Activity 3

Aidan Harrington

9/23/2020

Question 1) At high humidity and low temperature, high humidity and high temperature, low humidity and low temperature and low humididty and high temperature, there is more uncertainty in the measurements. The uncertainty in the instrument is minimized between 10 and 70 degrees C and between 15-90% relative humidity. Values outside of these ranges result in more uncertainty.

Question 2) The ATMOS 41 will reject a wind speed measurement that is 8x greater than the running average over the last ten measurements, taken every ten seconds. This prevents spikes in wind speed from random events being recorded (the provided example is a bumblebee hanging out in the sensor). However, in cases where unpredictable large increases in wind speed might be common or possible, this may result in the sensor falsely marking that as an erroneous measurement.

```
# Preparing the data
datW<-read.csv("data/bewkes_weather.csv",na.strings = c("#N/A"),skip=3,header=F)
sensorInfo<-read.csv("data/bewkes_weather.csv",na.strings = c("#N/A"),nrows=2)
colnames(datW)<-colnames(sensorInfo)
```

Question 3) Skip tells R to ignore the specified number of rows. Skip=3 means that the first three rows will not be included in the dataframe. Nrows tells R to only include the specified number of rows, nrows=2 includes only the first two rows of the csv in the dataframe. Header=false tells R not to use the firs row of the csv files as the column names.

```
# Loading in the package
library(lubridate)

# Standardizing date formats
dates<-mdy_hm(datW$timestamp,tz="America/New_York")
datW$doy<-yday(dates)
datW$hour<-hour(dates)+(minute(dates)/60)
datW$DD<-datW$doy+(datW$hour/24)

# Check for missing values with a function
na.chek<- function(x) is.na(x) %>% which() %>% length()
na.chek(datW$air.temperature)
```

```
## [1] 0
```

```
na.chek(datW$wind.speed)
```

```
## [1] 0
```

```
na.chek(datW$precipitation)
```

```
## [1] 0
```

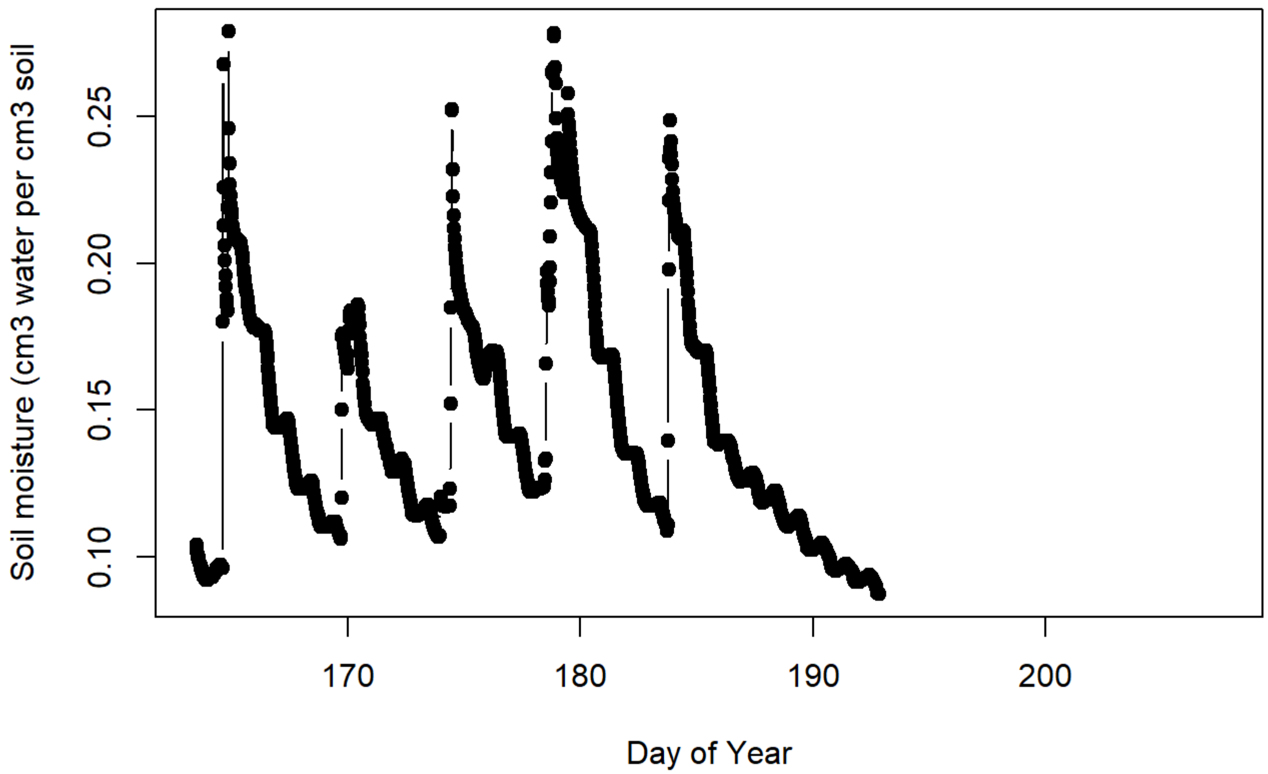
```
na.chek(datW$soil.temp)
```

```
## [1] 707
```

