Leaf Color Change Date in Harvard Forest

Kevin Buck, Kyle Elliott, Sydney Higgins

2023-05-05

Introduction

Starting in 1991, Researcher John O'Keefe and his team collected various phenological statistics for a number of species of trees in the Prospect Hill Tract Harvard Forest in Petersham, MA. He has continued to record the statistics and compile them in an online data archive in order to provide researchers with data to investigate if certain trends are correlated with certain species. In this data archive, one particular data set has statistics regarding the mean day of the year in which each tree of various given species has 50% of their leaves change color for autumn. We will be investigating how this day (listed in Julian days within the data set) is correlated with species type and investigate the annual trends for each species throughout the last 30 years.

Current literature agrees that there is likely a trend of later fall coloring within deciduous species due to longer growing seasons as a result of climate change. However, it is unknown if differences between species or this climactic variation describe more change in fall leaf phenology. Because of climate change, we expect warmer temperatures to last longer into the year. Regardless, our initial hypothesis is that species will still be an important defining factor because the varying species have different physiological traits. This impacts the leaf color-changing process, which we predict will have a net result of later Julian days for the 50% quota to be reached by each species. Leaf color change affects nutrient cycling, so understanding the relationship between climate change and leaves changing color in different species is crucial for predicting if the long-term effects of climate change or different species are the dominant factor when leaves change. If climate change is the only defining factor, we would see a time-dependent differentiating change from year to year. In contrast, if species were another defining factor, we would see an asymmetric time-dependent differentiating change that isn't as dependent on the year-to-year as it would be on each species.

Hypotheses

Null: There is no species effect on leaf color change date. Climactic variation over the years will explain most variation

HA1:The leaf color change date would be different for each species, but not change from year to year.

HA2:The leaf color change date would be later as years progress, with each species' leaf color change date being different as well.

Methods

Data Acquisition/ O'Keefe and his team's data was collected through weekly observations of percent leaf coloration and percent leaf fall started in September each year. Researchers observed when 50% of the leaves of trees of different species had changed color and started to fall, and recorded those dates. They recorded this data for 33 unique types of species from 1991 to the present day, with the most recent data coming from 2022. By 2001, O'Keefe and his team had seen that of the 33 species they had been recording, only 14 had shown a statistically significant variety in average Julian day of leaf drop. Because of this, from 2002 onward they decided to only record data and inquire further with those 14 species, which are the 14 that we are focusing on in this project.

Analysis/ From our desired data set, we filtered down to the 14 species that had recorded mean Julian dates for 50% of the leaves changing colors for every year from 1991 until present. To assess the variation in date of leaf color change in each species per year, we first created a scatter plot of the Julian date vs. year and had each species of tree which were represented by the different colors. We added trend lines for each

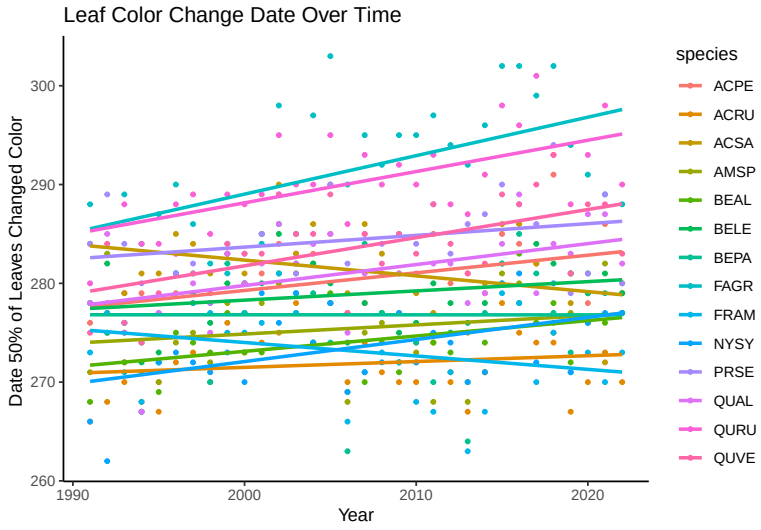
species to better make a preliminary visualization the overall trend for each species. We also made a graph of residuals, and confirmed the data fit the assumptions of linear models (linear, independence, equal variance and normally distributed residuals).

```
pheno <- read.csv(file="hf003-08-fall-mean-spp.csv") #read in the data
pheno_trimmed <- pheno %>% #pipe in data
  #filter in species present for every year from 1991 to present
  filter(species=="ACRU"|species=="ACSA"|species=="ACPE"|species=="AMSP"|
           species=="BEAL"|species=="BELE"|species=="BEPA"|species=="FAGR"|
           species=="FRAM"|species=="NYSY"|species=="PRSE"|species=="QUAL"|
           species=="QURU"|species=="QUVE")

pheno_trimmed %>%
  ggplot(aes(x = year, y = lc_doy)) + #set year as x axis and julian day of change as the y
  geom_point(aes(color=species, fill=species), size=1) +
  ggtitle("Leaf Color Change Date Over Time")+ #add title
  #adding in trend lines by species
  geom_smooth(aes(group=species,color=species), method = "lm", formula = y ~ x, se = F)+
  xlab("Year") + #label x axis
  ylab("Date 50% of Leaves Changed Color ") + #label y axis
  theme_classic() #changing background
```

```
## Warning: Removed 2 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 2 rows containing missing values (`geom_point()`).
```



