Boston Buoy Data Analysis Project

9/25/2020

1 Introduction

1.1 Overview

For this project, we use 30-year data collected by a single weather buoy (Station 44013) in the NOAA National Data Buoy Center to explore whether this is evidence of global warming. To explore this relationship, we study the effects of global warming, such as increases in average temperatures[1], ocean wave power[2], air pressure[3]. Therefore, we want to explore the evidence from these three aspects.

1.2 Outline

The outline of this report is as follows. Firstly, we read the 30-year data and clean the data. Next, we plot the line chart between temperature and years. Finally, we interpret our model and discuss the conclusions.

2 Date Cleaning

First of all, we read the 30-year data from 1987 to 2016 from the website.

```
#make URLs
url1 <- "http://www.ndbc.noaa.gov/view_text_file.php?filename=mlrf1h"
url2 <- ".txt.gz&dir=data/historical/stdmet/"
years <- c(1987:2016)
urls <- str_c(url1, years, url2, sep = "")
filenames <- str_c("mr", years, sep = "")

#read data
N <- length(urls)
for (i in 1:N){
  assign(filenames[i], read_table2(urls[i], col_names = TRUE))}
```

Then, we follow these steps to clean the data.
- Because the number of columns of tables is different, we respectively remove columns of mm, WD, BAR, DEWP, WDIR, PRES, and add the column of TIDE. Besides, because the existing value of TIDE is constant, we assign 99 to TIDE and make substitutions for NA data.
- we transform data type from char to num
- Based on related data, the hottest part of the day during the summer is usually between 3 p.m. and 4:30 p.m.[4] Besides, the record high temperature in Boston is 104 °F (40 °C), recorded on July 4, 1911. The record low temperature in Boston is -18 °F (-28 °C), recorded on February 9, 1934[5]. And based on historical data over ten years, the warmest sea in this day in Boston was recorded in 2018 and was 21°C/69.8°F, and the coldest was recorded in 2007 at 18.7°C/65.7°F[6]. Therefore, we select the data of 3 p.m. every day and select -18°F~104 °F as a temperature range to remove outlier.
- We unify the header of tables and merge information among tables.
- Since the data of air pressure doesn't exist and the data of ocean wave power is constant, we select the data of average temperature as the effect of global warming. We compute the average of air temperature and

water temperature by years and add the new column of the difference between air temperature and water temperature.