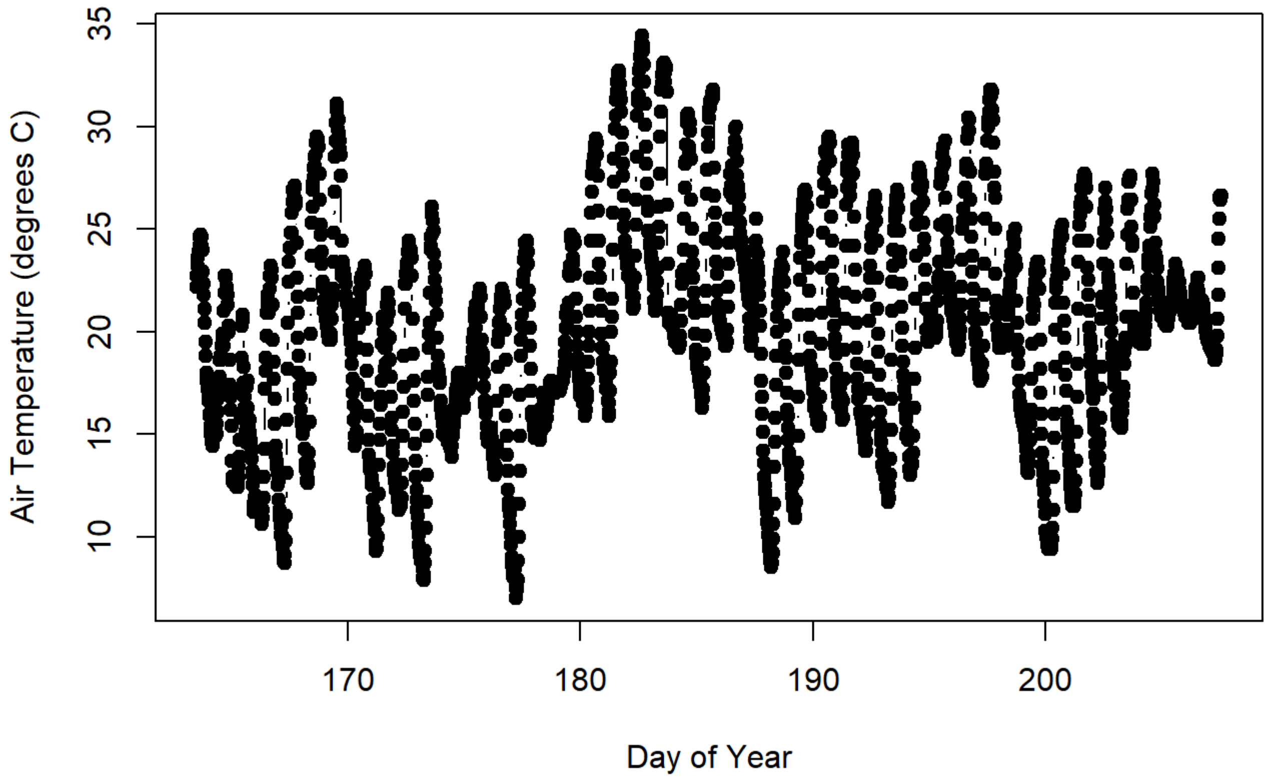
```
na.chek(datW$soil.moisture)
```

```
## [1] 707
```

```
# Plotting data for QA/QC
plot(datW$DD,datW$soil.moisture,pch=19,type="b",xlab="Day of Year",ylab="Soil moisture (cm3 water per cm3 soil")
```




```
plot(datW$DD,datW$air.temperature,pch=19,type="b",xlab="Day of Year",ylab="Air Temperature (degrees C)")
```




```
#Realistic Values
datW$air.tempQ1 <- ifelse(datW$air.temperature < 0, NA, datW$air.temperature)
quantile(datW$air.tempQ1)
```

```
##   0%  25%  50%  75% 100%
##  7.0 16.2 20.1 23.3 34.4
```

```
#Highs and lows in temperature
datW[datW$air.tempQ1<8,]
```

```
##          timestamp solar.radiation precipitation lightning.acvitivy
## 470 6/22/2018 6:00              12             0                  0
## 660 6/26/2018 5:00               0             0                  0
## 661 6/26/2018 5:30               1             0                  0
## 662 6/26/2018 6:00               7             0                  0
## 663 6/26/2018 6:30              17             0                  0
## 664 6/26/2018 7:00              26             0                  0
##     lightning.distance wind.dir wind.speed gust.speed air.temperature
## 470                  0      329       0.31       0.85             7.9
## 660                  0      318       0.35       0.61             7.5
## 661                  0      318       0.38       0.67             7.0
## 662                  0      317       0.39       0.61             7.4
## 663                  0      326       0.31       0.64             7.5
## 664                  0      336       0.27       0.46             7.9
##     relative.humid atmospheric.pressure soil.moisture soil.temp doy hour
## 470           0.85                96.00     0.1160974      9.61 173  6.0
## 660           0.85                96.93     0.1409786      9.50 177  5.0
## 661           0.88                96.94     0.1409786      9.37 177  5.5
## 662           0.86                96.95     0.1409786      9.33 177  6.0
## 663           0.86                96.96     0.1411936      9.45 177  6.5
## 664           0.88                96.96     0.1411936      9.67 177  7.0
##           DD air.tempQ1
## 470 173.2500        7.9
## 660 177.2083        7.5
## 661 177.2292        7.0
## 662 177.2500        7.4
## 663 177.2708        7.5
## 664 177.2917        7.9
```

```
datW[datW$air.tempQ1>33,]
```