To get an idea whether climatic variation, species effects, or an interaction between the two had important effects on trends in leaf color change date, we decided to run an ANCOVA with the data based upon a linear model. Explanatory variables (groups) would include year(a continous variable to act as a proxy for climactic variation interannually), species, and the interaction between the two. Due to autocorrelation of climactic trends as well as strong influence of one year's leaf color change data to the next, we selected ANCOVA over ANOVA intentionally, with year as a covariate with leaf change date. The F and p values of the ANCOVA will be informative of whether there are imporaatnt differences between years and species (or the interaction between the two). Pairwise comparisons were run post-hoc using a Tukey Test to look at specific comparisons between species. The results of that can be seen in the supplementary information.

```
options(contrasts = c("contr.sum","contr.ply")) #options for later post hoc comparison
model_1 <- lm(lc_doy ~ year * species,data=pheno_trimmed) #build linear model
#summary(model_1) #this prints out long, so see the supplementary info
```

```
#resid_panel(model_1,plot="all") #check on residuals, see SI
Anova(model_1, type = 3) #run an ANCOVA
```

```
## Anova Table (Type III tests)
##
## Response: lc_doy
##              Sum Sq  Df F value    Pr(>F)
## (Intercept)     2.9   1  0.1725    0.6781
## year          643.6   1 37.8977 1.750e-09 ***
## species       842.7  13  3.8169 7.958e-06 ***
## year:species  861.6  13  3.9027 5.367e-06 ***
## Residuals    7098.7 418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
pwc <- pheno_trimmed %>% tukey_hsd(lc_doy ~ species) #pair wise comparisons
#pwc #look at the PWC of species, also see SI
```

Results

Extremely low p-values for comparisons between different species, different years, and the interaction between the two show year, species, and the year-species interaction show that all of these factors have an important effect on variation in leaf change date ($p<.0001$ for all). However, the F-value for year that is 10 times larger ($F=37.9$) than that for species ($F=3.8$) and the year:species interaction (F-3.9) show that year is the dominating contributor to the trend in leaf color change date shift (which itself has $R^2=.69$ and $p<.0001$). Pairwise comparisons reveal varying levels of similarity between species in how much and in what direction then influence the date shift trend. One of the further tests we can do is compare different genera and different species within genera to see if the leaf change date is more dependent on species or genus.

Conclusion

Based on the analysis, we can conclude leaf turn date is more climate dependent than species-dependent, but species-dependent is still significant. We also see the interaction between species and year is significant, which means different species aren't responding in the same way to climate changes as years go on. This is not what we hypothesized, but makes intuitive sense, as different tree species respond differently to changing climate. Further testing will help clarify what effect species differences affect the leaf drop date, and help us understand which species are more or less affected by climate change from year to year. One of the further tests we can do is compare different genera and different species within genera to see if the leaf change date is more dependent on species or genus. Another further test we could do is adding back in the species with partial data sets and only looking at the data as far as we can use all species, assessing the differences between all species.

