Analysis for Hoecker and Higuera XXXX:

Setup

Load the requisite packages.

```
library(tidyverse) # For data manipulation and plotting
library(zoo) # For rolling statistics
library(Hmisc) # For correlation statistics
library(knitr) # For formatting this file
```

Specify directory on your machine containing data (ends with 'nowitna/data')

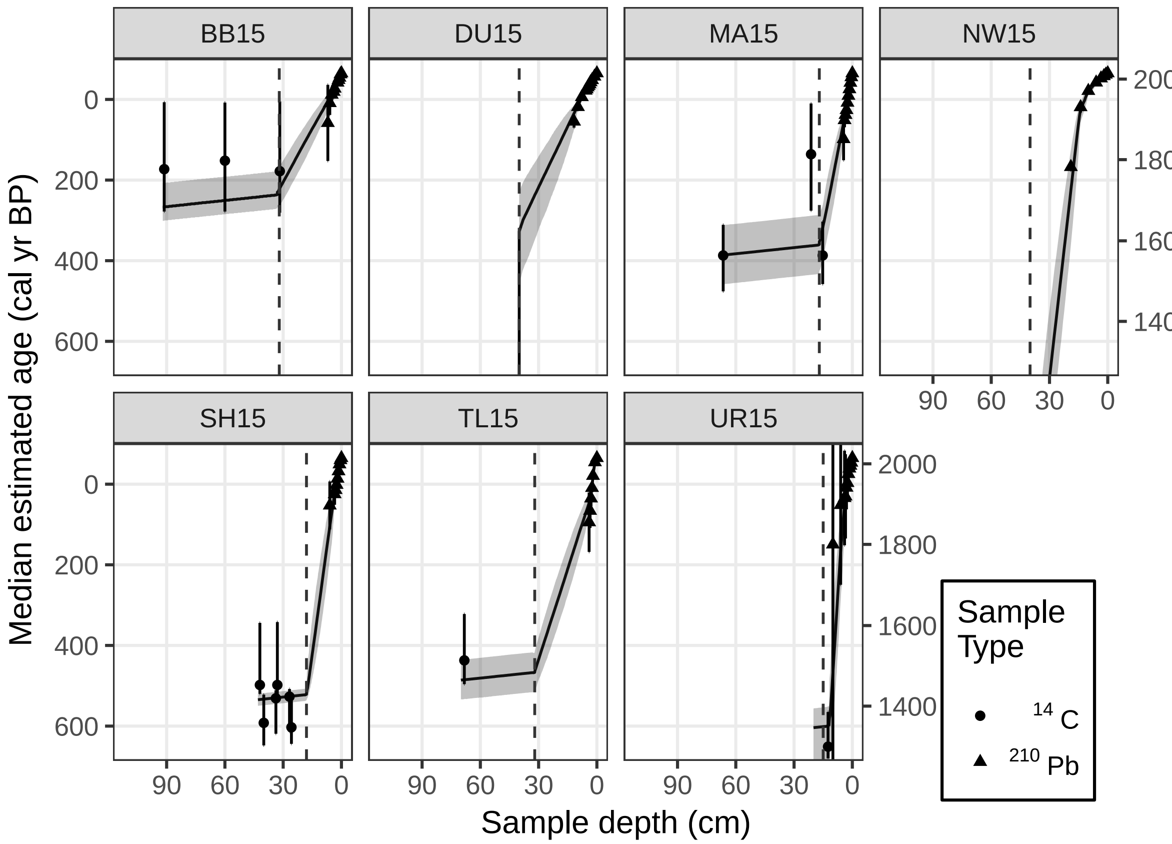
```
dataDir <- "/Users/tylerhoecker/GitHub/nowitna/data/"
```

Specify lakes and time period of analysis.

```
lakes <- c('BB15','DU15','MA15','NW15','SH15','TL15','UR15')
studyPeriod <- c(1550,2015)
```

Load chronological data. Age-depth models were built in Bacon v2.2 (http://www.chrono.qub.ac.uk/blaauw/bacon.html), using the parameters described in the text (Table 2).

```
dates <- read_csv(paste0(dataDir,'/ageData_sample_inventory.csv'))

ageModel.df <- map(paste0(dataDir,'ageModel_',lakes,'.csv'), read_csv) %>%
  `names<-` (lakes) %>%
  bind_rows(.id = 'lake')
```



Plot age-depth models.

Load charcoal data. These data have undergone peak analysis and interpolation in CharAnalysis (https://github.com/phiguera/CharAnalysis), using the parameters described in the text (Table 2).

```
char.ldf <- map(paste0(dataDir,'charResults_',lakes,'.csv'),read_csv)

char.df <- char.ldf %>%
  bind_rows() %>%
  mutate(ageCE = 1950 - ageTop,
         peakYr = ifelse(peaksFinal == 1,ageCE,NA),
         peakInsigYr = ifelse(peaksInsig == 1,ageCE,NA),
         peakYr1 = ifelse(peaks1 == 1,ageCE,NA),
         peakYr2 = ifelse(peaks2 == 1,ageCE,NA),
         peakYr3 = ifelse(peaks3 == 1,ageCE,NA)) %>%
  filter(ageCE >= studyPeriod[1]) %>%
  group_by(lake) %>%
         mutate(length = max(ageCE) - min(ageCE))
```

