

How are temperatures increasing and what are the implications?

Marc Los Huertos

1 Introduction

According to the IPCC, the temperature has been changing about 0.X degrees per XX years – but how do these values "map" onto our communities? Can we find out how these changes will affect specific people in communities we care about?

1.1 Learning Goals

For this project, you will evaluate determine if the Earth's temperature has in fact changed, and if so, by how much?

1.2 Driving Question

Is my region's climate changing?
How is climate change affecting my community?

1.3 Public Product

Narrative Blog...
with professional graphics and statistics.

1.4 Approach

2 Procedures

2.1 How is temperature data collected?

2.2 How are the data store, curated and checked for quality?

3 Data Source

3.1 Compressed Files

```
# Uncompress the files.
# ghcnd_all
source("summarySE.R")

tarfile = "C:\\workspace\\GitHub\\RTricks\\300_Global_Warming\\Raw Data\\ghcnd_all.tar.gz"

#ftpsource = ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/v3/ghcnm.tmax.latest.qca.tar.gz

#ghcnm.tmax.latests.qca.tar.qz
tarfile = "C:\\workspace\\GitHub\\RTricks\\300_Global_Warming\\Raw Data\\ghcnm.tmax.latest.q
# untar(tarfile)
```

```
stationfile = "/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/Data/ghcnd-stations.t
```

3.2 Obtain Locations

```
# read.table(stationfile, header=F, fill=T, row.names=NULL); head(stations)
stations = (read.fwf(stationfile, fill=T, widths= c(11, 9, 10, 7, 3, 32, 3, 4, 9), ))
names(stations)= c("ID", "LAT", "LONG", "ELEV", "STATE", "NAME", "GSN", "HCN_CRN", "WHOID")

head(stations)

##           ID      LAT      LONG  ELEV STATE
## 1 ACW00011604 17.1167 -61.7833  10.1
## 2 ACW00011647 17.1333 -61.7833  19.2
## 3 AEO00041196 25.3330  55.5170  34.0
## 4 AEM00041194 25.2550  55.3640  10.4
## 5 AEM00041217 24.4330  54.6510  26.8
## 6 AEM00041218 24.2620  55.6090 264.9
##                                NAME GSN HCN_CRN WHOID
## 1 ST JOHNS COOLIDGE FLD                      NA
## 2 ST JOHNS                      NA
## 3 SHARJAH INTER. AIRP                      GSN    41196
## 4 DUBAI INTL                      41194
## 5 ABU DHABI INTL                      41217
## 6 AL AIN INTL                      41218

str(stations)

## 'data.frame': 100747 obs. of  9 variables:
## $ ID      : Factor w/ 100747 levels "ACW00011604",...: 1 2 3 4 5 6 7 8 9 10 ...
## $ LAT     : num  17.1 17.1 25.3 25.3 24.4 ...
```

```
## $ LONG : num -61.8 -61.8 55.5 55.4 54.7 ...
## $ ELEV : num 10.1 19.2 34 10.4 26.8 ...
## $ STATE : Factor w/ 76 levels " ", " AB", " AK",...: 1 1 1 1 1 1 1 1 1 1 ...
## $ NAME : Factor w/ 93968 levels " 'S HEERENHOEK ",...: 79235 79234 762
## $ GSN : Factor w/ 3 levels "", " ", "GSN": 2 2 3 2 2 2 3 2 2 2 ...
## $ HCN_CRN: Factor w/ 4 levels "", " ", " CRN",...: 2 2 2 2 2 2 2 2 2 2 ...
## $ WHOID : num NA NA 41196 41194 41217 ...
```

Example of data:

AG000060680 22.8000 5.4331 1362.0 TAMANRASSET GSN 60680

subsectionSelecting and Example Location

Here's what the data look like:

ID 1-11 Character YEAR 12-15 Integer MONTH 16-17 Integer ELEMENT
 18-21 Character VALUE1 22-26 Integer MFLAG1 27-27 Character QFLAG1 28-
 28 Character SFLAG1 29-29 Character VALUE2 30-34 Integer MFLAG2 35-35
 Character QFLAG2 36-36 Character SFLAG2 37-37 Character
 . VALUE31 262-266 Integer MFLAG31 267-267 Character QFLAG31 268-268
 Character SFLAG31 269-269 Character

