INF 550 Section 3.7

Natasha Wesely

2022-09-28

3.7 USA-NPN Coding Lab

#1

For the purposes of this exercise we will be focusing on two NEON sites: HARV and CPER. Save these two sites into your workplace so that you can feed them into functions and packages.

```
sitesOfInterest <- c("HARV", "CPER")
```

#2

Define AGGD and write the equation using LaTeX. What is an appropriate time interval over which we should calculate AGGD?

AGGD is the Accumulated Growing Degree Day, which uses the “accumulated” temperature in an ecosystem to predict phenological change.

$GDD = ((T_{max} + T_{min})/2) - T_{base}$

An appropriate time interval over which we should calculate AGGD could be the growing season, which varies based on location.

#3

Use the neonUtilities package to pull plant phenology observations (DP1.10055.001). We will work with the statusintensity data

```
#TOS Phenology Data

dpid <- as.character('DP1.10055.001') #phe data

pheDat <- loadByProduct(dpID="DP1.10055.001",
                     site = sitesOfInterest,
                     package = "basic",
                     check.size = FALSE,
                     token=NEON_TOKEN)
```

```
## Finding available files
##   |                                                                      |
## Downloading files totaling approximately 93.742811 MB
## Downloading 165 files
##   |                                                                              |
```

```
#NEON sends the data as a nested list, so I need to undo that
# unlist all data frames
list2env(pheDat ,.GlobalEnv)
```

```
## <environment: R_GlobalEnv>
```

```
summary(phe_perindividualperyear)
```

```
##      uid            namedLocation        domainID            siteID         
##  Length:1623        Length:1623        Length:1623        Length:1623       
##  Class :character   Class :character   Class :character   Class :character  
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character  
##                                                                             
##                                                                             
##                                                                             
##                                                                             
##     plotID               date                          editedDate                  
##  Length:1623        Min.   :2013-09-05 00:00:00   Min.   :2013-08-20 00:00:00  
##  Class :character   1st Qu.:2015-09-14 00:00:00   1st Qu.:2015-05-26 00:00:00  
##  Mode  :character   Median :2018-07-11 00:00:00   Median :2018-03-15 00:00:00  
##                     Mean   :2018-02-07 03:32:03   Mean   :2017-09-24 09:43:19  
##                     3rd Qu.:2020-07-23 00:00:00   3rd Qu.:2020-06-30 00:00:00  
##                     Max.   :2022-06-20 00:00:00   Max.   :2022-05-24 00:00:00  
##                                                   NA’s   :11                   
##  individualID       patchOrIndividual  canopyPosition     plantStatus       
##  Length:1623        Length:1623        Length:1623        Length:1623       
##  Class :character   Class :character   Class :character   Class :character  
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character  
##                                                                             
##                                                                             
##                                                                             
##                                                                             
##   stemDiameter    measurementHeight maxCanopyDiameter ninetyCanopyDiameter
##  Min.   :  0.00   Min.   : 10       Min.   : 0.000    Min.   : 0.000      
##  1st Qu.: 11.93   1st Qu.:130       1st Qu.: 0.200    1st Qu.: 0.125      
##  Median : 23.45   Median :130       Median : 0.400    Median : 0.400      
##  Mean   : 26.65   Mean   :117       Mean   : 4.188    Mean   : 3.243      
##  3rd Qu.: 40.88   3rd Qu.:130       3rd Qu.: 7.700    3rd Qu.: 5.600      
##  Max.   :100.00   Max.   :150       Max.   :21.700    Max.   :20.000
##  NA's   :1001     NA's   :1001      NA's   :393       NA's   :393
##    patchSize       percentCover       height       diseaseType
##  Min.   :0.0630   Min.   : 0.10   Min.   : 0.000   Length:1623
##  1st Qu.:0.0630   1st Qu.: 9.00   1st Qu.: 0.100   Class :character
##  Median :0.0630   Median :18.00   Median : 0.300   Mode  :character
##  Mean   :0.1433   Mean   :24.38   Mean   : 6.615
##  3rd Qu.:0.2500   3rd Qu.:33.00   3rd Qu.:13.825
##  Max.   :0.2500   Max.   :99.00   Max.   :56.000
##  NA's   :1481     NA's   :1306    NA's   :47
##  samplingProtocolVersion  measuredBy         recordedBy
##  Length:1623             Length:1623        Length:1623
##  Class :character        Class :character   Class :character
##  Mode  :character        Mode  :character   Mode  :character
##
##
##
##
##    remarks            dataQF          publicationDate      release
##  Length:1623        Length:1623        Length:1623        Length:1623
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
```

```
summary(phe_statusintensity)
```

```
##      uid            namedLocation       domainID           siteID
##  Length:289666      Length:289666      Length:289666      Length:289666
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##     plotID               date                          editedDate
##  Length:289666      Min.   :2013-08-23 00:00:00   Min.   :2015-03-19 00:00:00
##  Class :character   1st Qu.:2016-03-23 00:00:00   1st Qu.:2016-05-09 00:00:00
##  Mode  :character   Median :2018-04-23 00:00:00   Median :2018-04-23 00:00:00
##                     Mean   :2018-04-12 16:55:33   Mean   :2018-05-31 18:52:43
##                     3rd Qu.:2020-09-03 00:00:00   3rd Qu.:2020-09-03 00:00:00
##                     Max.   :2022-08-11 00:00:00   Max.   :2022-08-15 00:00:00
##                                                   NA's   :450
##    dayOfYear    individualID       phenophaseName     phenophaseStatus
##  Min.   :  2    Length:289666      Length:289666      Length:289666
##  1st Qu.:121    Class :character   Class :character   Class :character
##  Median :178    Mode  :character   Mode  :character   Mode  :character
##  Mean   :185
##  3rd Qu.:251
##  Max.   :364
##  NA's   :3861
##  phenophaseIntensityDefinition phenophaseIntensity samplingProtocolVersion
##  Length:289666                Length:289666      Length:289666
##  Class :character             Class :character   Class :character
##  Mode  :character             Mode  :character   Mode  :character
##
##
##
##
##   measuredBy         recordedBy          remarks          dataEntryRecordID
##  Length:289666      Length:289666      Length:289666      Length:289666
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##
##     dataQF          publicationDate      release
##  Length:289666      Length:289666      Length:289666
##  Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character
##
##
##
##
```

```
#remove duplicate records
phe_statusintensity <- select(phe_statusintensity, -uid)
phe_statusintensity <- distinct(phe_statusintensity)

#Format dates
phe_statusintensity$date <- as.Date(phe_statusintensity$date, "%Y-%m-%d")
```

```
## Warning in as.POSIXlt.POSIXct(x, tz = tz): unknown timezone '%Y-%m-%d'
```

```
phe_statusintensity$editedDate <- as.Date(phe_statusintensity$editedDate, "%Y-%m-%d")
```

```
## Warning in as.POSIXlt.POSIXct(x, tz = tz): unknown timezone '%Y-%m-%d'
```

```
phe_statusintensity$year <- as.numeric(substr(phe_statusintensity$date, 1, 4))
phe_statusintensity$month <- as.numeric(format(phe_statusintensity$date, format="%m"))

df = phe_statusintensity %>%
  left_join(phe_perindividual, by = "individualID") %>%
  filter(phenophaseName == "Colored leaves",
         taxonID == "QURU",
         phenophaseStatus == "yes") %>%
  select(date.x, year, month, dayOfYear, siteID.x, individualID, phenophaseIntensity) %>%
  na.omit()
```