Load observed (historic) fire event data.

```
observed.df <- read_csv(paste0(dataDir,'/observedfireData.csv'))
obs.1km <- filter(observed.df,distance == 1.0) %>%
  rename(obsCE = ageCE)
```

Load tree demography data.

```
tree.df <- read_csv(paste0(dataDir,'treeData.csv'))
```

Load modeled and proxy climate data, bin into 5-year means.

```
goa.df <- read_csv(paste0(dataDir,'climateData_GOA.csv'))
cru.df <- read_csv(paste0(dataDir,'climateData_CRU.csv'))
# Extract growing season temperatures from CRU data
cru.gs.df <- cru.df %>%
  filter(month %in% c(4:9)) %>%
  group_by(yearCE,variable) %>%
  summarise(mean = mean(value))
```

Summarize fire event metrics

Mean signal to noise indices of records during study period:

```
kable(
  char.df %>%
    summarise(meanSNI = mean(SNI, na.rm = T)),
  caption = 'Mean SNI'
)
```

Table 1: Mean SNI

lake	meanSNI
BB15	9.8072
DU15	3.3015
MA15	8.6999
NW15	6.0621
SH15	11.1230
TL15	14.1624

lake	meanSNI
UR15	4.7611

Study-wide mean FRI (mean of all FRI pooled, rather than mean of means):

```
FRI <- char.df %>%
  select(lake,peakYr,length) %>%
  filter(!is.na(peakYr)) %>%
  mutate(FRI =  (lead(peakYr, 1) - peakYr) * -1)

kable(
  FRI %>%
    ungroup(lake) %>%
    summarise(min = min(FRI, na.rm = T),
              max = max(FRI, na.rm = T),
              mean = mean(FRI, na.rm = T),
              median = median(FRI, na.rm = T),
              sd = sd(FRI, na.rm = T)),
  caption = 'Study-wide mean FRI'
)
```

Table 2: Study-wide mean FRI

min	max	mean	median	sd
25	195	90	67.5	59.82819

Individual mean FRI:

```
kable(
  FRI %>%
  group_by(lake) %>%
  summarise(nFires = n(),
            mFRI = mean(FRI, na.rm = T),
            sdFRI = sd(FRI, na.rm = T),
            length = length[1],
            FF = nFires/length * 100) ,
  caption = 'Mean FRI'
)
```

Table 3: Mean FRI

lake	nFires	mFRI	sdFRI	length	FF
BB15	4	91.66667	57.95113	285	1.4035088
DU15	6	52.00000	22.80351	360	1.6666667
MA15	3	115.00000	84.85281	385	0.7792208
NW15	1	NaN	NA	465	0.2150538
SH15	4	128.33333	83.26664	465	0.8602151
TL15	4	115.00000	85.44004	465	0.8602151
UR15	3	62.50000	10.60660	465	0.6451613

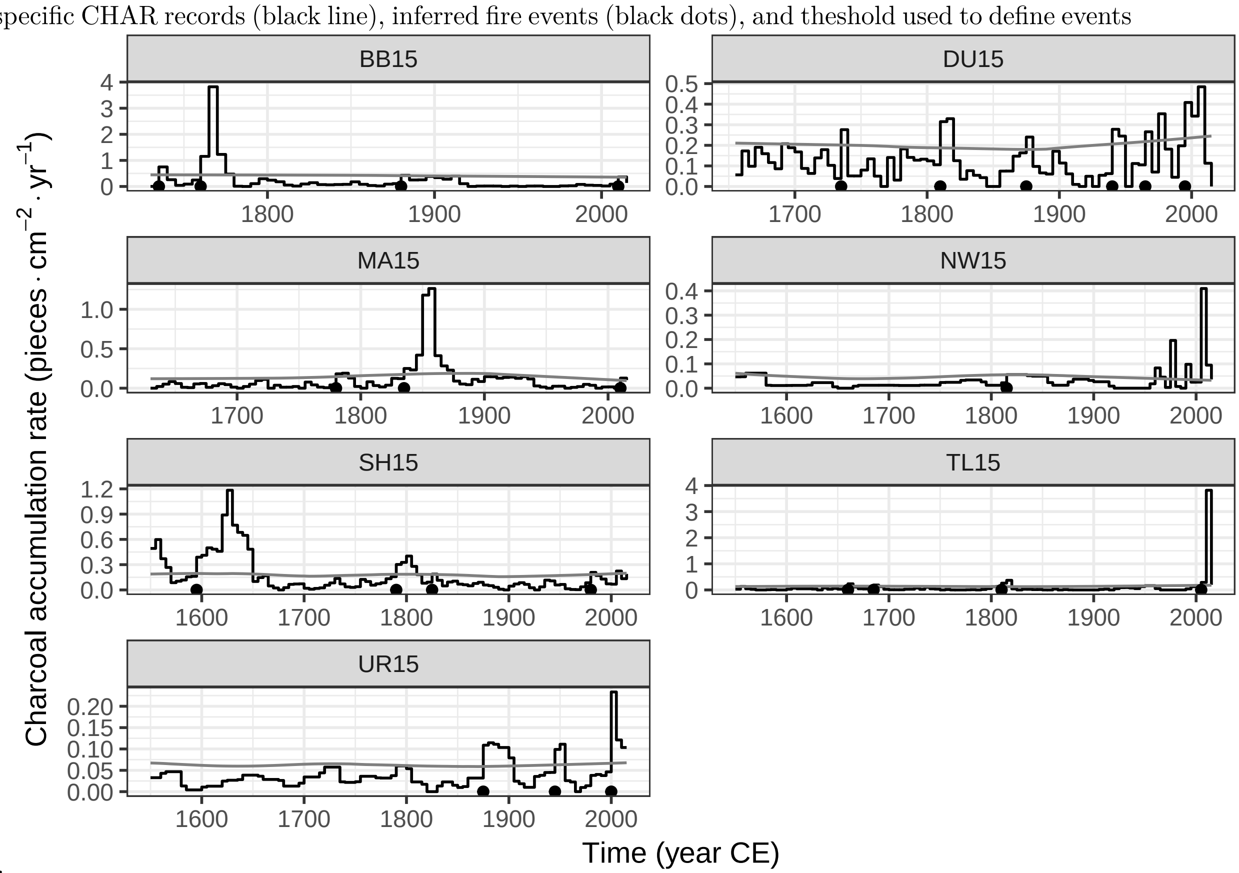
Calculate the difference (in years) between observed fire years within 1km of lakes and the most recent fire event detected in charcoal record

```
char.df <- char.df %>%
  group_by(lake) %>%
  right_join(.,obs.1km, by = 'lake')

kable(
  char.df %>%
  group_by(lake) %>%
  summarise(observed = obsCE[1],
            lastPeak = max(peakYr, na.rm = T),
            difference = abs(obsCE[1] - max(peakYr, na.rm = T))),
  caption = 'Inferred vs. observed fire event timing.'
)
```