Contribution Statement

We extend our appreciation to everyone who made this project possible, specifically Mr. O'Keefe and his team at Harvard that collected the data that we were able to use as well as Dr. McLachlan and his team of TAs (specifically Gretchen and Eva) that helped to advise and educate us on biostatistics. This project would not have been possible without your work as well, so we thank you for making this opportunity for our group! As far as the work was delegated for this project, Kevin set the foundation for most of the code and formatted the document, Kyle helped work on the graphs as well as the introduction through analysis, and Sydney completed the results, conclusions, as well as helped to organize and format the completed project.

Citations

O'Keefe J. 2023. Phenology of Woody Species at Harvard Forest since 1990. Harvard Forest Data Archive: HF003 (v.35). Environmental Data Initiative: https://doi.org/10.6073/pasta/eb0dd36c6ec62a918340b6bda38be832.

Supplementary Information

Linear Model Results and Coefficients

```
summary(model_1)
```

```
##
## Call:
## lm(formula = lc_doy ~ year * species, data = pheno_trimmed)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.8127  -2.6427   0.1924   2.6173  12.0243
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)       17.68169   42.57334   0.415 0.678119
## year               0.13061    0.02122   6.156 1.75e-09 ***
## species1         -95.44268  152.62876  -0.625 0.532099
## species2         135.77451  152.62876   0.890 0.374207
## species3         583.58224  152.62876   3.824 0.000152 ***
## species4          72.06813  152.62876   0.472 0.637045
## species5         -55.41739  152.62876  -0.363 0.716724
## species6          73.29907  152.62876   0.480 0.631305
## species7         259.86633  152.62876   1.703 0.089386 .
## species8        -507.42989  158.62931  -3.199 0.001485 **
## species9         528.32143  152.62876   3.461 0.000593 ***
## species10       -192.45020  152.62876  -1.261 0.208047
## species11         28.44716  152.62876   0.186 0.852236
## species12       -162.02406  152.62876  -1.062 0.289050
## species13       -363.03852  158.62931  -2.289 0.022601 *
## year:species1      0.04791    0.07607   0.630 0.529169
## year:species2     -0.07159    0.07607  -0.941 0.347136
## year:species3     -0.29007    0.07607  -3.813 0.000158 ***
## year:species4     -0.03805    0.07607  -0.500 0.617146
## year:species5      0.02481    0.07607   0.326 0.744437
## year:species6     -0.03695    0.07607  -0.486 0.627352
## year:species7     -0.13098    0.07607  -1.722 0.085826 .
## year:species8      0.25878    0.07904   3.274 0.001149 **
## year:species9     -0.26661    0.07607  -3.505 0.000506 ***
## year:species10     0.09281    0.07607   1.220 0.223100
## year:species11    -0.01184    0.07607  -0.156 0.876338
## year:species12     0.08145    0.07607   1.071 0.284901
## year:species13     0.18614    0.07904   2.355 0.018983 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 4.121 on 418 degrees of freedom
##   (2 observations deleted due to missingness)
## Multiple R-squared:  0.7089, Adjusted R-squared:  0.6901
## F-statistic: 37.71 on 27 and 418 DF,  p-value: < 2.2e-16
```

