Enviro data testing

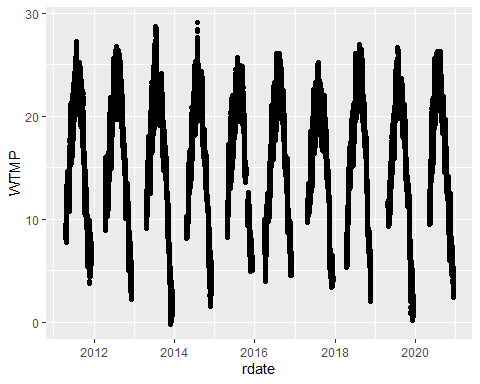
#What is this document?

This R markdown file displays the environmental data from each site that we used to calculate different environmental parameters. We include this document to 1) visually display each site’s data and its temporal range and quality, and 2) show the evidence for choosing certain environmental data sources over others based on data quality.

# Great Bay

### Jackson marine lab data will take too long. Here’s NERRS data from center bay, pretty close by

gb<-read.csv(here::here("data/test\_env/gb/GRBGBWQ.csv"))  
gb<-gb[,c(3,7)]  
gb$rdate<-as.POSIXct(gb$DateTimeStamp,tz="", "%m/%d/%Y%H:%M")  
gb<-na.omit(gb)  
gb$WTMP<-gb$Temp  
gb<-gb[,c(3,4)]  
gb$site<-"gb"  
  
ggplot(gb,aes(x=rdate,y=WTMP))+geom\_point()



mean(gb$WTMP)

## [1] 15.99839

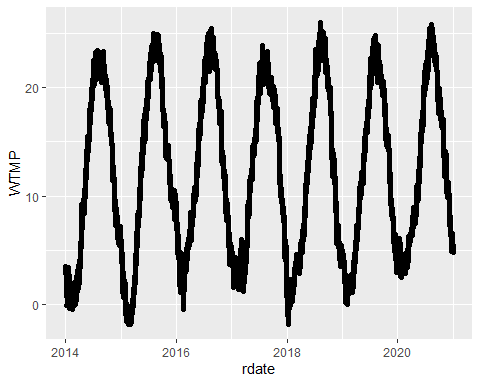
max(gb$WTMP)

## [1] 29.1

# Woods Hole

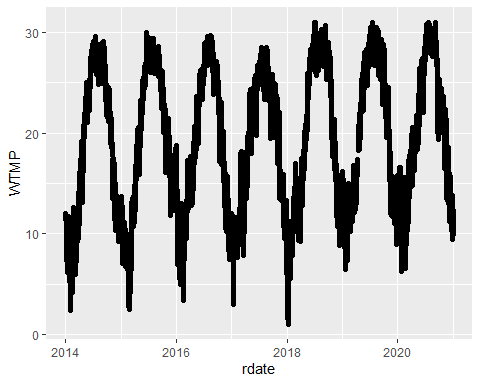
### Woods hole long term data is clean and useable. Same source as original analysis.

## Value  
## 1 26



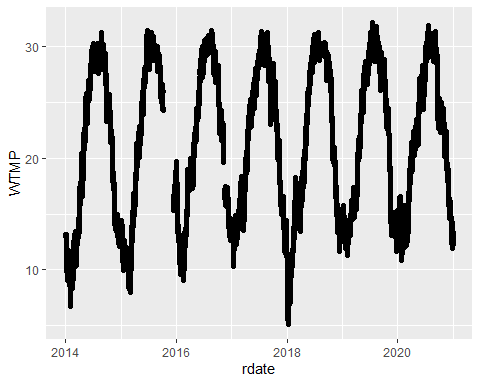
# Beaufort

### Beaufort long term data is clean and useable. Same source as original analysis.



# Folly beach

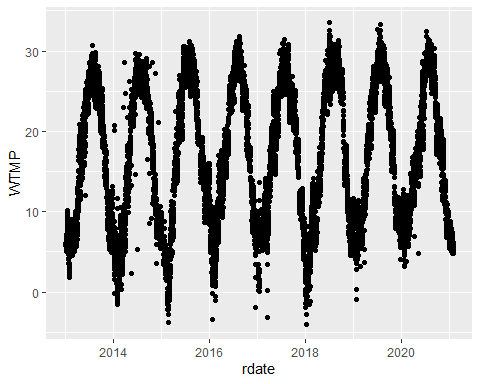
### Folly Beach long term data is clean and useable. Same source as original analysis.