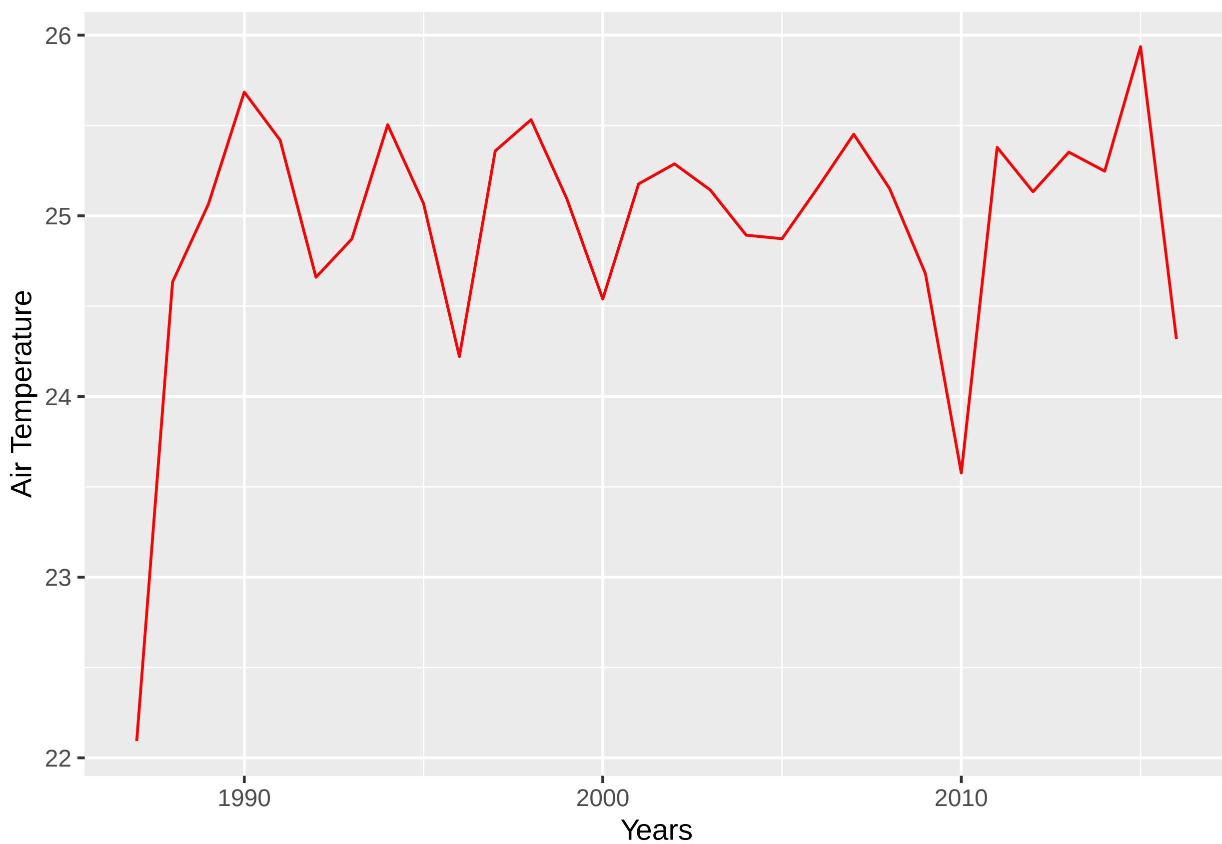
```
#remove the column of mm
for (i in 19:30){
  file<-get(filenames[i])
  file<-subset(file,select=-c(mm))
  assign(filenames[i],file)}

#add column of TIDE
for (i in 1:13){
  file<-get(filenames[i])
  file<-transform(file,TIDE=99)
  assign(filenames[i],file)}

#make substitutions for NA data in the column of TIDE and transform data type from char to num
for (i in 1:30){
  file<-get(filenames[i])
  file$TIDE[is.na(file$TIDE)] <- 99
  file<-as.data.frame(lapply(file,as.numeric))
  assign(filenames[i],file)}

#remove the first row of text
for (i in 21:30){
  file<-get(filenames[i])
  file<-file[-1,]
  assign(filenames[i],file)}

#unify the header, select the time of highest temperature in one day, and remove outlier of the column
for (i in 1:30){
  file<- get(filenames[i])
  colnames(file)[1] <-"YYYY"
  file<-transform(file,YYYY=i+1986)
  file<-filter(file,hh==15,ATMP>=-18&ATMP<=104,WTMP>=-18&WTMP<=104)
  assign(filenames[i],file)}
for (i in 1:20){
  file<- get(filenames[i])
  file<-subset(file,select=-c(WD,BAR,DEWP))
  assign(filenames[i],file)}
for (i in 21:30){
  file<- get(filenames[i])
  file<-subset(file,select=-c(WDIR,PRES,DEWP))
  assign(filenames[i],file)}

#merge information among tables
for (i in 1:30){
  file<- get(filenames[i])
  if(i == 1){
  MR <- file
  }
  else{
  MR <- rbind.data.frame(MR, file)
  }}
```

```r
#compute the average of air temperature and water temperature by years, and add the new column of the d
MR1<-MR%>%
  group_by(YYYY)%>%
  summarize(
    ATMP=mean(ATMP,na.rm=TRUE),
    WTMP=mean(WTMP,na.rm=TRUE),
  )
MR1<-transform(MR1,TMP=WTMP-ATMP)
```