Yes, there are ways to extract numerical values for string data that could be used for plotting. For example, you could count how many observations there are for each string type and make some kind of density visual. You could also subset the string to grab the numberical values and then convert those to numerical objects and use them for plotting directly.

#4

Using dpid DP1.00002.001 Single Aspirated Air Temperature calculate AGGD based on NEON tower data over the time period you decided upon in question 1. To save you time and frustration I've placed some mostly complete example code for one height on the tower just for Harvard. You will need to determine which height you think it best and complete these calculations for both sites. You will also need to consider things like filtering your temperature data for quality flags, and converting from GMT (Greenwich Mean Time) to your location's time:

```
dpid <- as.character('DP1.00002.001')  ##single aspirated air temperature

tempDat <- loadByProduct(dpID=dpid,
                        site = sitesOfInterest,
                        startdate = "2017-01",
                        enddate="2017-12",
                        avg=30,
                        package = "basic",
                        check.size = FALSE)
```

```
## Input parameter avg is deprecated; use timeIndex to download by time interval.
## Finding available files
##   |                                                                              |
##
## Downloading files totaling approximately 12.78088 MB
## Downloading 100 files
##   |                                                                              |
```

```
SAAT <- tempDat$SAAT_30min

# GDD typically reported in F
# convert df temps
SAAT$meanTempF=SAAT$tempSingleMean*1.8+32
SAAT$endDateTime = with_tz(SAAT$endDateTime, tzone = "America/New_York")


#pull date value from dateTime
SAAT$date <- substr(SAAT$endDateTime, 1, 10)

select(tempDat$sensor_positions_00002, c(HOR.VER, zOffset))
```

```
##    HOR.VER zOffset
## 1: 000.010    0.19
## 2: 000.020    5.29
## 3: 000.030   16.26
## 4: 000.040   22.52
## 5: 000.050   29.60
## 6: 000.010    0.16
## 7: 000.020    1.81
## 8: 000.030    3.87
```