Table 4: Inferred vs. observed fire event timing.

lake	observed	lastPeak	difference
BB15	1991	2010	19
DU15	1991	1995	4
MA15	2013	2010	3
NW15	NA	1815	NA
SH15	1986	1980	6
TL15	2013	2005	8
UR15	1991	2000	9

Plot lake-specific CHAR records (black line), inferred fire events (black dots), and theshold used to define events

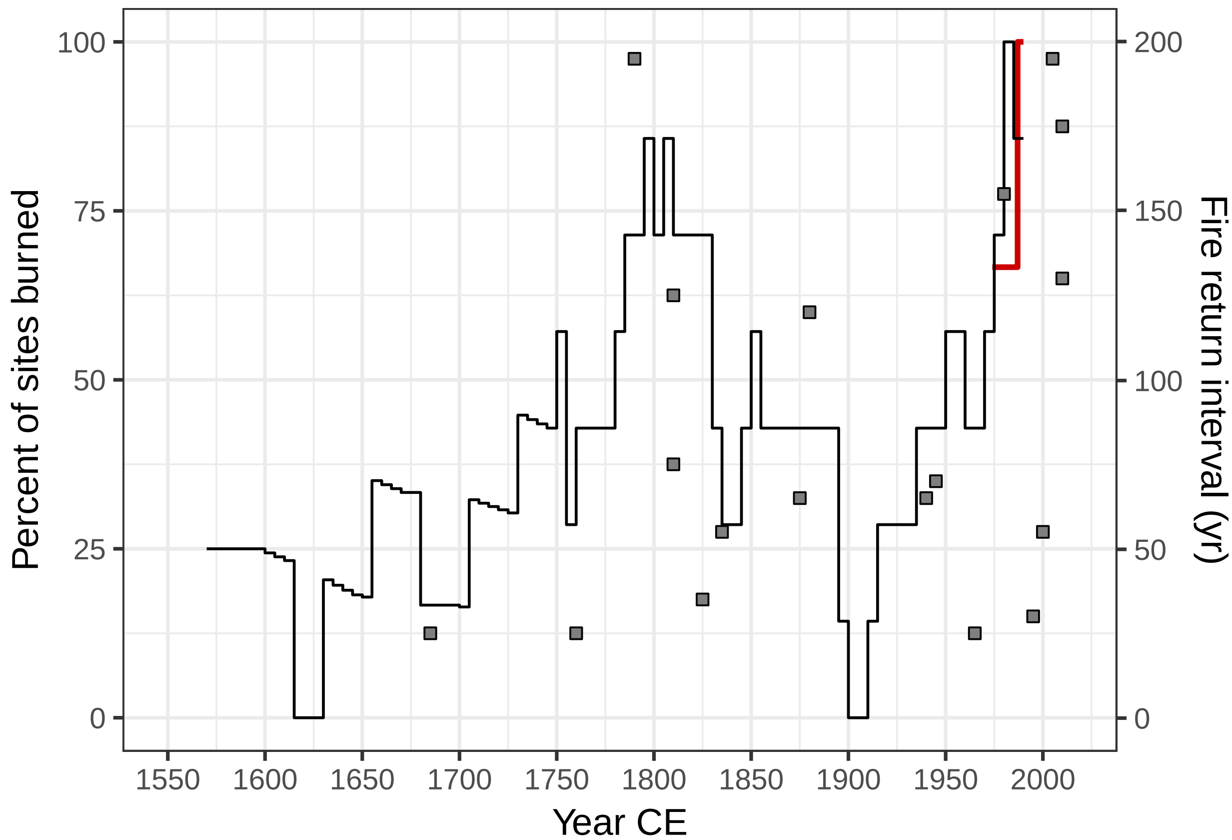


(grey line).

