

```
## Error: <text>:6:0: unexpected end of input
## 4:  png_private =
## 5:  gif_public =
##    ^
```

```
get_locationid <- function(FIPS){
  fips = ncdc_locs(locationcategoryid='ST', limit=55)
  temp <- data.frame(State = fips$data$name[FIPS],
                     id = fips$data$id[FIPS])
  temp$id <- as.character(temp$id)
  temp$State <- as.character(temp$State)
  return(temp)
}
```

```
## 'data.frame': 1 obs. of 2 variables:
## $ State: chr "Alabama"
## $ id : chr "FIPS:01"
```

Weather Station Records and Communicating Climate Change—Alabama

Marc Los Huertos

May 29, 2022

Contents

1	Evaluating Terrestrial Meteorological Data	3
1.1	Selected History of Climate Science	3
1.2	Why Look at Individual Stations?	4
1.3	Approach	4
1.3.1	NOAA Data Records	4
1.3.2	rNOAA Package and R	4
1.4	Selecting Weather Records by State	5
1.4.1	Identify List of State IDs (FIPS)	5
1.4.2	Selection Stations	5
2	Gathering Weather Record Datasets	6
2.1	Main Datasets of Interest	6
2.2	Functions to Collect and Clean GSOM	6
2.2.1	Functions to Report Probabilities	7
2.3	GSOM: Retrieve and Clean Data	9
2.4	CHCND: Retrieve and Clean Data	10
3	Data Analysis Processes	12
3.1	Mapping the Location of Weather Station	12
3.2	Using a Linear Model Monthly Trends	14
3.2.1	Trends in Tabular Formats	15
3.2.2	Defining TMAXmonth and TMINmonth	16
4	Communicating Long-term Weather Records	16
4.1	Complete Records vs. Post 1975 Trends	16
4.2	Filtering Seasonal Effect	18
4.2.1	Method 1 Filtering by Monthly Mean	20
4.2.2	Method 2: Polynomial Filter	22
4.3	Extreme Events—Using Daily Records	23
4.3.1	Complicated Nature of Rainfall Patterns	23
4.3.2	Drought	25

4.4	Record Setting Temperature Records	29
4.5	Iterate TMAX vs. Month Boxplots	39
4.6	Four Plots Compelling Figures	40
4.7	KISS	96
4.7.1	Change Point Analysis	96
4.8	Temp & Precipitation Probability	99
4.9	Using library densEstBayes	102
5	Animated GIFs	107
5.1	Probability Distributions	107
5.2	4 Weather Trend Plots	109
6	OLD Stuff	112
7	Other attempts...	112
7.1	Evaluating Records	113
7.2	Export Options	113
8	Sea Surface Temperature Data – SURP PROJECT WAITING TO HAPPEN	113
9	Satellite Data	114
10	Ice-Core Data	114
11	Conclusions	114

1 Evaluating Terrestrial Meteorological Data

1.1 Selected History of Climate Science

Geologists have known the climate has been changing over the Earth’s history. But what causes these changes has been a major research area for over 100 years. There are numerous drivers that contribute to changing climates – including the arrangement of the continents on the planet, the distance to the sun, energy generated by the sun, volcanic activity, and the composition of the Earth’s atmosphere.

It’s the last one that we’ll spend time because the Earth’s temperature are changing pretty dramatically over the last 100 years and the cause is no mystery – the human activity that has released carbon dioxide (CO₂) into the atmosphere. The two main sources of CO₂ is from land use change, e.g. deforestation, and the burning of fossil fuels, e.g. coal, oil, and natural gas.

The first to propose the role of CO₂ on the Earth’s atmosphere was a Swedish scientist Svante Arrhenius, who figured out that CO₂ absorbs infrared light. Moreover, he deduced that the Earth’s temperature was actually warmer than it might otherwise be if CO₂ was not part of the Earth’s atmosphere.

1.2 Why Look at Individual Stations?

I don't think there is a single, perfect way to analyze and communicate climate change. But the beauty of the network of stations in the USA and around the world is that these stations record weather as experienced by local people. And while individual stations may not represent the overall regional and global patterns well, this gives us a mechanism to connect local experiences to regional or global processes.

Of course, some may fixate on the local pattern and remain unconvinced of the larger context and for those folks, there may be better ways to communicate climate data.

However, I would be remiss in failing to mention that some may fixate on local patterns and use these patterns to ignore or to dismiss the patterns in other regions.

Finally, the impacts of climate change are highly specific to the region in question. Thus, once someone understands the impacts on climate change in their region, they may not be able to appreciate how different the climate impacts might affect other peoples, who may be more vulnerable, around the globe.

Thus, with these weaknesses in mind, I will pursue this project with an eye to address these other issues at later stages.

1.3 Approach

1.3.1 NOAA Data Records

The US National Oceanic and Atmospheric Administration (NOAA) maintains several sources of digital weather data from the USA and beyond. These data have been collected from stations around the country to support a wide range of human activities that include farming, aviation, shipping, and even armed conflict.

At various times, these records have been used to evaluate long-term climate change with varying success. Without a doubt, these data are not perfect, but they remain that foundation of an effective and professionally maintained environmental monitoring program that engenders integrity, even when facing budget cuts.

I will use these data to select for a station with a long record for each state in the USA. Future projects might evaluate the record for stations around the world, but we will see about that.

1.3.2 rNOAA Package and R

R is an open source programming environment that has become one of the most popular tools for statisticians and data scientists. Capitalizing on the open source framework, a wide range of libraries or packages have been developed to facilitate data processing, analysis, and graphical displays. One such package is rNOAA developed to collect and display climate records stored on NOAA servers.

Using the package requires the use of a key. To maintain the integrity of the key, it's best to avoid posting the key in a public repository and to encryp the key to ensure it's not abused.

1.4 Selecting Weather Records by State

There are numerous ways to analyze temperature records, where stations can be analyzed individually or records could be sampled and analyzed in spatially in grids. Each of these are valid approaches depending on the question to be addressed.

In this case the question is “Based on the longest state meterological record, is there a temperature trend?”

1.4.1 Identify List of State IDs (FIPS)

Using the rNOAA library in R, we can queary NOAA’s database to identify station codes (FIPS) by state. With the states and some territories, there are 55 FIPS for US weather stations.

rNOAA has a simple function to list for each of the states and the weather stations in each. I use `ncdc_locs()` functions to select each state and `ncdc_station()` to obtain the station ids with the longest records.

```
# List of States (alpha beta)
ncdc_locs(locationcategoryid='ST', limit=55)
```

The function queries the NOAA website and retrieves state codes, “FIPS:XX”. Each state has a number of weather stations,¹ some with a long record, some with a short record, and some with numerous interruptions. Our goal is to select a long record with few missing data.

1.4.2 Selection Stations

With the state ids, we can evaluate the metadata for all the weather stations, which will work to get the longest records, using `ncdc_stations()`.

First, we subset the data for stations that actively collecting data. Then we’ll sort to the active stations to find the one with the longest records. We will use these stations for our analysis.

```
GSOM_Stations <- ncdc_stations(datasetid='GSOM',
                               datatypeid = c("TMAX", "TMIN"),
                               locationid=fips$id, limit=1000,
                               sortfield = 'maxdate', sortorder='desc')

GSOM_Recent =
  GSOM_Stations$data[GSOM_Stations$data$maxdate>='2021-11-01',]
```

¹Project Idea: It would be nice to make a map of how concentrated the stations spatially.

```

GSOM_Coverage =
  GSOM_Recent[GSOM_Recent$datacoverage > 0.92,]
GSOM_Sorted = GSOM_Coverage[order(GSOM_Coverage$mindate),]

GSOM_Longest = GSOM_Sorted[1,] #Pick longest
# Second and Third for Comparisons
# GSOM_Longest = GSOM_Sorted[2,]
# GSOM_Longest = GSOM_Sorted[3,]

```

The record selected has the following metadata associated with it, which will be used for naming, labeling, and mapping.

```

elevation mindate maxdate latitude name datacoverage 101 136.6 1888-02-
01 2022-03-01 33.4163 TALLADEGA, AL US 0.9236 id elevationUnit longitude
101 GHCND:USC00018024 METERS -86.135 [1] 1888

```

2 Gathering Weather Record Datasets

2.1 Main Datasets of Interest

GSOM

CHCND

CHCNM

2.2 Functions to Collect and Clean GSOM

To collect the data, I used a short function, but the download time is painfully slow because only 1 year can be obtained at a time. Might want to get a work around for this at some point.

```

get_GSOM <- function(stid, datatype) {
  wtr<-list() # create an empty list
  for (i in startyear:2021) {
    start_date <- paste0(i, "-01-01")
    end_date <- paste0(i, "-12-31")

#save data portion to the list (elements named for the year
    wtr[[as.character(i)]] <- ncdc(datasetid='GSOM',
      stationid=stid, datatypeid=datatype, startdate =
        start_date, enddate = end_date, limit=400)$data
  }
#return the full list of data frames
  return(wtr)
}

```

```

stid = substr(GSOM_Longest$id, 7, 17)

get_GSOM2 <- function(stid){
  http.csv <- "https://www.ncei.noaa.gov/data/global-summary-of-the-month/access/"
  read.csv(paste(http.csv, stid, ".csv", sep=""))
}

# GSOM <- get_GSOM2(stid)

```

Functions to bin data into decades and scores.

```

floor_decade <- function(value){
  return(value - value %% 10) }
# Test Function
floor_decade(c(1883,1988))

## [1] 1880 1980

floor_score <- function(value){
  return(value - value %% 20) }

# Test Function
floor_score(c(1883,1893,1910, 1932,1940))

## [1] 1880 1880 1900 1920 1940

```

The function relies on two inputs, the station id and the measured parameter – TMAX and TMIN in this case. After that, the data needs to be clean up quite a bit.

Furthermore, I have converted units to Farenheit, which is not my favorite, but important for US consumption.

2.2.1 Functions to Report Probabilities

```

report_prob <-function(pvalue){
  if(pvalue > 0.05) return("> 0.05 (Not Significant)")
  if(pvalue < 0.05 & pvalue >= 0.001) return(
    paste("=", round(pvalue, 3), "(Statistically Significant)"))
  #if(pvalue < 0.01) print(round(pvalue, 4))
  if(pvalue < 0.001) return("< 0.001 (Statisically Significant)")
}

#test function
report_prob(0.0032)

```

```

report_prob2 <-function(lm){
  # lm=GSOM.lm
  if(anova(lm)$'Pr(>F)'[1] > 0.05){
    return("p-value > 0.05 (Not Significant)")
  }
  if(anova(lm)$'Pr(>F)'[1] < 0.05 &
    anova(lm)$'Pr(>F)'[1] >= 0.001){
    return(paste("Change ", round(coef(lm)[2]*356.25*100, 1),
      "/100 years, ", "p-value = ", round(anova(lm)$'Pr(>F)'[1], 3),
      "(Statistically Significant)", sep=""))
  }
  if(anova(lm)$'Pr(>F)'[1] < 0.001) {
    return(paste("Change ", round(coef(lm)[2]*325.25*100, 1),
      "/100 years, ", "p-value < 0.001 (Statistically Significant)",
      sep=""))
  }
}

report_prob3 <-function(lm){
  #lm=Drought.run.lm
  temp = data.frame(change = NA, p_value=NA, note=NA)
  if(anova(lm)$'Pr(>F)'[1] > 0.05){
    temp[,1] <- "";
    temp[,2] <- "p-value > 0.05";
    temp[,3] <- "(Not Significant)"
    return(temp)
  }
  if(anova(lm)$'Pr(>F)'[1] < 0.05 &
    anova(lm)$'Pr(>F)'[1] >= 0.001){
    temp[,1] <- paste("Change: ",
      round(coef(lm)[2]*356.25*100, 1), "F/100 years", sep="");
    temp[,2] <- paste("p-value = ",
      round(anova(lm)$'Pr(>F)'[1], 3), sep="");
    temp[,3] <- " (Statistically Significant)"
    return(temp)
  }
  if(anova(lm)$'Pr(>F)'[1] < 0.001) {
    temp[,1] <- paste("Change: ", round(coef(lm)[2]*356.25*100, 1), "F/100 years", sep="");
    temp[,2] <- "p-value < 0.001";
    temp[,3] <- " (Statistically Significant)"
    return(temp)
  }
}

```


2.3 GSOM: Retrieve and Clean Data

```
GSOM_TMAX <- get_GSOM(GSOM_Longest$id, 'TMAX')
GSOM_TMIN <- get_GSOM(GSOM_Longest$id, 'TMIN')
GSOM_PPT  <- get_GSOM(GSOM_Longest$id, 'PRCP')

# Bind the dataframes in the list
# together into one large dataframe

tbl_TMAX <- dplyr::bind_rows(GSOM_TMAX)
tbl_TMIN <- dplyr::bind_rows(GSOM_TMIN)
tbl_PPT  <- dplyr::bind_rows(GSOM_PPT)

class(tbl_TMAX) # [1] "tbl_df"      "tbl"      "data.frame"

## [1] "tbl_df"      "tbl"      "data.frame"

dfTbl_TMAX = as.data.frame(tbl_TMAX)
dfTbl_TMIN = as.data.frame(tbl_TMIN)
dfTbl_PPT  = as.data.frame(tbl_PPT)

class(dfTbl_TMAX) # [1] "data.frame"

## [1] "data.frame"

dfTbl_TMAX$TMAX = dfTbl_TMAX$value*9/5+32
dfTbl_TMIN$TMIN = dfTbl_TMIN$value*9/5+32
dfTbl_PPT$PPT  = dfTbl_PPT$value

dfTbl_TMAX$Date = as.Date(dfTbl_TMAX$date)
dfTbl_TMIN$Date = as.Date(dfTbl_TMIN$date)
dfTbl_PPT$Date  = as.Date(dfTbl_PPT$date)

dfTbl_TMAX <- subset(dfTbl_TMAX, select=c(Date, station, TMAX))
dfTbl_TMIN <- subset(dfTbl_TMIN, select=c(Date, TMIN))
dfTbl_PPT  <- subset(dfTbl_PPT,  select=c(Date, PPT))

dfTbl_TMAX[1,]

##           Date           station TMAX
## 1 1893-09-01  GHCND:USC00018024 83.75

GSOM <- merge(dfTbl_TMAX, dfTbl_TMIN, by="Date")
GSOM <- merge(GSOM, dfTbl_PPT, by="Date")

GSOM$Month = as.numeric(format(as.Date(GSOM$Date), format = "%m"))
```

```

GSOM$Year = as.numeric(format(as.Date(GSOM$Date), format = "%Y"))
GSOM$Decade = floor_decade(GSOM$Year)
GSOM$Score = floor_score(GSOM$Year)

#seq(min(GSOM$Year))

str(GSOM)

## 'data.frame': 1369 obs. of 9 variables:
## $ Date : Date, format: "1894-08-01" "1898-02-01" ...
## $ station: chr "GHCND:USC00018024" "GHCND:USC00018024" "GHCND:USC00018024" "GHCND:USC00018024" ...
## $ TMAX : num 87.8 56.2 69 68.5 87.2 ...
## $ TMIN : num 69.8 34.5 49.5 45.7 59.7 ...
## $ PPT : num 213.4 19.6 99.8 127.6 7.6 ...
## $ Month : num 8 2 3 4 5 6 7 8 9 10 ...
## $ Year : num 1894 1898 1898 1898 1898 ...
## $ Decade : num 1890 1890 1890 1890 1890 1890 1890 1890 1890 1890 ...
## $ Score : num 1880 1880 1880 1880 1880 1880 1880 1880 1880 1880 ...

```

2.4 CHCND: Retrieve and Clean Data

CHCND have been bias corrected...

```

GSOM_Longest$id

## [1] "GHCND:USC00018024"

stid = substr(GSOM_Longest$id, 7, 17)

CHCND.https <- "https://www.ncei.noaa.gov/data/global-historical-climatology-network-daily/"

get_CHCND <- function(stid) {
  #stid = "USC00013511"
  import <- read.csv(paste(CHCND.https, stid, ".csv", sep=""))
  selected = subset(import, select=c("DATE", "TMAX", "TMIN", "PRCP"))
  selected$TMAX = selected$TMAX/10*(9/5)+32
  selected$TMIN = selected$TMIN/10*(9/5)+32
  selected$Date = as.Date(selected$DATE)
  selected = subset(selected, select=-c("DATE"))
  #selected = selected[complete.cases(selected$TMAX),]
  selected
}

CHCND <- get_CHCND(stid); nrow(CHCND)

## [1] 46584

```

```

str(CHCND)

## 'data.frame': 46584 obs. of 4 variables:
## $ TMAX: num NA NA NA NA NA NA NA NA NA NA ...
## $ TMIN: num NA NA NA NA NA NA NA NA NA NA ...
## $ PRCP: int 0 0 0 0 274 0 538 0 0 3 ...
## $ Date: Date, format: "1888-02-01" "1888-02-02" ...

# Fill in Missing Dates
start_date = as.Date(paste0(min(year(CHCND$Date)), "-01-01"), format="%Y-%m-%d")
end_date = today()
dates = seq(start_date, end_date, by=1)
continuous_dates<-data.frame(Date=dates, Year=year(dates), yday = yday(dates))

str(continuous_dates)

## 'data.frame': 49092 obs. of 3 variables:
## $ Date: Date, format: "1888-01-01" "1888-01-02" ...
## $ Year: num 1888 1888 1888 1888 1888 ...
## $ yday: num 1 2 3 4 5 6 7 8 9 10 ...

CHCND <- merge(continuous_dates, CHCND, by="Date", all.x=TRUE)
names(CHCND)

## [1] "Date" "Year" "yday" "TMAX" "TMIN" "PRCP"

#CHCND$Year = as.numeric(format(as.Date(CHCND$Date), format = "%Y"))
CHCND$YearDay = CHCND$Year + yday(CHCND$Date)/366
CHCND$mmd <- format(CHCND$Date, "%m-%d")
CHCND$Month = as.numeric(format(as.Date(CHCND$Date), format = "%m"))
CHCND$Month.name = factor(format(as.Date(CHCND$Date), format = "%b"),
  levels = c("Jan", "Feb", "Mar", "Apr", "May", "Jun", "Jul",
    "Aug", "Sep", "Oct", "Nov", "Dec"))
CHCND$Season = "Winter"
CHCND$Season[CHCND$Month==4 | CHCND$Month==5 | CHCND$Month==6] = "Spring"
CHCND$Season[CHCND$Month==7 | CHCND$Month==8 | CHCND$Month==9] = "Summer"
CHCND$Season[CHCND$Month==10 | CHCND$Month==11 | CHCND$Month==12] = "Fall"

range(CHCND$TMAX, na.rm=T)

## [1] 15.08 109.04

spread = sd(CHCND$TMAX, na.rm=T)*4
TMAX_mean = mean(CHCND$TMAX, na.rm=T)

#CHCND$TMAX[complete.cases(CHCND$TMAX) & CHCND$TMAX > TMAX_mean+spread] <-NA
#CHCND$TMAX[complete.cases(CHCND$TMAX) & CHCND$TMAX < TMAX_mean-spread] <-NA
range(CHCND$TMAX, na.rm=T)

```

```
## [1] 15.08 109.04

head(CHCND)

##           Date Year yday TMAX TMIN PRCP YearDay mmdd Month Month.name Season
## 1 1888-01-01 1888     1   NA   NA   NA 1888.003 01-01     1      Jan Winter
## 2 1888-01-02 1888     2   NA   NA   NA 1888.005 01-02     1      Jan Winter
## 3 1888-01-03 1888     3   NA   NA   NA 1888.008 01-03     1      Jan Winter
## 4 1888-01-04 1888     4   NA   NA   NA 1888.011 01-04     1      Jan Winter
## 5 1888-01-05 1888     5   NA   NA   NA 1888.014 01-05     1      Jan Winter
## 6 1888-01-06 1888     6   NA   NA   NA 1888.016 01-06     1      Jan Winter
```

3 Data Analysis Processes

3.1 Mapping the Location of Weather Station

```
GSOM_Longest$name

## [1] "TALLADEGA, AL US"

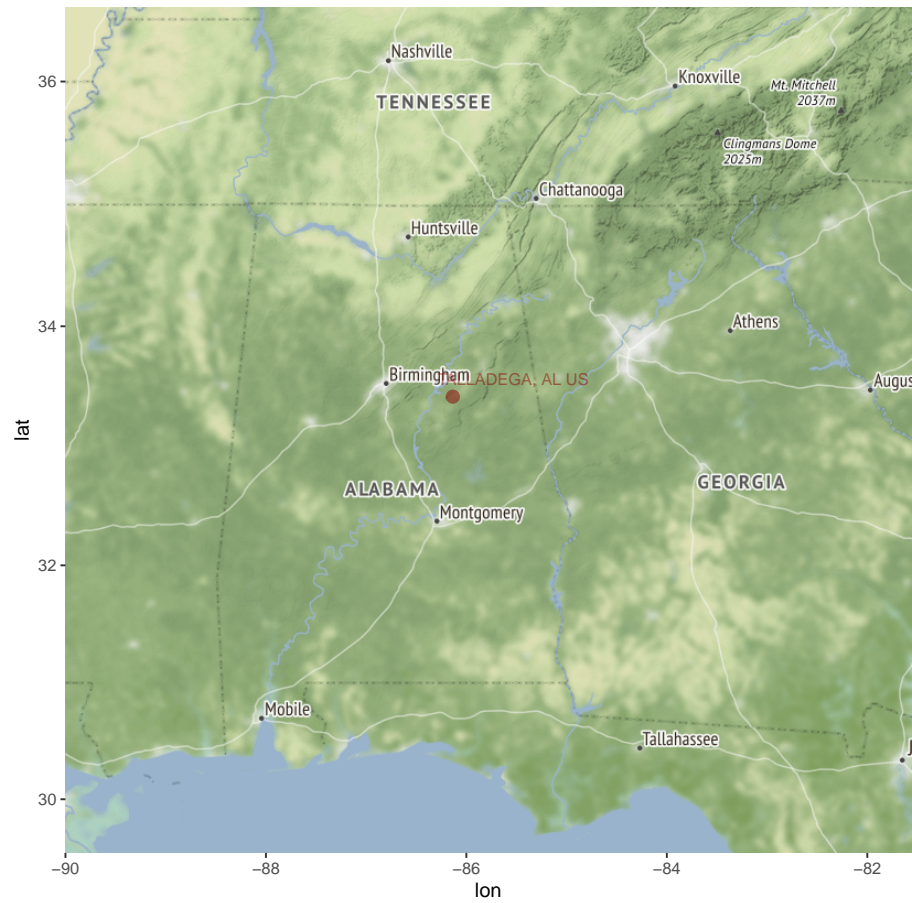
lat = GSOM_Longest$latitude
lon = GSOM_Longest$longitude
station = c(lon, lat)
station.df <- data.frame(lon = GSOM_Longest$longitude,
                          lat = GSOM_Longest$latitude,
                          Station = GSOM_Longest$name);

str(station.df)

## 'data.frame': 1 obs. of 3 variables:
## $ lon      : num -86.1
## $ lat      : num 33.4
## $ Station: Factor w/ 1 level "TALLADEGA, AL US": 1

myMap <- get_map(location=station, zoom=7, scale =2,
source="stamen", matype="terrain", messaging = FALSE, crop=FALSE)

ggmap(myMap) + geom_point(aes(x = lon, y = lat), data = station.df, alpha = .5, color="darkblue")
```



```
#zoom = 11, scale = 2, mapttype = 'watercolor',

png(paste0("png/", fips$State, "-", stid, "-MAP.png"),
    width = 480, height = 480, units = "px",
    pointsize = 12, bg = "white")

# For PNG?
ggmap(myMap)+
geom_point(aes(x = lon, y = lat), data = station.df,
  alpha = .5, color="darkred", size = 3) +
  geom_text(aes(x = lon, y = lat, label=Station),
    data = station.df, alpha = .5, color="darkred",
    size = 3, hjust=.1, vjust=-1)
dev.off()

## pdf
```

3.2 Using a Linear Model Monthly Trends

I used a linear model (`lm()`) to evaluate the long term trend for each month to determine which, if any, have long-term trends. At somepoint, I'll have to the stats correcting for the autocorrelation using a autoregressive model.

Evaluate both TMAX and TMIN in GSOM by Year using `MonthEvalStats()` function.

```
MonthEvalStatsOLD <- function(GSOM) {
  sumstats = NA
  for (m in 1:12){
    TMIN.lm = lm(TMIN~Date, GSOM[GSOM$Month==m,])
    TMAX.lm = lm(TMAX~Date, GSOM[GSOM$Month==m,])
    PPT.lm = lm(PPT~Date, GSOM[GSOM$Month==m,])

    sumstats = rbind(sumstats,
      data.frame(Month = m, Param="TMIN", Slope = coef(TMIN.lm)[2],
        r2 = summary(TMIN.lm)$r.squared, p_value= anova(TMIN.lm)$'Pr(>F)'[1]),
      data.frame(Month = m, Param="TMAX", Slope = coef(TMAX.lm)[2],
        r2 = summary(TMAX.lm)$r.squared, p_value= anova(TMAX.lm)$'Pr(>F)'[1]),
      data.frame(Month= m, Param="PPT", Slope = coef(PPT.lm)[2],
        r2 = summary(PPT.lm)$r.squared, p_value= anova(PPT.lm)$'Pr(>F)'[1]))
  }

  sumstats=data.frame(sumstats)[-1,]
  rownames(sumstats)<-NULL

  sumstats$Symbol = ""
  sumstats$Symbol[sumstats$p_value < 0.05] = "*"
  sumstats$Symbol[sumstats$p_value < 0.01] = "**"
  sumstats$Symbol[sumstats$p_value < 0.001] = "***"
  sumstats[,c(7,9)]
  return(sumstats)
}

MonthEvalStats <- function(GSOM) {
  sumstats = NA
  for (m in 1:12){
    TMIN.lm = lm(TMIN~Date, GSOM[GSOM$Month==m,])
    TMAX.lm = lm(TMAX~Date, GSOM[GSOM$Month==m,])
    PPT.lm = lm(PPT~Date, GSOM[GSOM$Month==m,])
```

```

sumstats = rbind(sumstats,
  data.frame(Month = m, Param="TMIN", Slope = coef(TMIN.lm)[2],
    r2 = summary(TMIN.lm)$r.squared, p_value= anova(TMIN.lm)$'Pr(>F)'[1]),
  data.frame(Month = m, Param="TMAX", Slope = coef(TMAX.lm)[2],
    r2 = summary(TMAX.lm)$r.squared, p_value= anova(TMAX.lm)$'Pr(>F)'[1]),
  data.frame(Month= m, Param="PPT", Slope = coef(PPT.lm)[2],
    r2 = summary(PPT.lm)$r.squared, p_value= anova(PPT.lm)$'Pr(>F)'[1]))

} #end loop

sumstats=data.frame(sumstats)[-1,]
rownames(sumstats)<-NULL
head(sumstats)

sumstats$Symbol = ""
sumstats$Symbol[sumstats$p_value < 0.05] = "*"
sumstats$Symbol[sumstats$p_value < 0.01] = "**"
sumstats$Symbol[sumstats$p_value < 0.001] = "***"
return(sumstats)
}

# test function
sumstats = MonthEvalStats(GSOM[500:4000,])

```

3.2.1 Trends in Tabular Formats

Admittedly, determining the months with the biggest changes isn't a very good approach for hypothesize testing – it's more like a fishing expedition, but as long as we understand the difference between an a priori hypothesis and an exploratory analysis, we should be okay if we make appropriate conclusions.

```

# Selecting Most Important Monthly Changes (TMAX overwrites)
#sumstats = MonthEvalStats(GSOM)

TMIN_Increase_month = with(sumstats[sumstats$Param=="TMIN",],
  Month[Slope==max(Slope, na.rm=T)])
TMIN_Decrease_month = with(sumstats[sumstats$Param=="TMIN",],
  Month[Slope==min(Slope, na.rm=T)])
TMAX_Increase_month = with(sumstats[sumstats$Param=="TMAX",],
  Month[Slope==max(Slope, na.rm=T)])
TMAX_Decrease_month = with(sumstats[sumstats$Param=="TMAX",],
  Month[Slope==min(Slope, na.rm=T)])
PPT_Increase_month = with(sumstats[sumstats$Param=="PPT",],
  Month[Slope==max(Slope, na.rm=T)])
PPT_Decrease_month = with(sumstats[sumstats$Param=="PPT",],

```

```
Month[Slope==min(Slope, na.rm=T)])
```

For this section, we'll look to see what months had the greatest changes for both TMIN and TMAX. By looking at significant slopes in whatever direction, we might learn if warming is really the dominant pattern.