Arizona, let's check import process for the sites...

```
stations[stations$ID=="US1AZMR0019",]

##           ID      LAT      LONG  ELEV STATE
## 48124 US1AZMR0019 33.5902 -111.9712 418.5   AZ
##
##           NAME GSN HCN_CRN WHOID
## 48124 SCOTTSDALE 8.8 SW              NA

# head(stations[stations$HCN_CRN==" CRN",])
```

Let's get the arizona data into R

```
# Read the file
dlyfile = "/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/AGM00060515"
test = read.fwf(dlyfile,widths = c(11, 4, 2, 4, rep(c(5, 1, 1, 1),31)))

## Warning in file(file, "rt"): cannot open file '/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/AGM00060515'
## No such file or directory
## Error in file(file, "rt"): cannot open the connection

str(test)

## Error in str(test): object 'test' not found
```

```
# practicing loops
for (year in c(2010,2011,2012,2013,2014,2015)){
  print(paste("The year is", year))
}
```

```

}

## [1] "The year is 2010"
## [1] "The year is 2011"
## [1] "The year is 2012"
## [1] "The year is 2013"
## [1] "The year is 2014"
## [1] "The year is 2015"

# Create New Variable Names
MFLAG=NA; QFLAG=NA; SFLAG=NA; VALUE=NA
for (i in 1:31){
  VALUE[i] = paste("DATE", i, sep="")
  MFLAG[i] = paste("MFLAG", i, sep="")
  QFLAG[i] = paste("QFLAG", i, sep="")
  SFLAG[i] = paste("SFLAG", i, sep="")
}

#print(QFLAG)

# Vector of variable names converted from a transposed matrix
tmp = as.vector(t(matrix(data=c(VALUE, MFLAG, QFLAG, SFLAG), ncol=4)))
Names = c("ID", "YEAR", "MONTH", "ELEMENT", tmp); length(Names)

## [1] 128

names(test)= Names; #test

## Error in names(test) = Names: object 'test' not found

head(test)

## Error in head(test): object 'test' not found

```

3.3 Process Selected Data Files

```

setwd("/home/CAMPUS/mwl04747/github/Climate_Change_Narratives/Data")

dly_list = list.files(pattern="*.dly"); head(dly_list)

## [1] "AGM00060515.dly" "US1AZCN0021.dly"

#for (i in 1:length(dly_list))

```

```

for (i in 1:1){
tmp <- read.fwf(dly_list[i], widths = c(11, 4, 2, 4, rep(c(5, 1, 1, 1),31)))
names(tmp) <- Names
assign(dly_list[i], subset(tmp, ELEMENT=="TMAX", select=c(1:4, seq(5, by = 4, length.out=31),
}))

library(tidyr)
library(dplyr)

##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(stringr)
#str(AGM00060515.dly)
#gather(US1AZCN0021.dly, "Temp", VALUE1)

library(reshape2)
tmp1 = melt(AGM00060515.dly, id=c("ID", "YEAR", "MONTH", "ELEMENT"))
head(tmp1)

##           ID YEAR MONTH ELEMENT variable value
## 1 AGM00060515 1984     3    TMAX   DATE1 -9999
## 2 AGM00060515 1984     4    TMAX   DATE1   190
## 3 AGM00060515 1984     5    TMAX   DATE1 -9999
## 4 AGM00060515 1984     6    TMAX   DATE1 -9999
## 5 AGM00060515 1984     7    TMAX   DATE1   430
## 6 AGM00060515 1984     8    TMAX   DATE1 -9999

tmp1$Day = as.numeric(str_sub(tmp1$variable,6,7)); head(tmp1)

##           ID YEAR MONTH ELEMENT variable value Day
## 1 AGM00060515 1984     3    TMAX   DATE1 -9999 NA
## 2 AGM00060515 1984     4    TMAX   DATE1   190 NA
## 3 AGM00060515 1984     5    TMAX   DATE1 -9999 NA
## 4 AGM00060515 1984     6    TMAX   DATE1 -9999 NA
## 5 AGM00060515 1984     7    TMAX   DATE1   430 NA
## 6 AGM00060515 1984     8    TMAX   DATE1 -9999 NA

tmp1$value[tmp1$value==--9999] = NA; head(tmp1)