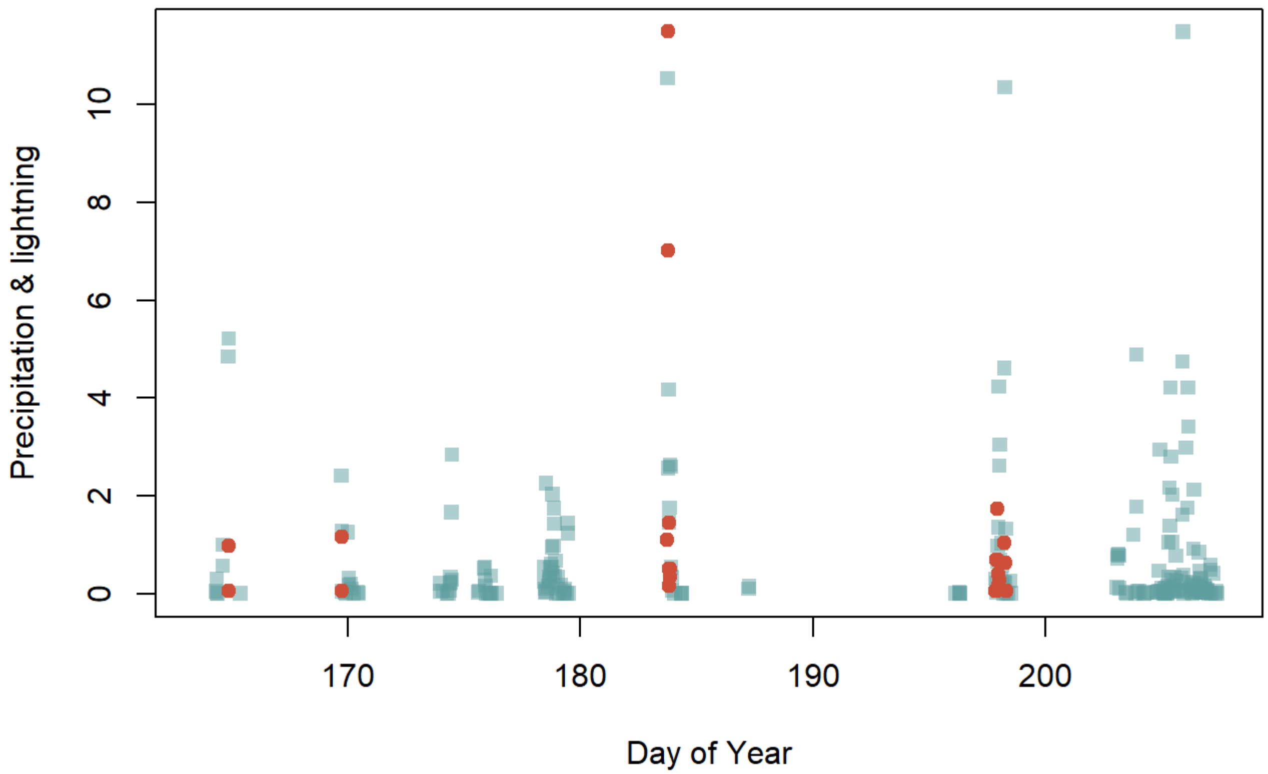
```
##          timestamp solar.radiation precipitation lightning.acvitivy
## 918 7/1/2018 14:00             915             0                  0
## 919 7/1/2018 14:30             896             0                  0
## 920 7/1/2018 15:00             870             0                  0
## 921 7/1/2018 15:30             826             0                  0
## 922 7/1/2018 16:00             761             0                  0
## 923 7/1/2018 16:30             691             0                  0
## 924 7/1/2018 17:00             611             0                  0
## 966 7/2/2018 14:00             804             0                  0
##     lightning.distance wind.dir wind.speed gust.speed air.temperature
## 918                  0       57       0.56       1.20            33.4
## 919                  0       70       0.67       1.83            33.6
## 920                  0       74       0.61       1.88            34.0
## 921                  0       77       0.64       2.33            34.1
## 922                  0       66       0.56       1.52            34.4
## 923                  0       71       0.56       1.33            34.0
## 924                  0       51       0.55       1.96            33.7
## 966                  0       56       0.86       3.33            33.1
##     relative.humid atmospheric.pressure soil.moisture soil.temp doy hour
## 918           0.52                96.29     0.1276925     24.55 182 14.0
## 919           0.51                96.26     0.1265884     24.72 182 14.5
## 920           0.47                96.26     0.1254821     24.81 182 15.0
## 921           0.45                96.23     0.1239293     24.77 182 15.5
## 922           0.43                96.22     0.1230399     24.85 182 16.0
## 923           0.44                96.21     0.1221491     24.95 182 16.5
## 924           0.46                96.19     0.1212568     24.94 182 17.0
## 966           0.50                96.30     0.1133860     25.22 183 14.0
##           DD air.tempQ1
## 918 182.5833       33.4
## 919 182.6042       33.6
## 920 182.6250       34.0
## 921 182.6458       34.1
## 922 182.6667       34.4
## 923 182.6875       34.0
## 924 182.7083       33.7
## 966 183.5833       33.1
```