Table ?? summarizes the monthly trends for TMAX.

```
## Error in print.xtable(TMIN.xtbl, type = "latex", comment = FALSE,
caption = "Monthly TMIN Trends", : argument 4 matches multiple formal
arguments
```

Table ?? summarizes the monthly trends for TMAX.

```
## Error in print.xtable(TMAX.xtbl, type = "latex", comment = FALSE,
caption = "TMAX Trends by Month", : argument 4 matches multiple formal
arguments
```

PPT changes are tricky to capture and I'll have to keep working on this (Table ??).

```
## Error in print.xtable(PPT.xtbl, type = "latex", caption = "Precipitation
Trends by Month", : argument 3 matches multiple formal arguments
```

3.2.2 Defining TMAXmonth and TMINmonth

```
TMINSlopeMax = max(abs(sumstats$Slope)[sumstats$Param=="TMIN"])
TMAXSlopeMax = max(abs(sumstats$Slope)[sumstats$Param=="TMAX"])

TMINmonthMax = as.numeric(subset(sumstats, select=Month, subset=abs(Slope)==TMINSlopeMax))
TMAXmonthMax = as.numeric(subset(sumstats, select=Month, subset=abs(Slope)==TMAXSlopeMax))
```

The greatest changes for Station GHCND:USC00018024

4 Communicating Long-term Weather Records

4.1 Complete Records vs. Post 1975 Trends

Communicating climate change based on station records is tricky. The long-term record would on the surface to be the most robust, but several issues arise with a naive analytical approach – my favorite!


```

setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/")

GSOM_1975.png = paste0(fips$State, "-", stid, "-GSOM_1975.png")

png(paste0(png_public, GSOM_1975.png), width = 480, height = 320, units = "px", pointsize =

## Error in paste0(png_public, GSOM_1975.png): object 'png_public'
not found

par(las=1, mfrow=c(1,1))
plot(TMAX~Date, GSOM, pch=20, cex=.5, col="grey", ylab="F", main=paste0(fips$State, "-", stid, "-GSOM_1975.png"))
GSOM.lm = lm(TMAX~Date, GSOM)
pred_dates <-data.frame(Date = GSOM$Date);
nrow(pred_dates);# pred_dates

## [1] 1369

#Predicts the values with confidence interval
ci <- predict(GSOM.lm, newdata = pred_dates,
              interval = 'confidence')
lines(pred_dates$Date, as.numeric(ci[,1]), col="gray50")

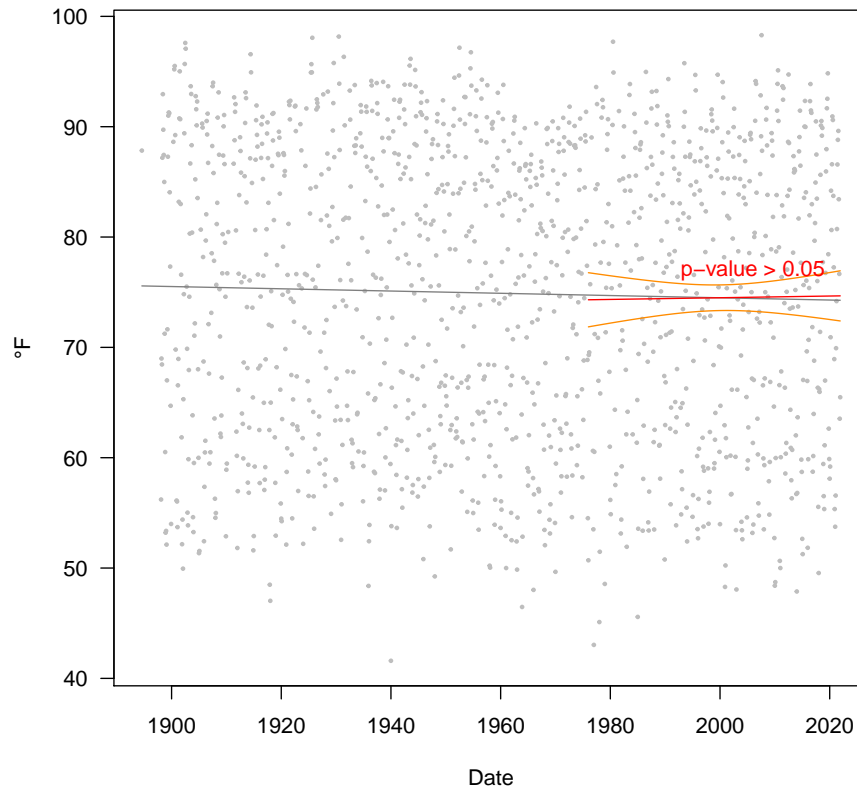
# Post 1975
GSOM.lm = lm(TMAX~Date, GSOM[GSOM$Year>1975,])
pred_dates <-data.frame(Date = GSOM$Date[GSOM$Year>1975]);
nrow(pred_dates); #pred_dates

## [1] 510

#Predicts the values with confidence interval
ci <- predict(GSOM.lm, newdata = pred_dates,
              interval = 'confidence')
lines(pred_dates$Date, as.numeric(ci[,1]), col="red")
lines(pred_dates$Date, as.numeric(ci[,2]), col="darkorange")
lines(pred_dates$Date, ci[,3], col="darkorange")
location_index = round(length(GSOM[GSOM$Year>1975,]$Date) * 0.99,0)
text(pred_dates$Date[location_index], ci[location_index,3],
     paste(report_prob3(GSOM.lm))[2], pos=2, cex=1.0, col="red")

```

Alabama-USC00018024



```
#abline(coef(lm(TMAX~Date, GSOM)), col="black")
#abline(coef(lm(TMAX~Date, GSOM[GSOM$Year>1975,])), col="red")
dev.off()

## null device
##      1
```

The noise in the data may suggest that no trend is present (Figure 1). It's tricky because the seasonal variation dominates the source of variation. In other words the intra-annual variation exceeds the inter-annual variation, making signal detection very difficult. Is there a way to "filter" out the seasonal effect? Yes, let's see how that works next.

4.2 Filtering Seasonal Effect

There are several ways to filter out seasonal effects. The easiest way is subtract the mean value for each date, but that's tricky because every four years there is

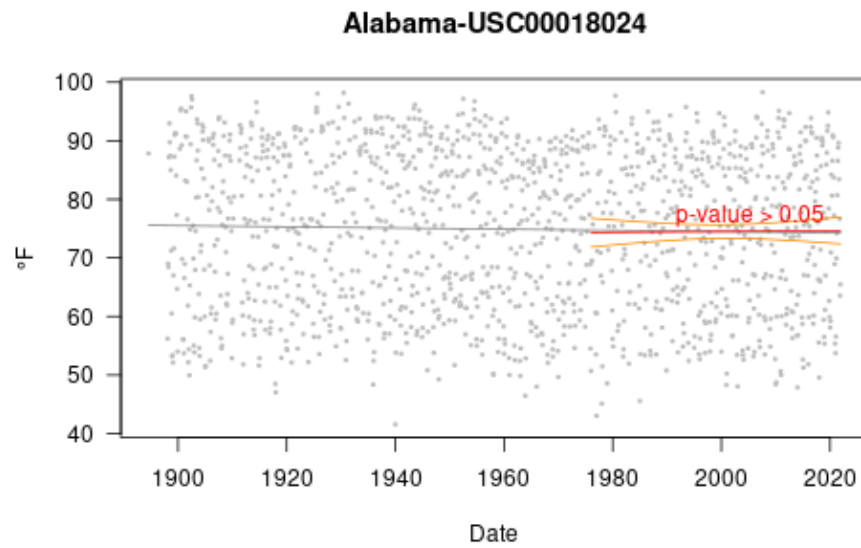


Figure 1: The climate trends from full record and post 1975 data. These data have a high level of variability due to seasonality effects that have not been filtered out.

an extra day in February – although there are ways to deal with this, a more straight forward way is to use mean monthly values to capture the seasonality for each month. With 12 months, this is a pretty good approach because there is pretty good resolution.

4.2.1 Method 1 Filtering by Monthly Mean

```
TMAX.Monthly.means = aggregate(TMAX~Month, data=GSOM, mean)
names(TMAX.Monthly.means)=c("Month", "TMAXmean")
GSOM2 = merge(GSOM, TMAX.Monthly.means, by="Month")
GSOM2$TMAX.anom = GSOM2$TMAX - GSOM2$TMAXmean

TMIN.Monthly.means = aggregate(TMIN~Month, GSOM, mean)
names(TMIN.Monthly.means)=c("Month", "TMINmean")
GSOM2 = merge(GSOM2, TMIN.Monthly.means, by="Month")
GSOM2$TMIN.anom = GSOM2$TMIN - GSOM2$TMINmean

PPT.Monthly.means = aggregate(PPT~Month, GSOM, mean)
names(PPT.Monthly.means)=c("Month", "PPTmean")
GSOM2 = merge(GSOM2, PPT.Monthly.means, by="Month")
GSOM2$PPT.anom = GSOM2$PPT - GSOM2$PPTmean

# Sort by date
GSOM2 <- GSOM2[order(GSOM2$Date),]

GSOM_anomaly.png = paste0(fips$State, "-", stdid, "-GSOM_anomaly1975.png")

png(paste0(png_public, GSOM_anomaly.png),
    width = 480, height = 320, units = "px",
    pointsize = 12, bg = "white")

## Error in paste0(png_public, GSOM_anomaly.png): object 'png_public'
## not found

par(las=1, mfrow=c(1,1))
par(las=1)
plot(TMAX.anom~Date, GSOM2, pch=20, cex=.5,
     col="grey", ylab="Max. Temp (anomaly) F",
     main=paste0("Seasonally Filtered", fips$State, " (", stdid, ")",
                 report_prob3(GSOM.lm)[3]))
GSOM.lm = lm(TMAX.anom~Date, GSOM2)
pred_dates <-data.frame(Date = GSOM2$Date);
nrow(pred_dates); #pred_dates

## [1] 1369
```

```

#Predicts the values with confidence interval
ci <- predict(GSOM.lm, newdata = pred_dates,
              interval = 'confidence')
lines(pred_dates$Date, as.numeric(ci[,1]), col="gray50")

ymax=max(GSOM2$TMAX.anom) - (max(GSOM2$TMAX.anom)-min(GSOM2$TMAX.anom))*0.3
ymax2 <- ymax - (max(GSOM2$TMAX.anom)-min(GSOM2$TMAX.anom))*0.1

location_index = round(length(GSOM2[GSOM2$Year>1975,]$Date) * 0.99,3)
text(pred_dates$Date[location_index], ymax,
      paste(report_prob3(GSOM.lm))[1], pos=2, cex=.9)
text(pred_dates$Date[location_index], ymax2,
      paste(report_prob3(GSOM.lm))[2], pos=2, cex=.9)

# Post 1975
GSOM.lm = lm(TMAX.anom~Date, GSOM2[GSOM2$Year>1975,])
pred_dates <-data.frame(Date = GSOM2$Date[GSOM2$Year>1975]);
nrow(pred_dates); #pred_dates

## [1] 510

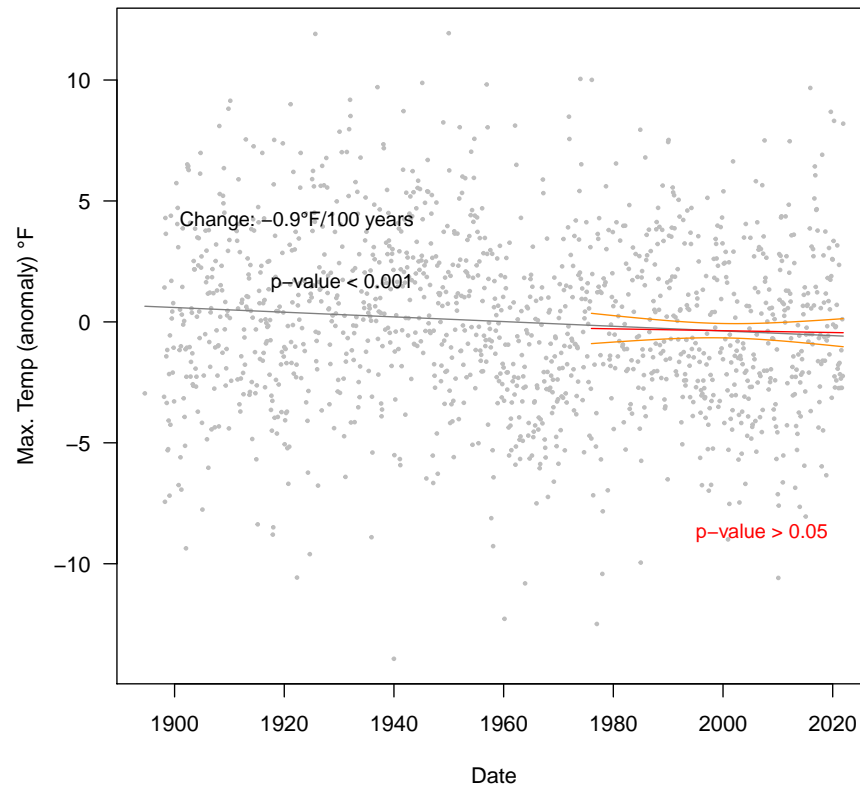
#Predicts the values with confidence interval
ci <- predict(GSOM.lm, newdata = pred_dates,
              interval = 'confidence')
lines(pred_dates$Date, as.numeric(ci[,1]), col="red")
lines(pred_dates$Date, as.numeric(ci[,2]), col="darkorange")
lines(pred_dates$Date, ci[,3], col="darkorange")

ymax=max(GSOM2$TMAX.anom) - (max(GSOM2$TMAX.anom)-min(GSOM2$TMAX.anom))*0.7
ymax2 <- ymax - (max(GSOM2$TMAX.anom)-min(GSOM2$TMAX.anom))*0.1

location_index = round(length(GSOM2[GSOM2$Year>1975,]$Date) * 0.99,0)
text(pred_dates$Date[location_index], ymax,
      paste(report_prob3(GSOM.lm))[1], pos=2, cex=.9, col="red")
text(pred_dates$Date[location_index], ymax2,
      paste(report_prob3(GSOM.lm))[2], pos=2, cex=.9, col="red")

```

Seasonally Filtered Alabama (USC00018024)(Not Significant)



```
dev.off()

## null device
##      1
```

And to see what we created, see Figure 2.

4.2.2 Method 2: Polynomial Filter

Project to be followed up with.

```
# fit polynomial:  $x^2 \cdot b_1 + x \cdot b_2 + \dots + b_n$ 

# create time series object
#X = [i%365 for i in range(0, len(series))]
# y = series.values
```

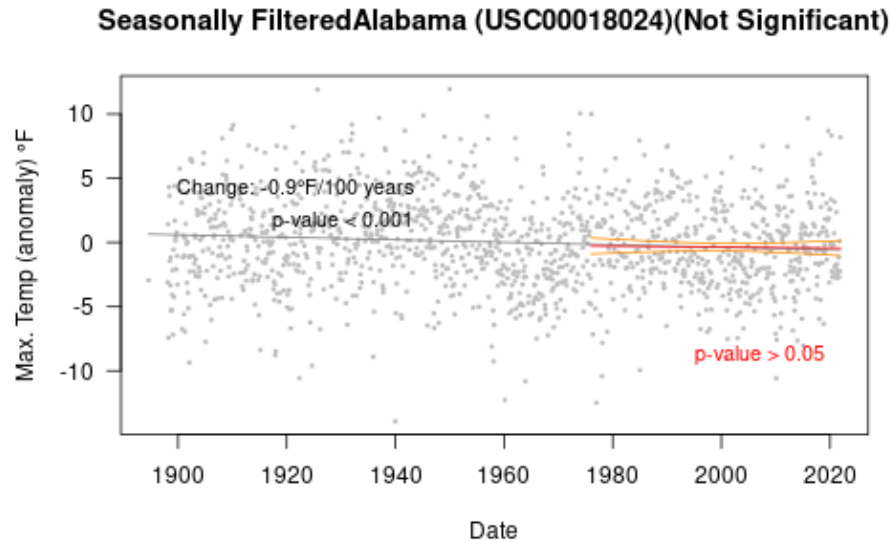


Figure 2: The changing in monthly temperature data.

```
# degree = 4
#coef = polyfit(X, y, degree)
# print('Coefficients: %s' % coef)
# create curve
```

4.3 Extreme Events—Using Daily Records

4.3.1 Complicated Nature of Rainfall Patterns

Rainfall trends are tough. Extreme events can occur in 24 hours or over long periods that might result in floods or droughts. Each region might have different patterns, so developing a consistent approach is tough.

We can look for trends in monthly averages, number of days without rain (important in tropics), and/or extreme events based on daily or hourly data.

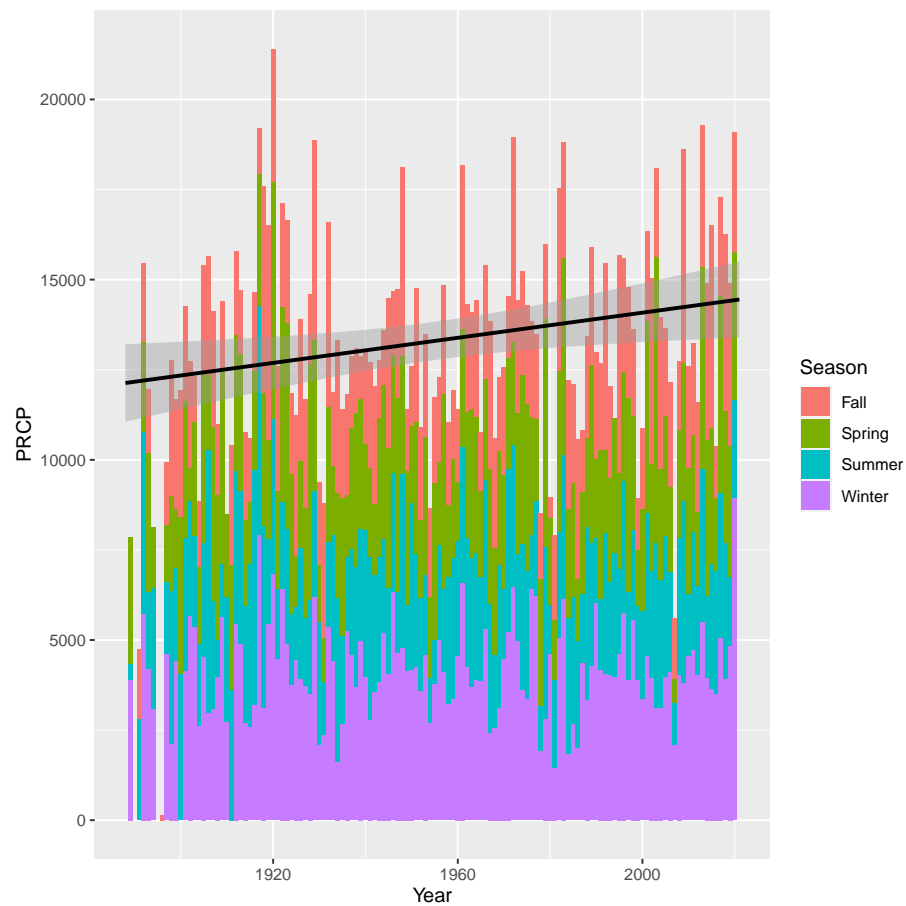
I don't know of a robust way to look at this for the entire globe.

```
PRCP.Total = aggregate(PRCP~Year, data=CHCND, sum, na.rm=T)
PRCP.Season.Total = aggregate(PRCP~Season+Year, data=CHCND, sum, na.rm=T)
```

Rainfall totals by season might be a useful way to think about changes, because the rainfall is often seasonal, I wonder if we can see patterns by season.

```
ggplot( ) +
  geom_bar(data = PRCP.Season.Total,
    aes(x=Year, y=PRCP, fill=Season), stat="identity") +
    xlim(min(ChCND$Year), max(ChCND$Year)-1) +
    #ylab("Number of Extreme Temps") + # for the y axis label
  geom_smooth(data = PRCP.Total,
    aes(y=PRCP, x=Year), method = "lm",
    se = T, color= "black")

## 'geom_smooth()' using formula 'y ~ x'
```



```
# + geom_smooth(data= PRCP.Season.Total, aes(x=Year, y = PRCP, color = Season, group=Season,
```


4.3.2 Drought

Days without rain...within a calendar year... bleed over between years isn't captured..

```
CHCND$PRCP.count = sequence(rle(CHCND$PRCP)$lengths)
Drought.run.temp <- data.frame(Year = NA, lengths=NA, values=NA)
for(i in min(CHCND$Year):max(CHCND$Year)){
  # print(i)
  run.length = rle(CHCND[CHCND$Year==i,]$PRCP)
  run.length.df = data.frame(Year = rep(i, length(run.length$values)),
    lengths = run.length$lengths,
    values = run.length$values)

  Drought.run.temp <- rbind(Drought.run.temp,
    run.length.df[run.length.df$values==0,])
}
Drought.run <- Drought.run.temp[-1,]
str(Drought.run)

## 'data.frame': 10689 obs. of 3 variables:
## $ Year : int NA NA NA NA NA NA NA NA NA NA NA ...
## $ lengths: int NA NA NA NA NA NA NA NA NA NA NA ...
## $ values : int NA NA NA NA NA NA NA NA NA NA NA ...

names(Drought.run)

## [1] "Year" "lengths" "values"

# What is a drought 10 days, 20 days, 40 days?

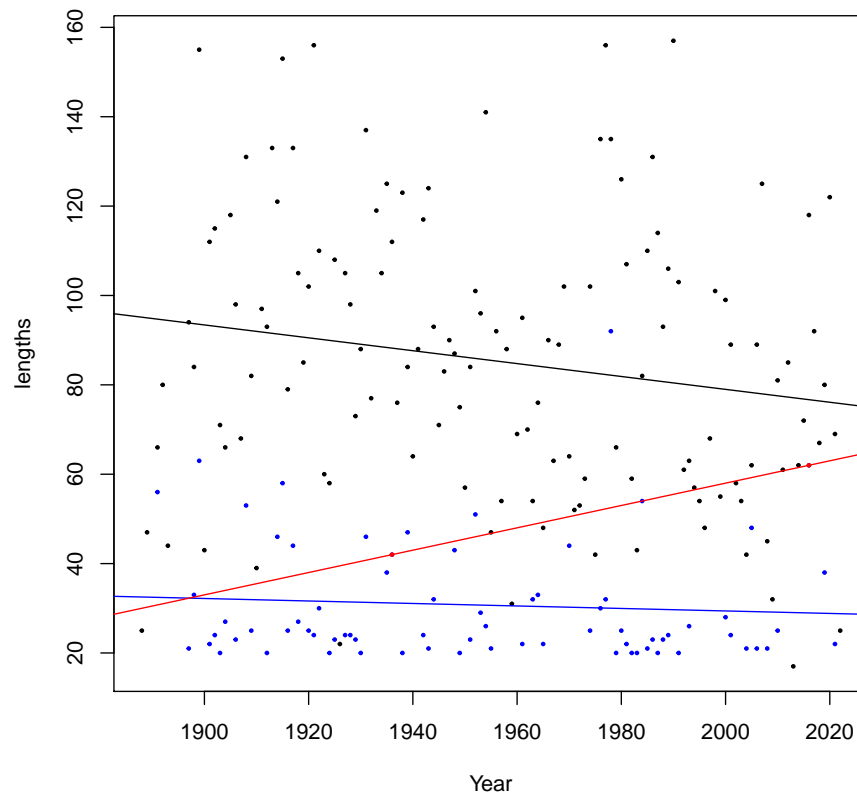
Drought.run.10 = aggregate(lengths~Year,
  data=Drought.run[Drought.run$lengths>=10,], sum)
Drought.run.20 = aggregate(lengths~Year,
  data=Drought.run[Drought.run$lengths>=20,], sum)
Drought.run.40 = aggregate(lengths~Year,
  data=Drought.run[Drought.run$lengths>=40,], sum)
Drought.run.100 = aggregate(lengths~Year,
  data=Drought.run[Drought.run$lengths>=100,], sum)

## Error in aggregate.data.frame(mf[1L], mf[-1L], FUN = FUN, ...):
## no rows to aggregate

plot(lengths~Year, Drought.run.10, pch=20, cex=.5)
points(lengths~Year, Drought.run.20, pch=20, col="blue", cex=.5)
points(lengths~Year, Drought.run.40, pch=20, col="red", cex=.5)
points(lengths~Year, Drought.run.100, pch=20, col="purple", cex=.5)
```

```
## Error in eval(m$data, eframe): object 'Drought.run.100' not found

abline(lm(lengths~Year, Drought.run.10))
abline(lm(lengths~Year, Drought.run.20), col="blue")
abline(lm(lengths~Year, Drought.run.40), col="red")
```



```
abline(lm(lengths~Year, Drought.run.100), col="purple")

## Error in is.data.frame(data): object 'Drought.run.100' not found

summary(lm(lengths~Year, Drought.run.100))

## Error in is.data.frame(data): object 'Drought.run.100' not found

plot(lengths~Year, Drought.run[Drought.run$lengths>30,], pch=20)
plot(lengths~Year, Drought.run[Drought.run$lengths>30,], pch=20)
```

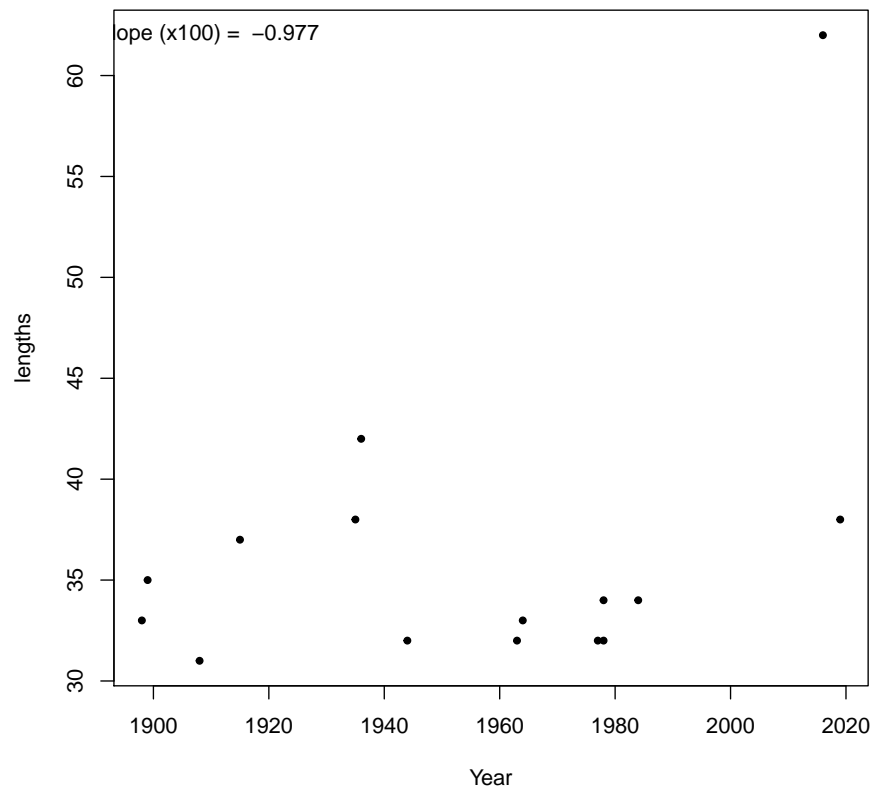
```

Drought.run.lm <- lm(lengths~Year, Drought.run[Drought.run$lengths>10,])
summary(Drought.run.lm)

##
## Call:
## lm(formula = lengths ~ Year, data = Drought.run[Drought.run$lengths >
##     10, ])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.029  -3.489  -1.682   1.520  47.212
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 34.486403  10.929405   3.155  0.00168 **
## Year        -0.009771   0.005589  -1.748  0.08093 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.208 on 628 degrees of freedom
## (3528 observations deleted due to missingness)
## Multiple R-squared:  0.004842, Adjusted R-squared:  0.003258
## F-statistic: 3.056 on 1 and 628 DF,  p-value: 0.08093

text(min(Drought.run$Year, na.rm=T), max(Drought.run$lengths, na.rm=T),
      paste("Slope (x100) = ", round(coef(Drought.run.lm)[2]*100, 3)), pos=4)

```



```
#plot(PRCP.count ~ Year, data=CHCND[CHCND$PRCP==0,])
```

Rainfall Probability Distributions by decade... to be developed.

```
CHCND$Decade <- floor_decade(CHCND$Year)

PRCP.Decade <- aggregate(PRCP~Month+Decade, data=CHCND, sum)
head(PRCP.Decade)

x <- PRCP.Decade$PRCP[PRCP.Decade$Decade==1900]
df <- approxfun(density(x))
plot(1:12, density(x))
xnew <- c(0.45, 1.84, 2.3)
points(xnew, df(xnew), col=2)
```

```
CHCND$Score <- floor_score(CHCND$Year)
```

4.4 Record Setting Temperature Records

In many cases, people seem to "feel" how temperature has been changing over time, and new records seem to capture the attention in the media. So, we'll create a updated record of maximum temperatures and display them.