Calculate proportion of sites burned through time, using 50-year windows in continuous 5-year time steps.

```
window = 50
timeStep = 5

sitesByYear <- char.df %>%
  group_by(ageCE) %>%
  summarise(sites = n())

pctBurned <- char.df %>%
  filter(!is.na(peakYr)) %>%
  group_by(peakYr) %>%
  summarise(n.burned = n()) %>%
  rename(ageCE = peakYr) %>%
  full_join(sitesByYear,.) %>%
  mutate(n.burned = ifelse(!is.na(n.burned),n.burned,0)) %>%
  # Window
  mutate(win.total = rollapply(sites, window/timeStep, fill= NA,
                               FUN = mean, na.rm =T)) %>%
  mutate(win.burn = rollsum(n.burned, window/timeStep, fill= NA)) %>%
  mutate(win.pct = win.burn/win.total*100)


modernTime <- seq(1950,2015,1)
modernSites = data.frame('ageCE' = modernTime, 'sites' = 6)

pctModern <- observed.df %>%
  filter(distance == 1) %>%
  group_by(ageCE) %>%
  summarise(n.burned = n()) %>%
  full_join(modernSites,.) %>%
  mutate(n.burned = ifelse(!is.na(n.burned),n.burned,0)) %>%
  # Window
  mutate(win.total = rollapply(sites, window, fill= NA,
                               FUN = mean, na.rm =T)) %>%
  mutate(win.burn = rollsum(n.burned, window, fill= NA)) %>%
  mutate(win.pct = win.burn/win.total*100)
```

Plot the result, a time series of FRI (grey squares), percent sites burned based on the charcoal record (black line), and percent sites burned based on observed fire data since 1950 (red line).



Build composite biomass burning recrod from individual records (no peak analysis or interpolation, raw charcoal count data).

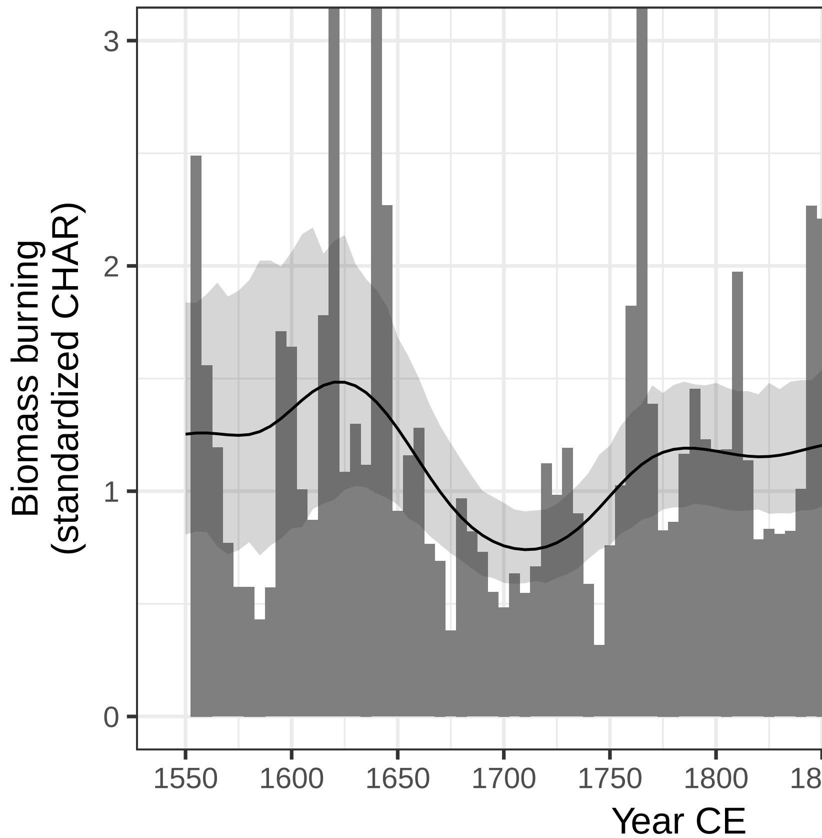
Import raw charcoal count data; derive and standardize charcoal accumulation rates (# cm2 yr-1, "CHAR").

```
# Function for standardizing non-zero charcoal accumlations rates
transFunc <- function(x) {
  x = ifelse(x > 0, x, NA)
  logX = log(x)
  zX = (logX - mean(logX, na.rm = T)) / sd(logX, na.rm = T)
  expX = exp(zX)
  expX[is.na(expX)] = 0
  return(expX)
}
# Calculate and standardize CHAR
charData <- map(paste0(dataDir,'charData_',lakes,'.csv'),read_csv) %>%
  `names<-` (lakes) %>%
  bind_rows(.id = 'lake') %>%
  group_by(lake) %>%
  mutate(sedAcc = (cmTop - cmBot) / (ageTop - ageBot),
         rawChar = (charCount / charVol) * sedAcc,
         char = transFunc(rawChar)) %>%
  rowwise() %>%
  mutate(age = round(mean(c(ageTop,ageBot)))) %>%
  select(char, age, lake)
```