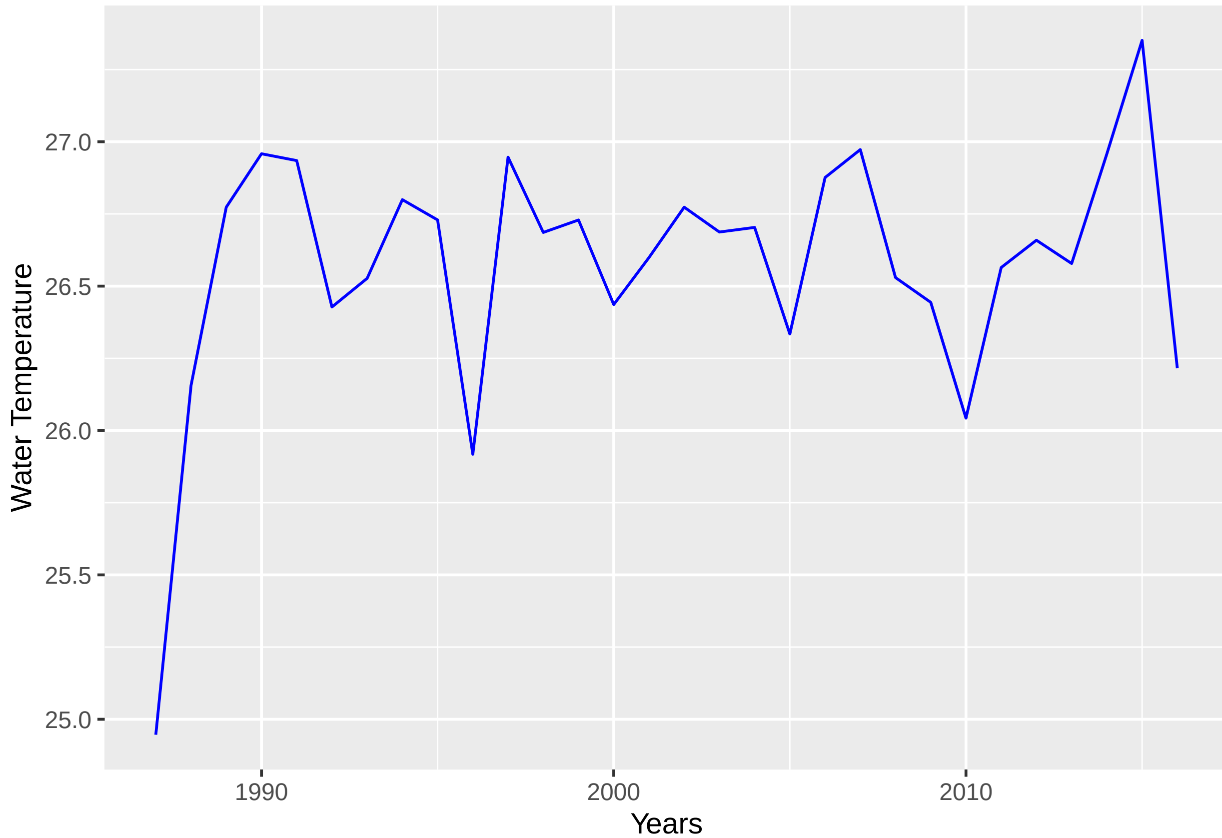
3 Plots

Finally, we plot line charts between years and air temperature, water temperature, the temperature difference between air and water.