Pairwise Comparisons

```
print(pwc, n=100)
```

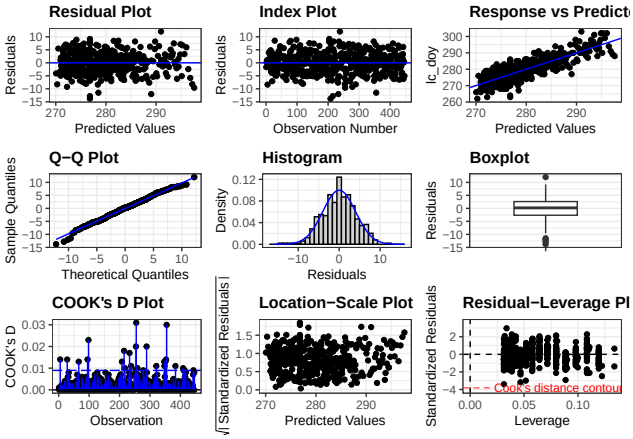
```
## # A tibble: 91 x 9
##    term    group1 group2 null.value estimate conf.low conf.high    p.adj
##  * <chr>   <chr>  <chr>       <dbl>    <dbl>    <dbl>     <dbl>    <dbl>
```

```
##  1 species ACPE   ACRU          0   -8.56   -12.3     -4.80   0
##  2 species ACPE   ACSA          0    0.875   -2.88     4.63   1   e+ 0
##  3 species ACPE   AMSP          0   -4.97    -8.73    -1.21   8.68e- 4
##  4 species ACPE   BEAL          0   -6.31   -10.1     -2.55   2.40e- 6
##  5 species ACPE   BELE          0   -1.53    -5.29     2.23   9.84e- 1
##  6 species ACPE   BEPA          0   -3.63    -7.38     0.133  7.14e- 2
##  7 species ACPE   FAGR          0   11.3      7.52    15.1    0
##  8 species ACPE   FRAM          0   -7.31   -11.1     -3.55   1.37e- 8
##  9 species ACPE   NYSY          0   -6.91   -10.7     -3.15   1.20e- 7
## 10 species ACPE   PRSE          0    4.00     0.242    7.76   2.49e- 2
## 11 species ACPE   QUAL          0    0.719   -3.04     4.48   1   e+ 0
## 12 species ACPE   QURU          0    9.92     6.13    13.7    0
## 13 species ACPE   QUVE          0    3.19    -0.570    6.95   1.99e- 1
## 14 species ACRU   ACSA          0    9.44     5.68    13.2    0
## 15 species ACRU   AMSP          0    3.59    -0.164    7.35   7.74e- 2
## 16 species ACRU   BEAL          0    2.25    -1.51     6.01   7.55e- 1
## 17 species ACRU   BELE          0    7.03     3.27    10.8    6.22e- 8
## 18 species ACRU   BEPA          0    4.94     1.18     8.70   9.8 e- 4
## 19 species ACRU   FAGR          0   19.9     16.1     23.7    0
## 20 species ACRU   FRAM          0    1.25    -2.51     5.01   9.98e- 1
## 21 species ACRU   NYSY          0    1.66    -2.10     5.41   9.7 e- 1
## 22 species ACRU   PRSE          0   12.6      8.80    16.3    0
## 23 species ACRU   QUAL          0    9.28     5.52    13.0    0
## 24 species ACRU   QURU          0   18.5     14.7     22.3    0
## 25 species ACRU   QUVE          0   11.7      7.99    15.5    0
## 26 species ACSA   AMSP          0   -5.84    -9.60    -2.09   2.17e- 5
## 27 species ACSA   BEAL          0   -7.19   -10.9     -3.43   2.70e- 8
## 28 species ACSA   BELE          0   -2.41    -6.16     1.35   6.59e- 1
## 29 species ACSA   BEPA          0   -4.5     -8.26    -0.742  4.89e- 3
## 30 species ACSA   FAGR          0   10.4      6.64    14.2    0
## 31 species ACSA   FRAM          0   -8.19   -11.9     -4.43   6.60e-11
## 32 species ACSA   NYSY          0   -7.78   -11.5     -4.02   9.63e-10
## 33 species ACSA   PRSE          0    3.13    -0.633    6.88   2.26e- 1
## 34 species ACSA   QUAL          0   -0.156   -3.91     3.60   1   e+ 0
## 35 species ACSA   QURU          0    9.04     5.25    12.8    0
## 36 species ACSA   QUVE          0    2.31    -1.45     6.07   7.18e- 1
## 37 species AMSP   BEAL          0   -1.34    -5.10     2.41   9.95e- 1
## 38 species AMSP   BELE          0    3.44    -0.320    7.20   1.14e- 1
## 39 species AMSP   BEPA          0    1.34    -2.41     5.10   9.95e- 1
## 40 species AMSP   FAGR          0   16.3     12.5     20.1    0
## 41 species AMSP   FRAM          0   -2.34    -6.10     1.41   6.98e- 1
## 42 species AMSP   NYSY          0   -1.94    -5.70     1.82   9.01e- 1
## 43 species AMSP   PRSE          0    8.97     5.21    12.7    0
## 44 species AMSP   QUAL          0    5.69     1.93     9.45   4.36e- 5
## 45 species AMSP   QURU          0   14.9     11.1     18.7    0
## 46 species AMSP   QUVE          0    8.16     4.40    11.9    8.48e-11
## 47 species BEAL   BELE          0    4.78     1.02     8.54   1.77e- 3
## 48 species BEAL   BEPA          0    2.69    -1.07     6.45   4.73e- 1
## 49 species BEAL   FAGR          0   17.6     13.8     21.4    0
## 50 species BEAL   FRAM          0   -1       -4.76     2.76   1   e+ 0
## 51 species BEAL   NYSY          0   -0.594   -4.35     3.16   1   e+ 0
## 52 species BEAL   PRSE          0   10.3      6.55    14.1    0
## 53 species BEAL   QUAL          0    7.03     3.27    10.8    6.22e- 8
## 54 species BEAL   QURU          0   16.2     12.4     20.0    0
```