```

```
##           ID YEAR MONTH ELEMENT variable value Day
## 1 AGM00060515 1984     3    TMAX   DATE1     NA  NA
## 2 AGM00060515 1984     4    TMAX   DATE1    190  NA
## 3 AGM00060515 1984     5    TMAX   DATE1     NA  NA
## 4 AGM00060515 1984     6    TMAX   DATE1     NA  NA
## 5 AGM00060515 1984     7    TMAX   DATE1    430  NA
## 6 AGM00060515 1984     8    TMAX   DATE1     NA  NA

tmp1$Temperature = tmp1$value/10

drops <- c("variable", "value")
tmp1 <- tmp1[, !(names(tmp1) %in% drops)]
tmp1$DECADE = round(tmp1$YEAR, -1)
# names(tmp1)
```

4 Presenting the Results

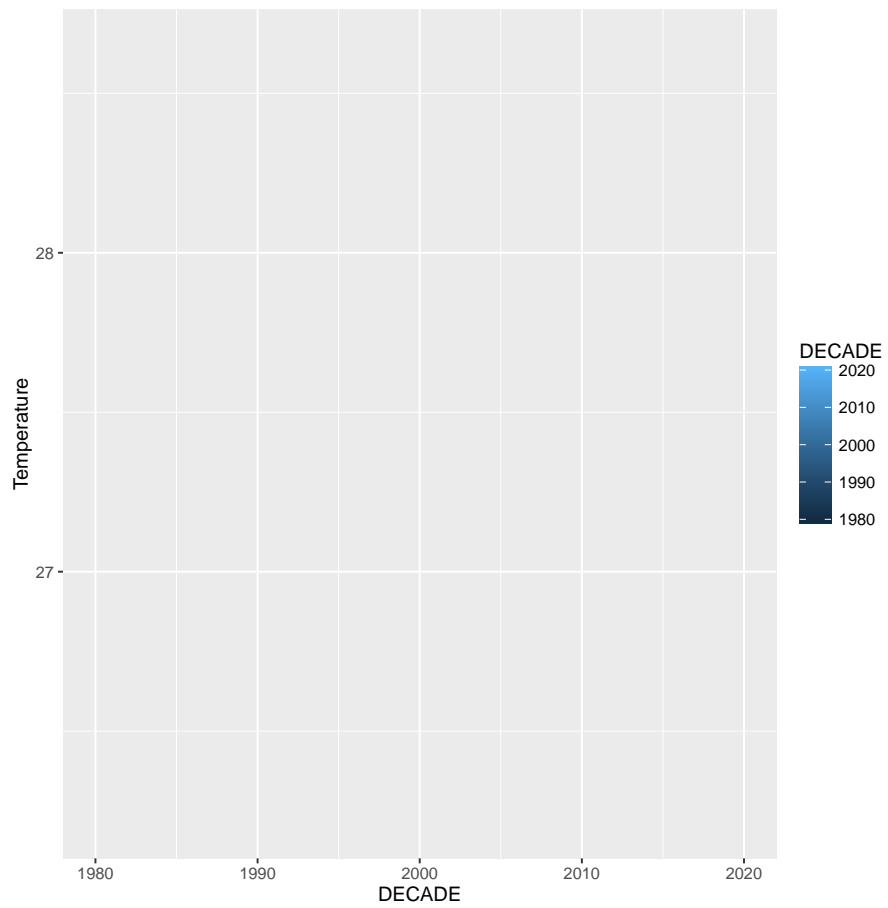
```
# call summarySE function....somehow...

library(ggplot2)

summarydf <- summarySE(tmp1, "Temperature", "DECADE", na.rm=T)

## -----
## You have loaded plyr after dplyr - this is likely to cause problems.
## If you need functions from both plyr and dplyr, please load plyr
## first, then dplyr:
## library(plyr); library(dplyr)
## -----
##
## Attaching package: 'plyr'
##
## The following objects are masked from 'package:dplyr':
##
##   arrange, count, desc, failwith, id, mutate, rename, summarise,
##   summarize

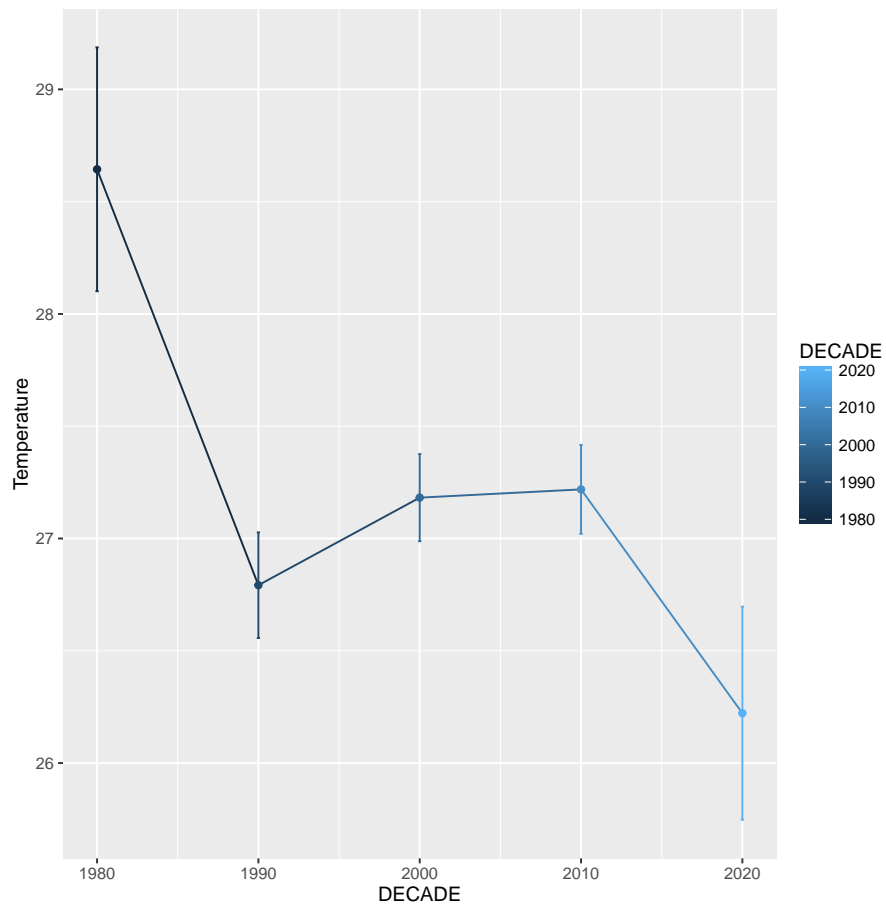
ggplot(summarydf, aes(y=Temperature, x=DECADE, color= DECADE))
```



```
+ geom_point() + geom_errorbar(limits, width=0.2)

## Error in +geom_point(): invalid argument to unary operator

ggplot(summarydf, aes(y=Temperature, x=DECADE, color= DECADE)) + geom_errorbar(aes(ymin=Temp
```