Question 4) The few random readings from the high and low values that I chose to cross-check with the nearest weather station were reasonably close. Considering that the nearest station is in Rome, NY, any minor differences between the Bewkes and there are hardly grounds to question the instrument.

```r
#plot precipitation and lightning strikes on the same plot
#normalize lighting strikes to match precipitation
lightscale <- (max(datW$precipitation)/max(datW$lightning.acvitivy)) * datW$lightning.acvitivy

#make the plot with precipitation and lightning activity marked
#make it empty to start and add in features
plot(datW$DD , datW$precipitation, xlab = "Day of Year", ylab = "Precipitation & lightning",
        type="n")
#plot precipitation points only when there is precipitation
#make the points semi-transparent
points(datW$DD[datW$precipitation > 0], datW$precipitation[datW$precipitation > 0],
        col= rgb(95/255,158/255,160/255,.5), pch=15)

#plot lightning points only when there is lightning
points(datW$DD[lightscale > 0], lightscale[lightscale > 0],
        col= "tomato3", pch=19)
```



Question 5) You can use “lightscale” to subset in “datW” because lightscale is a vector of the same length as datW and each row corresponds to the measurements it was calculated from. The code below shows that this is the case.

```
#The function here returns whatever the second statement is if the first is false.
#Ive modififed the function to return true statements
#In this case, the lengths of the two vectors are the same, so the function returns the assigned message.
assert.true<-function(statement,message){if(statement == TRUE){print(message)}}
assert.true(length(lightscale)==length(datW),"Equal length")
```

```r
#filter out storms in wind and air temperature measurements (using lightning and rain as proxy)
#filter out values with lightning that coincides with rainfall greater than 2mm or only rainfall over 5 mm.
#create a new air temp column
datW$air.tempQ2 <- ifelse(datW$precipitation  >= 2 & datW$lightning.acvitivy >0, NA,
                    ifelse(datW$precipitation > 5, NA, datW$air.tempQ1))
```

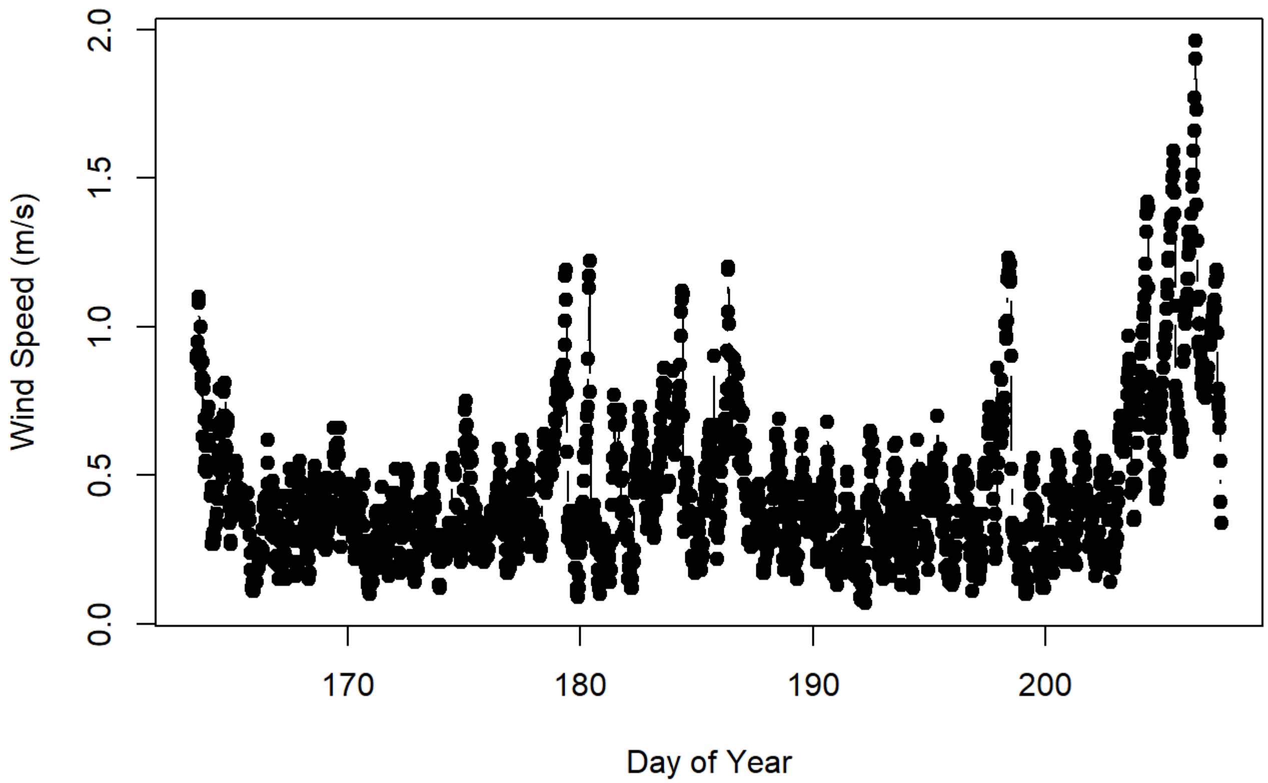
Question 6) The code below excludes the same thirteen rows of data as the filter done above on air temperature. The result is that we now have two new columns for air temperature and wind speed that have put NA values in place of the measurements that occurred on very rainy and stormy days. The assert.true function below returns “Same measurements removed” if the measurements selected for removal are actually the same, which was done by matching their timestamps.

```r
#same criteria as before but filtering out the wind values now
datW$wind.speedQ2 <- ifelse(datW$precipitation  >= 2 & datW$lightning.acvitivy >0, NA,
                    ifelse(datW$precipitation > 5, NA, datW$wind.speed))

#To check if this did what I wanted, I should see if the day and time of removed values is the same
#in both variables. I select the timestamps from the new air and wind columns where they have NA's and ask if they are
#identicial. As the assigned statement is returned, we know that the timestamps of the removed values were identical
assert.true(identical(select(filter(datW,is.na(datW$air.tempQ2==TRUE)),timestamp),select(filter(datW,is.na(datW$wind.speedQ2
==TRUE)),timestamp)),"Same Measurements Removed")
```