# Oyster

### Oyster long term data is clean and useable. Same source as original analysis.

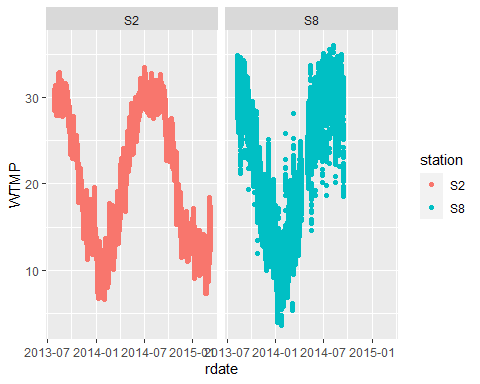
### Load and Analysis



# Skidaway

### Skidaway data requires some thought. Grove creek (right next to collection site) was the original data source I used, from 2014. This is a complete year of data. However, there are no other dates available from this location. A one-time data collection operation. The two stations reflect two slighly different locations within the creek - both are useable, but I previously used S2.

## Grove Creek

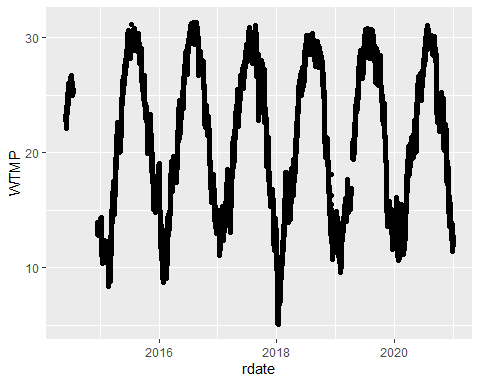


## [1] 20.6781

## [1] 33.478

## Savannah River Buoy

### Savannah river NDBC buoy is the next closes option, at 13 km. This data is pretty complete, but is further away (13km) and is seemingly affected by the Savannah River plume. Mean temperatures are lower by a degree, max temp by two degrees.

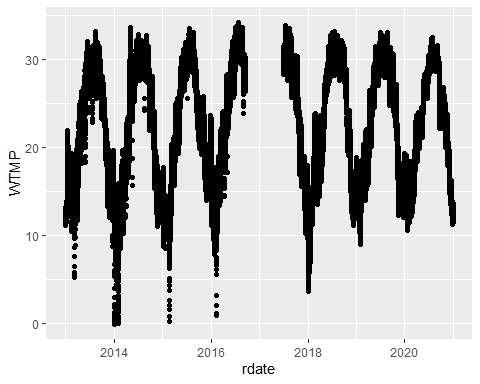


## [1] 21.28308

## [1] 31.3

## Sapelo

### Sapelo Island is a NERRS site to the south of Skidaway by about 60 km. This is the closest next buoy, NERRS or NDBC otherwise. Temps here are pretty much in line with grove creek estimates. This is because both are located within the Georgia barrier islands areas about the same distance up a tidal creek (3-4 km). If we can get over the fact this is 60 km away (maybe via corelation of temps), this is an option).

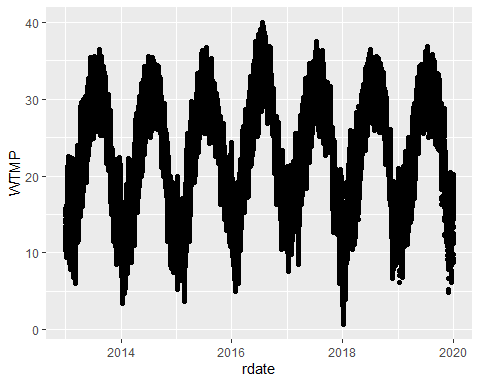


## [1] 22.30085

## [1] 34.2

## Sapleo 2

### The original sapelo site had freezing dates (low tide) and incomplete dates. Here’s anotehr logger from the same site.