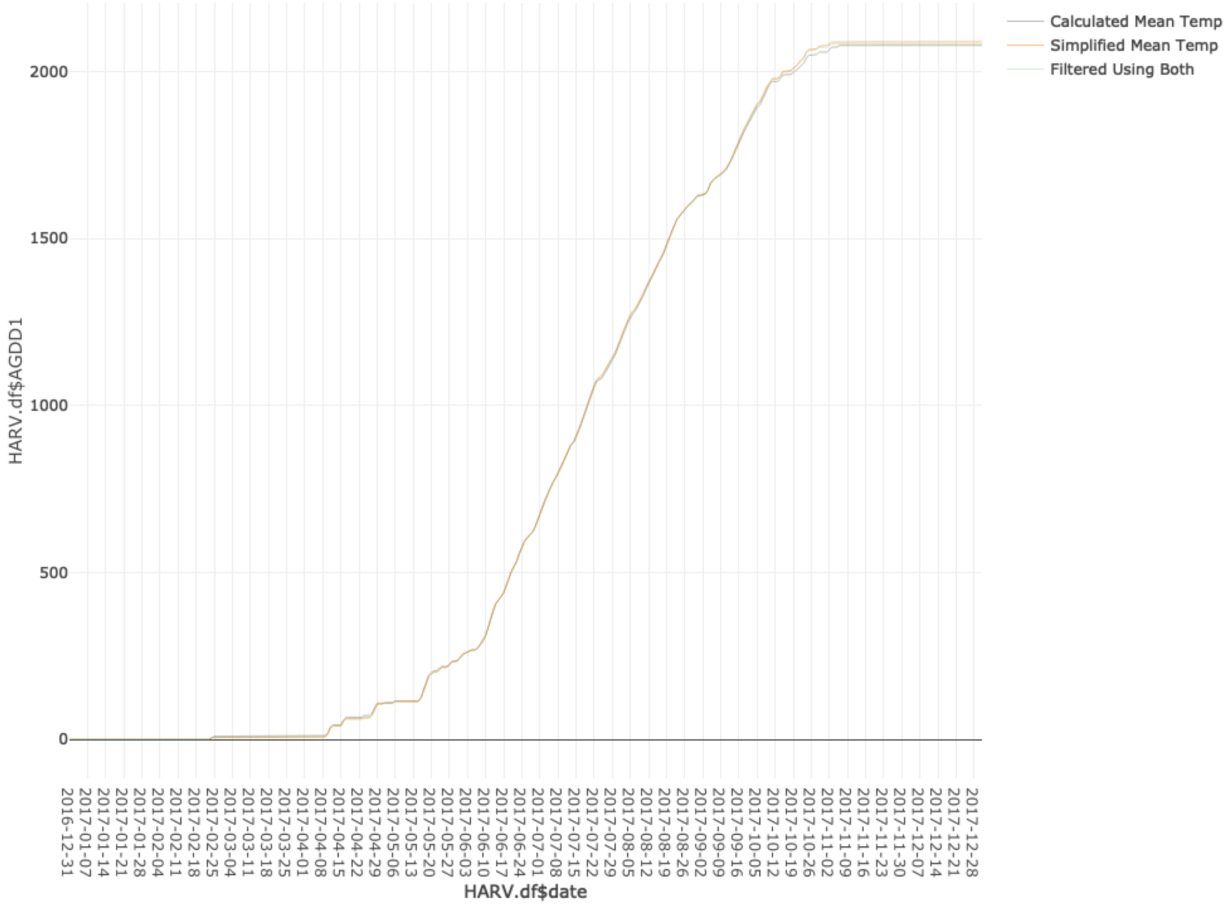
```
head(tempDat$sensor_positions_00002)
```

```
##    siteID HOR.VER         name
## 1:   HARV 000.010 CFGLOC100471
## 2:   HARV 000.020 CFGLOC100474
## 3:   HARV 000.030 CFGLOC100477
## 4:   HARV 000.040 CFGLOC100480
## 5:   HARV 000.050 CFGLOC100483
## 6:   CPER 000.010 CFGLOC100238
##                                           description                start end
## 1: Harvard Forest Single Aspirated Air Temperature L1 2010-01-01T00:00:00Z  NA
## 2: Harvard Forest Single Aspirated Air Temperature L2 2010-01-01T00:00:00Z  NA
## 3: Harvard Forest Single Aspirated Air Temperature L3 2010-01-01T00:00:00Z  NA
## 4: Harvard Forest Single Aspirated Air Temperature L4 2010-01-01T00:00:00Z  NA
## 5: Harvard Forest Single Aspirated Air Temperature L5 2010-01-01T00:00:00Z  NA
## 6: Central Plains Single Aspirated Air Temperature L1 2010-01-01T00:00:00Z  NA
##    referenceName referenceDescription       referenceStart referenceEnd xOffset
## 1:   TOWER100450 Harvard Forest Tower 2010-01-01T00:00:00Z           NA    5.36
## 2:   TOWER100450 Harvard Forest Tower 2010-01-01T00:00:00Z           NA    5.35
## 3:   TOWER100450 Harvard Forest Tower 2010-01-01T00:00:00Z           NA    5.35
## 4:   TOWER100450 Harvard Forest Tower 2010-01-01T00:00:00Z           NA    5.35
## 5:   TOWER100450 Harvard Forest Tower 2010-01-01T00:00:00Z           NA    5.35
## 6:   TOWER100223 Central Plains Tower 2010-01-01T00:00:00Z           NA    5.36
##    yOffset zOffset pitch roll azimuth referenceLatitude referenceLongitude
## 1:    2.40    0.19     0    0       0          42.53691          -72.17265
## 2:    2.36    5.29     0    0       0          42.53691          -72.17265
## 3:    2.36   16.26     0    0       0          42.53691          -72.17265
## 4:    2.36   22.52     0    0       0          42.53691          -72.17265
## 5:    2.36   29.60     0    0       0          42.53691          -72.17265
## 6:    2.40    0.16     0    0       0          40.81554         -104.74559
##    referenceElevation eastOffset northOffset xAzimuth yAzimuth  publicationDate
## 1:             348.13      -5.36       -2.40      270      180 20211211T013906Z
## 2:             348.13      -5.35       -2.36      270      180 20211211T013906Z
## 3:             348.13      -5.35       -2.36      270      180 20211211T013906Z
## 4:             348.13      -5.35       -2.36      270      180 20211211T013906Z
## 5:             348.13      -5.35       -2.36      270      180 20211211T013906Z
## 6:            1653.92      -5.36       -2.40      270      180 20211210T202950Z
```