Execute the method for building a composite biomass burning record used in Kelly et al. 2013 (doi: 10.1073/pnas.1305069110). The method estimates the parameters of a zero-inflated log-normal (ZIL) distribution of pooled charcoal counts in continuous moving windows of a user-defined width. In this analysis, 5-year and 50-year window widths are used (2.5 and 25 half-kernel widths, respectively).

This portion of the analysis is relegated to a separate R script for clarity of the workflow. Window widths and other parameters can be manipulated within the 'analysis_composite.r' script.

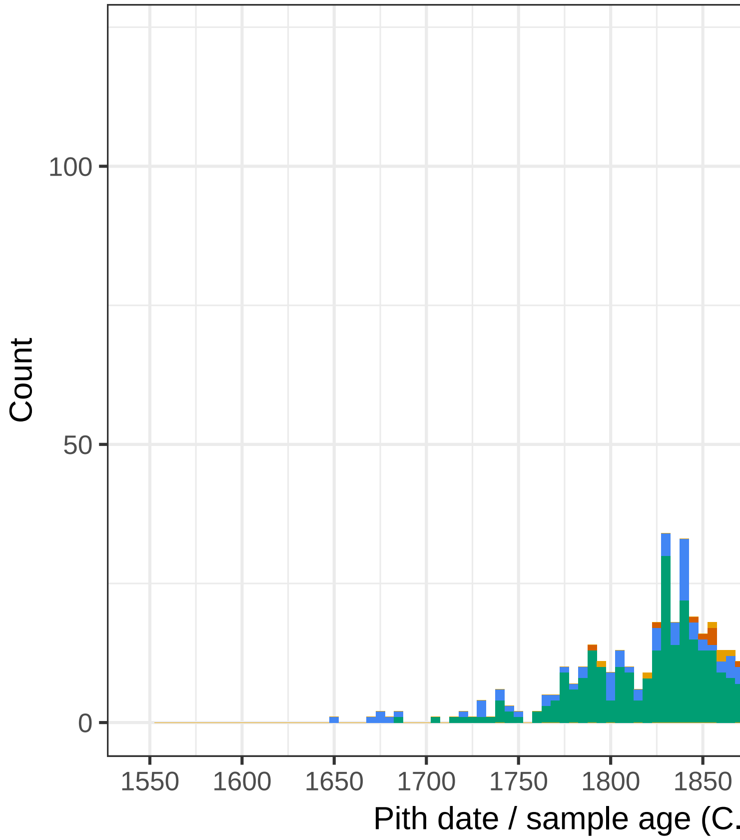
```
source(file = file.path("analysis_composite.R"))
```



Plot the result, a composite record of biomass burning.

Summarize tree demography data.

```
tree.df %>%
  mutate(ageCE = plyr::round_any(pith,5)) %>%
  group_by(sp) %>%
  summarise(count = n())
```



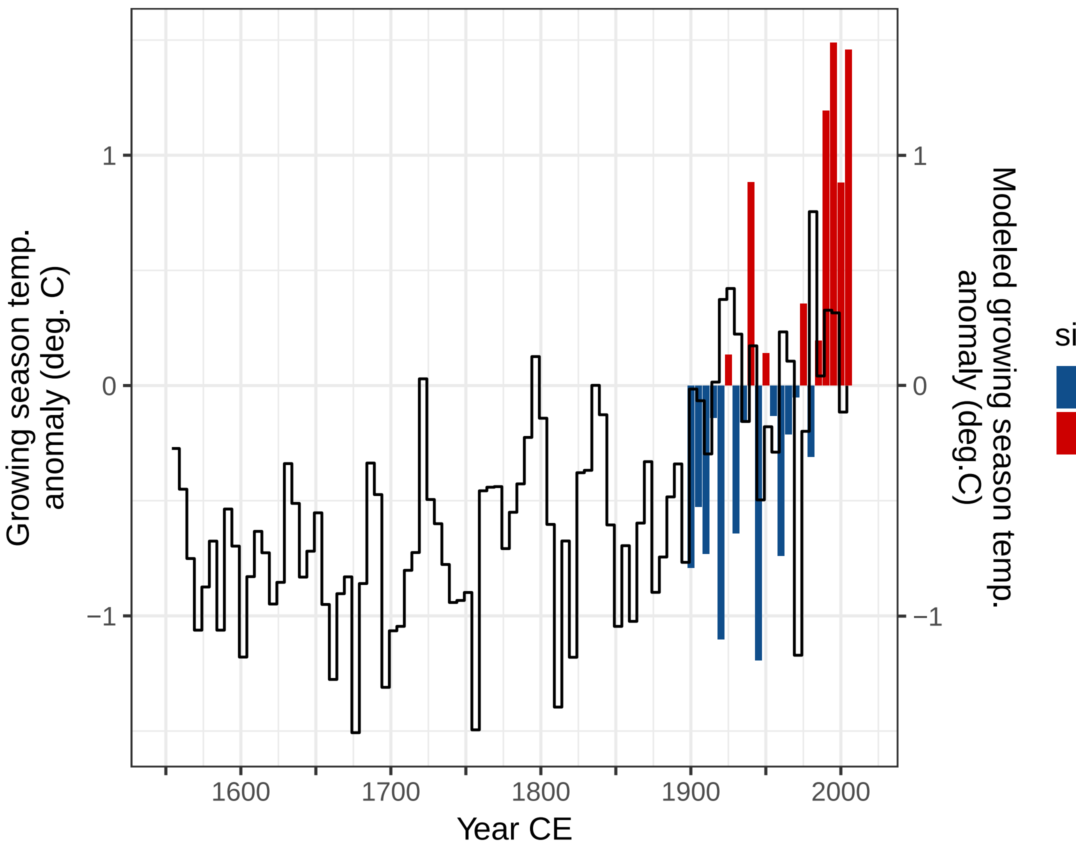
Plot the result, a time series of estimated pith dates by species.