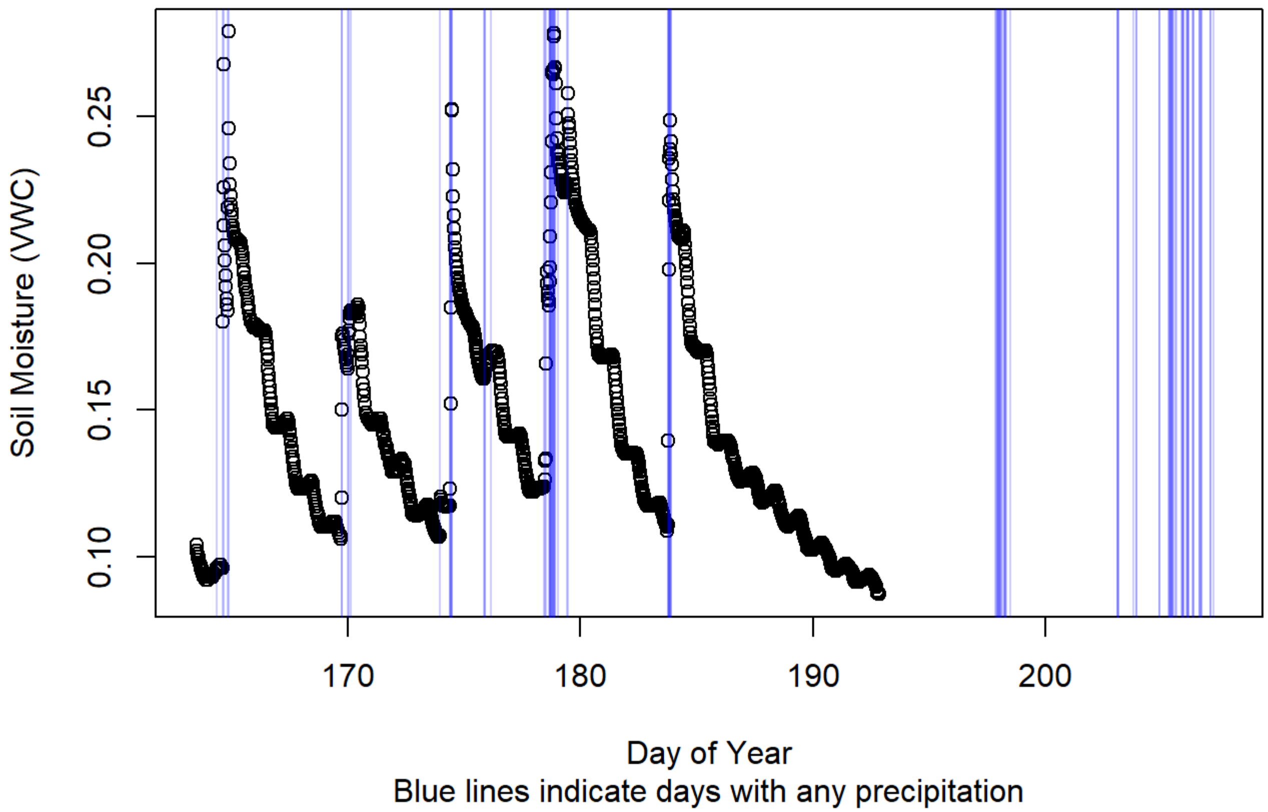
```
## [1] "Same Measurements Removed"
```

```r
#New plot of wind speed without the stormy days included
plot(datW$DD, datW$wind.speedQ2, pch=19, type="b", xlab = "Day of Year",
     ylab="Wind Speed (m/s)")
```



Question 7) The plot coded below shows that there is not an obvious reason for the sensor malfunction. We see that the sensor consistently responds to rainfall (the blue lines) with significant increases in soil moisture. In the days leading up to the malfunction there had been virtually no rain of appreciable amounts (greater than 0.2 mm), while all other days that saw an uptick in soil moisture were accompanied by rain above that cutoff. Furthermore, readings as low as we see around 190 days are probably not spurious considering they are in a similar range to values measured before the first recorded rain event on the plot. The resolution of the meter is 0.0008 VWC, which is far lower than where the measurements stopped. Storm damage also is not likely considering that notable winds occurred on days 183 and 198, which are not the day in question. Some other reason is responsible for sensor malfunction, but it appears that the measurements were accurate up until that point.

```
plot(datW$DD,datW$soil.moisture,xlab = "Day of Year",ylab="Soil Moisture (VWC)",sub = "Blue lines indicate days with any pre
cipitation")
abline(v=c(datW$DD[datW$precipitation > 0.2]),lwd=1,col=rgb(0, 0, 255, max = 255, alpha = 50))
points(datW$air.temperature[datW$air.temperature>25],datW$air.temperature[datW$air.temperature>25],type = "p",col="red")
```



Question 8) All code for the table shown below. Vectors for each part of the table are made individually and joined together with dataframe. The measurements are followed by the actual value and their uncertainty. The number of decimals may exceed the signficant figures despite my best efforts, R wants to add zeroes everywhere. Unertainties are displayed in the next column, followed by the number of samples and the start and end date of measurements. Sample sizes for the first two measures exclude the erroneous measurements made on stormy days, and the missing data from the soil probe.