```
day_temp <- SAAT%>%
  filter(verticalPosition=="030",
         finalQF == 0)%>%
  group_by(siteID, date)%>%
  mutate(dayMaxTemp=max(meanTempF), dayMinTemp=min(meanTempF),
         dayMeanTemp=mean(meanTempF))%>%
  select(siteID, date, dayMaxTemp, dayMinTemp, dayMeanTemp)%>%
  distinct()

##alternative, simplified mean, consistent with many GDD calculations
### does accumulation differ for true mean vs. simplified mean?
```

```
day_temp$mean2 <- (day_temp$dayMinTemp + day_temp$dayMaxTemp)/2

day_temp$GDD1 <- ifelse(day_temp$dayMeanTemp-50 < 0, 0, round(day_temp$dayMeanTemp-50, 0))
day_temp$GDD2 <- ifelse(day_temp$mean2-50 < 0, 0, round(day_temp$mean2-50, 0))
day_temp$GDD3 <- ifelse(day_temp$dayMeanTemp-50 < 0, 0, round(day_temp$mean2-50, 0))

# define year
day_temp$year <- substr(day_temp$date, 1, 4)

#function to add daily GDD values
sumr.2 <- function(x) {
    sapply(1:length(x), function(i) sum(x[1:i]))
}

#calculate Accumlated GDD
day_temp$AGDD3 <- sumr.2(x=day_temp$GDD3)
day_temp$AGDD2 <- sumr.2(x=day_temp$GDD2)
day_temp$AGDD1 <- sumr.2(x=day_temp$GDD1)
day_temp <- ungroup(day_temp)

library(plotly)


HARV.df = day_temp %>%
  filter(siteID == "HARV") %>%
  select(date, AGDD1, AGDD2, AGDD3)

CPER.df = day_temp %>%
  filter(siteID == "CPER") %>%
  select(date, AGDD1, AGDD2, AGDD3)

p1 = plot_ly() %>%
    add_trace(
      x= ~HARV.df$date,
      y = ~ HARV.df$AGDD1,
      type= 'scatter',
      mode = "lines",
      line = list(width = 1, color = "rgb(120,120,120)"),
      name = "Calculated Mean Temp",
      showlegend = TRUE,
      opacity=.5
    )%>%
  add_trace(
      data = HARV.df,
    x = ~ date,
    y = ~ AGDD2,
    name= 'Simplified Mean Temp',
    showlegend = TRUE,
    type = 'scatter',
    mode = 'lines',
    line = list(width = 1),
    opacity=.5)%>%
  add_trace(
      data = HARV.df,
    x = ~ date,
    y = ~ AGDD3,
    name= 'Filtered Using Both',
    showlegend = TRUE,
    type = 'scatter',
    mode = 'lines',
    line = list(width = 1),
    opacity=.2)

tmpFile <- tempfile(fileext = ".png")
export(p1, file = tmpFile)
```

