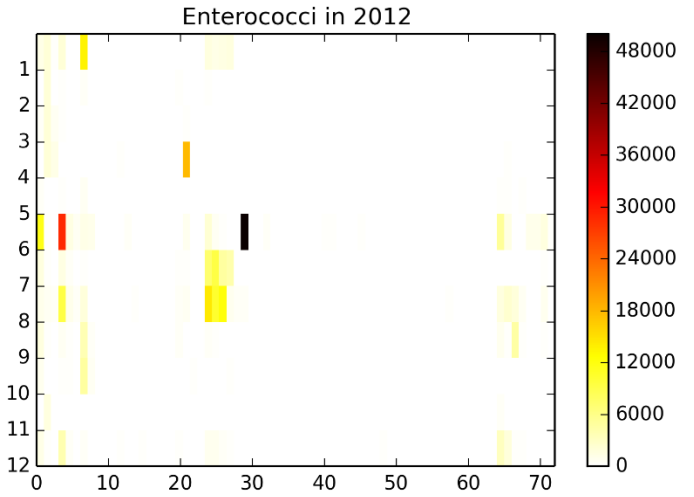

Figure 5, Enterococci monthly averaged values at each sampling point

What's interesting is that the first two largest values of Enterococci all happened in June. We can also see that the areas have high Enterococci values usually have high FC values too, which means these areas might be very dangerous, especially from June to August.

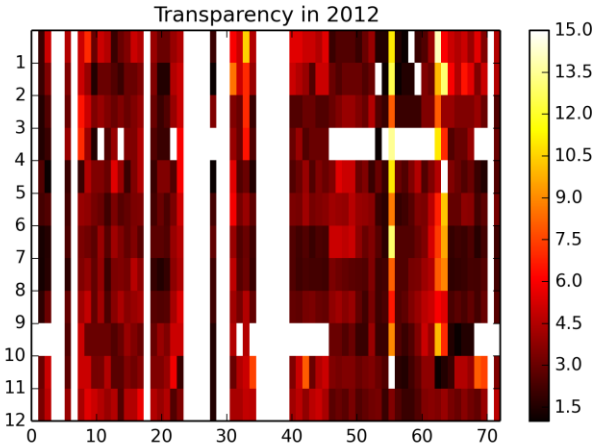

Figure 6, Transparency monthly averaged values at each sampling point

The pattern of Transparency is kind of stable through the year with each sampling location.

3. Best usages

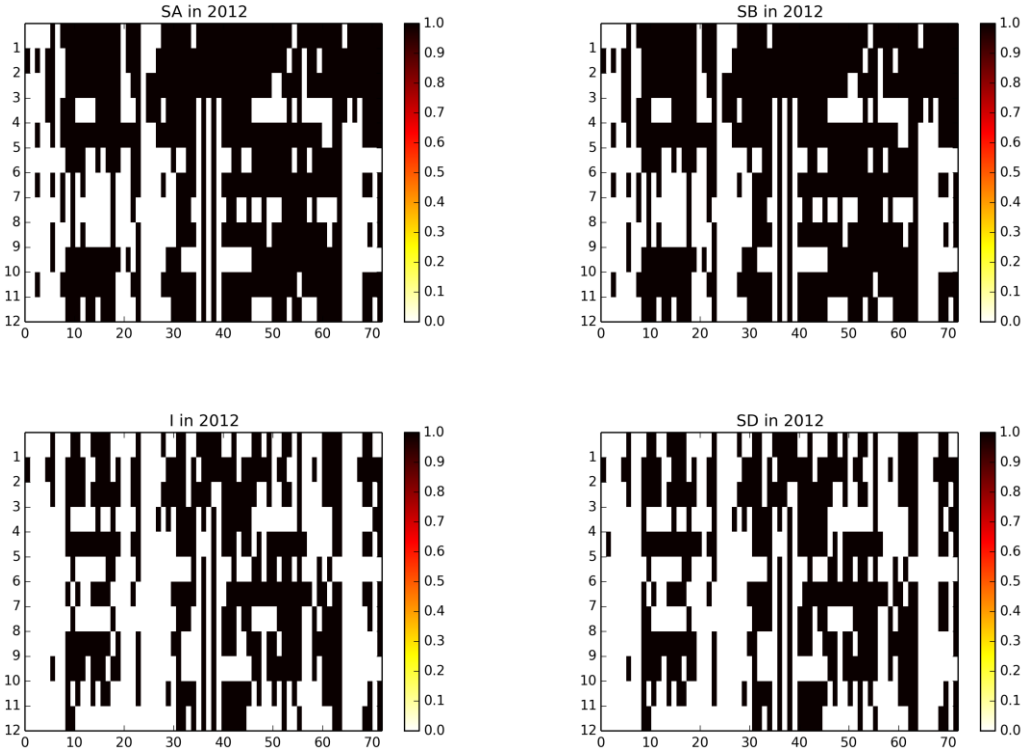

Figure 7, Best usages of waters

The black parts indicate the time and location of the SA, SB, I and SD classifications of waters in 2012. We can see that SA and SB, I and SD are very similar to each other, which means that there're no significant differences between FCs among SA and SB, I and SD in NYC in 2012, and the changes brought by the value of Enterococci from 10 to 35 is much more visible. From January to May, waters in most areas can be used for shellfishing or other recreational use.

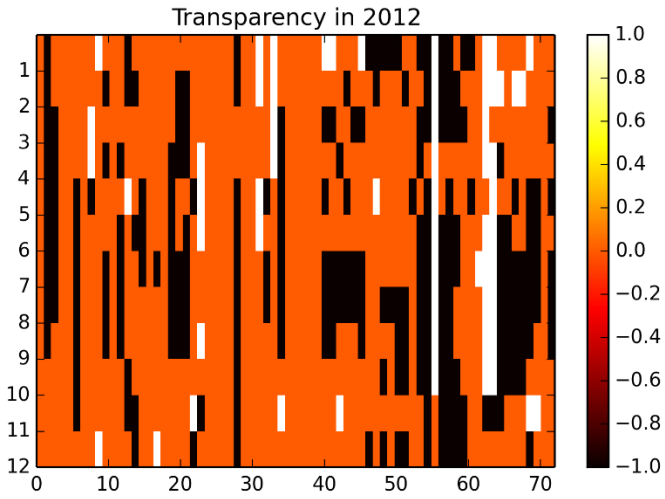

Figure 8, Transparency

The value 1(white) indicates the clear water, and -1(black) indicates the clear water. Actually, there's no area could keep clear through the year 2012.