4.1 NOAA dataset

New NOAA Directory – <ftp://ftp.ncdc.noaa.gov/pub/data/noaa/>

```
library(raster)

## Loading required package: sp
##
## Attaching package: 'raster'
##
## The following object is masked from 'package:dplyr':
##
##   select
##
## The following object is masked from 'package:tidyr':
```



```
##
##      extract

library(XML)

coords.fwt <- read.fwf("ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-history.txt",widths=c(6,1,
Names = c("USAF", "X1", "WBAN", "X2", "STATION_NAME", "X3", "CTRY", "X4", "ST", "X5", "CALL
Widths = c(6,      1,      5,      1,      29,      1,      2,      3,      2,      1,      4,

coords.fwt <- read.fwf("ftp://ftp.ncdc.noaa.gov/pub/data/noaa/isd-history.txt",widths=Widths

X4 ST X5

IA

ND
04
27
29

coords <- data.frame(ID=paste(as.factor(coords.fwt[,1])),WBAN=paste(as.factor(coords.fwt[,3]

## Warning in data.frame(ID = paste(as.factor(coords.fwt[, 1])), WBAN
= paste(as.factor(coords.fwt[, : NAs introduced by coercion
## Warning in data.frame(ID = paste(as.factor(coords.fwt[, 1])), WBAN
= paste(as.factor(coords.fwt[, : NAs introduced by coercion
```

NOAA Locations

```
plot(Lat ~ Lon, data=coords, xlim=c(-180, 180) )
```