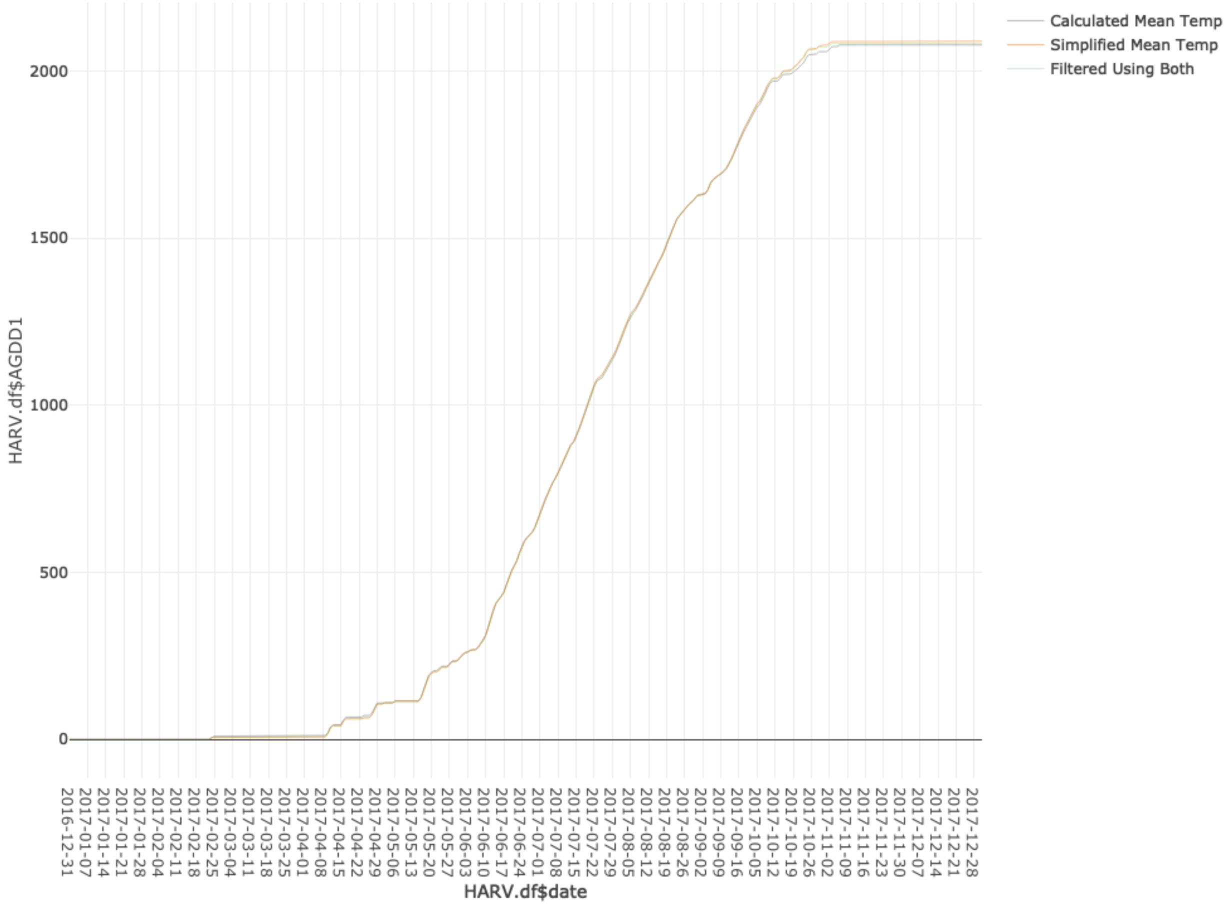
```
## Warning: 'export' is deprecated.
## Use 'orca' instead.
## See help("Deprecated")
```



```
p2 = plot_ly() %>%
    add_trace(
      x= ~CPER.df$date,
      y = ~ CPER.df$AGDD1,
      type= 'scatter',
```

```r
   mode = "lines",
   line = list(width = 1, color = "rgb(120,120,120)"),
   name = "Calculated Mean Temp",
   showlegend = TRUE,
   opacity=.5
  )%>%
 add_trace(
   data = CPER.df,
  x = ~ date,
  y = ~ AGDD2,
  name= 'Simplified Mean Temp',
  showlegend = TRUE,
  type = 'scatter',
  mode = 'lines',
  line = list(width = 1),
  opacity=.5)%>%
 add_trace(
   data = CPER.df,
  x = ~ date,
  y = ~ AGDD3,
  name= 'Filtered Using Both',
  showlegend = TRUE,
  type = 'scatter',
  mode = 'lines',
  line = list(width = 1),
  opacity=.2)

tmpFile <- tempfile(fileext = ".png")
export(p1, file = tmpFile)
```

```
## Warning: 'export' is deprecated.
## Use 'orca' instead.
## See help("Deprecated")
```