```
## 55 species BEAL   QUVE           0    9.5      5.74    13.3    0
## 56 species BELE   BEPA           0   -2.09    -5.85     1.66   8.37e- 1
## 57 species BELE   FAGR           0   12.8      9.05    16.6    0
## 58 species BELE   FRAM           0   -5.78    -9.54    -2.02   2.87e- 5
## 59 species BELE   NYSY           0   -5.37    -9.13    -1.62   1.68e- 4
## 60 species BELE   PRSE           0    5.53     1.77     9.29   8.64e- 5
## 61 species BELE   QUAL           0    2.25    -1.51     6.01   7.55e- 1
## 62 species BELE   QURU           0   11.4      7.66    15.2    0
## 63 species BELE   QUVE           0    4.72     0.961    8.48   2.23e- 3
## 64 species BEPA   FAGR           0   14.9     11.1     18.7    0
## 65 species BEPA   FRAM           0   -3.69    -7.45     0.0704 6.05e- 2
## 66 species BEPA   NYSY           0   -3.28    -7.04     0.477  1.63e- 1
## 67 species BEPA   PRSE           0    7.63     3.87    11.4    2.38e- 9
## 68 species BEPA   QUAL           0    4.34     0.586    8.10   8.34e- 3
## 69 species BEPA   QURU           0   13.5      9.75    17.3    0
## 70 species BEPA   QUVE           0    6.81     3.05    10.6    1.96e- 7
## 71 species FAGR   FRAM           0  -18.6    -22.4    -14.8    0
## 72 species FAGR   NYSY           0  -18.2    -22.0    -14.4    0
## 73 species FAGR   PRSE           0   -7.30   -11.1     -3.52   1.96e- 8
## 74 species FAGR   QUAL           0  -10.6    -14.4     -6.80   0
## 75 species FAGR   QURU           0   -1.39    -5.21     2.43   9.95e- 1
## 76 species FAGR   QUVE           0   -8.12   -11.9     -4.33   1.79e-10
## 77 species FRAM   NYSY           0    0.406   -3.35     4.16   1   e+ 0
## 78 species FRAM   PRSE           0   11.3      7.55    15.1    0
## 79 species FRAM   QUAL           0    8.03     4.27    11.8    2.06e-10
## 80 species FRAM   QURU           0   17.2     13.4     21.0    0
## 81 species FRAM   QUVE           0   10.5      6.74    14.3    0
## 82 species NYSY   PRSE           0   10.9      7.15    14.7    0
## 83 species NYSY   QUAL           0    7.62     3.87    11.4    2.38e- 9
## 84 species NYSY   QURU           0   16.8     13.0     20.6    0
## 85 species NYSY   QUVE           0   10.1      6.34    13.9    0
## 86 species PRSE   QUAL           0   -3.28    -7.04     0.477  1.63e- 1
## 87 species PRSE   QURU           0    5.92     2.13     9.71   1.92e- 5
## 88 species PRSE   QUVE           0   -0.813   -4.57     2.95   1   e+ 0
## 89 species QUAL   QURU           0    9.20     5.41    13.0    0
## 90 species QUAL   QUVE           0    2.47    -1.29     6.23   6.18e- 1
## 91 species QURU   QUVE           0   -6.73   -10.5     -2.94   3.94e- 7
## # i 1 more variable: p.adj.signif <chr>
```

Residuals

```
resid_panel(model_1,plot="all")
```


Residual Plot
Residuals
Predicted Values
Index Plot
Residuals
Observation Number
Response vs Predict
lc_doy
Predicted Values
Q−Q Plot
Sample Quantiles
Theoretical Quantiles
Histogram
Density
Residuals
Boxplot
Residuals
COOK's D Plot
COOK's D
Observation
Location−Scale Plot
√| Standardized Residuals |
Predicted Values
Residual−Leverage Pl
Standardized Residuals
Leverage
Cook's distance contour