This is a common way to communicate temperatures changes. I suspect we have a better sense of change when we notice "extreme" events...

```
names(CHCND)

minTMIN.length = aggregate(minTMIN~Year, data=CHCND, length)
minTMIN.length$group <- "Record Lows"
names(minTMIN.length) <- c("Year", "Num", "Group")
minTMIN.length$Num = -minTMIN.length$Num

maxTMAX.length = aggregate(maxTMAX~Year, data=CHCND, length);
maxTMAX.length$group <- "Record Highs"
names(maxTMAX.length) <- c("Year", "Num", "Group")

records = rbind(minTMIN.length, maxTMAX.length); # records

ggplot( ) +
  geom_point(data = CHCND, aes(y=TMIN, x=YearDay),
    size=.05, color="gray") +
  geom_bar(data = records, aes(x=Year, y=Num, fill=Group),
    stat="identity", position="identity") +
  xlim(min(CHCND$Year), max(CHCND$Year)-1) +
  #ylab("Number of Extreme Temps") + # for the y axis label
  scale_fill_manual("Legend",
    values = c("Record Highs" = "red", "Record Lows" = "blue")) +
  geom_smooth(data = CHCND, aes(y=TMIN, x=YearDay), method = "lm", se = FALSE)

ggplot( ) +
  geom_bar(data = records, aes(x=Year, y=Num, fill=Group),
    stat="identity", position="identity") +
  xlim(min(CHCND$Year), max(CHCND$Year)-1) +
  ylab("Number of Extreme Temps") + # for the y axis label
  scale_fill_manual("Legend",
    values = c("Record Highs" = "red", "Record Lows" = "blue"))
```

I tried to use a for loop and in then statements and it was painfully slow, so I converted the data to a matrix that can be used by barplots with much more efficiency!

Create the matrix

```
library(lubridate)
str(CHCND)

## 'data.frame': 49092 obs. of 13 variables:
## $ Date      : Date, format: "1888-01-01" "1888-01-02" ...
## $ Year      : num 1888 1888 1888 1888 1888 ...
## $ yday      : num 1 2 3 4 5 6 7 8 9 10 ...
## $ TMAX      : num NA NA NA NA NA NA NA NA NA NA ...
## $ TMIN      : num NA NA NA NA NA NA NA NA NA NA ...
## $ PRCP      : int NA NA NA NA NA NA NA NA NA NA ...
## $ YearDay   : num 1888 1888 1888 1888 1888 ...
## $ mmdd      : chr "01-01" "01-02" "01-03" "01-04" ...
## $ Month     : num 1 1 1 1 1 1 1 1 1 1 ...
## $ Month.name: Factor w/ 12 levels "Jan","Feb","Mar",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ Season    : chr "Winter" "Winter" "Winter" "Winter" ...
## $ PRCP.count: int 1 1 1 1 1 1 1 1 1 1 ...
## $ Score     : num 1880 1880 1880 1880 1880 1880 1880 1880 1880 1880 ...

TMAX.mat.noleap <- matrix(NA, nrow=366, ncol=max(CHCND$Year) - min(CHCND$Year))
TMIN.mat <- matrix(NA, nrow=366, ncol=max(CHCND$Year) - min(CHCND$Year))
#TMAX.mat.leap <- matrix(NA, nrow=1, ncol=max(CHCND$Year) - min(CHCND$Year))

# Dumb Method, fraction of year might be better...

CHCND.noleap = subset(CHCND, select=c(Date, Year, yday, TMAX, TMIN, PRCP),
                      subset=(mmdd!="02-29"))

## Add yday for leap year
CHCND.noleap$yday[CHCND.noleap$yday>=60 & !leap_year(CHCND.noleap$Date)]<-CHCND.noleap$yday

## Create leap year dataframe
CHCND.leap = subset(CHCND, select=c(Date, Year, yday, TMAX, TMIN, PRCP),
                    subset=(mmdd=="02-29"))
names(CHCND.noleap)

## [1] "Date" "Year" "yday" "TMAX" "TMIN" "PRCP"

years = seq(min(CHCND$Year), max(CHCND$Year), by=1)
year.seq = data.frame(Year = years, Col = seq_len(length(seq(min(CHCND$Year), max(CHCND$Year)))))

for(i in min(CHCND.noleap$Year):max(CHCND.noleap$Year)){
```

```

    for(j in c(1:59, 61:366)){
      # i=2016; j = 50;
      TMAX.mat.noleap[j, year.seq$Col[year.seq$Year==i]] <-
      CHCND.noleap$TMAX[CHCND.noleap$Year==i & CHCND.noleap$yday == j]
      TMIN.mat[j, year.seq$Col[year.seq$Year==i]] <-
      CHCND.noleap$TMIN[CHCND.noleap$Year==i & CHCND.noleap$yday == j]
    }
    if(Leap_year(i)){
      TMAX.mat.noleap[60, year.seq$Col[year.seq$Year==i]] <- CHCND.leap$TMAX[CHCND.leap$Year==i & CHCND.leap$yday == j]
      TMIN.mat[60, year.seq$Col[year.seq$Year==i]] <- CHCND.leap$TMIN[CHCND.leap$Year== i & CHCND.leap$yday == j]
      print(paste0("Added Leap Year", i))
    } else {print(paste0("Process non-leap year ", i))}
  }

## [1] "Added Leap Year1888"
## [1] "Process non-leap year 1889"
## [1] "Process non-leap year 1890"
## [1] "Process non-leap year 1891"
## [1] "Added Leap Year1892"
## [1] "Process non-leap year 1893"
## [1] "Process non-leap year 1894"
## [1] "Process non-leap year 1895"
## [1] "Added Leap Year1896"
## [1] "Process non-leap year 1897"
## [1] "Process non-leap year 1898"
## [1] "Process non-leap year 1899"
## [1] "Process non-leap year 1900"
## [1] "Process non-leap year 1901"
## [1] "Process non-leap year 1902"
## [1] "Process non-leap year 1903"
## [1] "Added Leap Year1904"
## [1] "Process non-leap year 1905"
## [1] "Process non-leap year 1906"
## [1] "Process non-leap year 1907"
## [1] "Added Leap Year1908"
## [1] "Process non-leap year 1909"
## [1] "Process non-leap year 1910"
## [1] "Process non-leap year 1911"
## [1] "Added Leap Year1912"
## [1] "Process non-leap year 1913"
## [1] "Process non-leap year 1914"
## [1] "Process non-leap year 1915"
## [1] "Added Leap Year1916"
## [1] "Process non-leap year 1917"
## [1] "Process non-leap year 1918"

```

```
## [1] "Process non-leap year 1919"
## [1] "Added Leap Year1920"
## [1] "Process non-leap year 1921"
## [1] "Process non-leap year 1922"
## [1] "Process non-leap year 1923"
## [1] "Added Leap Year1924"
## [1] "Process non-leap year 1925"
## [1] "Process non-leap year 1926"
## [1] "Process non-leap year 1927"
## [1] "Added Leap Year1928"
## [1] "Process non-leap year 1929"
## [1] "Process non-leap year 1930"
## [1] "Process non-leap year 1931"
## [1] "Added Leap Year1932"
## [1] "Process non-leap year 1933"
## [1] "Process non-leap year 1934"
## [1] "Process non-leap year 1935"
## [1] "Added Leap Year1936"
## [1] "Process non-leap year 1937"
## [1] "Process non-leap year 1938"
## [1] "Process non-leap year 1939"
## [1] "Added Leap Year1940"
## [1] "Process non-leap year 1941"
## [1] "Process non-leap year 1942"
## [1] "Process non-leap year 1943"
## [1] "Added Leap Year1944"
## [1] "Process non-leap year 1945"
## [1] "Process non-leap year 1946"
## [1] "Process non-leap year 1947"
## [1] "Added Leap Year1948"
## [1] "Process non-leap year 1949"
## [1] "Process non-leap year 1950"
## [1] "Process non-leap year 1951"
## [1] "Added Leap Year1952"
## [1] "Process non-leap year 1953"
## [1] "Process non-leap year 1954"
## [1] "Process non-leap year 1955"
## [1] "Added Leap Year1956"
## [1] "Process non-leap year 1957"
## [1] "Process non-leap year 1958"
## [1] "Process non-leap year 1959"
## [1] "Added Leap Year1960"
## [1] "Process non-leap year 1961"
## [1] "Process non-leap year 1962"
## [1] "Process non-leap year 1963"
```



```
## [1] "Added Leap Year1964"
## [1] "Process non-leap year 1965"
## [1] "Process non-leap year 1966"
## [1] "Process non-leap year 1967"
## [1] "Added Leap Year1968"
## [1] "Process non-leap year 1969"
## [1] "Process non-leap year 1970"
## [1] "Process non-leap year 1971"
## [1] "Added Leap Year1972"
## [1] "Process non-leap year 1973"
## [1] "Process non-leap year 1974"
## [1] "Process non-leap year 1975"
## [1] "Added Leap Year1976"
## [1] "Process non-leap year 1977"
## [1] "Process non-leap year 1978"
## [1] "Process non-leap year 1979"
## [1] "Added Leap Year1980"
## [1] "Process non-leap year 1981"
## [1] "Process non-leap year 1982"
## [1] "Process non-leap year 1983"
## [1] "Added Leap Year1984"
## [1] "Process non-leap year 1985"
## [1] "Process non-leap year 1986"
## [1] "Process non-leap year 1987"
## [1] "Added Leap Year1988"
## [1] "Process non-leap year 1989"
## [1] "Process non-leap year 1990"
## [1] "Process non-leap year 1991"
## [1] "Added Leap Year1992"
## [1] "Process non-leap year 1993"
## [1] "Process non-leap year 1994"
## [1] "Process non-leap year 1995"
## [1] "Added Leap Year1996"
## [1] "Process non-leap year 1997"
## [1] "Process non-leap year 1998"
## [1] "Process non-leap year 1999"
## [1] "Added Leap Year2000"
## [1] "Process non-leap year 2001"
## [1] "Process non-leap year 2002"
## [1] "Process non-leap year 2003"
## [1] "Added Leap Year2004"
## [1] "Process non-leap year 2005"
## [1] "Process non-leap year 2006"
## [1] "Process non-leap year 2007"
## [1] "Added Leap Year2008"
```

```

## [1] "Process non-leap year 2009"
## [1] "Process non-leap year 2010"
## [1] "Process non-leap year 2011"
## [1] "Added Leap Year2012"
## [1] "Process non-leap year 2013"
## [1] "Process non-leap year 2014"
## [1] "Process non-leap year 2015"
## [1] "Added Leap Year2016"
## [1] "Process non-leap year 2017"
## [1] "Process non-leap year 2018"
## [1] "Process non-leap year 2019"
## [1] "Added Leap Year2020"
## [1] "Process non-leap year 2021"

## Error in '[<-'(*tmp*', j, year.seq$Col[year.seq$Year == i], value
= CHCND.noleap$TMAX[CHCND.noleap$Year == : subscript out of bounds

TMAX.mat.noleap[,6]

##      [1]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [13]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [25]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [37]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [49]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [61]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [73]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [85]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [97]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [109]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [121]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [133]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [145]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [157]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [169]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [181]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [193]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [205]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [217]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [229]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [241]      NA      NA      NA      NA 89.96 86.00 86.00 87.08 78.98 86.00 86.00 86.00
##   [253] 80.96 80.06 78.98 78.98 82.04 86.00 86.00 89.06 84.92 87.08 84.02 84.92
##   [265] 87.98 86.00 86.00 84.92 84.02 82.94 78.08 77.00 75.92 78.08 84.02 80.96
##   [277] 80.06 80.96 77.00 82.04 77.00 75.92 82.04 84.92 80.96 80.06 78.98 68.00
##   [289] 64.94 68.00 66.92 66.02 73.04 78.08 78.98 71.96 78.98 78.98 82.04 80.06
##   [301] 80.06 71.06 59.00 60.08 62.96 68.00 73.94 77.00 75.02 59.00 62.06 62.06
##   [313] 62.96 66.92 64.94 66.02 69.08 69.08 69.98 69.98 57.02 71.06 71.96 64.04

```

```

## [325] 53.96 69.08 68.00 50.00 42.08 44.96 51.08 57.02 59.00 68.00 64.04 55.94
## [337] 60.08 62.06 32.00 46.94 51.98 53.96 57.92 57.92 57.92 62.96 66.02 60.98
## [349] 57.92 66.02 66.02 42.08 50.00 55.04 46.94 51.98 57.92 66.92 68.00 66.02
## [361] 66.02 57.92 62.06 60.98 64.04 42.08

TMAX.mat.noleap[,7]

##      [1]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [13]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##     [25]      NA      NA      NA      NA      NA      NA      NA      NA 62.06 55.04 71.96 39.92 44.96
##     [37] 55.04 62.96 68.00 73.04 60.98 53.96 60.08 50.00 50.00 37.04 42.98 57.92
##     [49] 69.98 66.02 66.92 62.96 50.00 35.96 35.96 35.06 42.08 59.00 55.04      NA
##     [61] 64.94 69.08 73.04 73.04 73.04 62.96 69.98 73.94 73.04 78.98 78.08 66.92
##     [73] 75.02 80.06 78.08 78.98 80.06 82.94 84.02 75.92 84.02 86.00 68.00 68.00
##     [85] 46.94 35.06 42.98 59.00 51.98 59.00 73.94 66.92 64.94 73.94 71.06 71.06
##     [97] 66.92 75.02 75.02 82.04 62.96 73.04 64.94 68.00 77.00 73.04 84.92 89.96
##    [109] 89.06 82.94 73.04 73.94 71.06 71.96 75.02 78.98 82.04 84.92 87.08 89.06
##    [121] 82.94 86.00 86.00 84.92 89.06 87.98 87.98 91.94 86.00 93.92 93.02 87.08
##    [133] 87.08 86.00 87.08 84.02 89.06 89.96 84.02 64.04 60.98 64.94 75.92 71.06
##    [145] 75.92 78.08 80.06 82.04 84.92 82.94 78.08 75.92 73.04 82.94 87.98 91.04
##    [157] 91.94 78.08 84.02 91.04 93.92 93.02 91.94 91.94 96.08 98.06 91.94 93.02
##    [169] 91.94 86.00 89.06 89.96 95.00 89.96 91.94 89.96 84.92 93.92 95.00 96.98
##    [181] 98.06 96.98 93.92 93.92 93.92 98.06 93.92 80.06 89.06 89.06 75.02 77.00
##    [193] 84.92 89.06 89.96 89.96 91.04 86.00 89.06 87.98 89.06 89.96 87.08 87.08
##    [205] 87.98 87.08 89.96 89.06 87.08 91.94 89.06 89.06 91.04 91.94 87.98 89.96
##    [217] 89.06 75.02 78.08 82.94 89.96 93.02 95.00 96.08 93.92 98.06 98.96 98.96
##    [229] 89.96 86.00 89.06 89.06 89.06 87.98 75.02 71.06 82.94 84.92 87.98 86.00
##    [241] 86.00 84.02 87.08 87.98      NA      NA      NA      NA      NA      NA      NA      NA
##    [253]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [265]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [277]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [289]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [301]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [313]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [325]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [337]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [349]      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA      NA
##    [361]      NA      NA      NA      NA      NA      NA      NA

TMAX.mat.noleap[60:61,120:134]

##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
## [1,]  NA 51.98  NA  NA  NA 64.04  NA  NA  NA 69.08  NA  NA
## [2,]  77 62.06 62.06 51.08 78.98 73.04 48.92 48.92 55.04 71.96 71.06 71.06
##      [,13] [,14] [,15]
## [1,]  NA 57.02  NA
## [2,] 66.92 64.94 80.96

```

```
TMIN.mat[60:61,120:134]
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12]
## [1,]    NA 19.04    NA    NA    NA 53.06    NA    NA    NA 28.94    NA    NA
## [2,] 33.08 32.00 30.92 24.08 44.96 62.06 33.08 19.94 51.98 44.96 53.06 55.94
##      [,13] [,14] [,15]
## [1,]    NA 24.98    NA
## [2,] 57.92 42.08 57.02
```

```
setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/")
png(paste0("Social_Media/png/", fips$State, "-", stid, "-CHCND_Records.png"), width = 480, height = 320)

results<-NULL
decades <- years[years/10 == floor(years/10)]
i = max(CHCND$Year)
# START LOOP
  j = which(years %in% i)
  if(sum(is.na(TMAX.mat.noleap[,j]))==366) next

## Error in TMAX.mat.noleap[, j]: subscript out of bounds

TMAX1 = apply(TMAX.mat.noleap[,1:j], 1, function (x) which.max(x));

## Error in TMAX.mat.noleap[, 1:j]: subscript out of bounds

is.na(TMAX1) <- lengths(TMAX1) == 0

## Error in lengths(TMAX1): object 'TMAX1' not found

TMAX1 <- unlist(TMAX1)

## Error in unlist(TMAX1): object 'TMAX1' not found

TMAX1 <- count(TMAX1)

## Error in count(TMAX1): object 'TMAX1' not found

#str(TMAX1)
names(TMAX1)=c("Year", "TMAX")

## Error in names(TMAX1) = c("Year", "TMAX"): object 'TMAX1' not found

TMAX_na = data.frame(Year=1:j)
TMAX <- merge(TMAX_na, TMAX1, all.x=TRUE, by="Year")

## Error in merge(TMAX_na, TMAX1, all.x = TRUE, by = "Year"): object
'TMAX1' not found

if(sum(is.na(TMIN.mat[,j]))==366) next
```

```

## Error in TMIN.mat[, j]: subscript out of bounds

  # Select Minimum and Change to Negative Value
TMIN1 = apply(TMIN.mat[,1:j], 1, function (x) which.min(x));

## Error in TMIN.mat[, 1:j]: subscript out of bounds

is.na(TMIN1) <- lengths(TMIN1) == 0

## Error in lengths(TMIN1): object 'TMIN1' not found

TMIN1 <- unlist(TMIN1)

## Error in unlist(TMIN1): object 'TMIN1' not found

TMIN1 <- count(TMIN1) # Max Counts Negative

## Error in count(TMIN1): object 'TMIN1' not found

#str(TMIN1)
names(TMIN1)=c("Year", "TMIN")

## Error in names(TMIN1) = c("Year", "TMIN"): object 'TMIN1' not found

TMIN_na <- data.frame(Year=1:j)
TMIN <- merge(TMIN_na, TMIN1, all.x=TRUE, by="Year")

## Error in merge(TMIN_na, TMIN1, all.x = TRUE, by = "Year"): object
'TMIN1' not found

R1 <- merge(TMAX, TMIN, by="Year")

## Error in merge(TMAX, TMIN, by = "Year"): object 'TMAX' not found

R1$Index = rep(j, nrow(R1))

## Error in nrow(R1): object 'R1' not found

#results = rbind(results, R)
R1$TMIN = -R1$TMIN

## Error in eval(expr, envir, enclos): object 'R1' not found

## Sorting out X Axis
tic.no <- 4
rowskip = round(nrow(R1)/tic.no, 0)

## Error in nrow(R1): object 'R1' not found

row_num <- seq_len(nrow(R1)) %% rowskip

## Error in nrow(R1): object 'R1' not found

row.sel = which(row_num %in% c(1))

```

```

## Error in row_numb %in% c(1): object 'row_numb' not found
index.year <- years[row.sel]
## Error in eval(expr, envir, enclos): object 'row.sel' not found
# switch to decades?

xticks = row.sel
## Error in eval(expr, envir, enclos): object 'row.sel' not found
xlabs = index.year
## Error in eval(expr, envir, enclos): object 'index.year' not found
yrange = range(c(R1$TMIN, R1$TMAX), na.rm=T)
## Error in eval(expr, envir, enclos): object 'R1' not found
ytics = floor(seq(yrange[1], yrange[2], length.out=tic.no))
## Error in seq(yrange[1], yrange[2], length.out = tic.no): object
'yrange' not found
ylabs = as.character(abs(ytics))
## Error in eval(expr, envir, enclos): object 'ytics' not found
par(las=1, xpd=TRUE)
plot(c(1,nrow(R1)), c(yrange[1], yrange[2]), ty="n", xaxt='n', yaxt='n', xlab="Year", ylab=
## Error in nrow(R1): object 'R1' not found
axis(2,at=ytics, labels=ylabs)
## Error in axis(2, at = ytics, labels = ylabs): object 'ytics' not
found
axis(1,at=xticks, labels=xlabs)
## Error in axis(1, at = xticks, labels = xlabs): object 'xticks' not
found
barplot(height = R1$TMAX, space=0, add = TRUE, axes = FALSE, col="red")
## Error in barplot(height = R1$TMAX, space = 0, add = TRUE, axes =
FALSE, : object 'R1' not found
barplot(height = R1$TMIN, space=0, add = TRUE, axes = FALSE, col="blue")
## Error in barplot(height = R1$TMIN, space = 0, add = TRUE, axes =
FALSE, : object 'R1' not found
# END LOOP

dev.off()

## pdf
## 2

```

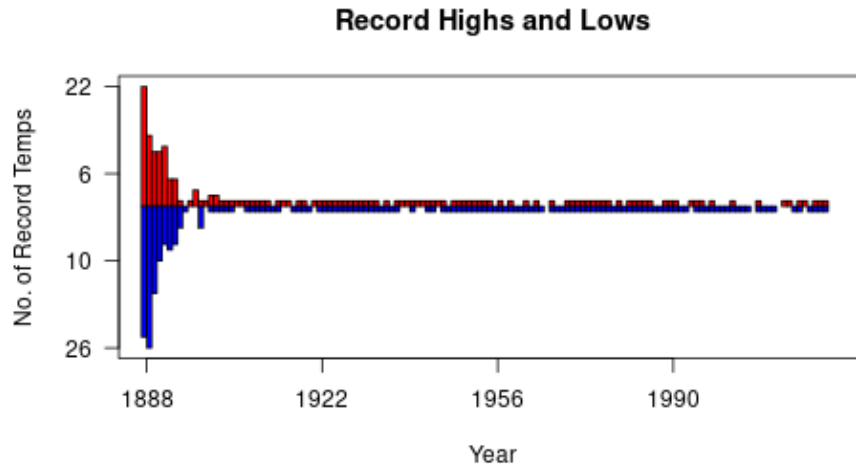
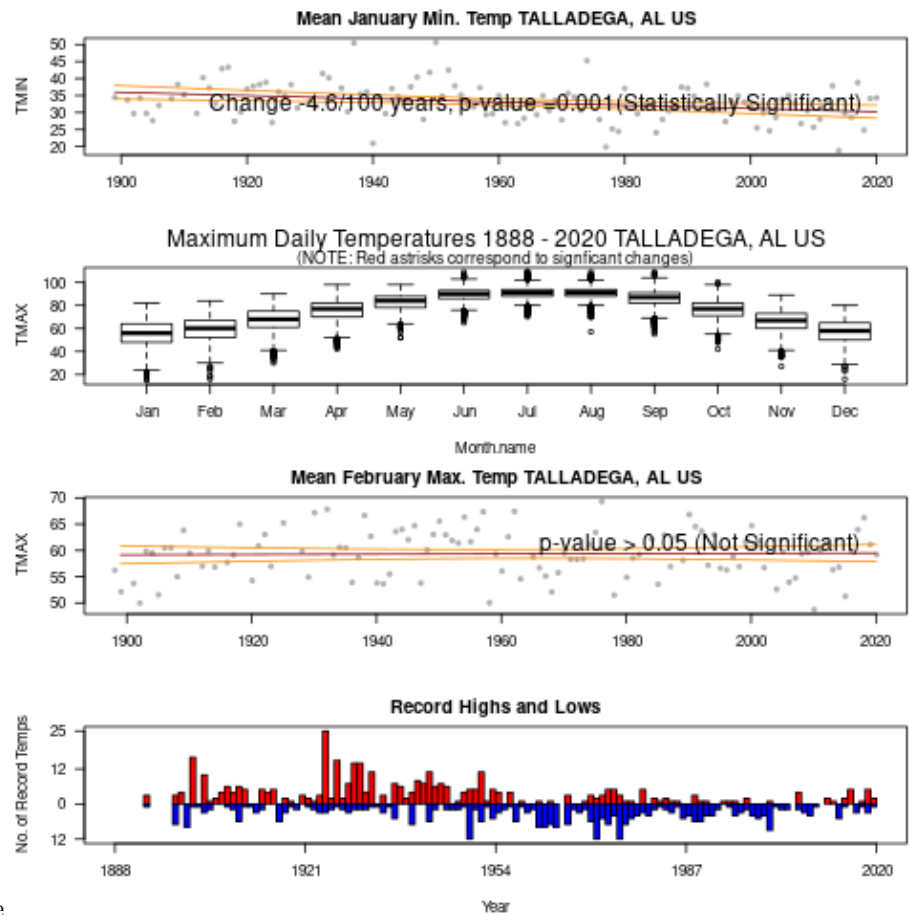


Figure 3: Daily temperatures that have been the highest on record (in red) and lowest on record (in blue). In some cases, climate change has created more records in the recent decades, while other stations seem don't show that trend.

The patterns of record temperatures often shows increasing number of new high temperature records and fewer record low temperatures more recently, but as usual, it depends on the location (Figure 3).

4.5 Iterate TMAX vs. Month Boxplots

Evaluating the changes in TMAX and Monthly temperatures might be useful, but for now, I think it's hard to see the patterns.



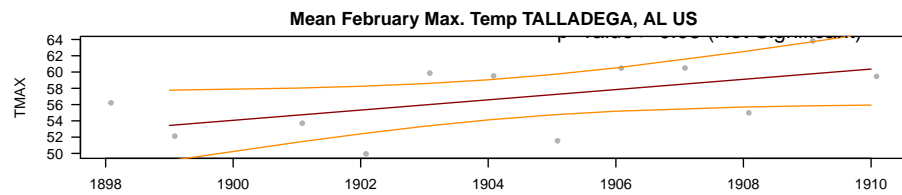
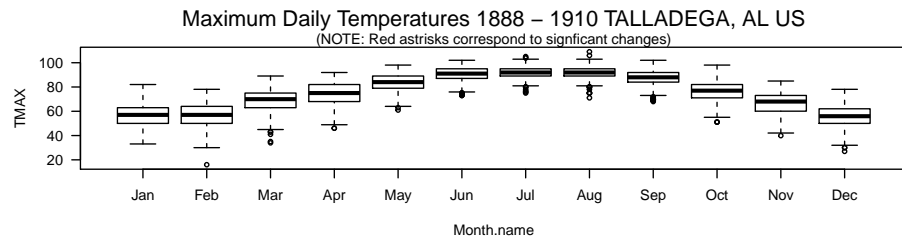
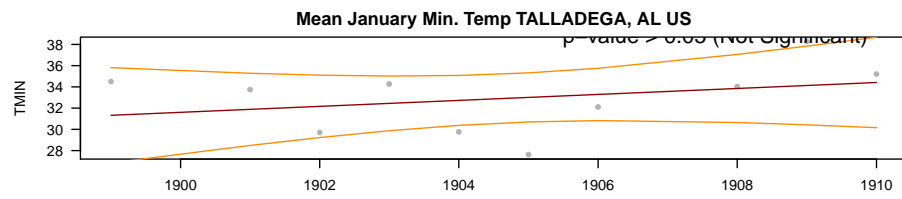
beginfigure

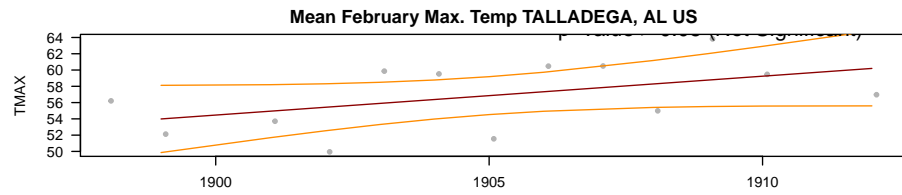
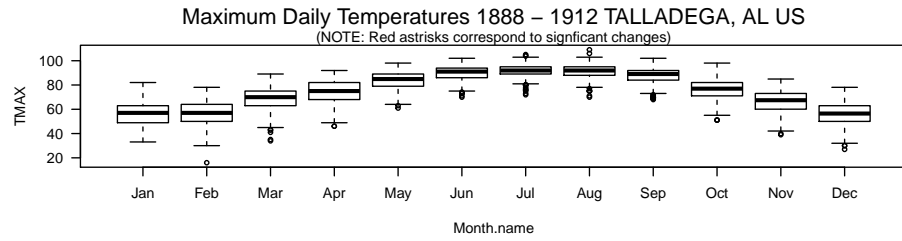
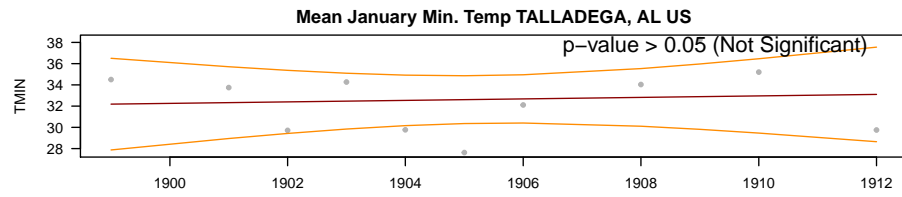
Climate can be analyzed using several types of lenses. In this case, we have analyzed show the months with the greatest changes. The first figure is monthly average of TMINs (daily low temperatures) with a best fit line. The second figure shows the monthly TMAX range and asterisks indicate significant changes over the station record and the third figure is the trend for these TMAXs over time and includes the best fit line. The final figure shows the daily temperatures that have been the highest on record (in red). In some cases, climate change has created more records in the recent decades, while other stations seem don't show that trend.