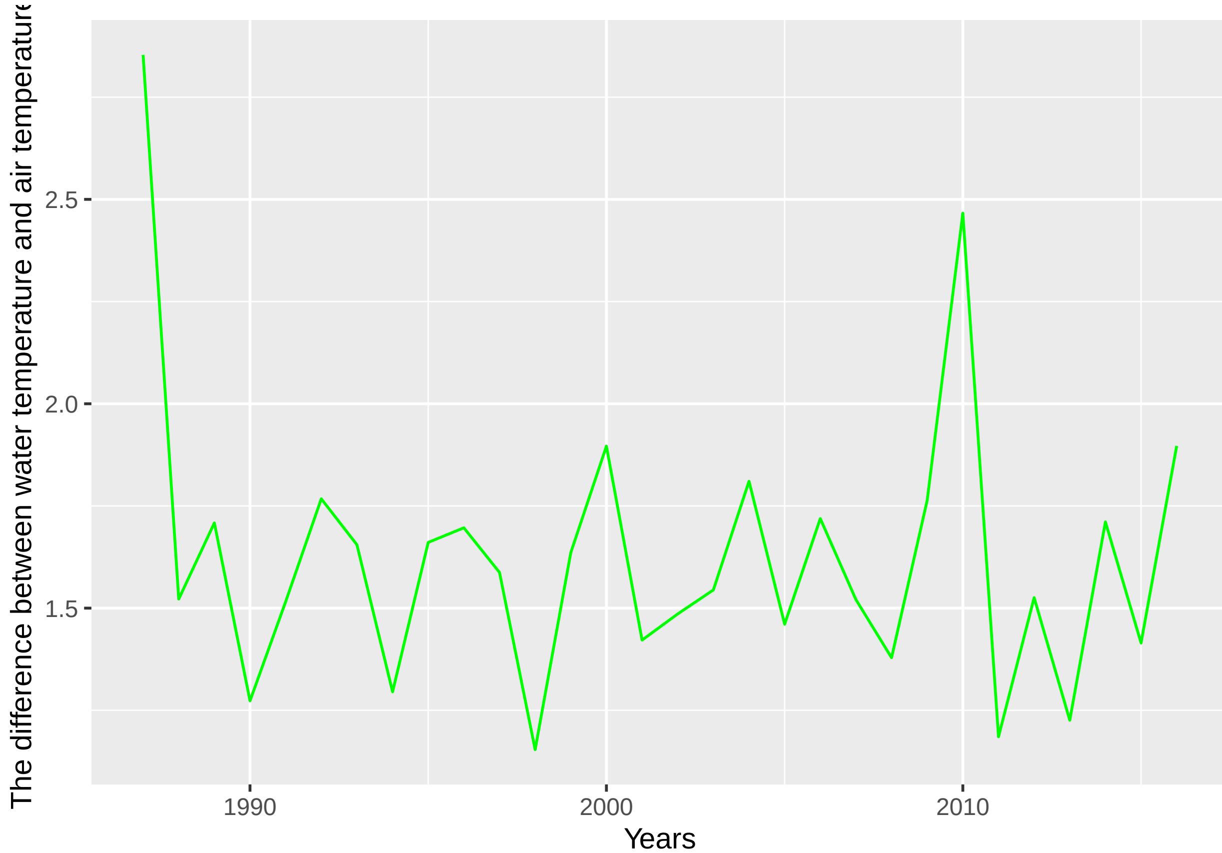
```r
#plot line charts between temperature and years
ggplot(MR1)+xlab("Years")+ylab("Air Temperature")+
  geom_line(aes(x=MR1[,1],y=MR1[,2]),color="red")
```