## [1] 22.39

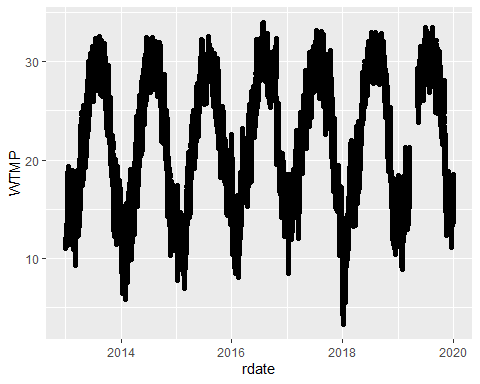
## [1] 40

## Sapelo 3

### Trying one more option

sap3<-read.csv(here::here("data/test\_env/sk/SAPCAWQ.csv"))  
sap3<-sap3[,c(3,7)]  
  
sap3$rdate<-as.POSIXct(sap3$DateTimeStamp,tz="", "%m/%d/%Y%H:%M")  
sap3$WTMP<-sap3$Temp  
sap3$WTMP<-as.numeric(sap3$WTMP)  
  
  
ggplot(sap3,aes(x=rdate,y=WTMP))+geom\_point()

## Warning: Removed 10275 rows containing missing values (geom\_point).



mean(sap3$WTMP,na.rm=T)

## [1] 21.65186

max(sap3$WTMP,na.rm=T)

## [1] 34

#metrics

sk\_metrics<-data.frame("site"=c("gcsk","sk","sap","sap2","sap3"),"quantile"=NA,"decile"=NA, "max"=NA, "mean"=NA,"summer mean"=NA,"seasonlength10"=NA,"seasonlength12"=NA,"initialspawn"=NA,"maxspawn"=NA)  
gcsk2\_avg<-na.omit(gcsk2\_avg)  
sap\_avg<-na.omit(sap\_avg)  
sap2\_avg<-na.omit(sap2\_avg)  
sap3\_avg<-na.omit(sap3\_avg)  
  
sk\_metrics[1,5]<-mean(gcsk2\_avg$gcsk2)  
sk\_metrics[2,5]<-mean(sk\_avg$sk)  
sk\_metrics[3,5]<-mean(sap\_avg$sap)  
sk\_metrics[4,5]<-mean(sap2\_avg$sap2)  
sk\_metrics[5,5]<-mean(sap3\_avg$sap3)  
  
sk\_metrics[1,4]<-max(gcsk2\_avg$gcsk2)  
sk\_metrics[2,4]<-max(sk\_avg$sk)  
sk\_metrics[3,4]<-max(sap\_avg$sap)  
sk\_metrics[4,4]<-max(sap2\_avg$sap2)  
sk\_metrics[5,4]<-max(sap3\_avg$sap3)  
  
s.gcsk2\_avg<-filter(gcsk2\_avg,rdate>"2014-06-01 00:00:00" & rdate< "2014-09-30 00:00:00")  
s.sk\_avg<-filter(sk\_avg,rdate>"2014-06-01 00:00:00" & rdate< "2014-09-30 00:00:00")  
s.sap\_avg<-filter(sap\_avg,rdate>"2014-06-01 00:00:00" & rdate< "2014-09-30 00:00:00")  
s.sap2\_avg<-filter(sap2\_avg,rdate>"2014-06-01 00:00:00" & rdate< "2014-09-30 00:00:00")  
s.sap3\_avg<-filter(sap3\_avg,rdate>"2014-06-01 00:00:00" & rdate< "2014-09-30 00:00:00")  
  
sk\_metrics[1,6]<-mean(s.gcsk2\_avg$gcsk2)  
sk\_metrics[2,6]<-mean(s.sk\_avg$sk)  
sk\_metrics[3,6]<-mean(s.sap\_avg$sap)  
sk\_metrics[4,6]<-mean(s.sap2\_avg$sap2)  
sk\_metrics[5,6]<-mean(s.sap3\_avg$sap3)  
  
  
sk\_metrics[1,2]<-quantile(s.gcsk2\_avg$gcsk2,0.75,type=1)  
sk\_metrics[2,2]<-quantile(s.sk\_avg$sk,0.75,type=1)  
sk\_metrics[3,2]<-quantile(s.sap\_avg$sap,0.75,type=1)  
sk\_metrics[4,2]<-quantile(s.sap2\_avg$sap2,0.75,type=1)  
sk\_metrics[5,2]<-quantile(s.sap3\_avg$sap3,0.75,type=1)  
  