Compare proxies

Standardize and bin data to allow for direct comparison.

```
# Modify tree dataframe
tree.cor.df <- tree.df %>%
  mutate(ageCE = plyr::round_any(pith,5)) %>%
  group_by(sp, ageCE) %>%
  summarise(count = n()) %>%
  tidyr::spread(key = sp, value = count) %>%
  mutate(bepa = ifelse(is.na(bepa),0,bepa),
         lala = ifelse(is.na(lala),0,lala),
         pigl = ifelse(is.na(pigl),0,pigl),
         pima = ifelse(is.na(pima),0,pima),
         potr = ifelse(is.na(potr),0,potr)) %>%
  mutate(tree.count = rowSums(.[c('bepa','lala','pigl','pima','potr')]))

# Use GOA mean from 1550-2010 to standardize both climate datasets.
goa.1900 <- goa.df %>%
  filter(yearCE >= 1900)
goa.mean <- mean(goa.1900$temp)

# Bin data into universal 5-year means
goa.df$yearBins <- cut(goa.df$yearCE,include.lowest = T,right = F,
                       breaks = seq(min(goa.df$yearCE),max(goa.df$yearCE),5),
                       labels = seq(min(goa.df$yearCE),2005,5))
```

```
goa.binned <- goa.df %>%
  group_by(yearBins) %>%
  summarise(bin.temp = mean(temp)) %>%
  mutate(zscore = (bin.temp - goa.mean)) %>%
  mutate(sign = ifelse(zscore >= 0,'positive','negative'))


cru.gs.df$yearBins <- cut(cru.gs.df$yearCE,include.lowest = T,right = F,
                          breaks = seq(1900,2010,5),
                          labels = seq(1900,2005,5))
cru.binned <- cru.gs.df %>%
  group_by(yearBins,variable) %>%
  summarise(bin.mean = mean(mean)) %>%
  group_by(variable) %>%
  mutate(zscore = (bin.mean - mean(bin.mean))) %>%
  mutate(sign = ifelse(zscore >= 0,'positive','negative'))

# Created combined dataframe of standardized, binned proxies for 1550-1895
goa.1550 <- goa.binned %>%
  mutate(ageCE = as.numeric(as.character(.$yearBins))) %>%
  select(-sign,-yearBins) %>%
  filter(ageCE < 1900)

combined.1550_1895.df <- composite.df %>%
  select(ageCE, lowMean, highMean) %>%
  inner_join(., pctBurned, by = 'ageCE') %>%
  inner_join(., goa.1550, by = 'ageCE') %>%
  left_join(., tree.cor.df, by = 'ageCE') %>%
  select(lowMean, highMean, win.pct, bin.temp, tree.count)

# Created combined dataframe of standardized, binned proxies for 1900-2005
goa.1900 <- goa.binned %>%
  mutate(ageCE = as.numeric(as.character(.$yearBins))) %>%
  select(-sign,-yearBins)

cru.combined <- cru.binned %>%
  mutate(ageCE = as.numeric(as.character(yearBins))) %>%
  select(ageCE,variable,bin.mean) %>%
  spread(key = variable, value = bin.mean)

combined.1900_2010.df <- composite.df %>%
  inner_join(., pctBurned, by = 'ageCE') %>%
  inner_join(., cru.combined, by = 'ageCE') %>%
  inner_join(., goa.1900, by = 'ageCE') %>%
  left_join(., tree.cor.df, by = 'ageCE') %>%
  select(lowMean, highMean, win.pct, cru.precip = precip, cru.temp = temp, goa.temp = bin.temp, tree.co
```



Plot standardized climate data.

Calculate Spearman rank correlation among fire, tree, and climate records.

```
corr.1550_1895.df <- rcorr(as.matrix(combined.1550_1895.df), type="pearson")
corr.1900_2010.df <- rcorr(as.matrix(combined.1900_2010.df), type="pearson")
```

Correlation coefficients and p-values for period 1550-1895 CE.