#5

Plot your calculated AGGD and comment on your calculations. Do you need to revise your time horizon or sensor height?

After doing the calculations and looking at my plots the first time, I went back and changed my sensor height. I realized after plotting that the sensor height I had picked was not available at the CPRE site. Because the taxon I picked is an oak, I wanted to use the highest sensor. But I had to picked the highest sensor height that was present at both sites.

#6

Now we're going to build a model to see how AGGD impacts phenological status. But Wait. Is phenology all driven by temperature? Should you consider any other variables? What about AGGD and just plain temperature? Also, we have one very temperate site, and another that is a semi-arid grassland. Should water availability of any sort be considered? Any other variables or data?

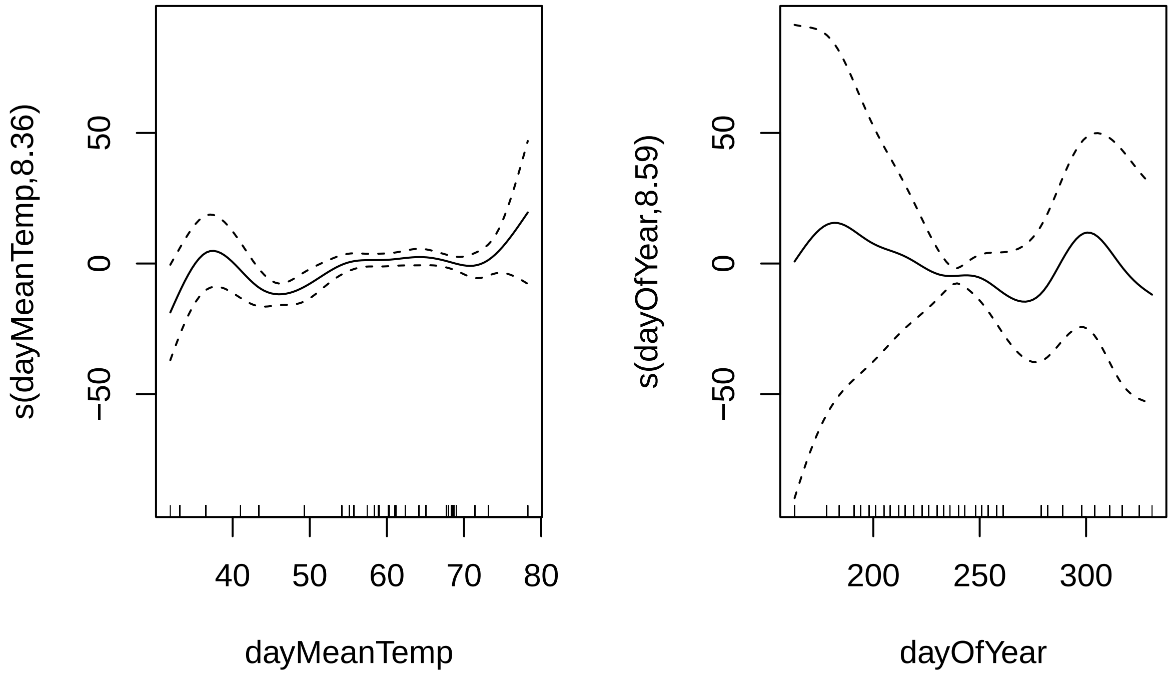
Yes, it has been widely documented that phenological change is driven by more than just growing degree days. It's well researched that temperature, solar radiation (photo period), and water availability all strongly impact phenology in addition to AGGD.