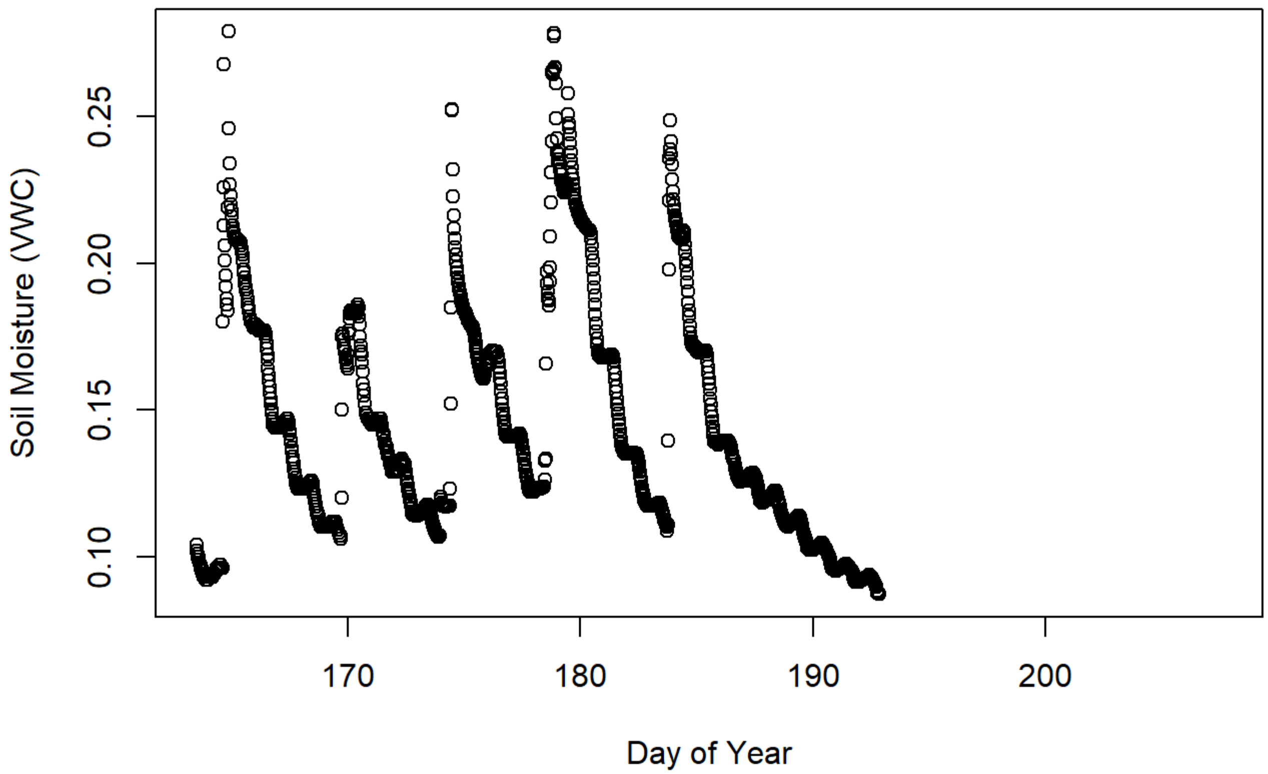
```
rain.error<-(datW$precipitation*0.05) #calculating error for rain
Measurement<-c("Mean Air Temperature (C)","Mean Wind Speed (m/s)","Mean Soil Moisture (VWC)","Mean Soil Temperature (C)",
               "Total Precipitation (mm)")
Value<-c(signif(mean(datW$air.tempQ2,na.rm=T),3),signif(mean(datW$wind.speedQ2,na.rm = T),2),
                     signif(mean(datW$soil.moisture,na.rm=T),2),signif(mean(datW$soil.temp,na.rm=T),2),sum(datW$precipitatio
n))
Uncertainty<-c(0.6,0.3,0.03,1,signif(sum(rain.error),4))
Sample_Size<-c(length(datW$air.tempQ2)-na.chek(datW$air.tempQ2),length(datW$wind.speedQ2)-na.chek(datW$wind.speedQ2),
                  length(datW$soil.moisture)-na.chek(datW$soil.moisture),length(datW$soil.temp)-na.chek(datW$soil.temp),
                  length(datW$precipitation)-na.chek(datW$precipitation))
Measurement_Start<-c("6/12/2018 11:30","6/12/2018 11:30","6/12/2018 11:30","6/12/2018 11:30","6/12/2018 11:30")
Measurement_End<-c("7/26/2018 14:00","7/26/2018 14:00","7/11/2018 21:00","7/11/2018 21:00","7/26/2018 14:00")
Summary_Table<-data.frame(Measurement,Value,Uncertainty,Sample_Size,Measurement_Start,Measurement_End)

print(Summary_Table)
```