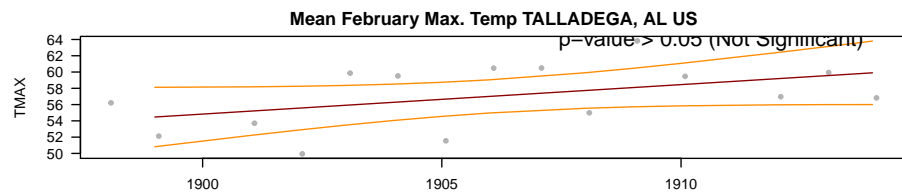
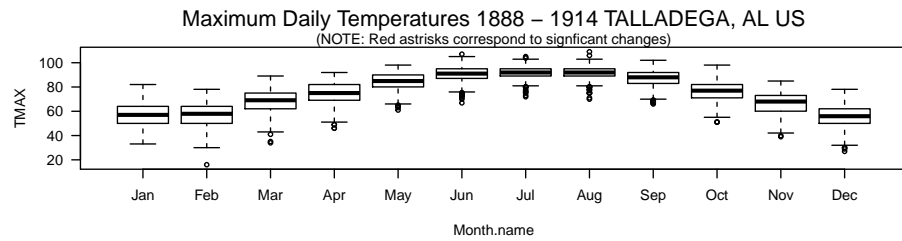
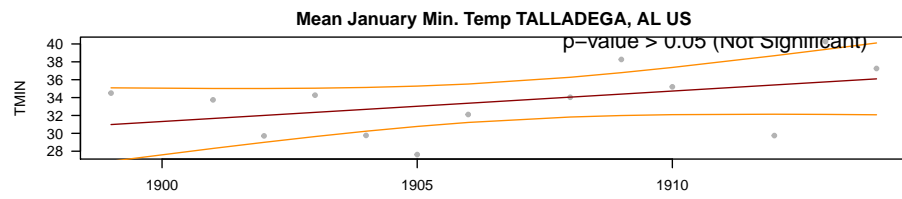
4.6 Four Plots Compelling Figures

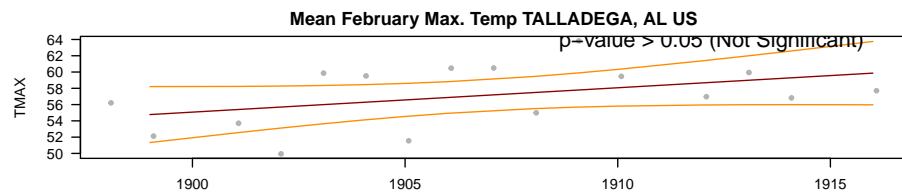
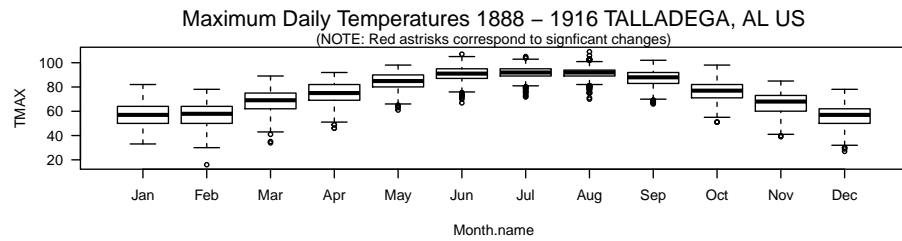
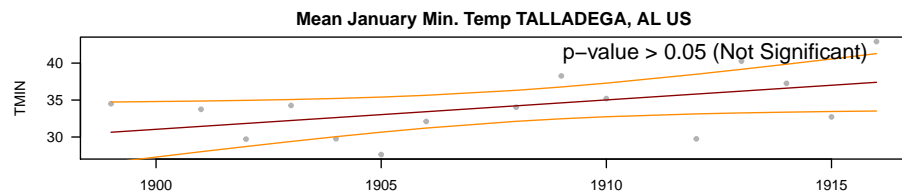
To test the code, I have created graphics that can then be used in the animation process, i.e. try to create code that doesn't get too complicated and then fail!

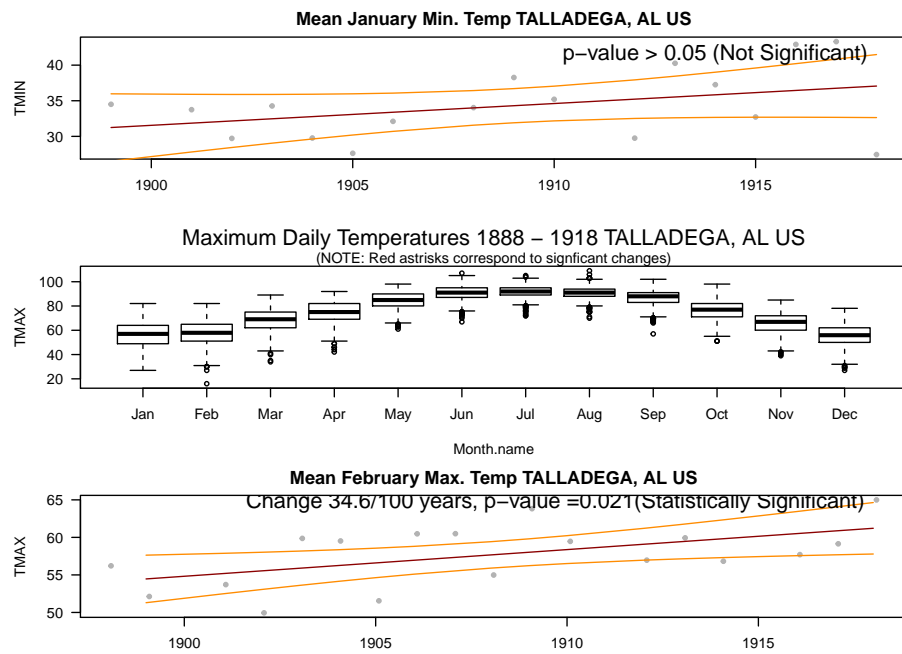
```
## Error in paste0(png_public, panel4.png): object 'png_public' not found
```

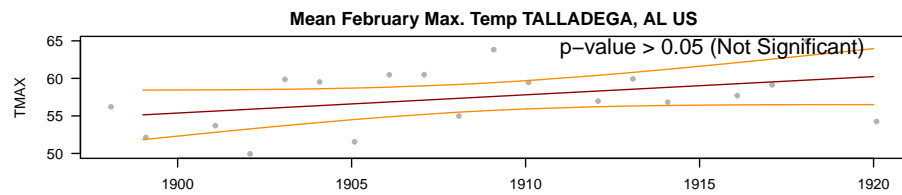
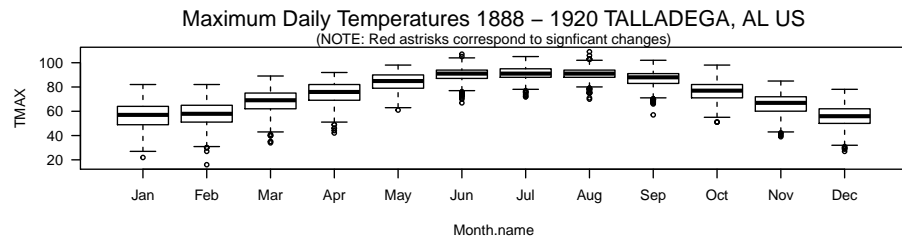
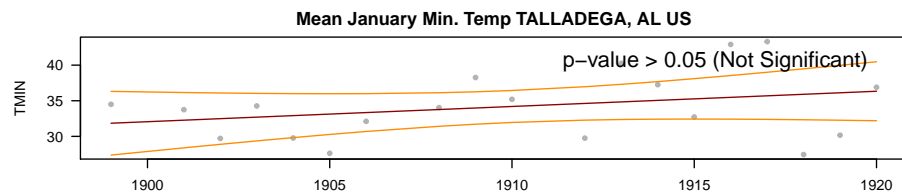



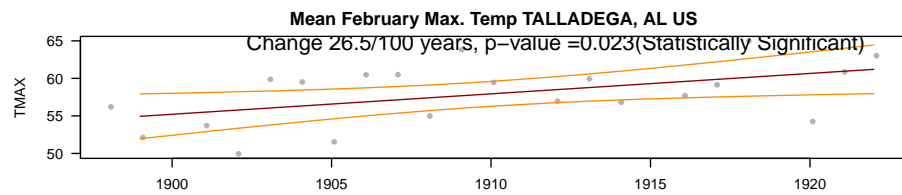
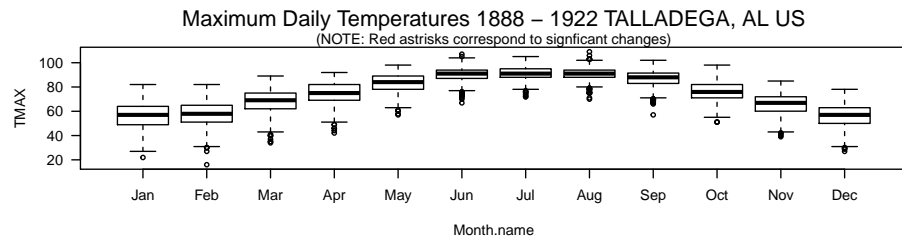
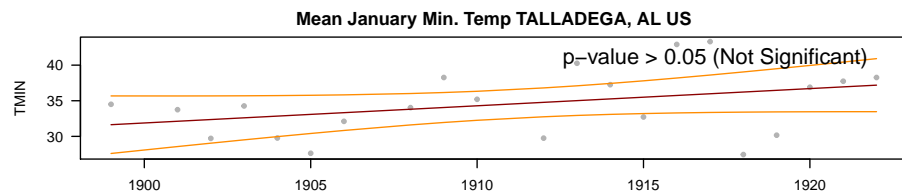


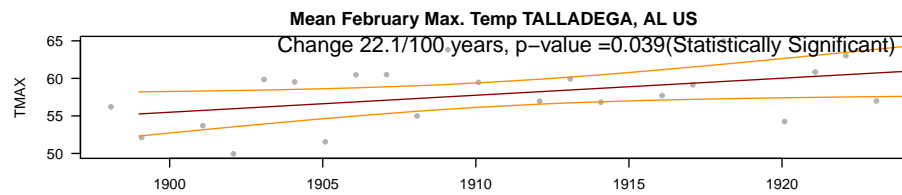
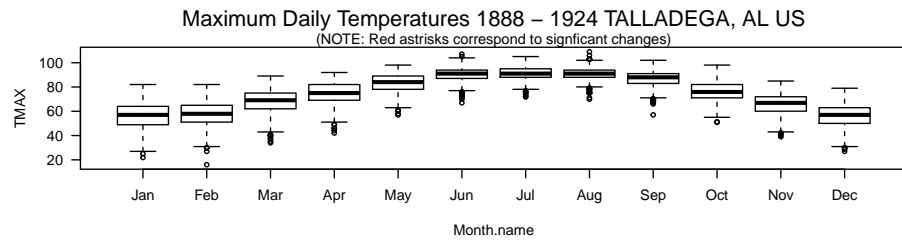
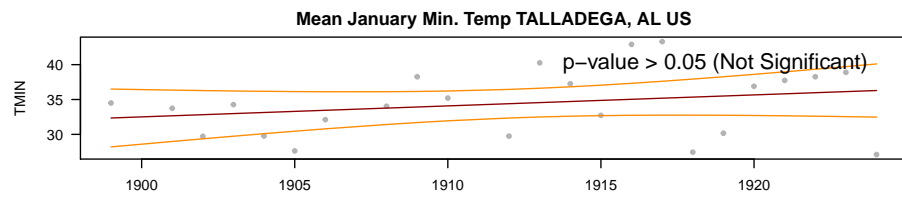


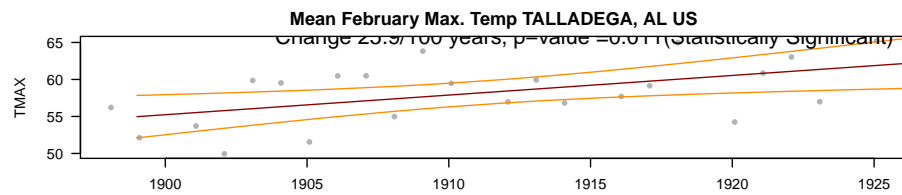
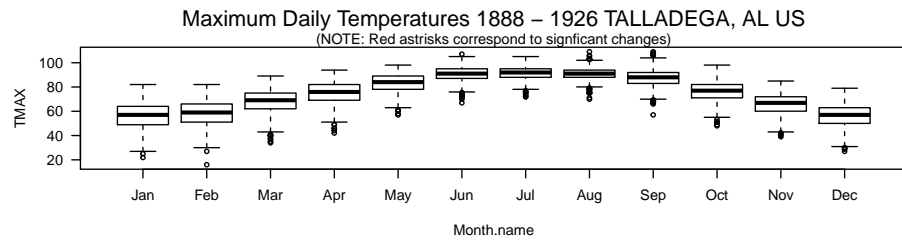
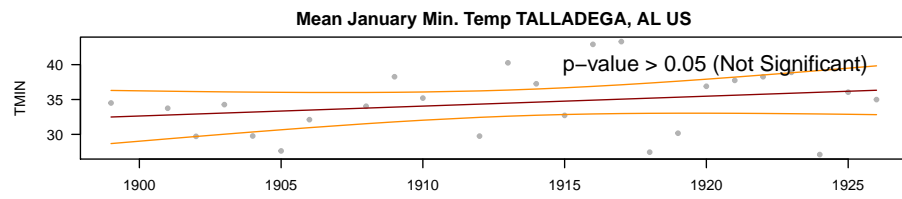


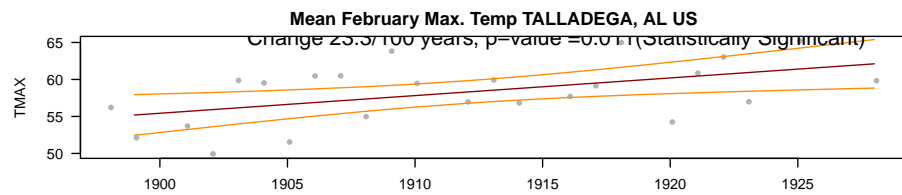
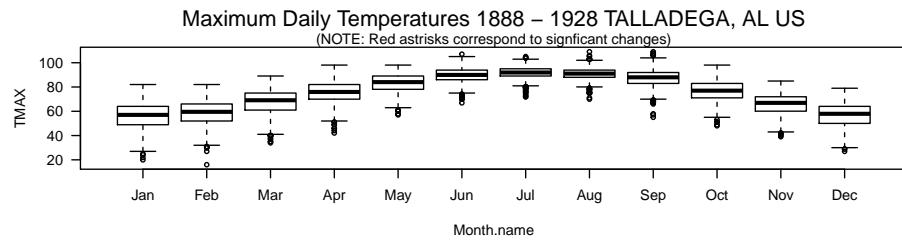
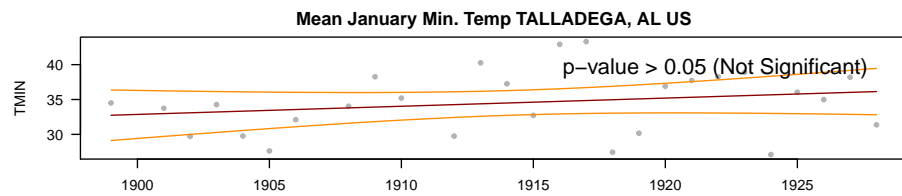


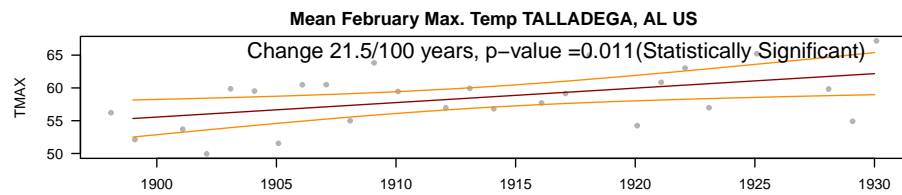
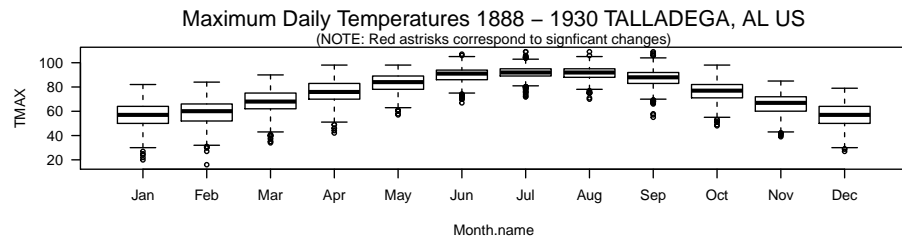
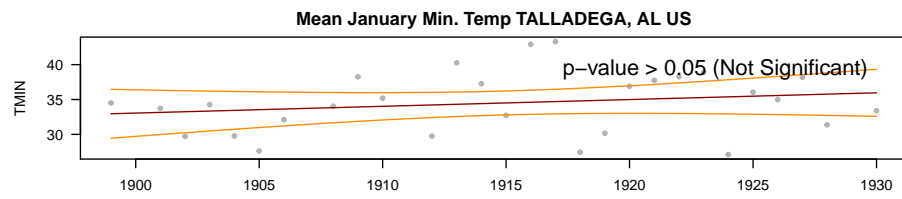


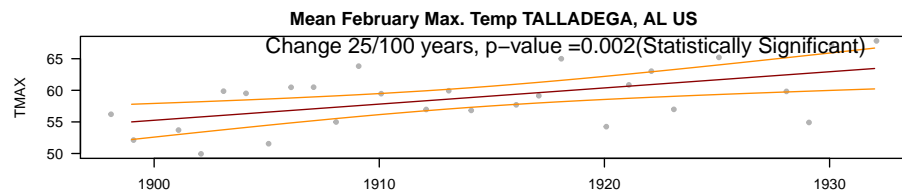
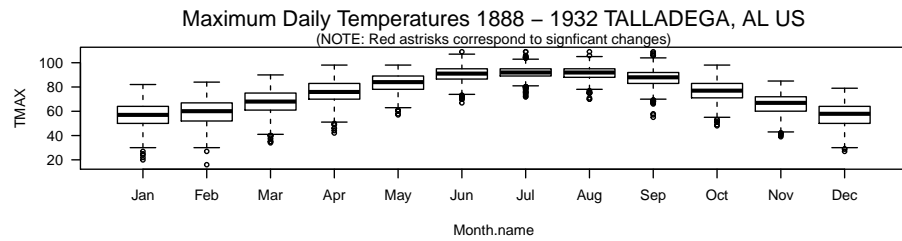
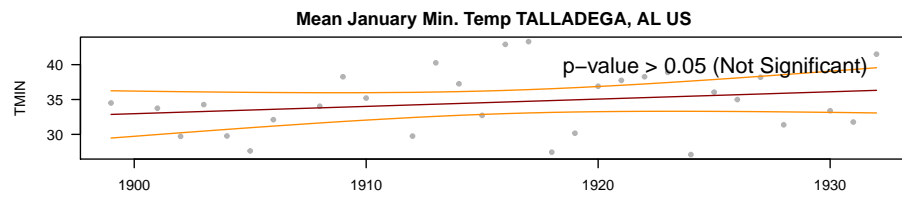


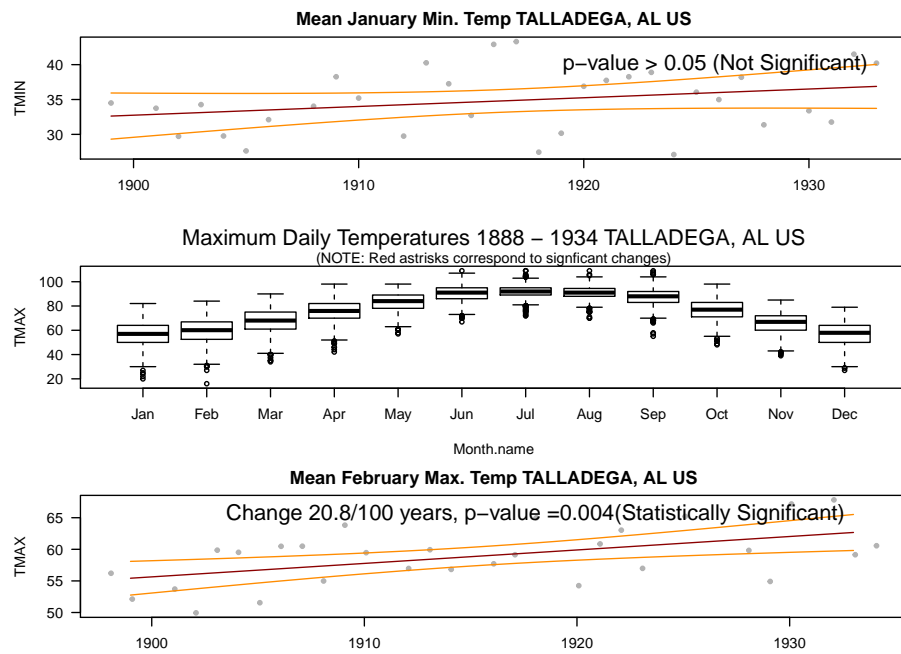


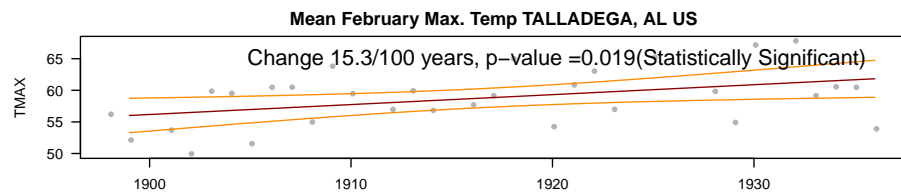
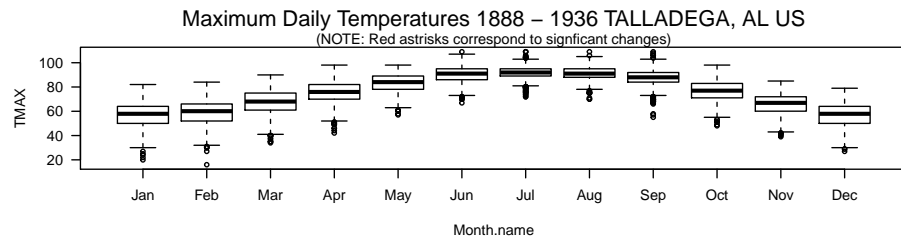
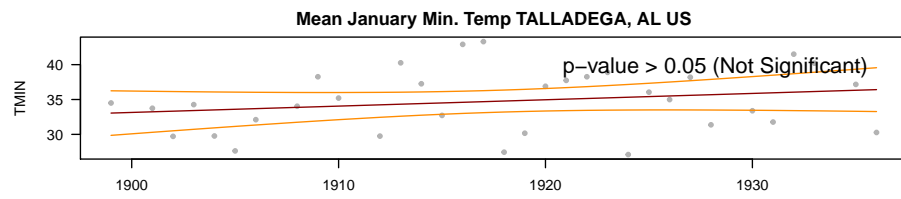


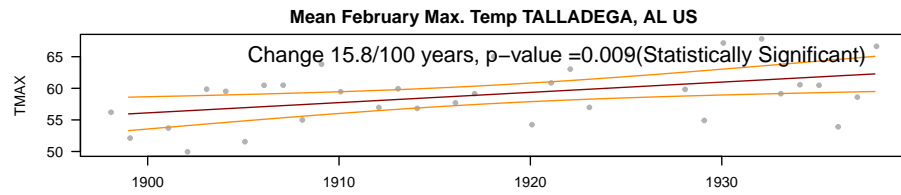
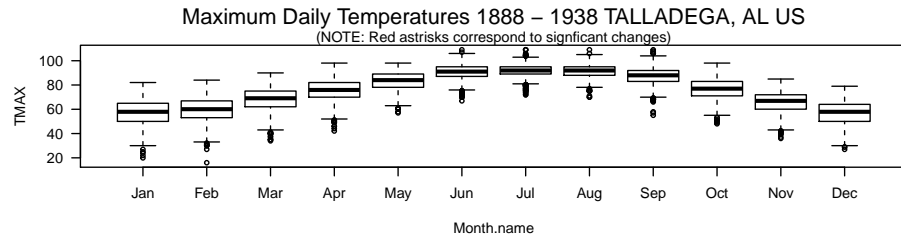
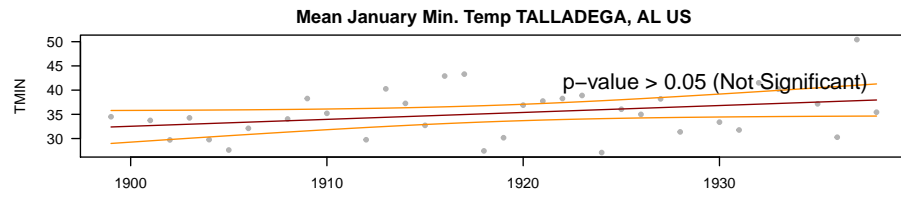


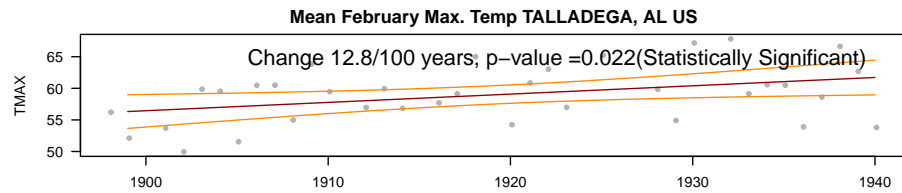
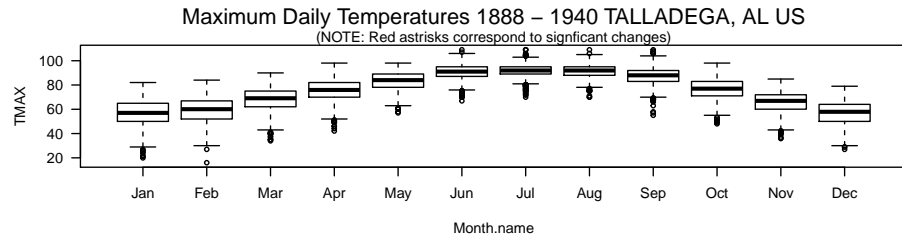
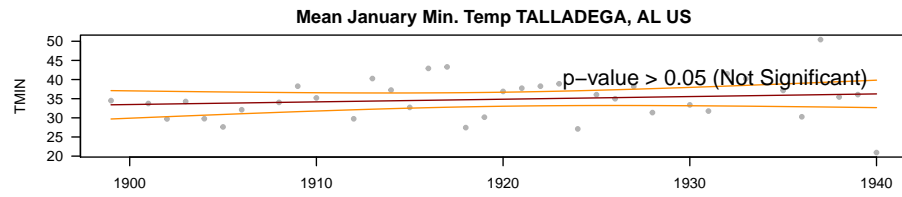


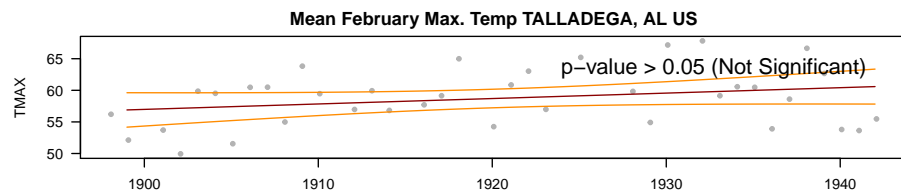
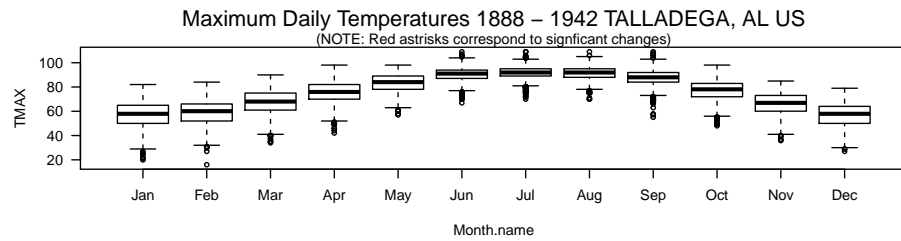
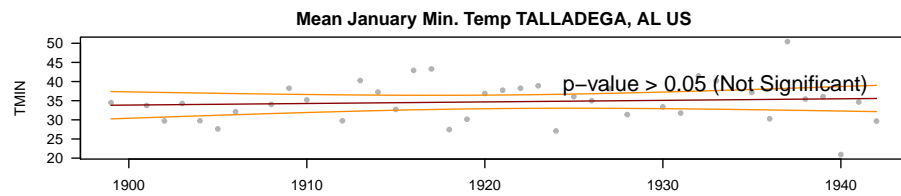


