Hurricane Irma Analysis

William Norfolk

11/5/2019

Based on the results of the exploratory and bivariate analyses a deeper dive into the specific trends of Hurricane Irma are warrented.

First load the libraries.

library(readxl)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(tidyverse)

## -- Attaching packages ------------------------------------------------- tidyverse 1.2.1 --

## v ggplot2 3.2.1 v readr 1.3.1  
## v tibble 2.1.3 v purrr 0.3.2  
## v tidyr 0.8.3 v stringr 1.4.0  
## v ggplot2 3.2.1 v forcats 0.4.0

## -- Conflicts ---------------------------------------------------- tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library(forcats)  
library(ggthemes)  
library(plotly)

##   
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':  
##   
## last\_plot

## The following object is masked from 'package:stats':  
##   
## filter

## The following object is masked from 'package:graphics':  
##   
## layout

library(knitr)  
library(naniar)  
library(broom)  
library(gridExtra)

##   
## Attaching package: 'gridExtra'

## The following object is masked from 'package:dplyr':  
##   
## combine

library(zoo)

##   
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

Next load the clean data and take a look.

WQ\_clean\_data <- readRDS("../../data/processed\_data/processeddata.rds")  
  
glimpse(WQ\_clean\_data)

## Observations: 522  
## Variables: 15  
## $ Month <chr> "01", "02", "02", "02", "02", "02", "02", "02...  
## $ Day <chr> "08", "08", "08", "08", "09", "09", "09", "12...  
## $ Year <chr> "16", "16", "16", "16", "16", "16", "16", "16...  
## $ military\_time <dbl> 1415, 1515, 1550, 1555, 1001, 1015, 1022, 103...  
## $ location <chr> "Boat Ramp", "Grecian Dry Rocks", "Grecian Dr...  
## $ instructor\_name <chr> "Katy, Sarah, Driver", "Chelsea", "Katy, Tomm...  
## $ group\_name <chr> "NA", "McLean High School", "McLean High Scho...  
## $ ph <dbl> 8.0, 8.4, 8.2, 8.4, 8.0, 8.0, 8.0, 8.0, 8.4, ...  
## $ ammonia <dbl> 0.00, 0.00, 0.00, 0.00, 0.25, 0.00, 0.00, 0.0...  
## $ dissolved\_oxygen <dbl> 5.0, 4.0, 4.0, 6.0, 8.0, 4.0, 5.0, 6.0, 6.0, ...  
## $ water\_temp <dbl> NA, 23.5, 21.0, 36.0, 18.0, 18.0, 18.0, 18.3,...  
## $ salinity <dbl> 36, 40, 44, 35, 33, 30, 33, 35, 40, 30, 35, 2...  
## $ equipment <chr> "kit", "kit", "kit", "kit", "kit", "kit", "ki...  
## $ island\_side <chr> "ocean", "ocean", "ocean", "ocean", NA, "bay"...  
## $ site\_type <chr> "Seagrass/Mangrove", "Coral Reef", "Coral Ree...

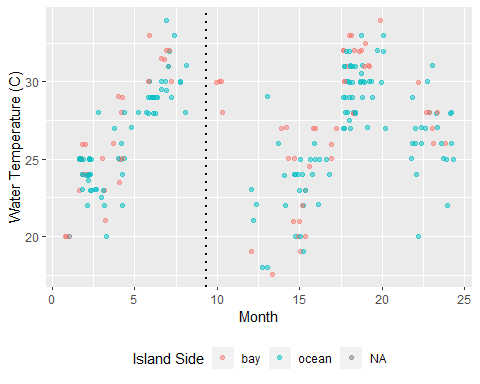
For this analysis we will subset the data to isolate the time frame close to Irma landfall. Landfall took place on September 10, 2017. We will selct data from 2017-2018 to view hurricane effects

year\_2017 <- subset(WQ\_clean\_data, Year == "17")  
year\_2018 <- subset(WQ\_clean\_data, Year == "18")  
  
#We also need to recode month for 2018 so it adds continiously.  
  
year\_2018 <- year\_2018 %>% dplyr::mutate(Month = recode(Month,  
 "01" = "13",  
 "02" = "14",  
 "03" = "15",  
 "04" = "16",  
 "05" = "17",  
 "06" = "18",  
 "07" = "19",  
 "08" = "20",  
 "09" = "21",  
 "10" = "22",  
 "11" = "23",  
 "12" = "24"))  
  
irma\_subset <- rbind(year\_2017, year\_2018)  
  