sk\_metrics[1,3]<-quantile(s.gcsk2\_avg$gcsk2,0.9,type=1)  
sk\_metrics[2,3]<-quantile(s.sk\_avg$sk,0.9,type=1)  
sk\_metrics[3,3]<-quantile(s.sap\_avg$sap,0.9,type=1)  
sk\_metrics[4,3]<-quantile(s.sap2\_avg$sap2,0.9,type=1)  
sk\_metrics[5,3]<-quantile(s.sap3\_avg$sap3,0.9,type=1)  
  
  
  
sk\_metrics[1,7]<-gcsk2\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(gcsk2))%>%na.omit()%>%filter(daily\_wtmp>12.5)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[2,7]<-sk\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sk))%>%na.omit()%>%filter(daily\_wtmp>12.5)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[3,7]<-sap\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap))%>%na.omit()%>%filter(daily\_wtmp>12.5)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[4,7]<-sap2\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap2))%>%na.omit()%>%filter(daily\_wtmp>12.5)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[5,7]<-sap3\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap3))%>%na.omit()%>%filter(daily\_wtmp>12.5)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[1,8]<-gcsk2\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(gcsk2))%>%na.omit()%>%filter(daily\_wtmp>10)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[2,8]<-sk\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sk))%>%na.omit()%>%filter(daily\_wtmp>10)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[3,8]<-sap\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap))%>%na.omit()%>%filter(daily\_wtmp>10)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[4,8]<-sap2\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap2))%>%na.omit()%>%filter(daily\_wtmp>10)%>%tally

## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[5,8]<-sap3\_avg%>%mutate(day=as.Date(rdate,format="%Y-%m-%d"))%>%group\_by(day)%>%dplyr::summarise(daily\_wtmp=mean(sap3))%>%na.omit()%>%filter(daily\_wtmp>10)%>%tally

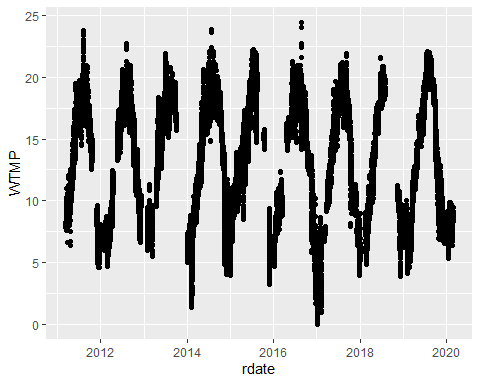
## `summarise()` ungrouping output (override with `.groups` argument)

sk\_metrics[1,9]<-mean((filter(gcsk2\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-03-30 00:00:00"))$gcsk2)  
sk\_metrics[2,9]<-mean((filter(sk\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-03-30 00:00:00"))$sk)  
sk\_metrics[3,9]<-mean((filter(sap\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-03-30 00:00:00"))$sap)  
sk\_metrics[4,9]<-mean((filter(sap2\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-03-30 00:00:00"))$sap2)  
sk\_metrics[5,9]<-mean((filter(sap3\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-03-30 00:00:00"))$sap3)  
  
sk\_metrics[1,10]<-mean((filter(gcsk2\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-05-30 00:00:00"))$gcsk2)  
sk\_metrics[2,10]<-mean((filter(sk\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-05-30 00:00:00"))$sk)  
sk\_metrics[3,10]<-mean((filter(sap\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-05-30 00:00:00"))$sap)  
sk\_metrics[4,10]<-mean((filter(sap2\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-05-30 00:00:00"))$sap2)  
sk\_metrics[5,10]<-mean((filter(sap3\_avg,rdate>"2014-03-01 00:00:00" & rdate< "2014-05-30 00:00:00"))$sap3)

# Willapa

## We originally used data from the Nahcotta buoy, the closest data option to our site. We stitched data from the first half of 2015 and the last half of 2016 together to get a full years of data - as you can see, the data available from this site is pretty patchy. Not really suitable to use as a long time series. We were able to get an extended dataset from PSI.