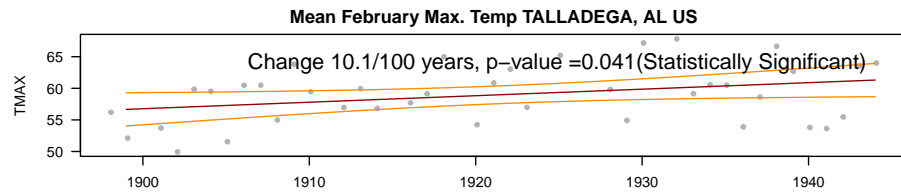
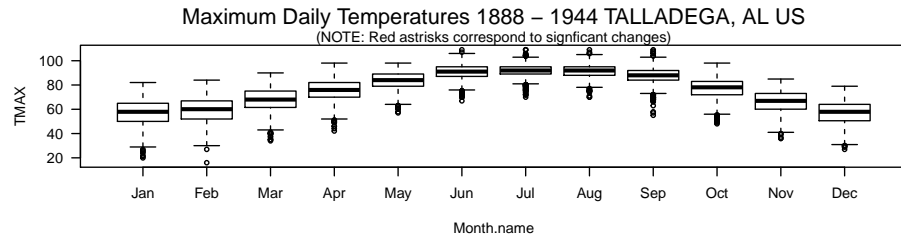
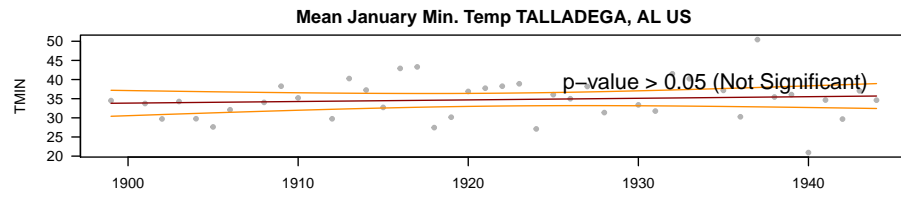


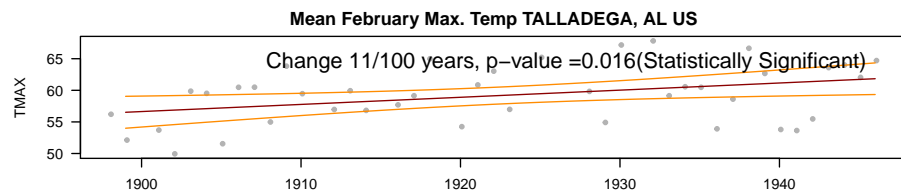
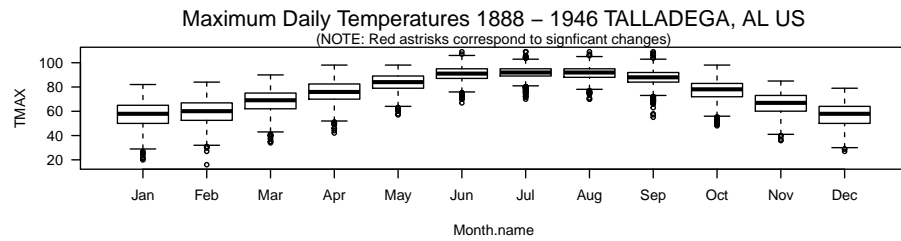
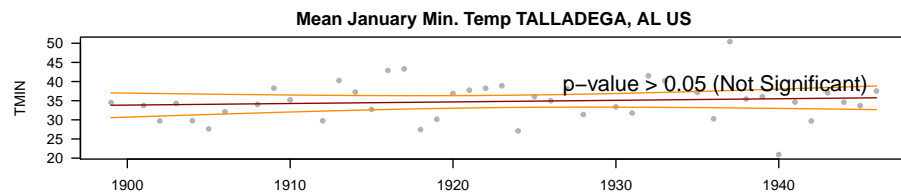


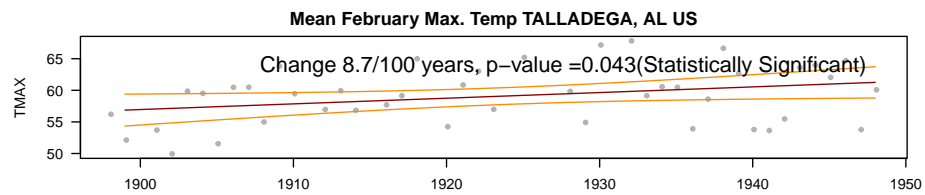
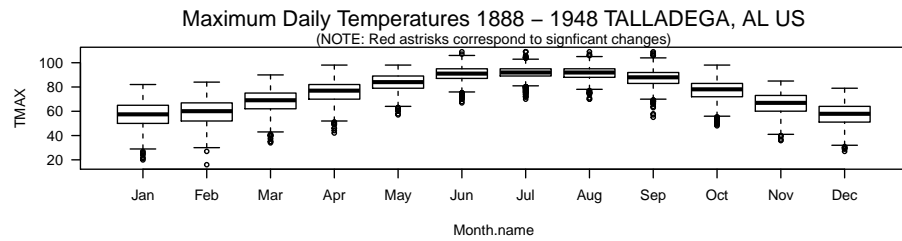
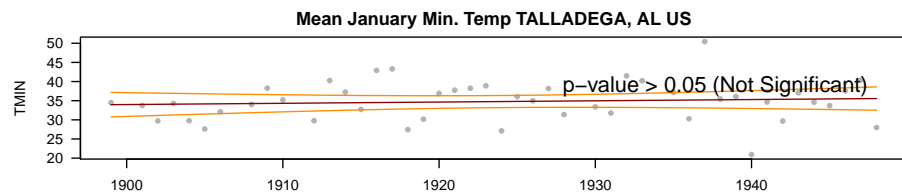


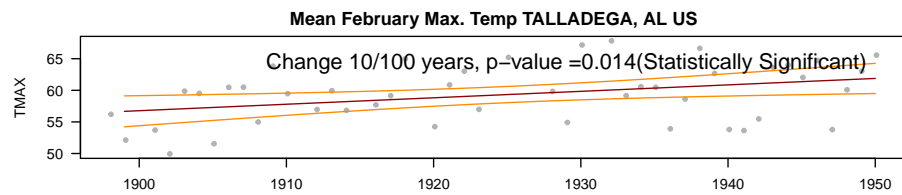
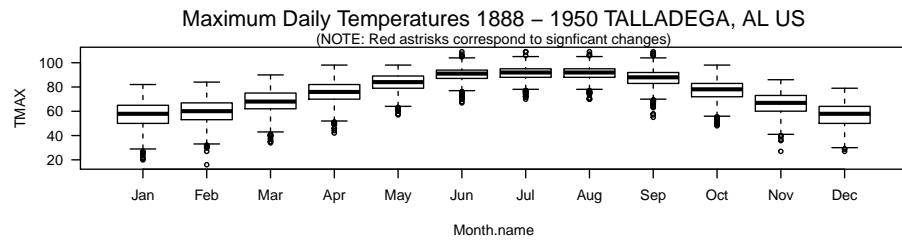
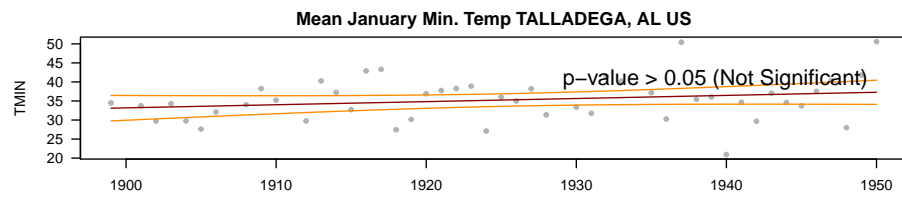


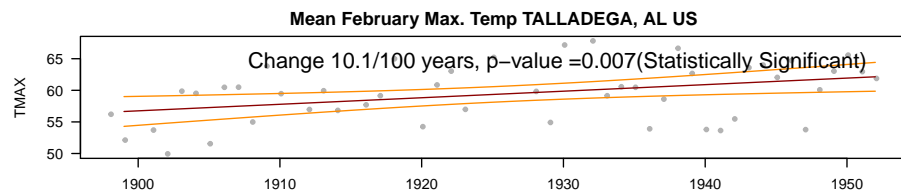
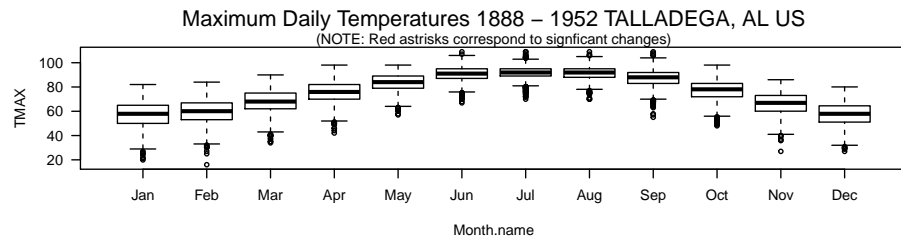
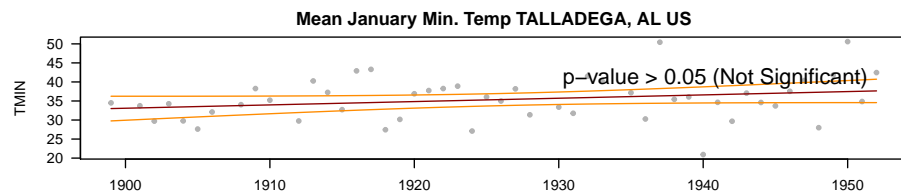


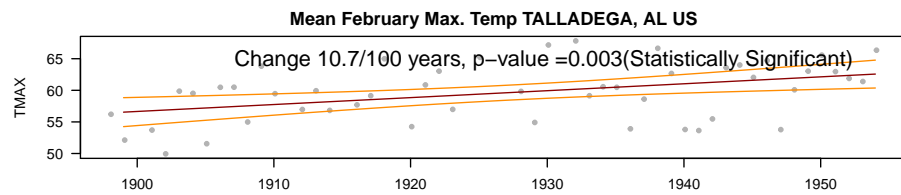
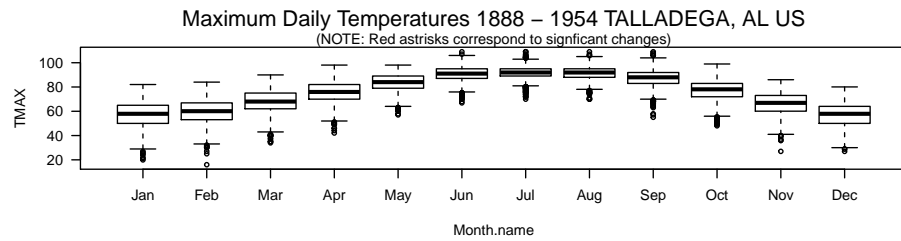
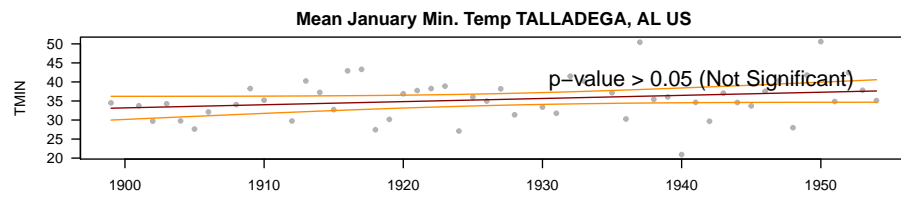


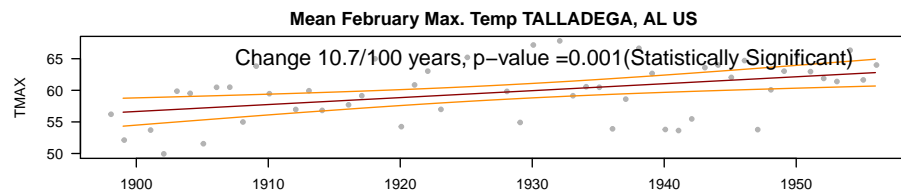
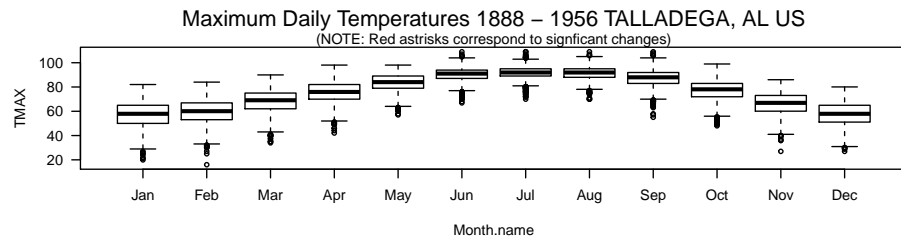
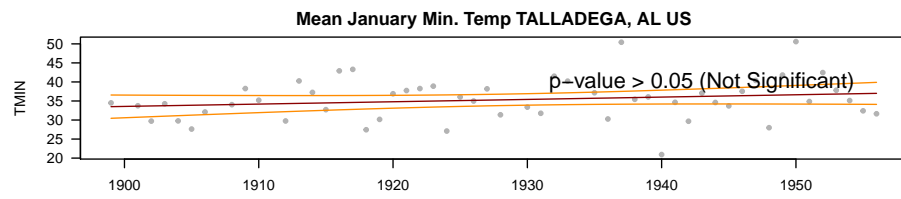


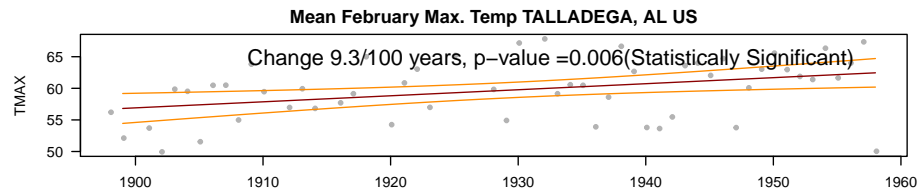
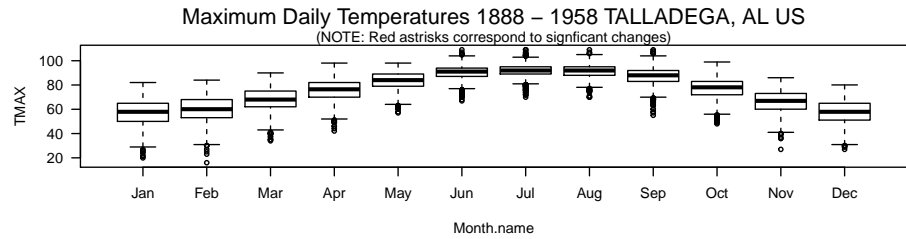
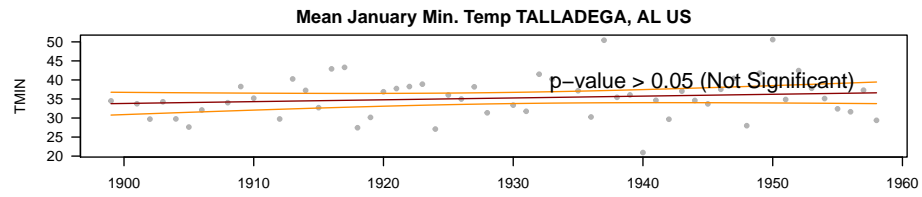


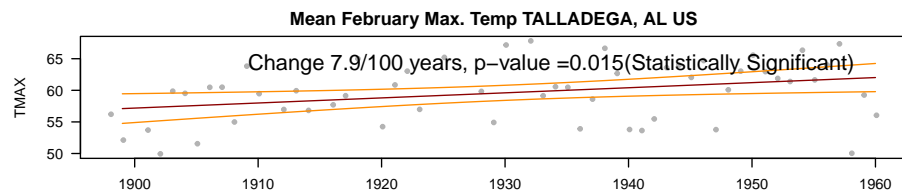
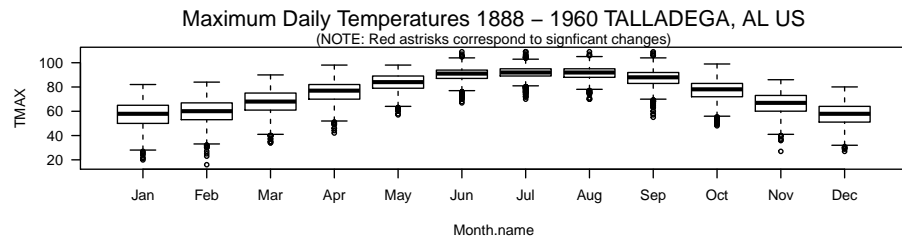
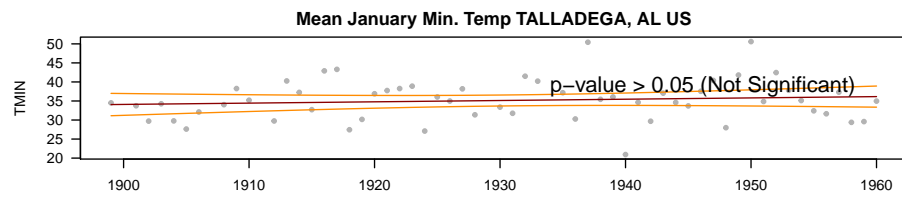


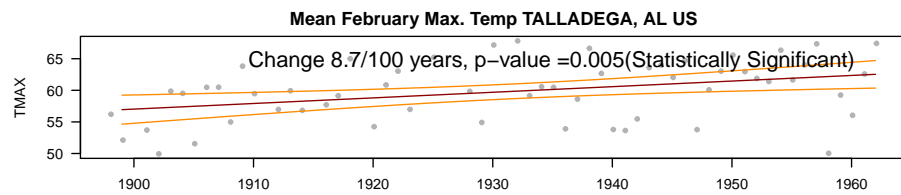
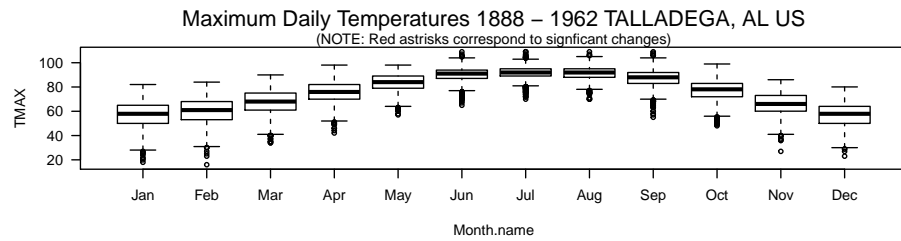
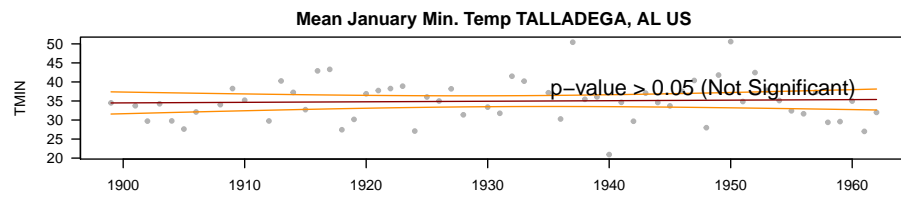


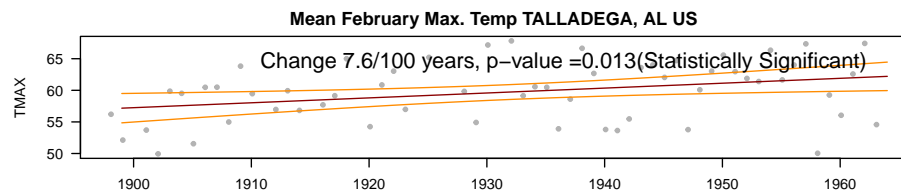
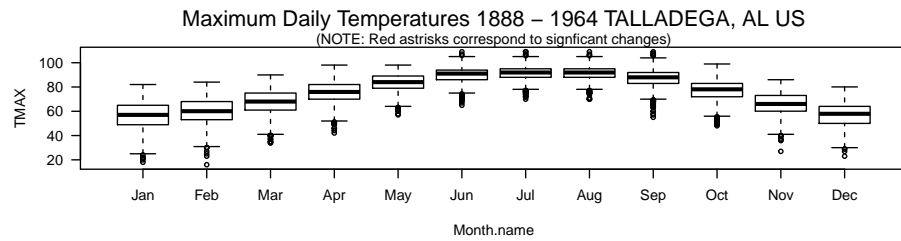
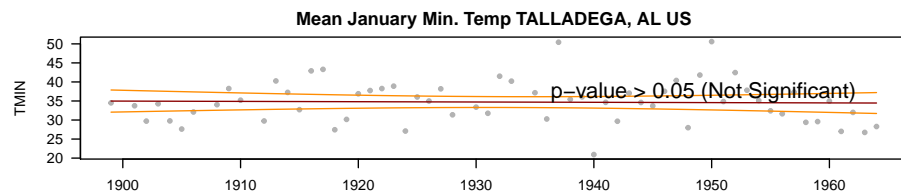


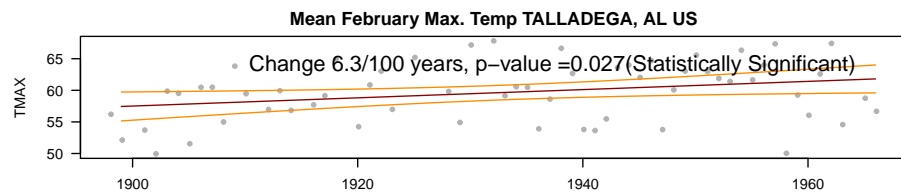
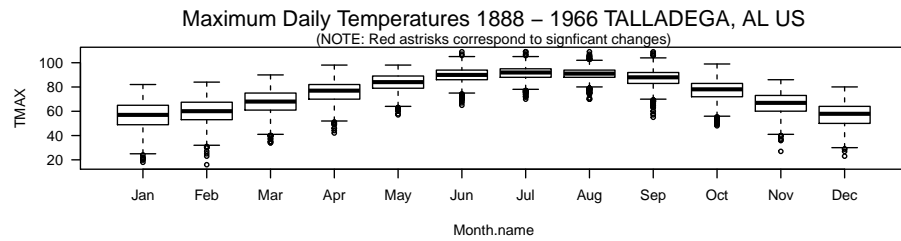
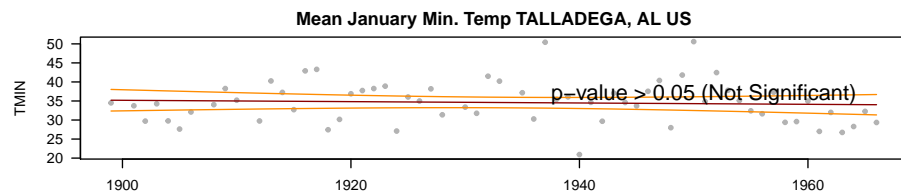


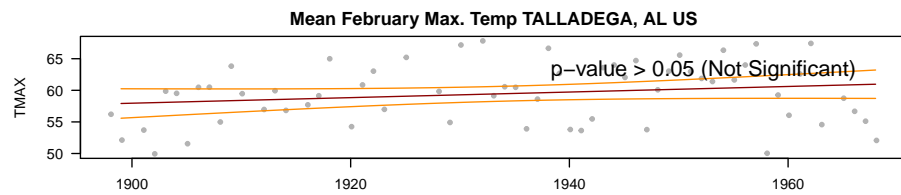
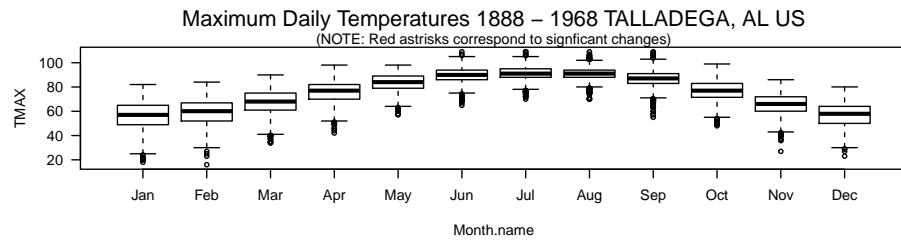
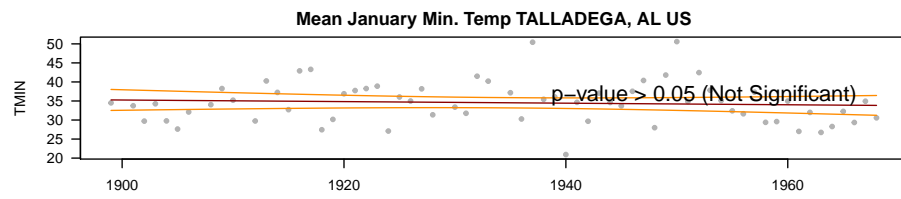


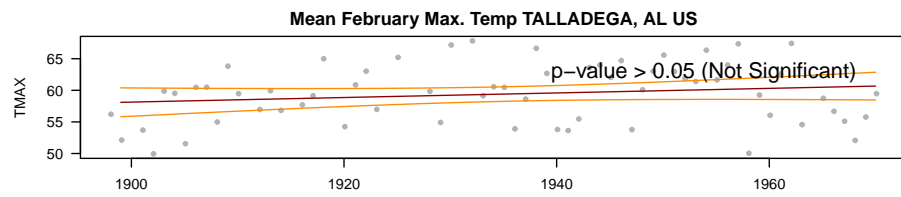
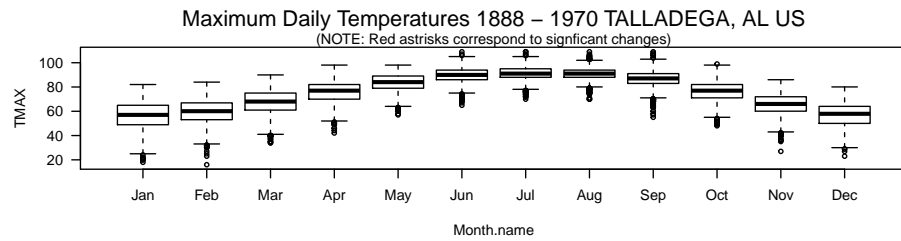
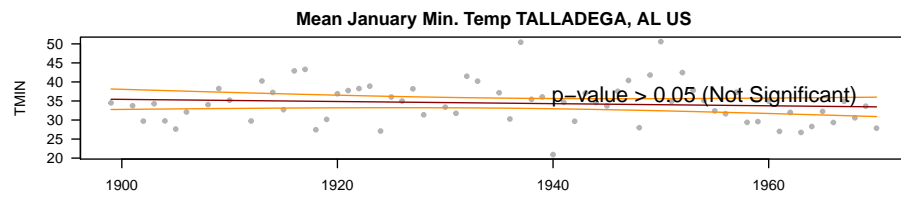


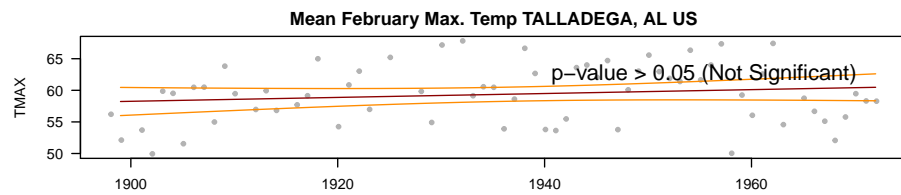
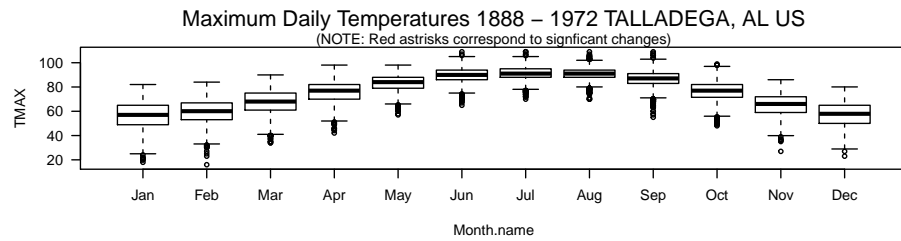
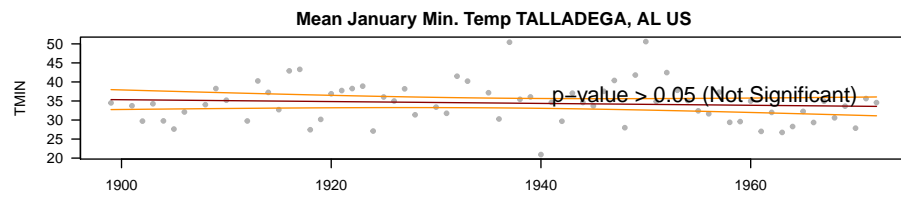