#Then convert Month back to numeric  
irma\_subset$Month <- as.numeric(as.character(irma\_subset$Month))

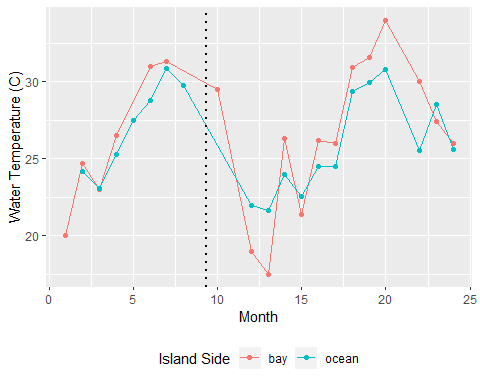
Lets take a quick look at the parameters for this subset.

irma\_temp\_plot <- ggplot(irma\_subset, aes(x = Month, y = water\_temp, color = island\_side)) + geom\_jitter(alpha = 0.5) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Water Temperature (C)") + labs(color = "Island Side")  
  
irma\_temp\_plot

## Warning: Removed 6 rows containing missing values (geom\_point).



mean\_temp\_irma <- aggregate(water\_temp ~ Month + Year +island\_side, irma\_subset, mean)  
  
irma\_mean\_temp\_plot <- ggplot(mean\_temp\_irma, aes(x = Month, y = water\_temp, color = island\_side)) + geom\_line(alpha = 1) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Water Temperature (C)") +labs(color = "Island Side")+ geom\_point()  
  
irma\_mean\_temp\_plot



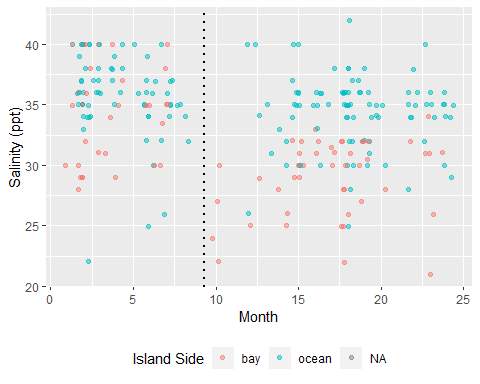
ggsave(filename = "../../results/imra\_mean\_temp\_plot.png",plot = irma\_mean\_temp\_plot)

## Saving 5 x 4 in image

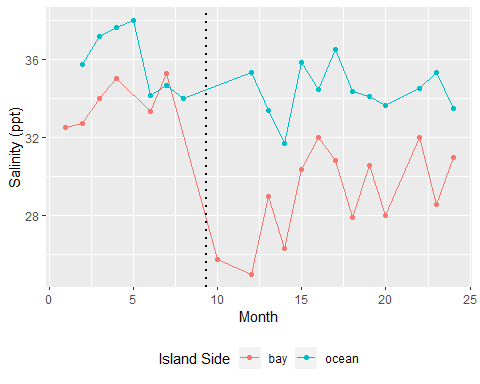
It looks like water temperature takes a dive following Irma landfall, however any temperature changes as a result of a hurricane we would expect to see quickly after landfall and these changes are delayed a few months in both ocean and bayside sites. There is a chance that hurricane influences could reduce the water temperature of the bayside due to heavy rainfall, but the trends visualized here are more likely a result of seasonal temperature change into winter months.

irma\_sal\_plot <- ggplot(irma\_subset, aes(x = Month, y = salinity, color = island\_side)) + geom\_jitter(alpha = 0.5) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Salinity (ppt)") + labs(color = "Island Side")  
  
irma\_sal\_plot

## Warning: Removed 6 rows containing missing values (geom\_point).



mean\_sal\_irma <- aggregate(salinity ~ Month + Year +island\_side, irma\_subset, mean)  
  
irma\_mean\_sal\_plot <- ggplot(mean\_sal\_irma, aes(x = Month, y = salinity, color = island\_side)) + geom\_line(alpha = 1) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Salinity (ppt)") + labs(color = "Island Side") +geom\_point()  
  