Table 5: R

	lowMean	highMean	win.pct	bin.temp	tree.count
lowMean	1.0000000	0.4281591	-0.0953683	0.0877348	0.5825862
highMean	0.4281591	1.0000000	-0.1060908	0.0639380	0.1618889
win.pct	-0.0953683	-0.1060908	1.0000000	0.1504365	0.3575777
bin.temp	0.0877348	0.0639380	0.1504365	1.0000000	0.2877709
tree.count	0.5825862	0.1618889	0.3575777	0.2877709	1.0000000

Table 6: P-Value

	lowMean	highMean	win.pct	bin.temp	tree.count
lowMean	NA	0.0002177	0.4462325	0.4701598	0.0000517
highMean	0.0002177	NA	0.3965416	0.5989920	0.3057057
win.pct	0.4462325	0.3965416	NA	0.2279422	0.0200780
bin.temp	0.4701598	0.5989920	0.2279422	NA	0.0646033
tree.count	0.0000517	0.3057057	0.0200780	0.0646033	NA

Correlation coefficients and p-values for period 1900-2010 CE.

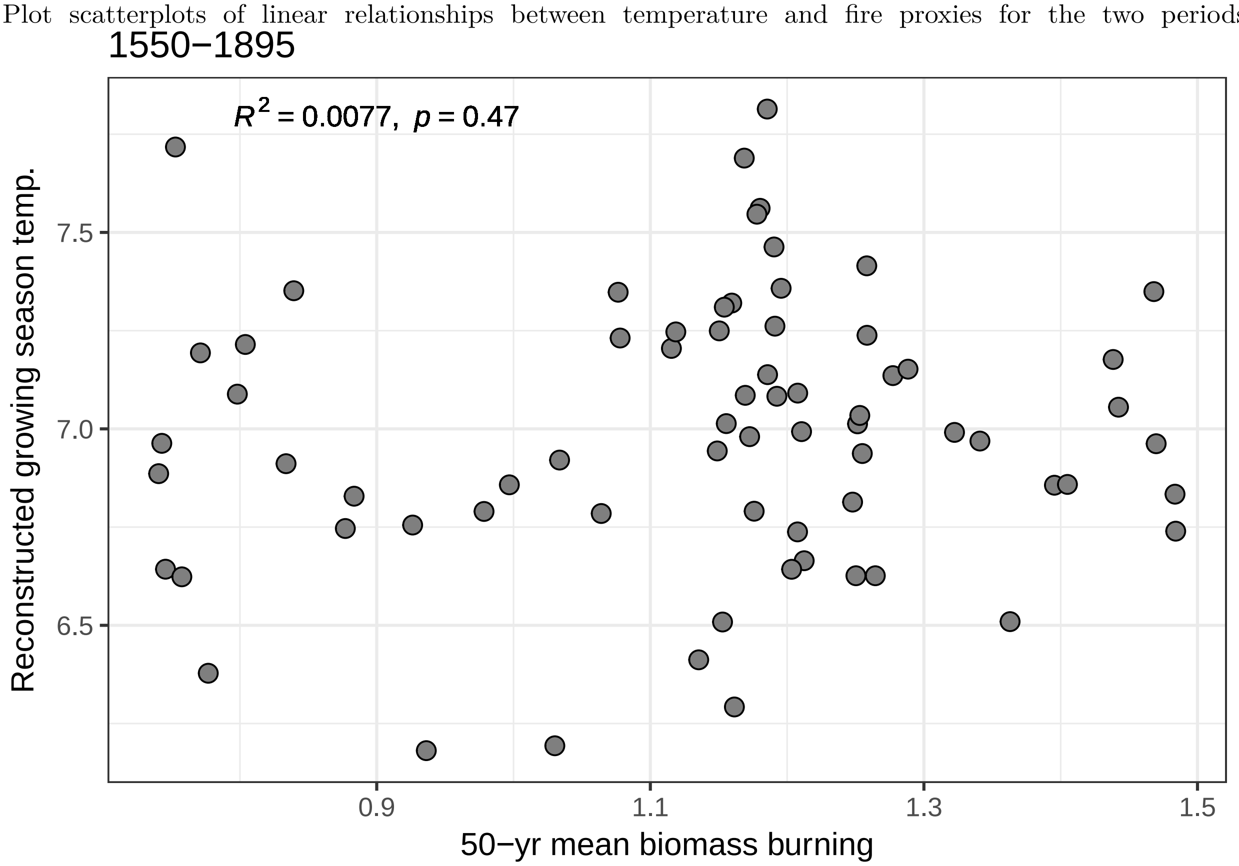
Table 7: R

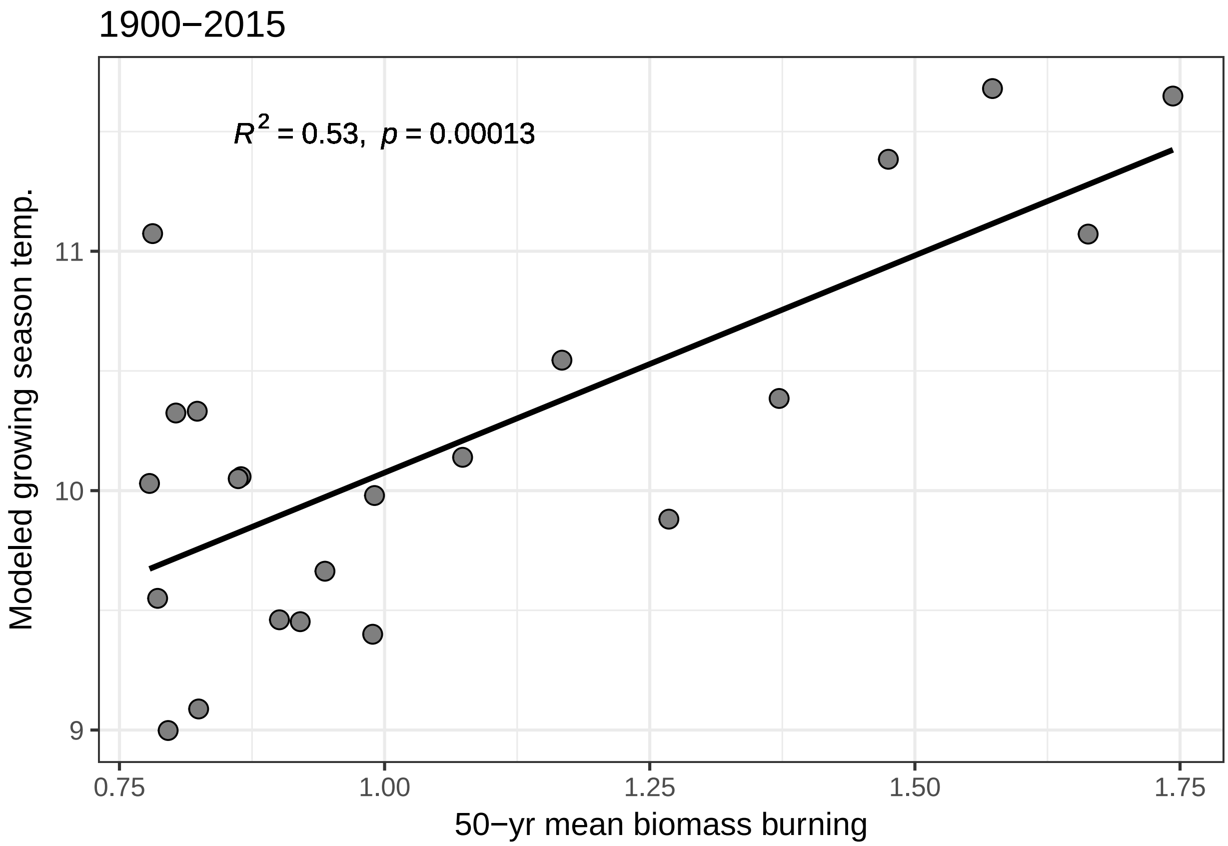
	lowMean	highMean	win.pct	cru.precip	cru.temp	goa.temp	tree.count
lowMean	1.0000000	0.6485720	0.7016179	0.0065759	0.7260299	0.1350489	-0.4466599
highMean	0.6485720	1.0000000	-0.0004314	0.1016308	0.5216335	-0.0203060	0.1628251
win.pct	0.7016179	-0.0004314	1.0000000	0.3219985	0.4962589	0.1905874	0.1032472