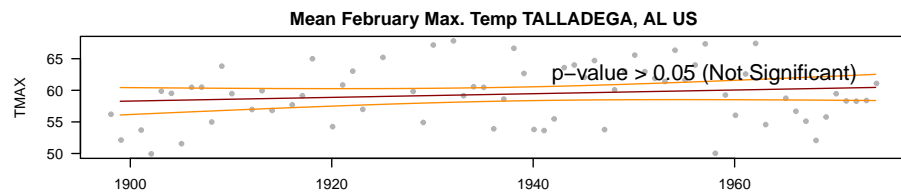
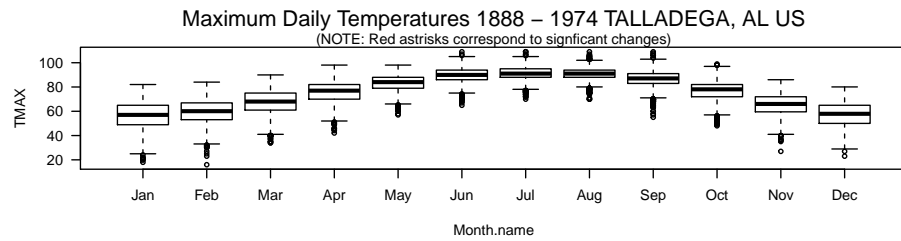
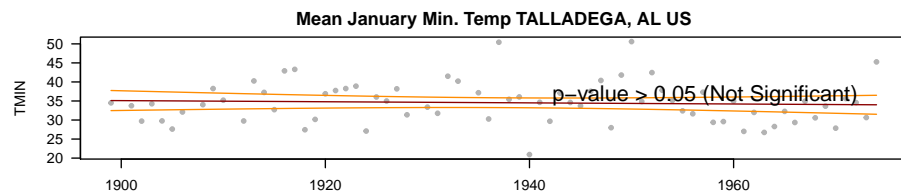


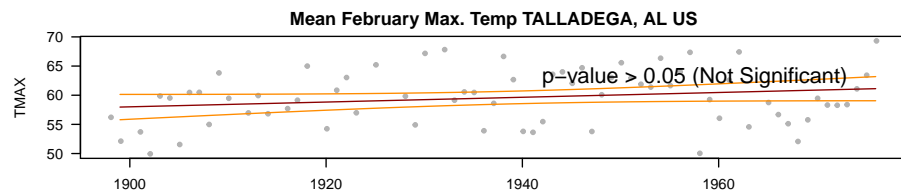
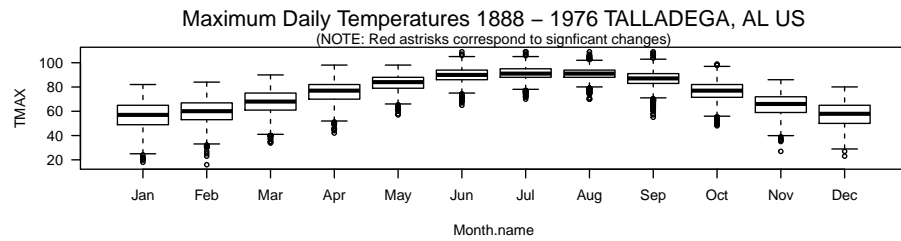
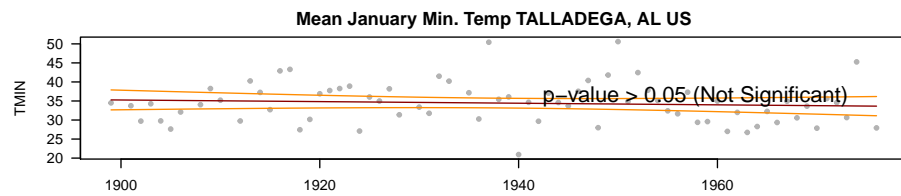


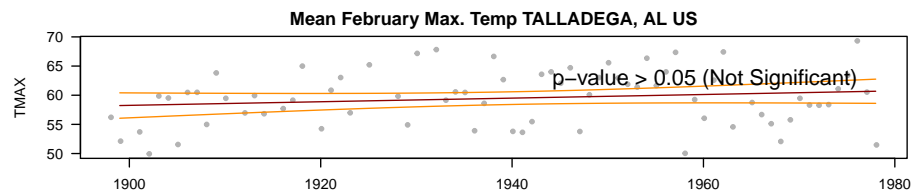
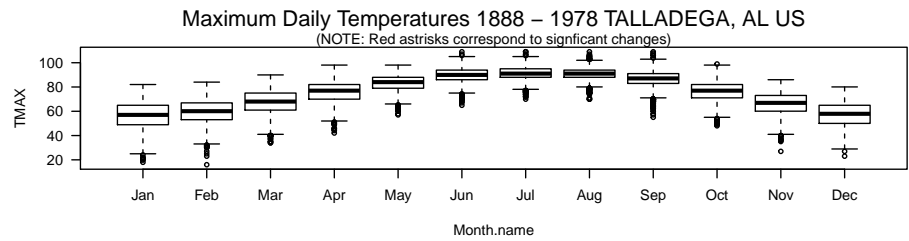
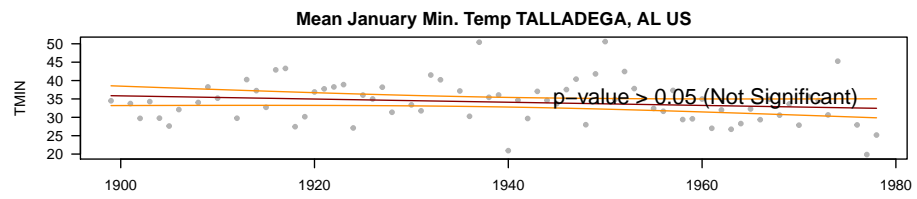


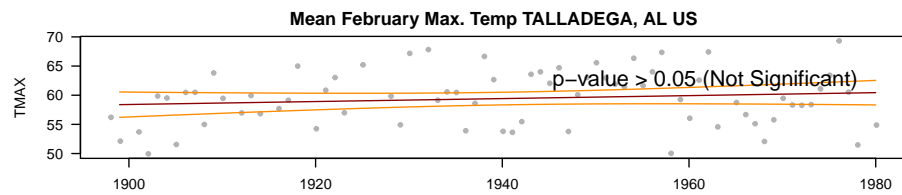
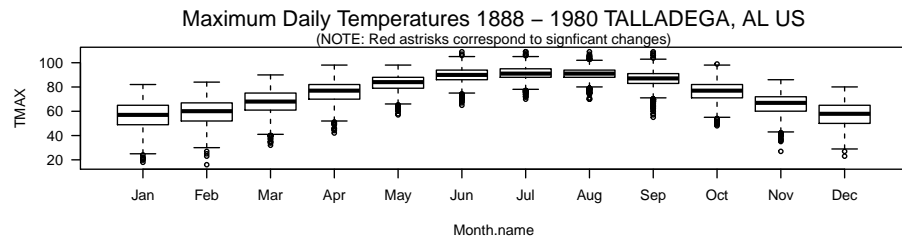
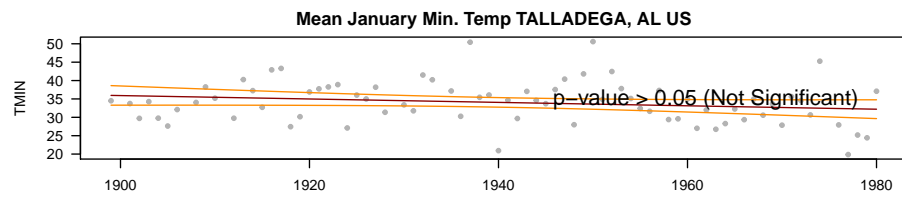


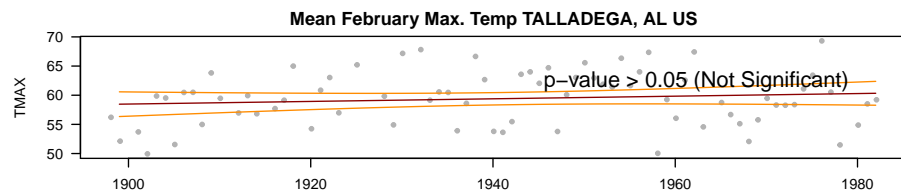
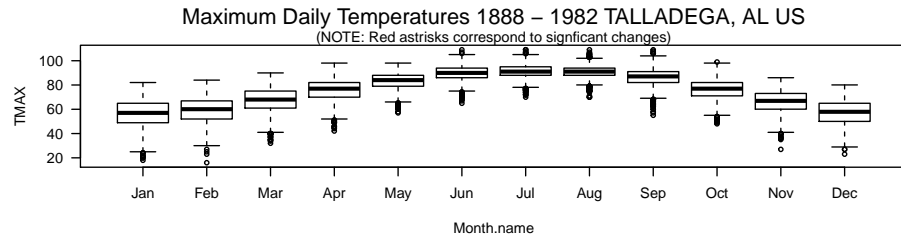
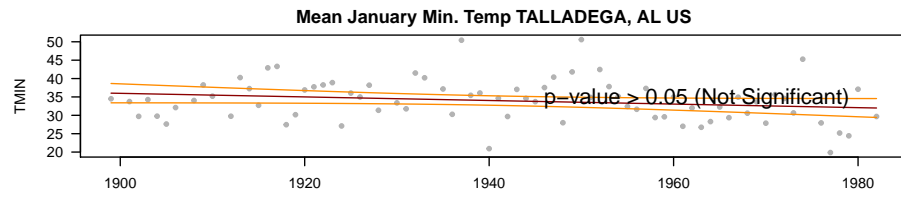


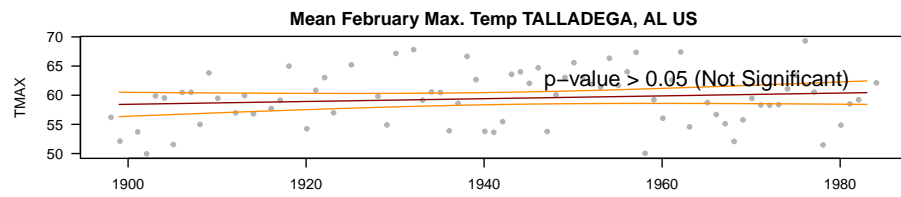
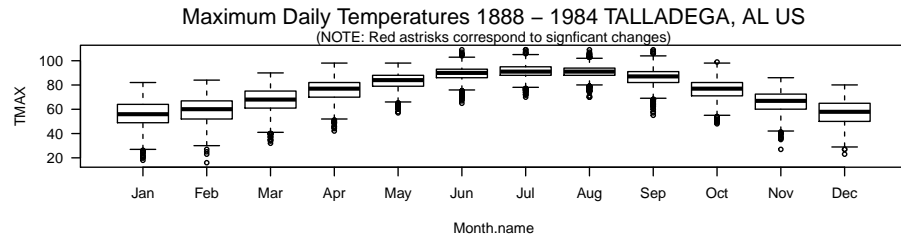
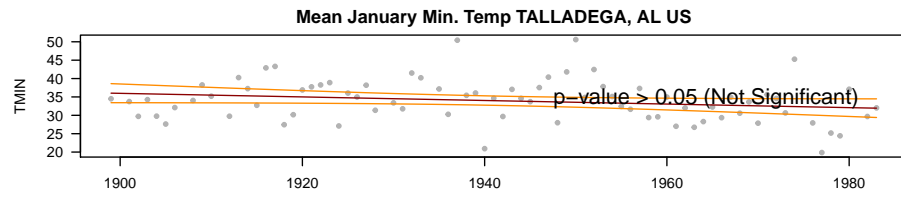


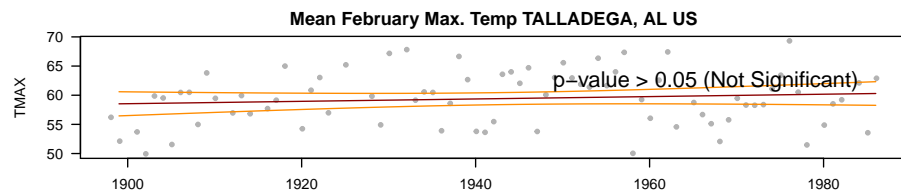
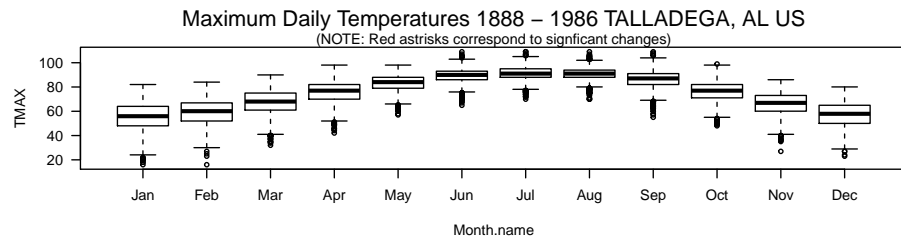
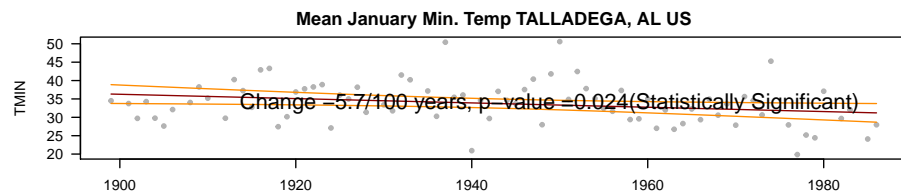


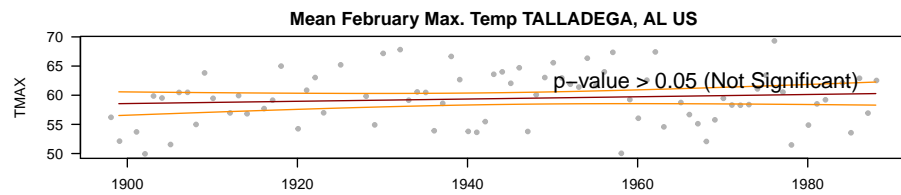
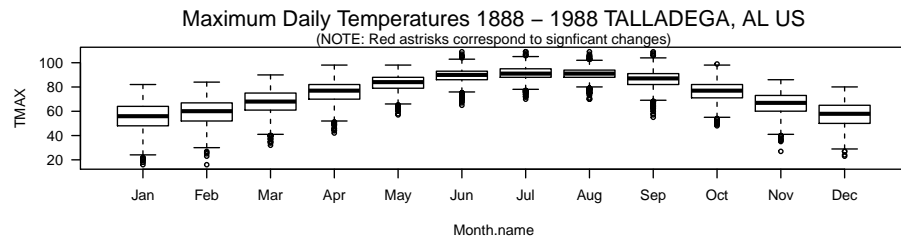
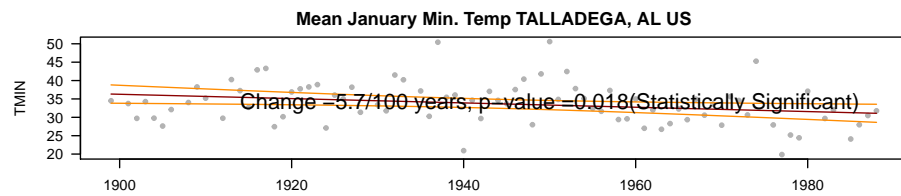


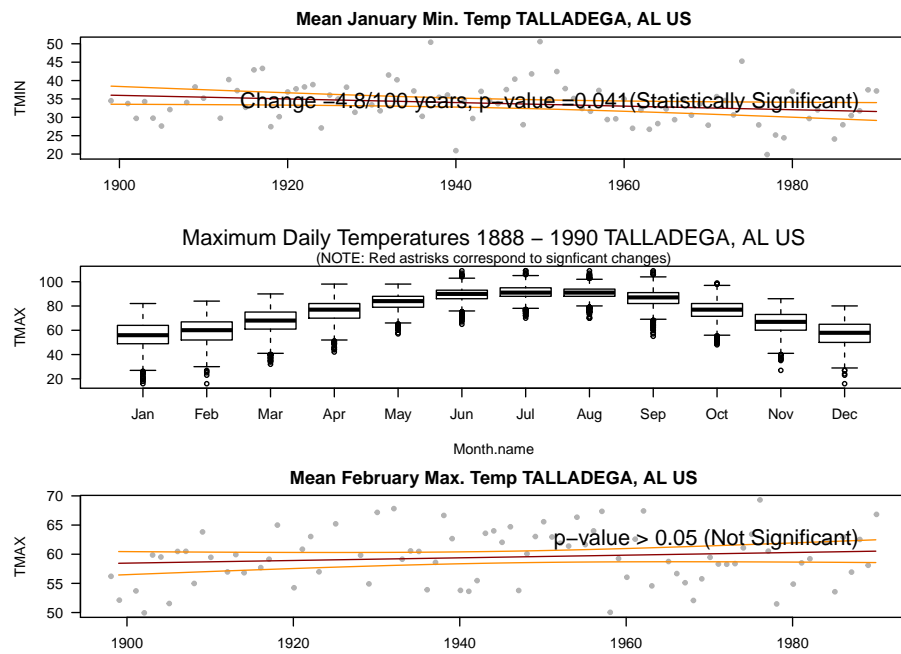


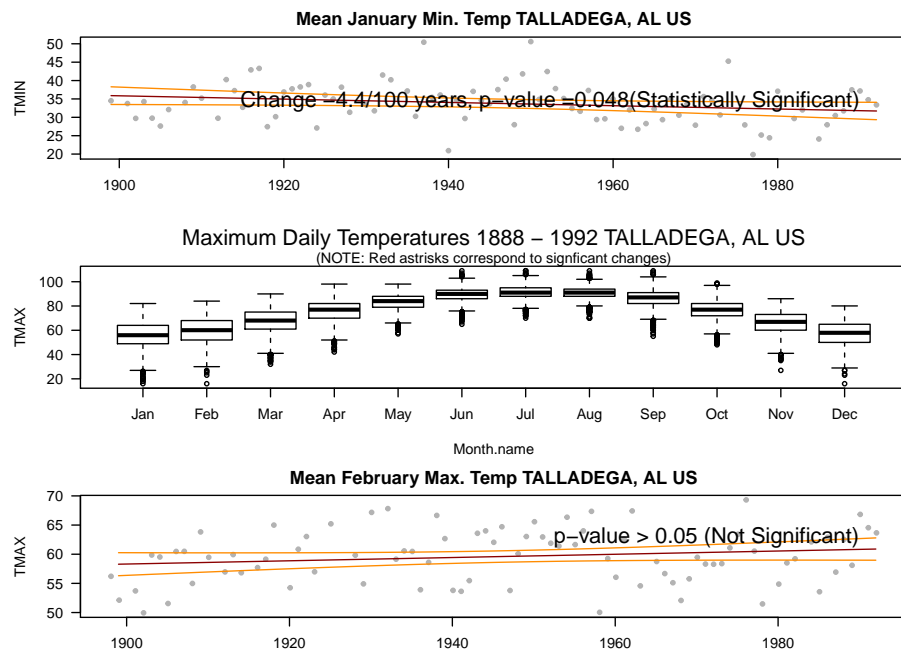


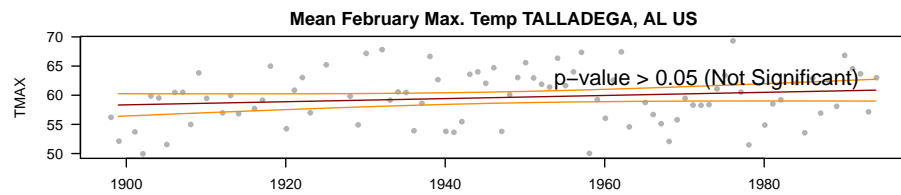
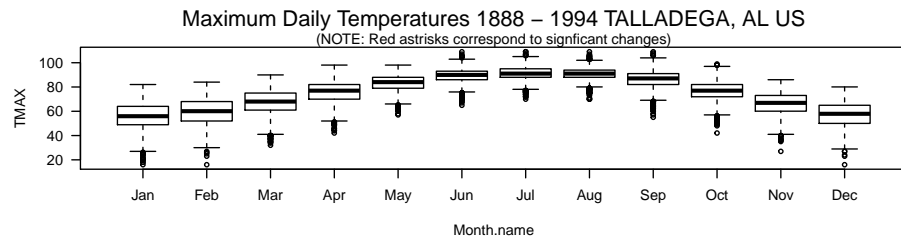
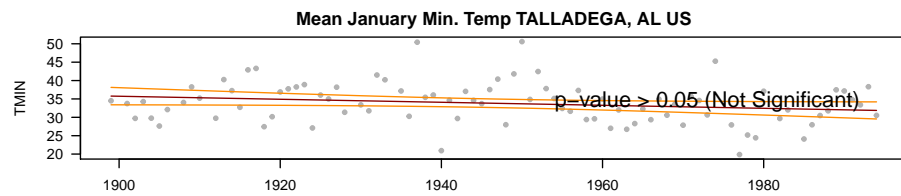


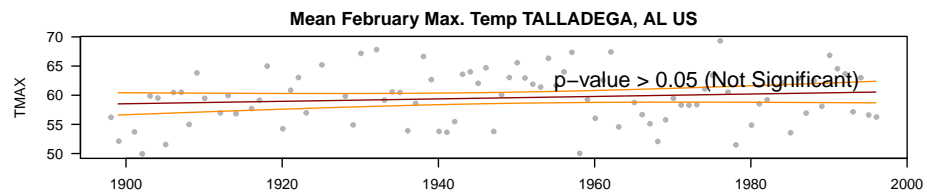
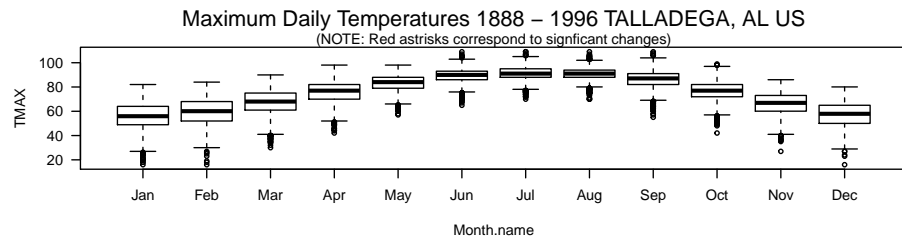
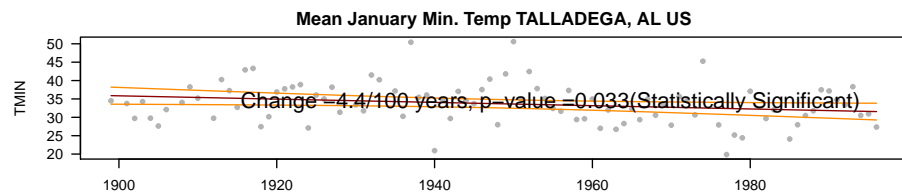


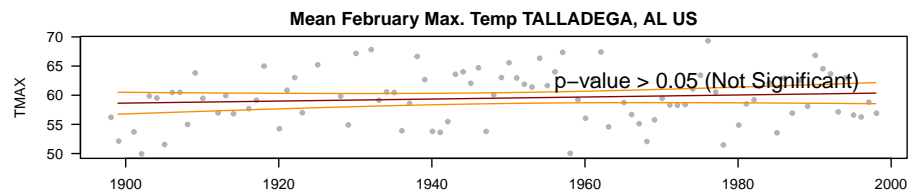
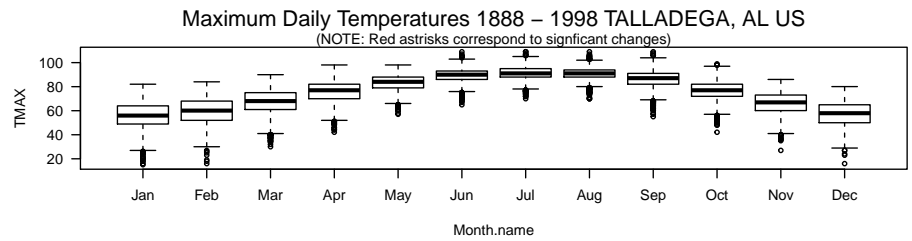
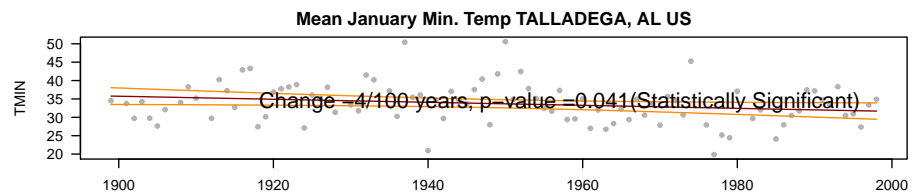


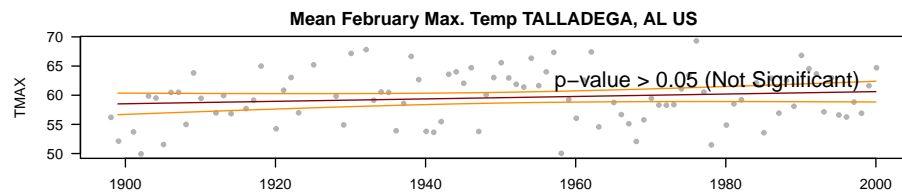
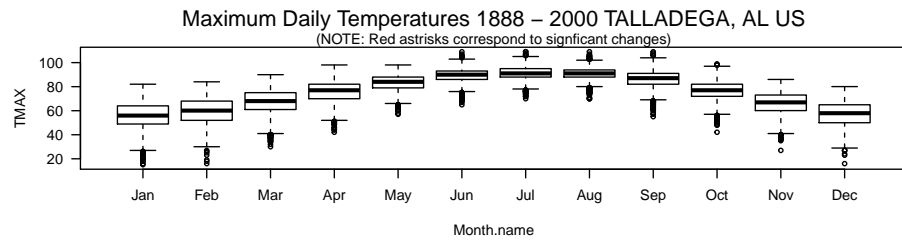
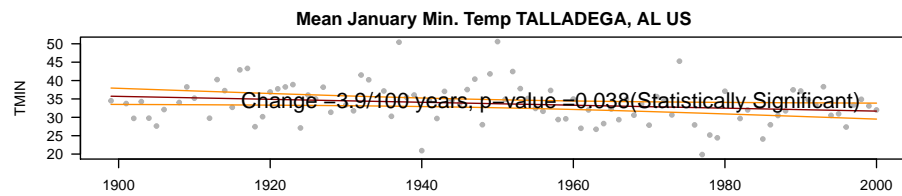


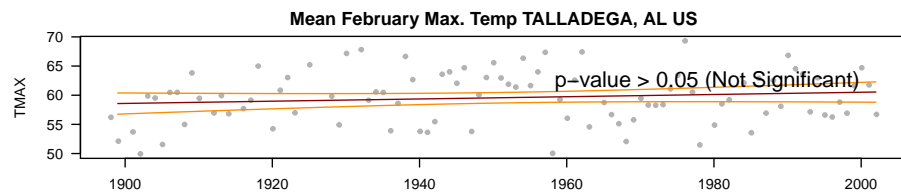
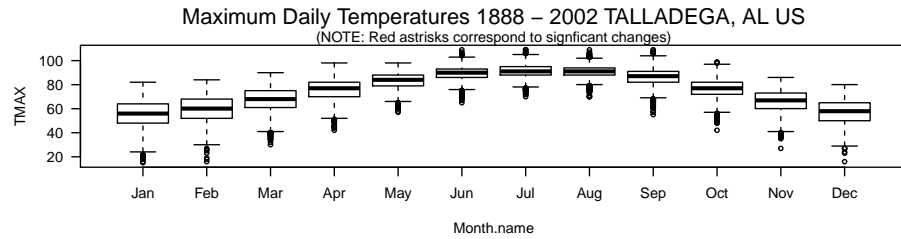
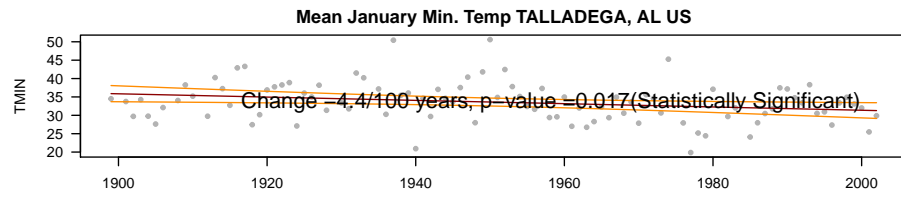


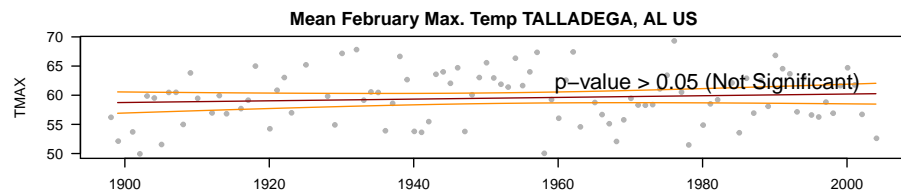
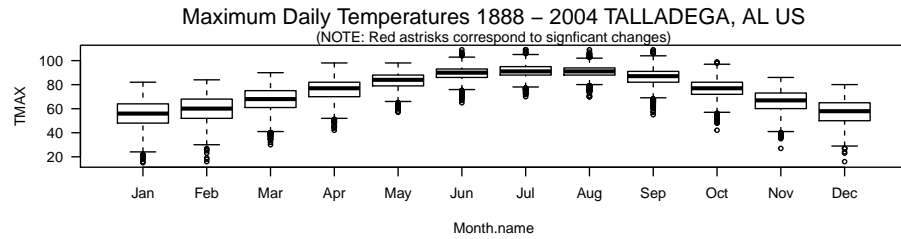
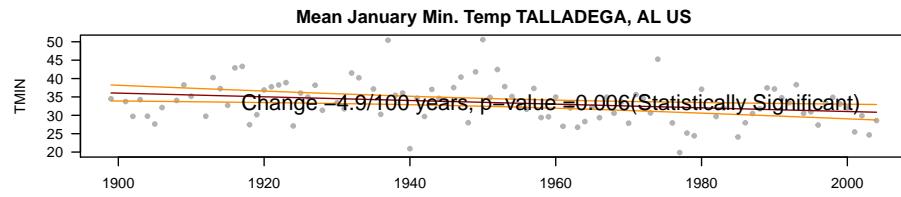