Create a GAM (Generalized Additive Model) for your phenological data including any variables you think might be relevant.

```r
# set up the data
day_temp = day_temp %>%
  mutate(
    date = ymd(date)
  )

gam.df = df %>%
  mutate(
    phenoInstNumb = case_when(
      phenophaseIntensity == "< 5%" ~ 5,
      phenophaseIntensity == "5-24%" ~ 15,
      phenophaseIntensity == "25-49%" ~ 37,
      phenophaseIntensity == "50-74%" ~ 62,
      phenophaseIntensity == "75-94%" ~ 85,
      phenophaseIntensity == ">= 95%" ~ 95,
    )
  ) %>%
  rename(date = date.x, siteID = siteID.x) %>%
  left_join(day_temp, by = c("date", 'siteID')) %>%
  filter(siteID == "HARV") %>%
  # get rid of any dates outside of 2017
  filter(year(date) == 2017)


library(mgcv)
model <- mgcv::gam(phenoInstNumb ~ AGDD3 + s(dayMeanTemp) + s(dayOfYear),
                   data = gam.df)
mgcv::summary.gam(model)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## phenoInstNumb ~ AGDD3 + s(dayMeanTemp) + s(dayOfYear)
##
## Parametric coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -23.13175   59.15861  -0.391    0.696
## AGDD3         0.02038    0.03741   0.545    0.586
##
## Approximate significance of smooth terms:
##                  edf Ref.df      F p-value
## s(dayMeanTemp) 8.357  8.823  6.439  <2e-16 ***
## s(dayOfYear)   8.592  8.882 17.006  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.407   Deviance explained = 42.5%
## GCV = 93.383  Scale est. = 90.338    n = 581
```

```r
mgcv::plot.gam(model, pages=1 )
```



I tried a several different GAMs with a variety variables and decided this was the best model.

7-8

7. Now that we have a model for NEON data, let's use the rnpn package to see how adding additional data could improve our fit. Use the taxonID that you selected at each NEON tower, and feed that to the rnpn package to grab observational data and increase your number of observations.

8. Pull AGGD from USA-NPN based on the observations you just pulled.

```
npn.df = npn_download_status_data(
  request_source = 'NAU',
  years = c('2017'),
  states = c("MA"),
  agdd_layer = 50,
  # get only observations for Quercus rubra
  species_ids = 102
)
```

```
## using a custom handler function.

## opening curl input connection.

##   |                                                                              |
##   |                                                                              |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                                 |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                             |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                                |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##   |                                                                               |
##  Found 5000 records...
```

```
## closing curl input connection.
```

```
## Service is currently unavailable. Please try again later!
```

```
npn.df = npn.df %>%
  filter(phenophase_description == "Colored leaves",
         intensity_value != -9999) %>%
  select(day_of_year, observation_date,
         update_datetime, intensity_value,
         genus, species, site_id, `gdd:agdd_50f`) %>%
  mutate(
    phenoInstNumb = case_when(
      intensity_value == "Less than 5%" ~ 5,
      intensity_value == "5-24%" ~ 15,
      intensity_value == "25-49%" ~ 37,
      intensity_value == "50-74%" ~ 62,
      intensity_value == "75-94%" ~ 85,
      intensity_value == "95% or more" ~ 95,
    )
  ) %>%
  rename(date = observation_date,
         dayOfYear = day_of_year,
         AGDD3 = `gdd:agdd_50f`)
```

#9

Combine your NEON and USA-NPN data into the same data.frame and re-fit your GAM.