cru.precip	0.0065759	0.1016308	0.3219985	1.0000000	0.0945258	0.5103515	0.4681770
cru.temp	0.7260299	0.5216335	0.4962589	0.0945258	1.0000000	0.1377940	-0.2031893
goa.temp	0.1350489	-0.0203060	0.1905874	0.5103515	0.1377940	1.0000000	-0.2416750
tree.count	-0.4466599	0.1628251	0.1032472	0.4681770	-0.2031893	-0.2416750	1.0000000

Table 8: P-Value

	lowMean	highMean	win.pct	cru.precip	cru.temp	goa.temp	tree.count
lowMean	NA	0.0010949	0.0008145	0.9768298	0.0001307	0.5490288	0.0552209
highMean	0.0010949	NA	0.9986016	0.6526860	0.0127801	0.9285311	0.5053973
win.pct	0.0008145	0.9986016	NA	0.1788163	0.0306865	0.4344750	0.6835042
cru.precip	0.9768298	0.6526860	0.1788163	NA	0.6756368	0.0152327	0.0432170
cru.temp	0.0001307	0.0127801	0.0306865	0.6756368	NA	0.5408600	0.4041067
goa.temp	0.5490288	0.9285311	0.4344750	0.0152327	0.5408600	NA	0.3188622
tree.count	0.0552209	0.5053973	0.6835042	0.0432170	0.4041067	0.3188622	NA

Plot scatterplots of linear relationships between temperature and fire proxies for the two periods.




1900−2015
R2 = 0.53, p = 0.00013
Modeled growing season temp.
9
10
11
0.75
1.00
1.25
1.50
1.75
50−yr mean biomass burning