irma\_mean\_sal\_plot



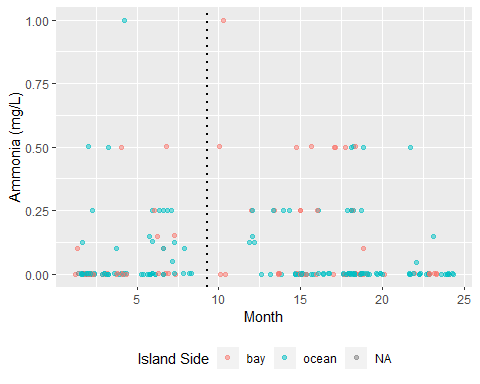
ggsave(filename = "../../results/imra\_mean\_sal\_plot.png",plot = irma\_mean\_sal\_plot)

## Saving 5 x 4 in image

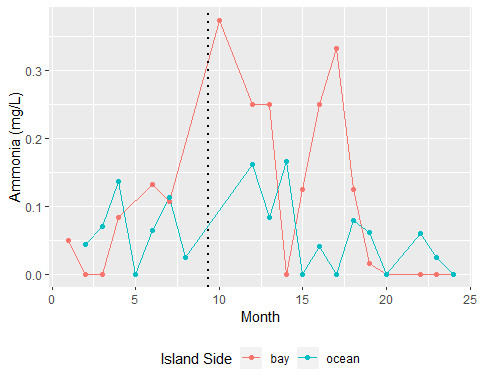
Salinity shows a distinct drop in bayside sites followinf Hurricane Irma. This is likely due to the storm surge and rainfall influences of the storm. These influences are more strongly noticed on the bayside since the water depth is shallow. It takes a substantial global even to reduce the salinity of the ocean waters so we would not expect to see a strong decrease at these sites.

irma\_amm\_plot <- ggplot(irma\_subset, aes(x = Month, y = ammonia, color = island\_side)) + geom\_jitter(alpha = 0.5) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Ammonia (mg/L)") + labs(color = "Island Side")  
  
irma\_amm\_plot

## Warning: Removed 2 rows containing missing values (geom\_point).



mean\_amm\_irma <- aggregate(ammonia ~ Month + Year +island\_side, irma\_subset, mean)  
  
irma\_mean\_amm\_plot <- ggplot(mean\_amm\_irma, aes(x = Month, y = ammonia, color = island\_side)) + geom\_line(alpha = 1) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Ammonia (mg/L)") + labs(color = "Island Side") + geom\_point()  
  
irma\_mean\_amm\_plot



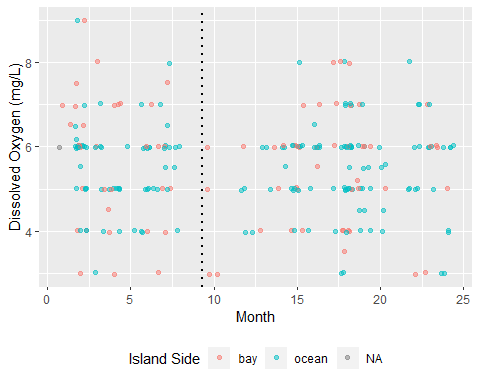
ggsave(filename = "../../results/imra\_mean\_amm\_plot.png",plot = irma\_mean\_amm\_plot)

## Saving 5 x 4 in image

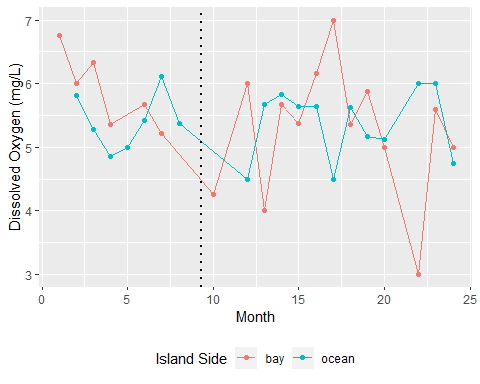
Ammonia shows a sunstantial spike in levels following Hurricane Irma landfall. This is likely due to storm-induced nutrint influx from surge, flooding, rainfall, septig/sewage overflow, etc. It is expected that these effects would be more strongly visualized in the bayside due to the close proximity to point and non-point sources of pollution and the silty substrate of the estuary. There is an interesting dip in the ammonia a few months following the spike, I am unsure as to the nature of this. It is possible there was a large number of measures from a particular site that showed a quicker recovery from the storm during this time frame.