## Nahcotta



## [1] 13.41351

## [1] 24.4529

## [1] 13.01989

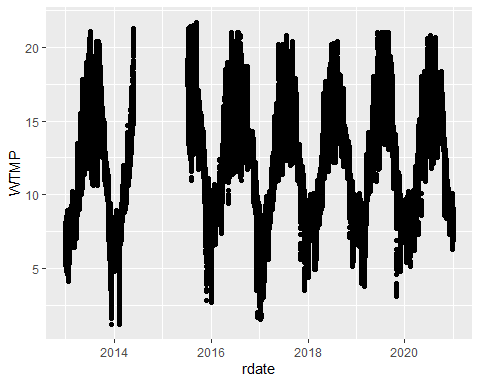
## [1] 13.3107

## [1] 13.07447

## [1] 13.55013

## Toke Point

### Toke Point is the next closest data soruce, about 30km away close to the mouth of Willapa Bay and thus is cooler (mean by 1.5C, max by 5C). However, this record is pretty complete.



## [1] 21.7

## [1] 12.35239

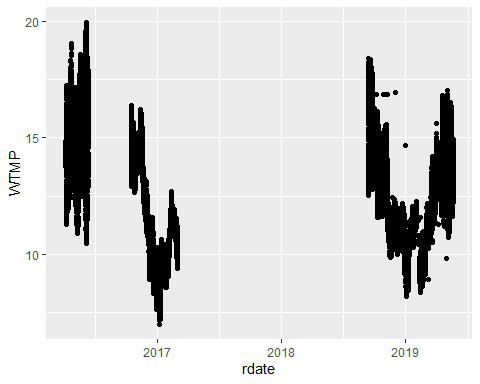
# Humboldt

## Tuluwat Island

### Tuluwat island (“indian” island) was one source that we used in our analysis, but this is very incomplete.

tul<-read.csv(here::here("data/test\_env/hm/ind.csv"))  
tul<-tul%>%unite("DateTimeStamp", c(1:2),sep="")  
tul$rdate<-as.POSIXct(tul$DateTimeStamp,tz="", "%m/%d/%y%H:%M")  
ggplot(tul,aes(x=rdate,y=WTMP))+geom\_point()

## Warning: Removed 4 rows containing missing values (geom\_point).



max(tul$WTMP)

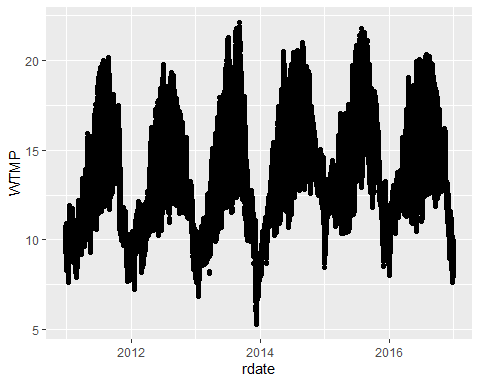
## [1] 19.97

mean(tul$WTMP)

## [1] 12.59349

## Wiyot tribe source of Tuluwat island

### This data is still form Tuluwat, but a different portal. Unsure of why this one is so complete but the one above is not. The mean temperature lines up with the paper, but the maximum temperature is four degrees lower.

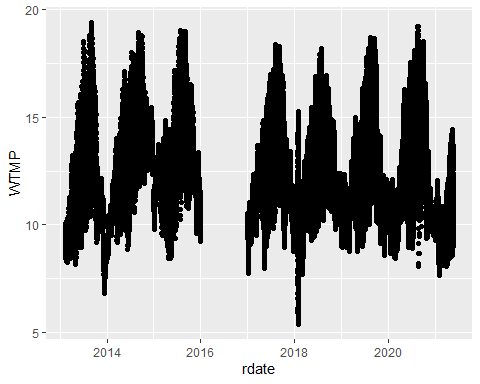


## [1] 13.63778

## [1] 22.13

## Humboldt University

### The humboldt university buoy is slightly farther out toward the bay mouth (12km instead of 7km), and has a pretty complete record. However, temperatures are alsoc cooler here than our initial analysis, mean by two degrees and max by 7 degrees.



## [1] 19.39

## [1] 11.94317