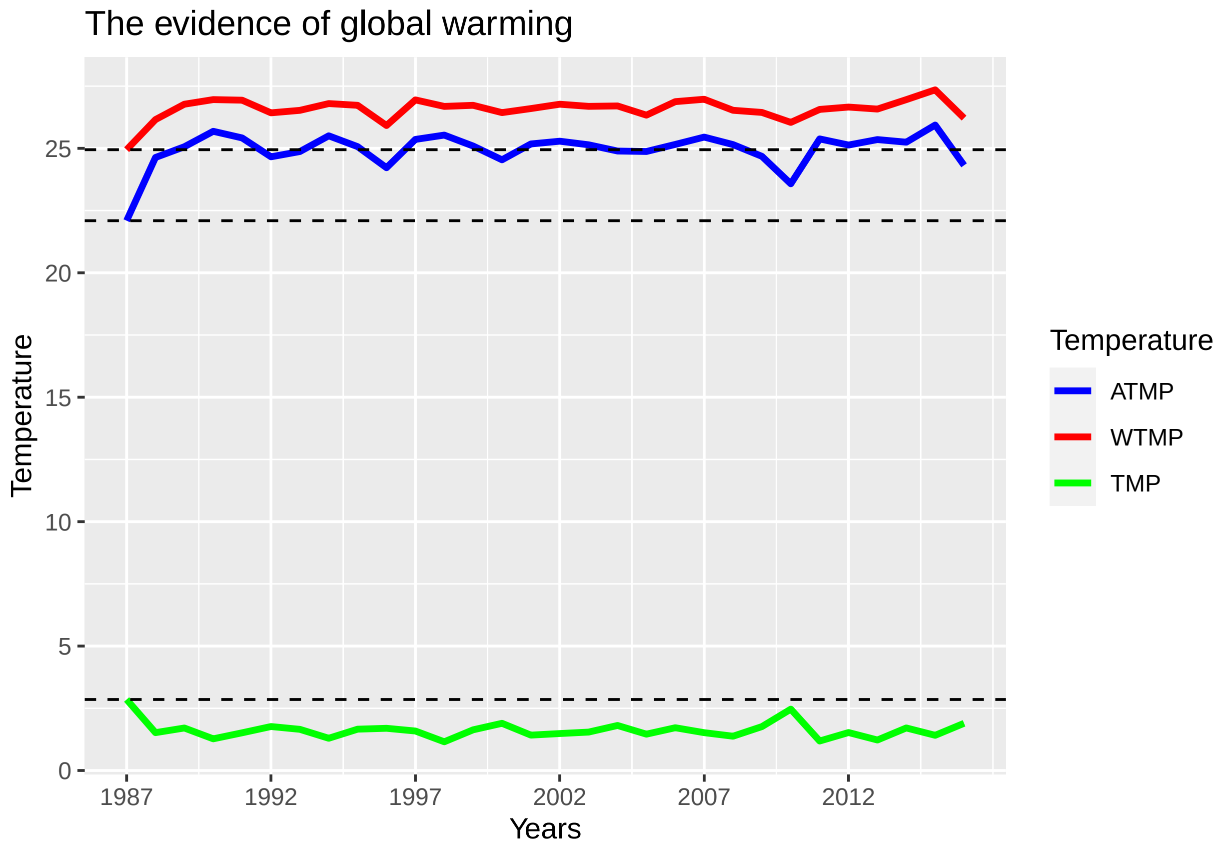

```r
ggplot(MR1)+xlab("Years")+ylab("Water Temperature")+
  geom_line(aes(x=MR1[,1],y=MR1[,3]),color="blue")
```




```
ggplot(MR1)+xlab("Years")+ylab("The difference between water temperature and air temperature")+
  geom_line(aes(x=MR1[,1],y=MR1[,4]),color="green")
```




```r
MR2=melt(MR1, id.vars = c("YYYY"))
ggplot(MR2)+
  geom_line(aes(x=MR2[,1],y=MR2[,3],color=MR2[,2]),size=1.2)+
  scale_color_manual(values = c("blue","red","green"))+
  labs(x="Years",y="Temperature",color="Temperature",title="The evidence of global warming")+
  scale_x_continuous(breaks = seq(1987,2016,5))+
  scale_y_continuous(breaks = seq(0,30,5))+
  geom_abline(intercept=MR2[1,3],slope=0,lty=2,col="black")+
  geom_abline(intercept=MR2[31,3],slope=0,lty=2,col="black")+
  geom_abline(intercept=MR2[61,3],slope=0,lty=2,col="black")
```



4 Discussion

From the above line charts, we can find:
- The air temperature and water temperature both increase fastest in the first year.
- Although air temperature and water temperature fluctuate among years, they both increase gradually in the 30 years.
- Although the temperature difference between air and water fluctuates among years, it decreases gradually in the 30 years. Therefore, we have found the effect of global warming that the air temperature and water temperature increase gradually in the 30 years.

References

[1] Alina, B., Stephanie, P., (2017). *Live Science* [online]. Available from: https://www.livescience.com/37057-global-warming-effects.html [accessed 24 September 2020].
[2] Reguero, B.G., Losada, I.J. & Méndez, F.J. A recent increase in global wave power as a consequence of oceanic warming. *Nat Commun* **10**, 205 (2019). https://doi.org/10.1038/s41467-018-08066-0
[3] Ball, P. Global greenhouse affects air pressure. *Nature* (2003). https://doi.org/10.1038/news030317-6
[4] Gerald, E., (2017). *SCIENCING* [online]. Available from: https://sciencing.com/calculate-solar-insolation-8435082.html [accessed 24 September 2020].
[5] *Wikipedia* [online]. Available from: https://en.wikipedia.org/wiki/Climate_of_Massachusetts [accessed 24 September 2020].
[6] *SEA TEMPERATURE INFO* [online]. Available from: https://seatemperature.info/boston-water-temperature.html#:~:text=Water%20temperature%20in%20Boston%20today,%C2%B0C%2F65.7%C2%B0F [accessed 24 September 2020].