irma\_do\_plot <- ggplot(irma\_subset, aes(x = Month, y = dissolved\_oxygen, color = island\_side)) + geom\_jitter(alpha = 0.5) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Dissolved Oxygen (mg/L)") + labs(color = "Island Side")  
  
irma\_do\_plot

## Warning: Removed 3 rows containing missing values (geom\_point).



mean\_do\_irma <- aggregate(dissolved\_oxygen ~ Month + Year +island\_side, irma\_subset, mean)  
  
irma\_mean\_do\_plot <- ggplot(mean\_do\_irma, aes(x = Month, y = dissolved\_oxygen, color = island\_side)) + geom\_line(alpha = 1) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("Dissolved Oxygen (mg/L)") + labs(color = "Island Side") + geom\_point()  
  
irma\_mean\_do\_plot



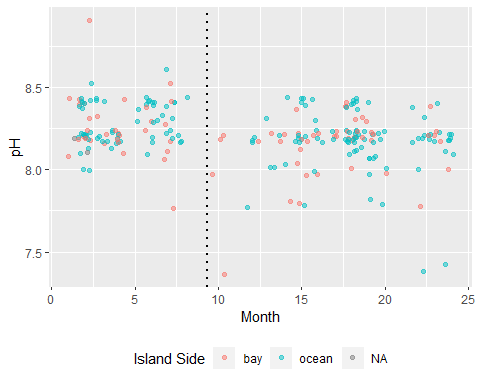
ggsave(filename = "../../results/imra\_mean\_do\_plot.png",plot = irma\_mean\_do\_plot)

## Saving 5 x 4 in image

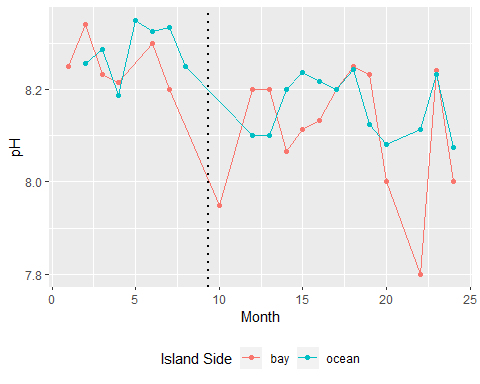
Dissolved oxgen shows a dip in assocation with the hurrincane, however this dip does not look terribly distinct from standard fluctuation. There is a sharp decline later in 2018, this may be associated with a strong algal bloom even in South Florida that summer, however I am unsure of the cause of this decline.

irma\_ph\_plot <- ggplot(irma\_subset, aes(x = Month, y = ph, color = island\_side)) + geom\_jitter(alpha = 0.5) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("pH") + labs(color = "Island Side")  
  
irma\_ph\_plot

## Warning: Removed 2 rows containing missing values (geom\_point).



mean\_ph\_irma <- aggregate(ph ~ Month + Year +island\_side, irma\_subset, mean)  
  
irma\_mean\_ph\_plot <- ggplot(mean\_ph\_irma, aes(x = Month, y = ph, color = island\_side)) + geom\_line(alpha = 1) + theme(legend.position = "bottom") + geom\_vline(xintercept = 9.33, color = "black", linetype = "dotted", size = 1) + ylab("pH") + labs(color = "Island Side") + geom\_point()  
  
irma\_mean\_ph\_plot



ggsave(filename = "../../results/imra\_mean\_ph\_plot.png",plot = irma\_mean\_ph\_plot)

## Saving 5 x 4 in image

Lastly, pH shows a distinct drop in bayside site location immediately following Irma. This is likely related to the salinity decrease due to heavy rainfall. Keys water tend to be basic (~8.0) and heavy rain fall of a neutral pH reduces the overall pH of smaller/shallow bodies of water.

Last Step to finish is statistical test, T-test?