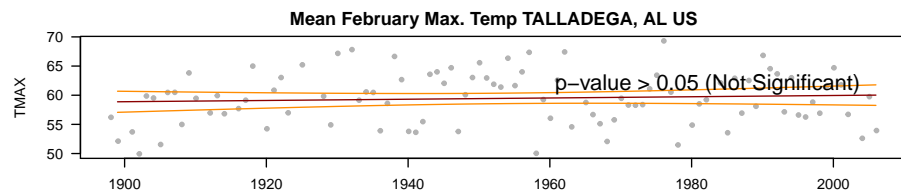
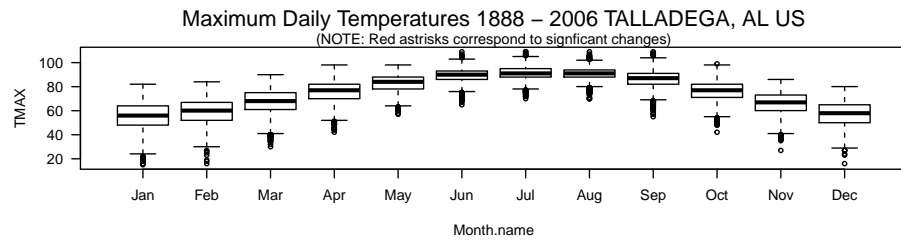
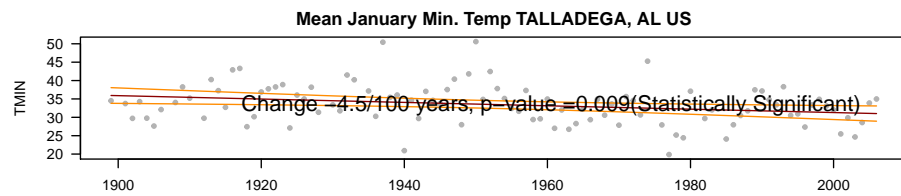


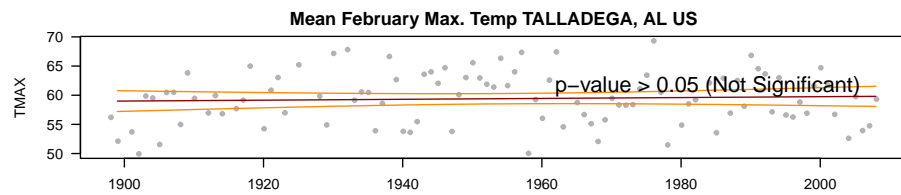
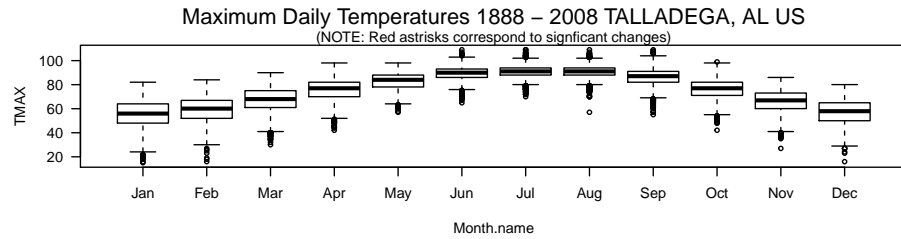
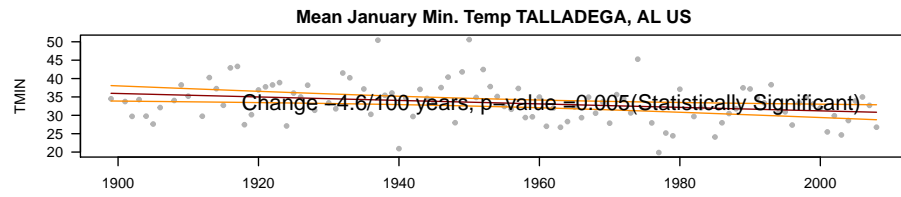


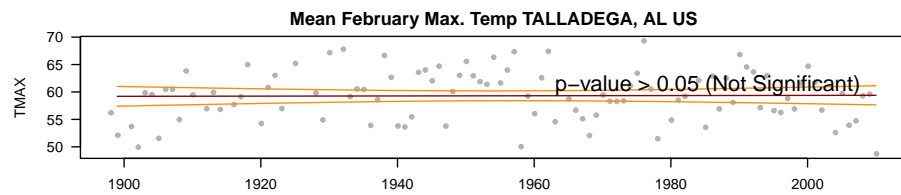
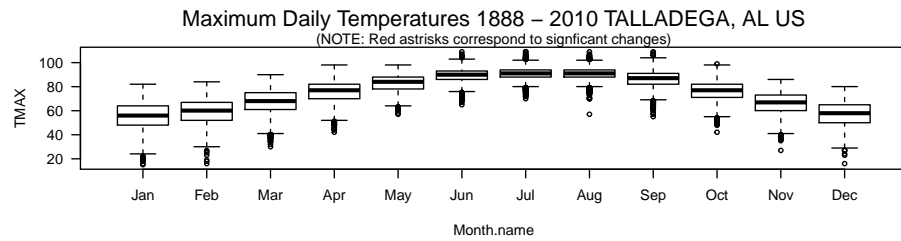
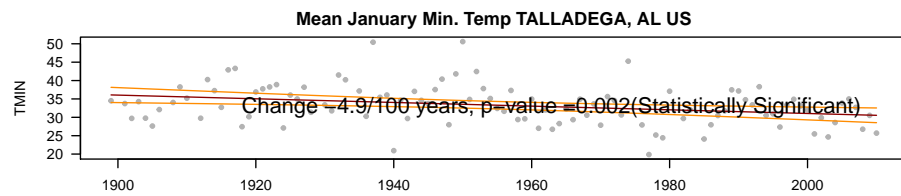


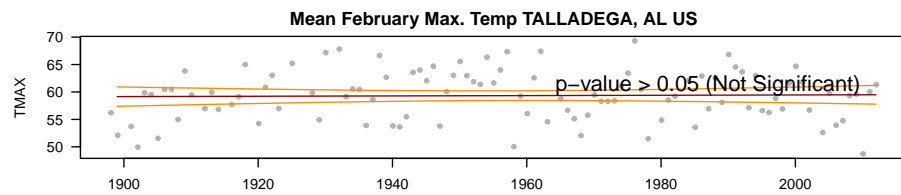
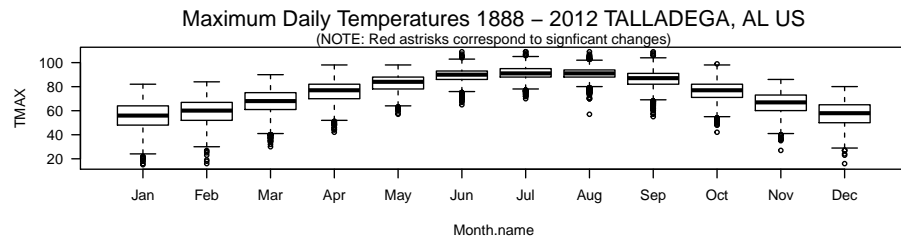
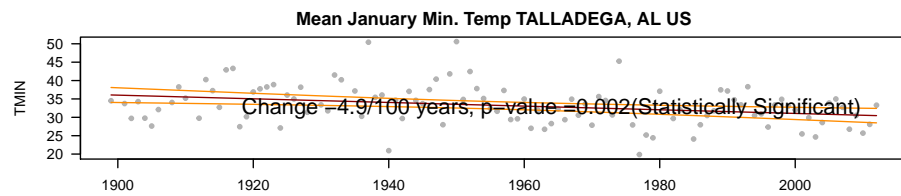


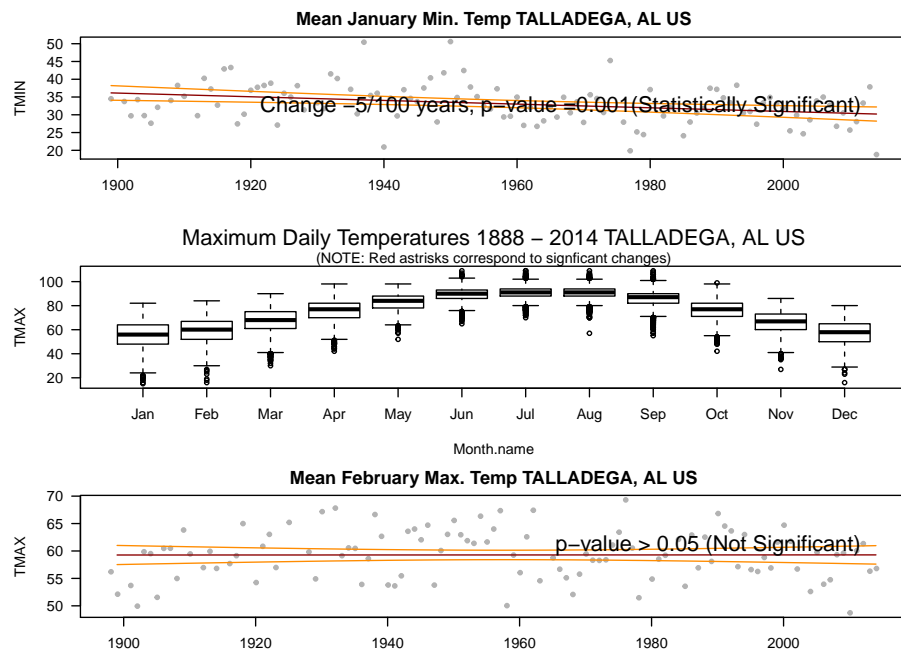


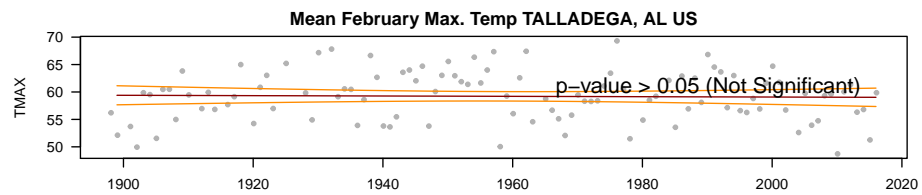
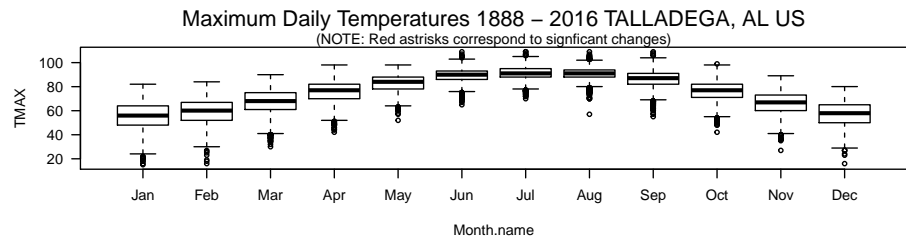
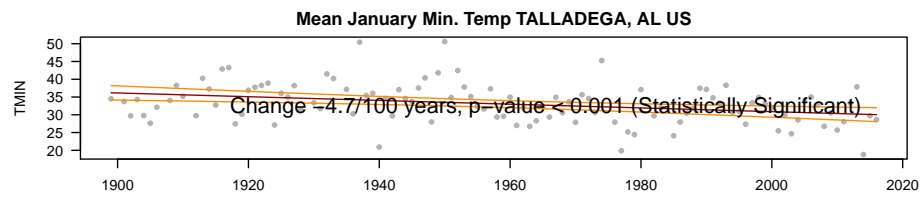


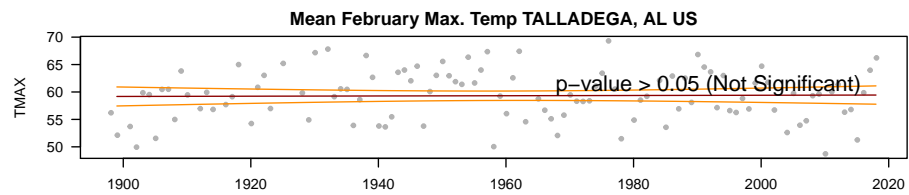
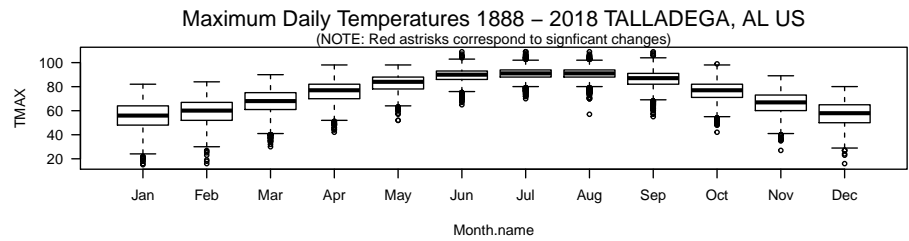
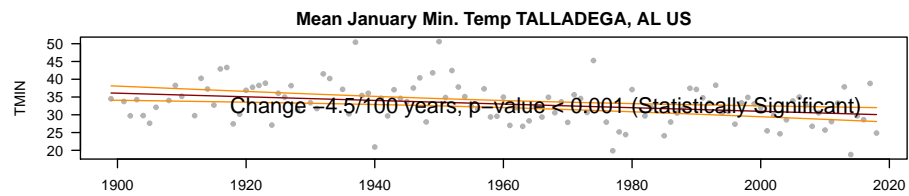


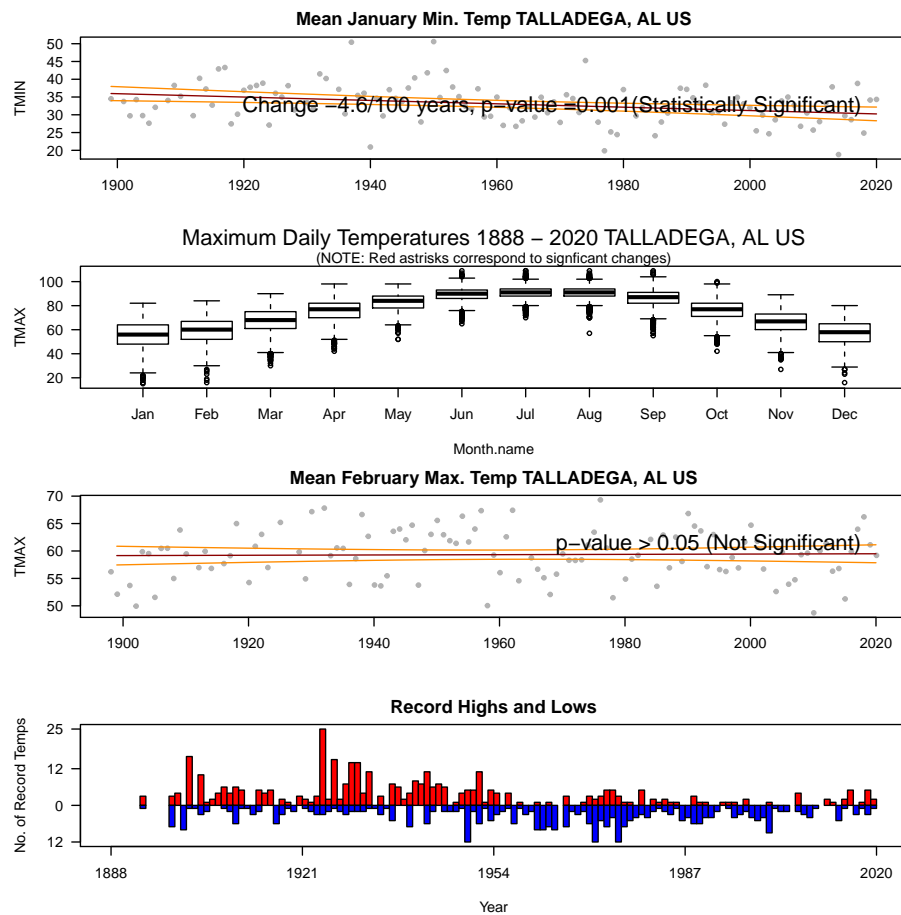












```
## null device
##          1
```

4.7 KISS

Keeping it simple is critical in communicating scientific information. In this section, I try to come up with a consistent message for every state and a simple graphic.

4.7.1 Change Point Analysis

First, TMIN and TMAX and change point analysis...

<https://cran.r-project.org/web/packages/mcp/readme/README.html>

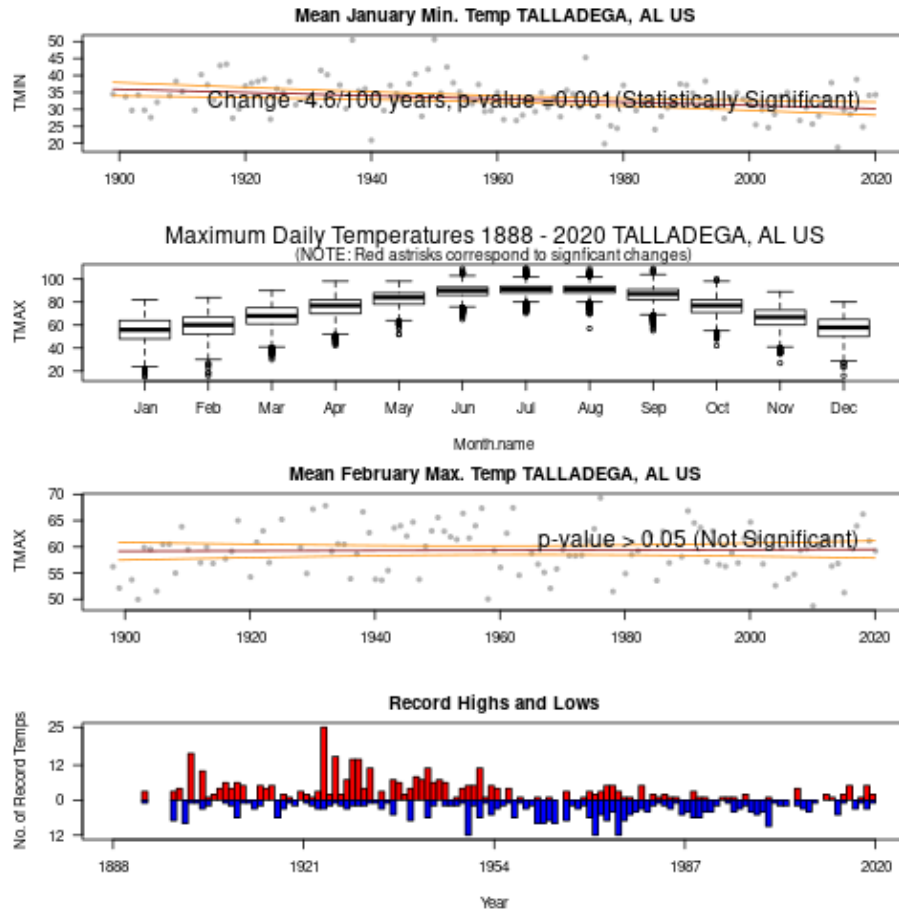


Figure 4: Climate can be analyzed using several types of lenses. In this case, we have analyzed show the months with the greatest changes. The first figure is monthly average of TMINs (daily low temperatures) with a best fit line. The second figure shows the monthly TMAX range and asterisks indicate singnificant changes over the station record and the third figure is the trend for these TMAXs over time and includes the best fit line. The final figure shows the daily temperatures that have been the highest on record (in red). In some cases, climate change has created more records in the recent decades, while other stations seem don't show that trend.

```

dyn.load('/opt/jags/4.3.1/lib/libjags.so.4')
dyn.load('/opt/jags/3.6/lib/libjags.so.4')
library(mcp)

# Define the model
model = list(
  response ~ 1, # plateau (int_1)
  ~ 0 + time, # joined slope (time_2) at cp_1
  ~ 1 + time # disjointed slope (int_3, time_3) at cp_2
)

# Get example data and fit it
ex = mcp_example("demo")
fit = mcp(model, data = ex$data)

dyn.load('/opt/jags/4.3.1/lib/libjags.so.4')
library(mcp)
# library(rlang)

# Simulate
set.seed(42) # I always use 42; no fiddling
df = data.frame(
  x = 1:100,
  y = c(rnorm(30, 2), rnorm(40, 0), rnorm(30, 1))
)

# Plot it
plot(df)
abline(v = c(30, 70), col="red")

model = list(y~1, 1~1, 1~1) # three intercept-only segments
fit_mcp = mcp(model, data = df, par_x = "x")

summary(fit_mcp)

library(patchwork)
plot(fit_mcp) + plot_pars(fit_mcp, pars = c("cp_1", "cp_2"), type = "dens_overlay")

model = list(
  price ~ 1 + ar(2),
  ~ 0 + time + ar(1)
)
ex = mcp_example("ar")
fit = mcp(model, ex$data)

```

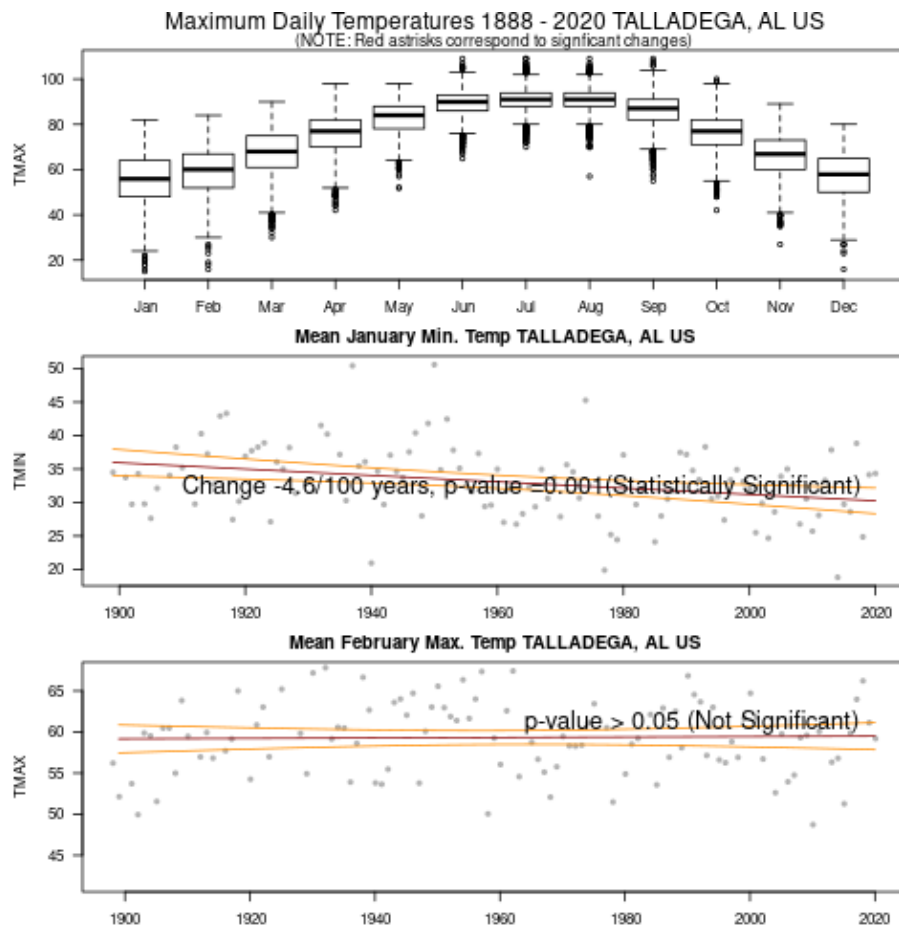


Figure 5: Keep it simple stupid!

```
summary(fit)
```

Let's create a figure that simplifies the narrative, if we can!

```
## pdf
## 2
```

4.8 Temp & Precipitation Probability

To highlight the patterns of change, it might be useful to analyze how the probability distribution might change – we can use a normal probability distribution as a theoretical distribution (and we can check if this distribution is appropriate).

prate with a Chi-Square test), or we can use the data to create a empirical distribution, which is my favored approach.

I started with decade bins, but used 20 years bins (scores) to simplify the graphics while keeping a pretty good temporal resolution.

```
library(wesanderson)

h.ramp <- rev(heat.colors(length(unique(GSOM2$Score))+1))[-1]
h.ramp <- wes_palette("Zissou1", length(unique(GSOM2$Score)),
  type = "continuous")[1:length(unique(GSOM2$Score))]
#TMAX.anomaly.Score = aggregate(TMAX.anom ~ Score, GSOM2, mean)
#TMAX.sd.anomaly.Score = aggregate(TMAX.anom ~ Score, GSOM2, sd)

# I hate list!
TMAX.anomaly.list = aggregate(TMAX.anom ~ Score, GSOM2,
  FUN = function(x) c(mean = mean(x), sd = sd(x)))
TMIN.anomaly.list = aggregate(TMIN.anom ~ Score, GSOM2,
  FUN = function(x) c(mean = mean(x), sd = sd(x)))
PPT.anomaly.list = aggregate(PPT.anom ~ Score, GSOM2,
  FUN = function(x) c(mean = mean(x), sd = sd(x)))

setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/Social_Media")
png(paste0("png//", fips$State, "-", stid, "-GSOM-normalPDF.png"),
  width = 480, height = 480, units = "px", pointsize = 12, bg = "white")
# TMIN
par(mfrow=c(2,2), las=1)
Anom.x = seq(min(GSOM2$TMIN.anom), max(GSOM2$TMIN.anom), by=.1)
Anom.y = max(dnorm(Anom.x,
  mean=TMIN.anomaly.list$TMIN.anom[1,1],
  sd=TMIN.anomaly.list$TMIN.anom[1,2]))*1.2

plot(Anom.x, dnorm(Anom.x,
  mean=TMIN.anomaly.list$TMIN.anom[1, 1],
  sd=TMIN.anomaly.list$TMIN.anom[1, 2]),
  ty="l", col=h.ramp[1], ylim=c(0, Anom.y),
  ylab="Density", xlab="TMIN Anomaly (F)", main="")

abline(v=TMIN.anomaly.list$TMIN.anom[1,1], col=h.ramp[1], lwd=2)
for(i in 2:nrow(TMIN.anomaly.list)){
  lines(Anom.x, dnorm(Anom.x,
    mean=TMIN.anomaly.list$TMIN.anom[i, 1],
    sd=TMIN.anomaly.list$TMIN.anom[i, 2]), col=h.ramp[i])
}
abline(v=TMIN.anomaly.list$TMIN.anom[i,1], col=h.ramp[i], lwd=2)
```

```

Delta = TMIN.anomaly.list$TMIN.anom[i,1] - TMIN.anomaly.list$TMIN.anom[1,1]

text(TMIN.anomaly.list$TMIN.anom[i,1], Anom.y*.9,
     paste0("Change ", round(Delta, 1), "F"), pos=4)

mtext(paste0(fips$State, " (", GSOM_Longest$name, ")"), side=3, line=2)

# TMAX
Anom.x = seq(min(GSOM2$TMAX.anom), max(GSOM2$TMAX.anom), by=.1)
Anom.y = max(dnorm(Anom.x,
  mean=TMAX.anomaly.list$TMAX.anom[1,1],
  sd=TMAX.anomaly.list$TMAX.anom[1,2]))*1.2

plot(Anom.x, dnorm(Anom.x,
  mean=TMAX.anomaly.list$TMAX.anom[1, 1],
  sd=TMAX.anomaly.list$TMAX.anom[1, 2]),
  ty="l", col=h.ramp[1], ylim=c(0, Anom.y),
  ylab="Density", xlab="TMAX Anomaly (F)")

abline(v=TMAX.anomaly.list$TMAX.anom[1,1], col=h.ramp[1], lwd=2)
for(i in 2:nrow(TMAX.anomaly.list)){
  lines(Anom.x, dnorm(Anom.x,
    mean=TMAX.anomaly.list$TMAX.anom[i, 1],
    sd=TMAX.anomaly.list$TMAX.anom[i, 2]), col=h.ramp[i])
}
abline(v=TMAX.anomaly.list$TMAX.anom[i,1], col=h.ramp[i], lwd=2)
Delta = TMAX.anomaly.list$TMAX.anom[i,1] - TMAX.anomaly.list$TMAX.anom[1,1]

text(TMAX.anomaly.list$TMAX.anom[i,1], Anom.y*.9,
     paste0("Change ", round(Delta, 1), "F"), pos=4)

# PPT
Anom.x = seq(min(GSOM2$PPT.anom), max(GSOM2$PPT.anom), by=.1)
Anom.y = max(dnorm(Anom.x,
  mean=PPT.anomaly.list$PPT.anom[1,1],
  sd=PPT.anomaly.list$PPT.anom[1,2]))*1.2

plot(Anom.x, dnorm(Anom.x,
  mean=PPT.anomaly.list$PPT.anom[1, 1],
  sd=PPT.anomaly.list$PPT.anom[1, 2]),
  ty="l", col=h.ramp[1], ylim=c(0, Anom.y),
  ylab="Density", xlab="PPT Anomaly")

abline(v=PPT.anomaly.list$PPT.anom[1,1], col=h.ramp[1], lwd=2)
for(i in 2:nrow(PPT.anomaly.list)){

```

```

lines(Anom.x, dnorm(Anom.x,
  mean=PPT.anomaly.list$PPT.anom[i, 1],
  sd=PPT.anomaly.list$PPT.anom[i, 2]), col=h.ramp[i])
}
abline(v=PPT.anomaly.list$PPT.anom[i,1], col=h.ramp[i], lwd=2)
Delta = PPT.anomaly.list$PPT.anom[i,1] - PPT.anomaly.list$PPT.anom[1,1]

text(PPT.anomaly.list$PPT.anom[i,1], Anom.y*.9,
  paste0("Change ", round(Delta, 1), " inches"), pos=4)

dev.off()

## pdf
## 2

```

This figure is pretty effective, but still needs work.