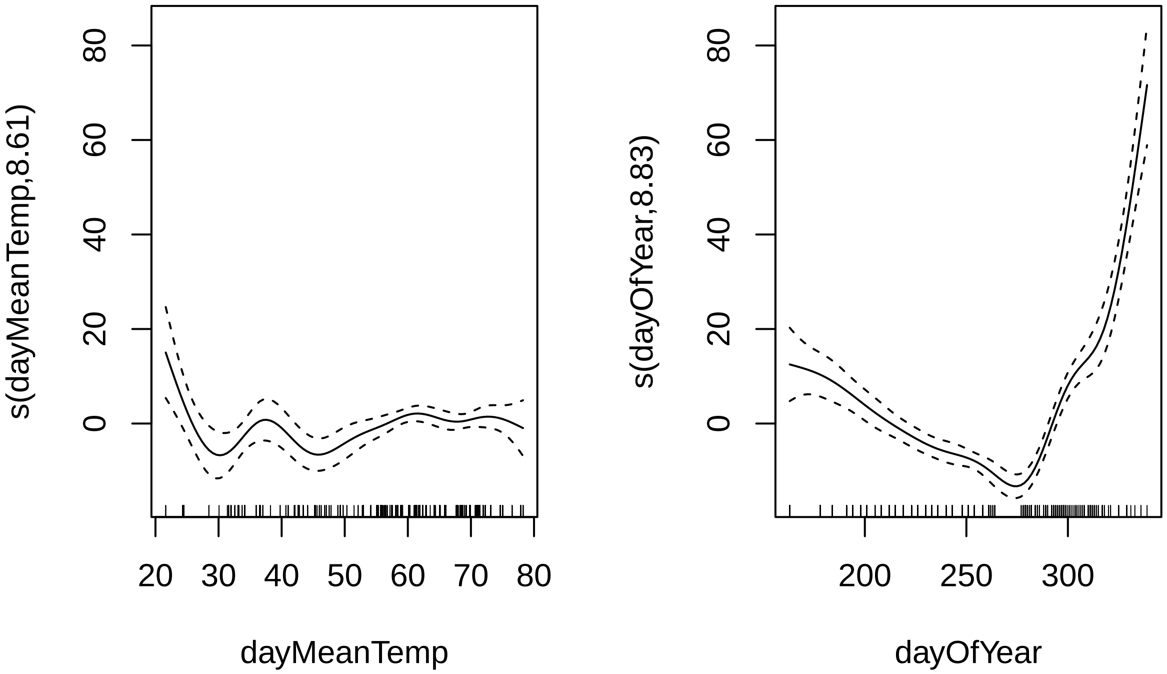
Summarize your new model

Plot your new model

```
# add the NPN data to the GAM dataframe
subgam1 = gam.df %>%
  # grab only the vars we need
  select(date, phenoInstNumb, AGDD3, dayOfYear)
subgam2 = npn.df %>%
  select(date, phenoInstNumb, AGDD3, dayOfYear)
subgam3 = rbind(subgam1, subgam2)
newgam.df = left_join(subgam3, day_temp, by = "date")

model <- mgcv::gam(phenoInstNumb ~ AGDD3.x + s(dayMeanTemp) + s(dayOfYear),
                   data = newgam.df)
mgcv::summary.gam(model)
```

```
##
## Family: gaussian
## Link function: identity
##
## Formula:
## phenoInstNumb ~ AGDD3.x + s(dayMeanTemp) + s(dayOfYear)
##
## Parametric coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -13.941570   2.406611  -5.793 8.65e-09 ***
## AGDD3.x       0.015109   0.001353  11.165  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Approximate significance of smooth terms:
##                  edf Ref.df      F  p-value
## s(dayMeanTemp) 8.607  8.950  3.787 0.000168 ***
## s(dayOfYear)   8.831  8.988 49.434  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## R-sq.(adj) =  0.509   Deviance explained = 51.6%
## GCV = 177.09  Scale est. = 174.5     n = 1328
```

```
mgcv::plot.gam(model, pages=1 )
```



Comment on your new model: was it improved? If so how?

Yes, my model did improve some. My R2 has increased and my residuals are smaller.