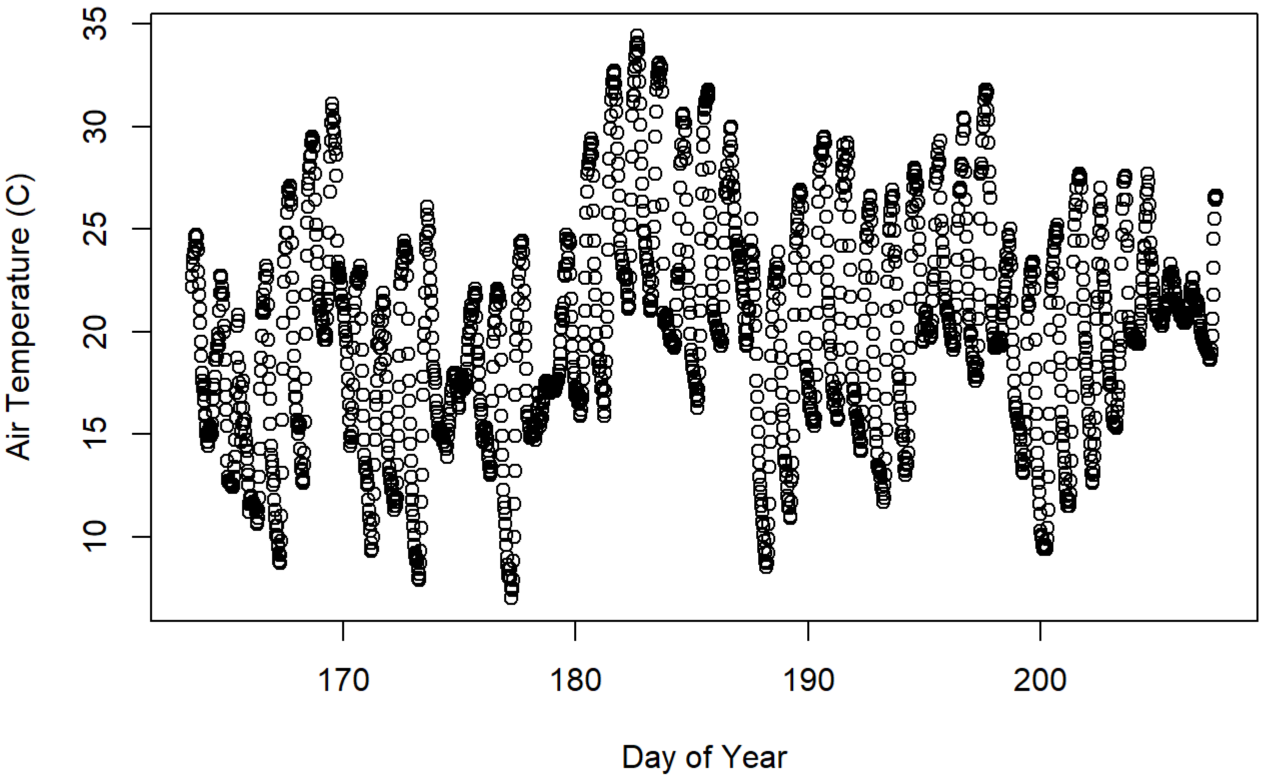
```
##                 Measurement   Value Uncertainty Sample_Size Measurement_Start
## 1  Mean Air Temperature (C)  20.000       0.600        2105   6/12/2018 11:30
## 2     Mean Wind Speed (m/s)   0.450       0.300        2105   6/12/2018 11:30
## 3  Mean Soil Moisture (VWC)   0.140       0.030        1411   6/12/2018 11:30
## 4 Mean Soil Temperature (C)  17.000       1.000        1411   6/12/2018 11:30
## 5  Total Precipitation (mm) 177.828       8.891        2118   6/12/2018 11:30
##   Measurement_End
## 1 7/26/2018 14:00
## 2 7/26/2018 14:00
## 3 7/11/2018 21:00
## 4 7/11/2018 21:00
## 5 7/26/2018 14:00
```

Question 9) The plots show a number of trends in these four vairables. Soil and air temperature are very closely correlated in the same direction. The air is warmer, but the soil follows it closely. Spikes in soil moisture occur whenever there is precipitation, for obvious reasons, but dries out quickly. Large dips in soil temperature are also associated with instances of rainfall. Other interesting trends during the study period include the relative consistency of rain every few days, predictable daily fluctuations in soil and air temperatures, and the speed at which soil appears to dry out.

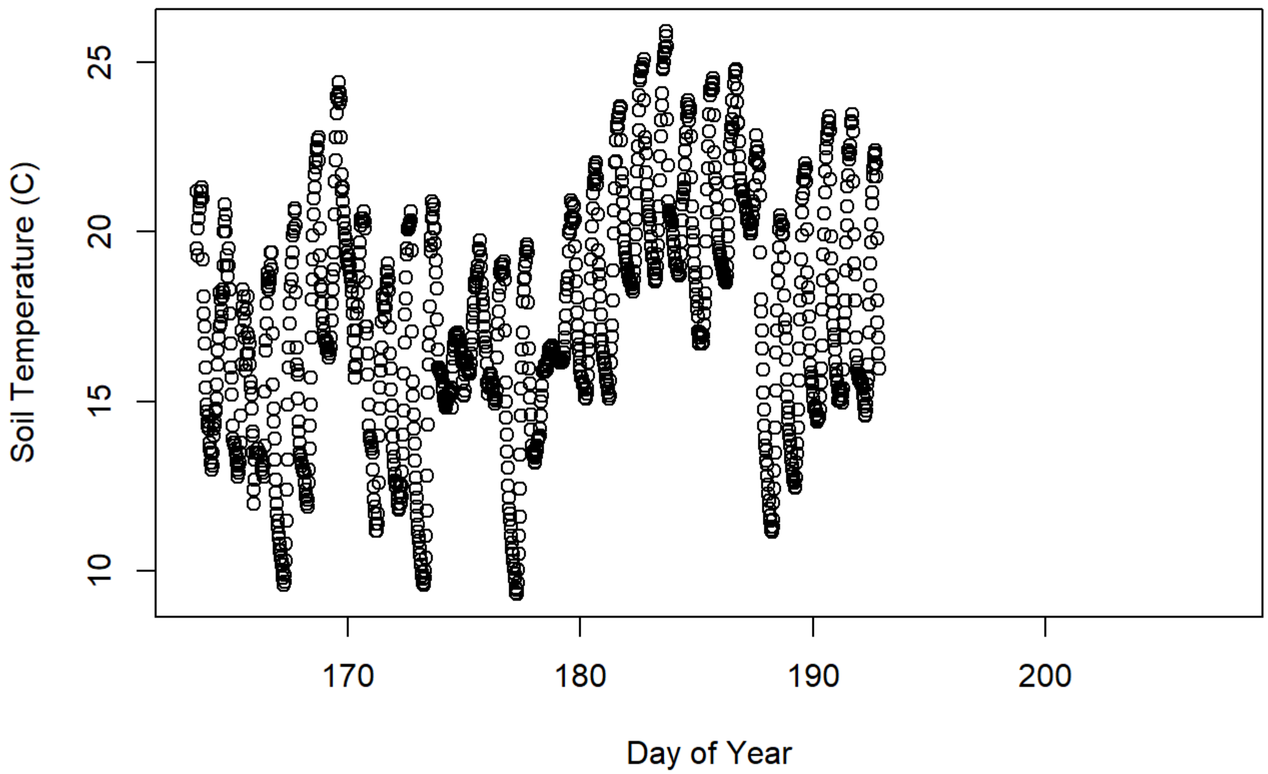
```
plot(datW$DD,datW$soil.moisture,xlab="Day of Year",ylab="Soil Moisture (VWC)")
```