4.9 Using library densEstBayes

Now, I used a screen split to look at the distribution of the temperate anomalies. First, we look at a simple histogram of the entire dataset.

```

par(mfrow=c(1,1))
hist(GSOM2$TMIN.anom, col = "gold",
  main = "", probability = TRUE,
  xlab = "Minimum Temperature Anomaly (F)")

```

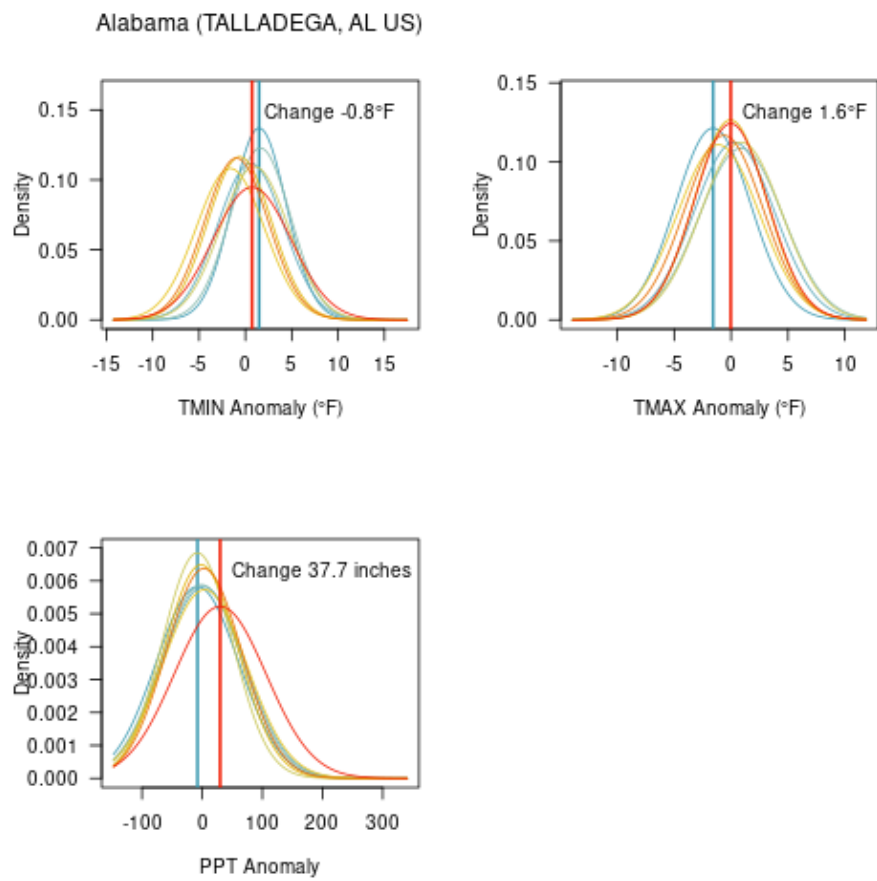
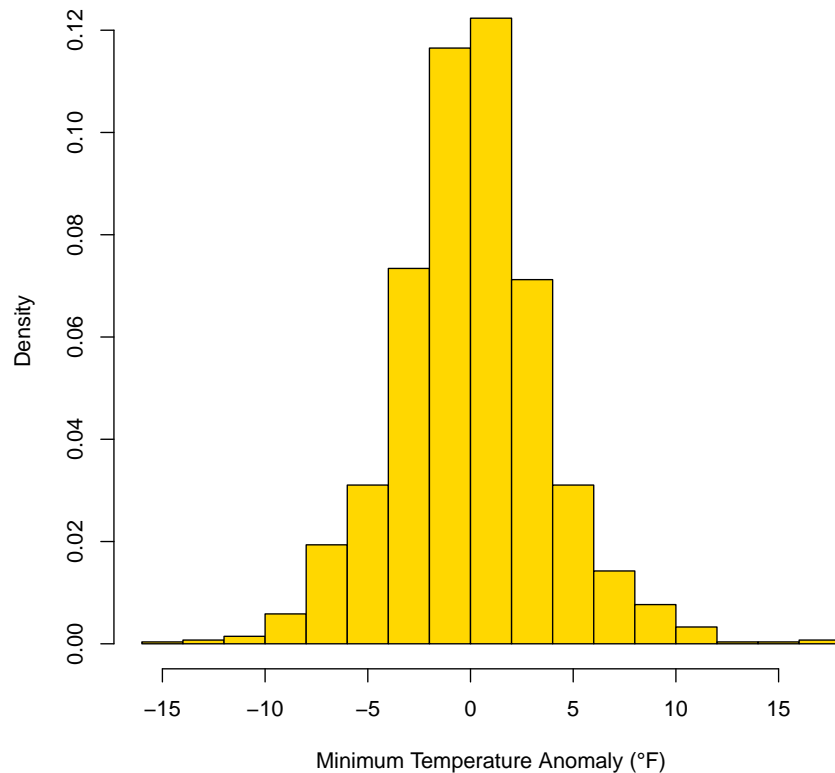


Figure 6: The changing in monthly temperature data, assuming a normal probability distribution.



The data center around zero, as expected, but are these normally distributed?

For TMAX there is a 0.031 probability that the distribution is the same as the normal distribution. For TMIN there is a 0 probability that the distribution is the same as the normal distribution. For PPT is a 0 probability that the distribution is the same as the normal distribution.

```
if(shapiro.test(GSOM2$TMAX.anom)$p.value<.05 |
  shapiro.test(GSOM2$TMIN.anom)$p.value<.05 |
  shapiro.test(GSOM2$PPT.anom)$p.value<.05) text="to avoid " else text="to use"
```

These values suggest that there is good reason to avoid the normal probability distribution.

Next we use a function to estimate the probability distribution using a markof chain the creates an estimated probability distribution. This doesn't always work when the distribution is not even and their only 10 years of data

per slot. I suspect, I should make this by every 20 years. Plus that will go way faster and I think the data visualization will be more robust.

```
setwd("/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/Social_Media")

if(!file.exists(paste0("png/", fips$State, "-", stid, "-GSOM-estDensity.png"))){
  print("Creating estimated density distribution")

png(paste0("png/", fips$State, "-", stid, "-GSOM-estDensity.png"),
    width = 480, height = 320, units = "px", pointsize = 12, bg = "white")

# Split Screen TMAX Legend TMIN
# screen with values for left, right, bottom, and top.
split.screen(rbind(c(.01, 0.99, 0.86, 0.95),
                    c(0.01, 0.45, 0.01, 0.85),
                    c(0.45, 0.55, 0.01, 0.85),
                    c(0.55, 0.99, 0.01, 0.85)))

screen(1)
par(mar=c(0,0,0,0))
plot(NA, xaxt='n', yaxt='n', bty='n', ylab='', xlab='', xlim=c(0,10), ylim=c(0, 10))
mtext(paste0(fips$State, " (", GSOM_Longest$name, ")"), side=3, line=-1, cex=1.4)

screen(2)

# Determine xg (range)
dest <- densEstBayes(GSOM2$TMIN.anom, method = "NUTS"); dest$range.x

control = densEstBayes.control(range.x = dest$range.x,
                               numBins = 401,
                               numBasis = 50, sigmabeta = 1e5, ssigma = 1000,
                               convToler = 1e-5, maxIter = 500, nWarm = NULL,
                               nKept = NULL, nThin = 1, msgCode = 1)

#destSMFVB <- densEstBayes(GSOM2$TMIN.anom, method = "SMFVB", control = control)
#plot(destSMFVB, plotIt=T, xlab = "TMIN", main = "", setCol=h.ramp[i])
par(las=1, mar=c(4, 4, 0, 0) + 0.1)
for(i in 1:length(unique(GSOM2$Score))){
  # i = 13
  GSOM2sub = GSOM2[GSOM2$Score==sort(unique(GSOM2$Score))[i],]
  dest <- densEstBayes(GSOM2sub$TMIN.anom, method = "NUTS", control = control)
  xg = plot(dest, plotIt=FALSE)$xg
  densEstg = plot(dest, plotIt=FALSE)$densEstg

  if(i==1) plot(0, type = "n", bty = "l",
                xlim=range(xg), ylim=c(0,0.25),
```

```

        xlab = "TMIN anomaly (F)", main = "", ylab="Density")
        lines(xg, densEstg, col=h.ramp[i])
    rug(jitter(GSOM2sub$TMIN.anom, amount = 0.2), col=h.ramp[i])
}

screen(3)
par(mar=c(0,0,1,0))
plot(NA, xaxt='n', yaxt='n', bty='n', ylab='', xlab='', xlim=c(0,10), ylim=c(0,10))
#
legend("topright", inset=c(0,0), bg="transparent", bty="n",
       legend=unique(GSOM2$Score),
       fill=h.ramp, horiz=FALSE, cex=0.85)

screen(4)
par(las=1, mar=c(4, 4, 0, 0) +0.1)
# Determine xg (range)
dest <- densEstBayes(GSOM2$TMAX.anom, method = "NUTS"); dest$range.x

control = densEstBayes.control(range.x = dest$range.x,
                               numBins = 401,
                               numBasis = 50, sigmabeta = 1e5, ssigma = 1000,
                               convToler = 1e-5, maxIter = 500, nWarm = NULL,
                               nKept = NULL, nThin = 1, msgCode = 1)

for(i in 1:length(unique(GSOM2$Score))){
    # i = 13
    GSOM2sub = GSOM2[GSOM2$Score==sort(unique(GSOM2$Score))[i],]
    dest <- densEstBayes(GSOM2sub$TMAX.anom, method = "NUTS", control = control)
    xg = plot(dest, plotIt=FALSE)$xg
    densEstg = plot(dest, plotIt=FALSE)$densEstg

    if(i==1) plot(0, type = "n", bty = "l",
                  xlim=range(xg), ylim=c(0,.25),
                  xlab = "TMAX anomaly (F)", main = "", ylab="Density")
    lines(xg, densEstg, col=h.ramp[i])
    rug(jitter(GSOM2sub$TMIN.anom, amount = 0.2), col=h.ramp[i])
}
#rug(jitter(GSOM2sub$TMIN.anom, amount = 0.2), col=h.ramp[i])

close.screen(all.screens = TRUE)
dev.off()
} else {
    print("Skipping estimated density distribution chunk")}

```

The process to create these figures is very time consuming, so in general, I

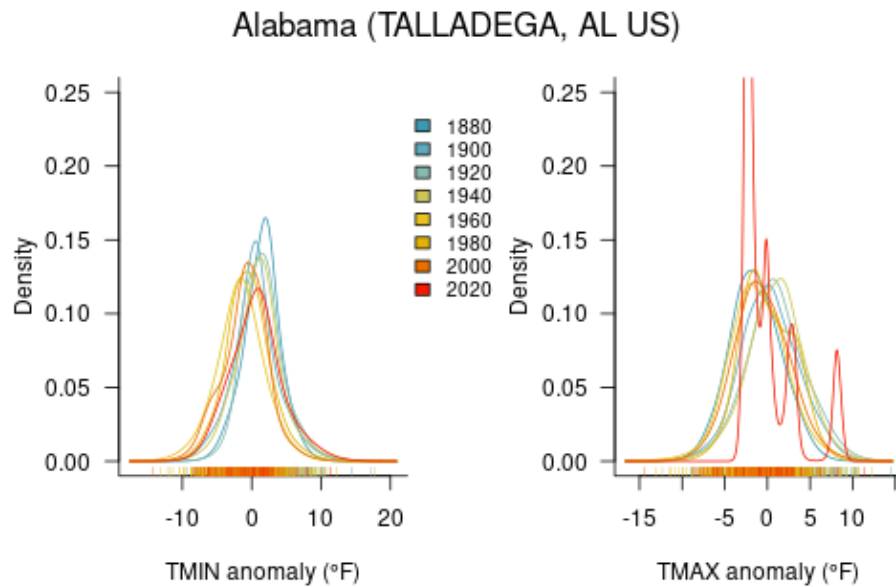


Figure 7: The changing in monthly temperature data.

need to come up with an if then statement to avoid creating these everytime!

5 Animated GIFs

So far, this creates a gif file, but I haven't been able get the gif in the pdf directly yet. I will need an additional package or create separate png that are combined. For now, we'll create a gif file to be used in separate documents.

5.1 Probability Distributions

```
setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/docs/")
if(!file.exists(paste0("Climate_gifs/",
  fips$State, "-", stid, "_GSOM-normalPDF.gif"))){
  print("Creating animated normal probability ")

# Define an image_graph size
img <- image_graph(600, 480, res = 96)

# START -----

ylim_new=NA
```

```

for(i in 1:length(unique(GSOM2$Decade)))
{
  # i = 9
  decade=(unique(GSOM2$Decade))[order(unique(GSOM2$Decade))==i]

  GSOM2sub <- GSOM2[GSOM2$Decade==decade,]
  h.ramp <- rev(heat.colors(length(unique(GSOM2$Decade))+1))[-1]

  # Determine Stats for PDFs
  TMAX.mean.anomaly.decade = aggregate(TMAX.anom ~ Decade, GSOM2sub, mean)
  TMAX.sd.anomaly.decade = aggregate(TMAX.anom ~ Decade, GSOM2sub, sd)
  names(TMAX.sd.anomaly.decade)=c("Decade", "TMAX.sd.anom")
  TMIN.mean.anomaly.decade = aggregate(TMIN.anom ~ Decade, GSOM2sub, mean)
  TMIN.sd.anomaly.decade = aggregate(TMIN.anom ~ Decade, GSOM2sub, sd)
  names(TMIN.sd.anomaly.decade)=c("Decade", "TMIN.sd.anom")

  TMAX.temp = merge(TMAX.mean.anomaly.decade, TMAX.sd.anomaly.decade, by="Decade")

  TMIN.temp = merge(TMIN.mean.anomaly.decade, TMIN.sd.anomaly.decade, by="Decade")

  GSOM.Monthly.Anom.mean.sd = merge(TMAX.temp, TMIN.temp, by="Decade")

  par(las=1, mfrow=c(1,2), mar= c(4, 4, 2, 1) + 0.1)

  Anom.x = seq(min(GSOM2$TMAX.anom), max(GSOM2$TMAX.anom), by=.1)
  plot(Anom.x, dnorm(Anom.x, mean=GSOM.Monthly.Anom.mean.sd$TMAX.anom[1],
    sd=GSOM.Monthly.Anom.mean.sd$TMAX.sd.anom[1]), ty="l", col=h.ramp[i], ylab="Density", xlab="Anomaly")
  abline(v=mean(GSOM2$TMAX.anom[GSOM2$Decade==min(GSOM$Decade)]))
  mtext(paste0(fips$State, " ", decade), side=3)

  Anom.x = seq(min(GSOM2$TMIN.anom), max(GSOM2$TMIN.anom), by=.1)
  plot(Anom.x, dnorm(Anom.x, mean=GSOM.Monthly.Anom.mean.sd$TMIN.anom[1],
    sd=GSOM.Monthly.Anom.mean.sd$TMIN.sd.anom[1]), ty="l", col=h.ramp[i], ylab="Density", xlab="Anomaly")
  abline(v=mean(GSOM2$TMIN.anom[GSOM2$Decade==min(GSOM$Decade)]))
  mtext(paste0(fips$State, " ", decade), side=3)
}

par(las=1, mfrow=c(1,2), mar= c(4, 4, 2, 1) + 0.1)

TMAX.anomaly.decade = aggregate(TMAX.anom ~ Decade, GSOM2,
  FUN = function(x) c(mean = mean(x), sd = sd(x)))
TMIN.anomaly.decade = aggregate(TMIN.anom ~ Decade, GSOM2,
  FUN = function(x) c(mean = mean(x), sd = sd(x)))

```

```

Anom.x = seq(min(GSOM2$TMIN.anom), max(GSOM2$TMIN.anom), by=.1)
plot(Anom.x, dnorm(Anom.x, mean=TMIN.anomaly.decade$TMIN.anom[[1,1]],
  sd=TMIN.anomaly.decade$TMIN.anom[[1,2]]), ty="l", col=h.ramp[1], ylab="Density", xlab="TMIN")
mtext(paste0(fips$State, " ", decade), side=3)
for(i in 2:nrow(TMIN.anomaly.decade)){
  lines(Anom.x, dnorm(Anom.x, mean=TMIN.anomaly.decade$TMIN.anom[[i,1]], sd=TMIN.anomaly.decade$TMIN.anom[[i,2]]), col=h.ramp[i])
}
abline(v=mean(GSOM2$TMIN.anom[GSOM2$Decade==min(GSOM$Decade)]), col="blue")
abline(v=mean(GSOM2$TMIN.anom[GSOM2$Decade==max(GSOM$Decade)]), col="red")

Anom.x = seq(min(GSOM2$TMAX.anom), max(GSOM2$TMAX.anom), by=.1)
plot(Anom.x, dnorm(Anom.x, mean=TMAX.anomaly.decade$TMAX.anom[[1,1]],
  sd=TMAX.anomaly.decade$TMAX.anom[[1,2]]), ty="l", col=h.ramp[1], ylab="Density", xlab="TMAX")
mtext(paste0(fips$State, " ", decade), side=3)
for(i in 2:nrow(TMAX.anomaly.decade)){
  lines(Anom.x, dnorm(Anom.x, mean=TMAX.anomaly.decade$TMAX.anom[[i,1]], sd=TMAX.anomaly.decade$TMAX.anom[[i,2]]), col=h.ramp[i])
}
abline(v=mean(GSOM2$TMAX.anom[GSOM2$Decade==min(GSOM$Decade)]), col="blue")
abline(v=mean(GSOM2$TMAX.anom[GSOM2$Decade==max(GSOM$Decade)]), col="red")

# END -----
dev.off()

GSOM_animation <- image_animate(img, fps = 1, loop=2, optimize = TRUE)
#print(GSOM_animation)
setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/docs/")
image_write(GSOM_animation, paste("Climate_gifs/",
  fips$State, "-", stid, "_GSOM-normalPDF.gif", sep=""))
} else {
  print("Skipping animated normal distribution chunk")}

```

The file is saved in the main directory.

5.2 4 Weather Trend Plots

```

setwd("/home/CAMPUS/mw104747/github/Climate_Change_Narratives/")

gif_public = "docs/Social_Media/States/Climate_gifs/"
panel4.anim = paste0(fips$State, "-", stid, "-4panel.gif")

if(!file.exists(paste0(gif_path, panel4.anim))){
  print("Creating animated 4panel.gif")
}

```

```

# if (!file.exists(paste0("Climate_gifs/", fipsLState, "-", stid, "_GSOM-4plots.gif"))){

img <- image_graph(600, 480, res = 96)
# START ----
ylim_new=NA
for(i in seq(min(GSOM$Year), max(GSOM$Year), by=2))
{
  par(las=1, mfrow=c(4,1), mar= c(4, 4, 2, 1) + 0.1)
  # TMINmonthMax
  GSOMsub <- GSOM[GSOM$Month==TMINmonthMax & GSOM$Year<=i,]
  if(nrow(GSOMsub)<10) next
  plot(TMIN~Date, GSOMsub[GSOMsub$Month==TMINmonthMax,],
       col='gray70', pch=20, xlab="",
       main=paste("Mean", format(GSOMsub$Date,"%B")[1],
                  "Min. Temp", GSOM_Longest$name))
  GSOM.lm = lm(TMIN~Date, GSOMsub)
  pred_dates <-data.frame(Date = GSOMsub$Date);
  nrow(pred_dates); pred_dates
  #Predicts the values with confidence interval
  ci <- predict(GSOM.lm, newdata = pred_dates,
               interval = 'confidence')
  lines(pred_dates$Date, as.numeric(ci[,1]), col="darkred")
  lines(pred_dates$Date, as.numeric(ci[,2]), col="darkorange")
  lines(pred_dates$Date, ci[,3], col="darkorange")
  location_index = round(length(GSOMsub$Date) * 0.99,0)
  text(pred_dates$Date[location_index], ci[location_index,3],
       paste(report_prob2(GSOM.lm)), pos=2, cex=1.5)

  # Box Plot of TMAX by Month
  CHCNDsub = subset(CHCND, CHCND$Year<=i,
                   select=c(Month, Month.name, TMAX, TMIN))
  boxplot(TMAX ~ Month.name, data=CHCNDsub, main="")
  symbol.y = (par()$yaxp[2])-diff(par()$yaxp[1:2])*0.99
  #symbol.y = (par()$yaxp[2])
  text(sumstats$Month, symbol.y, sumstats$TMAX_Symbol,
       col="red", cex=2)
  mtext(paste("Maximum Daily Temperatures", min(CHCND$Year),
             "-", i, GSOM_Longest$name), line=1)
  mtext("(NOTE: Red asterisks correspond to significant changes)",
       line=0, cex=0.7)

  # TMAXmonthMax
  GSOMsub <- GSOM[GSOM$Month==TMAXmonthMax & GSOM$Year<=i,]
  ylim = range(GSOMsub$TMAX)

```

```

# if(!is.na(ylim_new)) ylim[2]=ylim_new
plot(TMAX~Date, GSOMsub, col='gray70', pch=20, xlab="",
      ylim=ylim,
      main=paste("Mean", format(GSOMsub$Date,"%B")[1],
                  "Max. Temp", GSOM_Longest$name))
GSOM.lm = lm(TMAX~Date, GSOMsub)

ci <- predict(GSOM.lm, newdata = pred_dates,
              interval = 'confidence')
lines(pred_dates$Date, as.numeric(ci[,1]), col="darkred")
lines(pred_dates$Date, as.numeric(ci[,2]), col="darkorange")
lines(pred_dates$Date, ci[,3], col="darkorange")

text(pred_dates$Date[location_index], ci[location_index,3],
      paste(report_prob2(GSOM.lm)), pos=2, cex=1.5)

# Record High Temperatures
# START
j = which(years %in% i)
if(sum(!is.na(TMAX.mat.noleap[,j]))==366){
TMAX1 = apply(TMAX.mat.noleap[,1:j], 1, function (x) which.max(x));
is.na(TMAX1) <- lengths(TMAX1) == 0
TMAX1 <- unlist(TMAX1)
TMAX1 <- count(TMAX1)
str(TMAX1)
names(TMAX1)=c("Year", "TMAX")
TMAX_na = data.frame(Year=1:j)
TMAX <- merge(TMAX_na, TMAX1, all.x=TRUE, by="Year")
}

if(sum(is.na(TMIN.mat[,j]))==366) next
# Select Minimum and Change to Negative Value
TMIN1 = apply(TMIN.mat[,1:j], 1, function (x) which.min(x));
is.na(TMIN1) <- lengths(TMIN1) == 0
TMIN1 <- unlist(TMIN1)
TMIN1 <- count(TMIN1) # Max Counts Negative
#str(TMIN1)
names(TMIN1)=c("Year", "TMIN")
TMIN_na <- data.frame(Year=1:j)
TMIN <- merge(TMIN_na, TMIN1, all.x=TRUE, by="Year")

R1 <- merge(TMAX, TMIN, by="Year")
R1$Index = rep(j, nrow(R1))
#results = rbind(results, R)
R1$TMIN = -R1$TMIN

```

```

## Sorting out X Axis
tic.no <- 4
rowskip = round(nrow(R1)/tic.no, 0)
row_numb <- seq_len(nrow(R1)) %% rowskip
row.sel = which(row_numb %in% c(1))
index.year <- years[row.sel]
# switch to decades?

xtics = row.sel
xlabs = index.year

yrange = range(c(R1$TMIN, R1$TMAX), na.rm=T)
ytics = floor(seq(yrange[1], yrange[2], length.out=tic.no))
ylabs = as.character(ytics)

par(las=1, xpd=TRUE)
plot(c(1,nrow(R1)), c(yrange[1], yrange[2]), ty="n", xaxt='n', yaxt='n', xlab="Year", ylab="")
axis(2,at=ytics, labels=ylabs)
axis(1,at=xtics, labels=xlabs)
barplot(height = R1$TMAX, space=0, add = TRUE, axes = FALSE, col="red")
barplot(height = R1$TMIN, space=0, add = TRUE, axes = FALSE, col="blue")
# END
}

# STOP ----
dev.off()

GSOM_animation <- image_animate(img, fps = 1, loop=2, optimize = TRUE)
image_write(GSOM_animation, paste0(gif_path, panel4.anim))

# fipsLState, "-", stid, "_GSOM-4plots.gif"))

} else {
  print("Skipping animated GSOM-4plots chunk")}

## Error in paste0(gif_path, panel4.anim): object 'gif_path' not found

```

The file is saved in the main directory.

6 OLD Stuff

7 Other attempts...


```

ncdc_locs(locationcategoryid='CITY', sortfield='name',
           sortorder='desc')

# ncdc_locs(locationcategoryid='CITY',
# locationid='FIPS:01', sortfield='name', sortorder='desc')

#ncdc_datasets(locationcategoryid='CITY',
# locationid='FIPS:01', sortfield='name', sortorder='desc')

out <- ncdc(datasetid='NORMAL_DLY', stationid='GHCND:USW00014895',
            datatypeid='dly-tmax-normal', startdate = '2010-05-01',
            enddate = '2010-05-10')

```

```

with_units <- ncdc(datasetid='GHCND', stationid='GHCND:USW00014895',
                  datatypeid='TMAX', startdate = '2010-05-01',
                  enddate = '2010-10-31', limit=500, add_units = TRUE)
head( with_units$data )

## # A tibble: 6 x 9
##   date          datatype station    value fl_m fl_q fl_so fl_t units
##   <chr>         <chr>    <chr>    <int> <chr> <chr> <chr> <chr> <chr>
## 1 2010-05-01T00:00:00 TMAX    GHCND:USW0~  222 ""   ""    0    2400 celciu~
## 2 2010-05-02T00:00:00 TMAX    GHCND:USW0~  222 ""   ""    0    2400 celciu~
## 3 2010-05-03T00:00:00 TMAX    GHCND:USW0~  233 ""   ""    0    2400 celciu~
## 4 2010-05-04T00:00:00 TMAX    GHCND:USW0~  222 ""   ""    0    2400 celciu~
## 5 2010-05-05T00:00:00 TMAX    GHCND:USW0~  272 ""   ""    0    2400 celciu~
## 6 2010-05-06T00:00:00 TMAX    GHCND:USW0~  194 ""   ""    0    2400 celciu~

```

7.1 Evaluating Records

TBD

7.2 Export Options

TBD

8 Sea Surface Temperature Data – SURP PROJECT WAITING TO HAPPEN

In contrast to terrestrial data, sea surface temperature (SST) is quite difficult to obtain and process. There are numerous tools to access the data, but they

often require knowledge of complex software tools that are not easy to set up or programming experience with python or others.

<https://climexp.knmi.nl/select.cgi?id=someone@somewhere&field=ersstv5>

There are, however, a few tools build for R users that seem to accomplish all that we need.

https://rda.ucar.edu/index.html?hash=data_user&action=register

<https://rda.ucar.edu/datasets/ds277.9/>

Alternatively, we can download flat ascII tables of gridded data:

<https://www1.ncdc.noaa.gov/pub/data/cmb/ersst/v5/ascii/>

9 Satellite Data

TBD

10 Ice-Core Data

TBD

11 Conclusions

Developing a robust method to analyze weather stations is both time consuming and difficult to justify the outcome. In part because the data suggest that each station (region) requires different types of analysis, based on the expected patterns of temperature and rainfall. As climate scientists have known for decades, the terminology of global warming is not very useful. Not because scientists are trying to hide something or promote some biased agenda, but that even as warming of the global average is well documented, the impacts of climate change on each region is highly specific, requiring specificity in the analysis.

Hopefully, this little analysis has created some mechanism for others to appreciate this complexity.

```
## Error in data.frame(State = fips$State, Station = stid, gif_public  
= gif_public, : object 'png_public' not found  
## Error in is.data.frame(x): object 'dbase' not found
```

The document took 4.4 minutes to process and compile. My next goal will be to optimize the process and streamline the time to analyze.