Based on the visual differences in data it may be prudent to run a few simple significance tests comparing the data from 2017 to the others years of study. However since Irma made landfall in Sepetember, it is likely effects of the storm lingered for a few months after the event. So it would be prudent to create two new subsets of data and compare them.

month\_list <- c(9,10,11,12,13,14)  
filter\_irma\_fall <- irma\_subset %>% dplyr::filter(Month %in% month\_list)

year\_2016 <- subset(WQ\_clean\_data, Year == "16")  
year\_2017\_mod <- year\_2017  
year\_2017\_mod <- year\_2017\_mod %>% dplyr::mutate(Month = recode(Month,  
 "01" = "13",  
 "02" = "14",  
 "03" = "15",  
 "04" = "16",  
 "05" = "17",  
 "06" = "18",  
 "07" = "19",  
 "08" = "20",  
 "09" = "21",  
 "10" = "22",  
 "11" = "23",  
 "12" = "24"))  
  
not\_irma\_subset <- rbind(year\_2016, year\_2017\_mod)  
  
filter\_not\_irma\_fall <- not\_irma\_subset %>% dplyr::filter(Month %in% month\_list)

Now that we have our two fall subsets we will compare the individual water quality parameters between each using a Kolmogorov-Smirnov test. Since our data is likely to be seasonally be influenced as well as storm influenced we will use this test since it does not require a normal distribution.

ks.test(filter\_irma\_fall$ph, filter\_not\_irma\_fall$ph)

## Warning in ks.test(filter\_irma\_fall$ph, filter\_not\_irma\_fall$ph): cannot  
## compute exact p-value with ties

##   
## Two-sample Kolmogorov-Smirnov test  
##   
## data: filter\_irma\_fall$ph and filter\_not\_irma\_fall$ph  
## D = 0.25848, p-value = 0.334  
## alternative hypothesis: two-sided

ks.test(filter\_irma\_fall$dissolved\_oxygen, filter\_not\_irma\_fall$dissolved\_oxygen)

## Warning in ks.test(filter\_irma\_fall$dissolved\_oxygen,  
## filter\_not\_irma\_fall$dissolved\_oxygen): cannot compute exact p-value with  
## ties

##   
## Two-sample Kolmogorov-Smirnov test  
##   
## data: filter\_irma\_fall$dissolved\_oxygen and filter\_not\_irma\_fall$dissolved\_oxygen  
## D = 0.22727, p-value = 0.4995  
## alternative hypothesis: two-sided

ks.test(filter\_irma\_fall$ammonia, filter\_not\_irma\_fall$ammonia)

## Warning in ks.test(filter\_irma\_fall$ammonia, filter\_not\_irma\_fall$ammonia):  
## cannot compute exact p-value with ties

##   
## Two-sample Kolmogorov-Smirnov test  
##   
## data: filter\_irma\_fall$ammonia and filter\_not\_irma\_fall$ammonia  
## D = 0.51228, p-value = 0.001802  
## alternative hypothesis: two-sided

ks.test(filter\_irma\_fall$water\_temp, filter\_not\_irma\_fall$water\_temp)

## Warning in ks.test(filter\_irma\_fall$water\_temp,  
## filter\_not\_irma\_fall$water\_temp): cannot compute exact p-value with ties

##   
## Two-sample Kolmogorov-Smirnov test  
##   
## data: filter\_irma\_fall$water\_temp and filter\_not\_irma\_fall$water\_temp  
## D = 0.29365, p-value = 0.2273  
## alternative hypothesis: two-sided

ks.test(filter\_irma\_fall$salinity, filter\_not\_irma\_fall$salinity)

## Warning in ks.test(filter\_irma\_fall$salinity,  
## filter\_not\_irma\_fall$salinity): cannot compute exact p-value with ties

##   
## Two-sample Kolmogorov-Smirnov test  
##   
## data: filter\_irma\_fall$salinity and filter\_not\_irma\_fall$salinity  
## D = 0.51111, p-value = 0.002419  
## alternative hypothesis: two-sided

So it looks like our Two-sample Kolmogorov-Smirnov test shows a significant difference in the salinity and ammonia following Hurricane Irma landfall. This pattern is consistent with known influences of storm surge/heavy rainfall on the salinity and ammonia content of Key Largo waters. It should be noted that our statistical tests are limited strongly by the number of observations taken during the alloted timeframe. Few measurements are taken imediately following Irma landfall, thus grouping statistical tests by bay and ocean side sites (which would be preferrable) would produce n values for the filter\_irma\_fall dataset which are below 10 therefore we will keep the data as a group. It should be noted that there would likely be larger imapcts to the bayside sites.