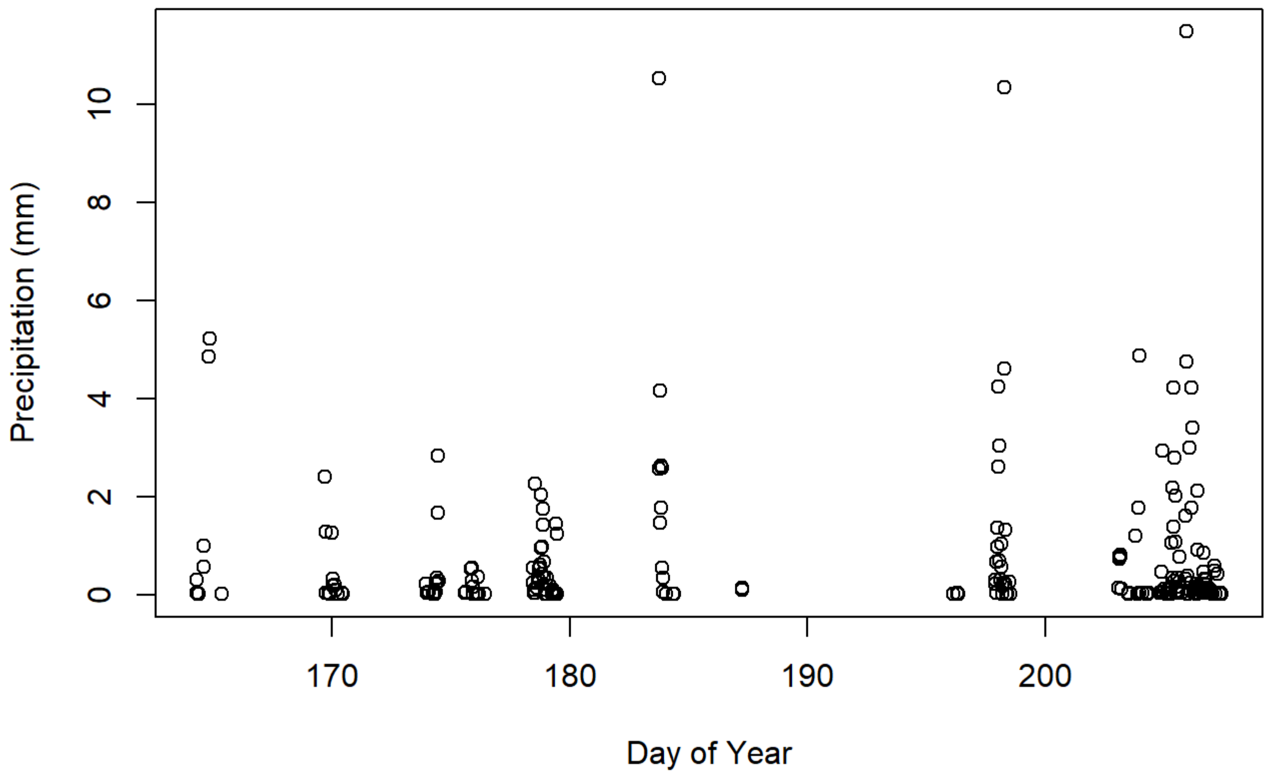

```
plot(datW$DD,datW$air.temperature,xlab="Day of Year",ylab="Air Temperature (C)")
```




```
plot(datW$DD,datW$soil.temp,xlab="Day of Year",ylab="Soil Temperature (C)")
```




```
plot(datW$DD[datW$precipitation>0],datW$precipitation[datW$precipitation>0],xlab="Day of Year",ylab="Precipitation (mm)")
```



Question 10) Link to github repository https://github.com/aharrington99/GEOG331/blob/master/Activity%203.rmd (https://github.com/aharrington99/GEOG331/blob/master/Activity%203.rmd)