Testing_temporal_subsampling

Alyssa Willson

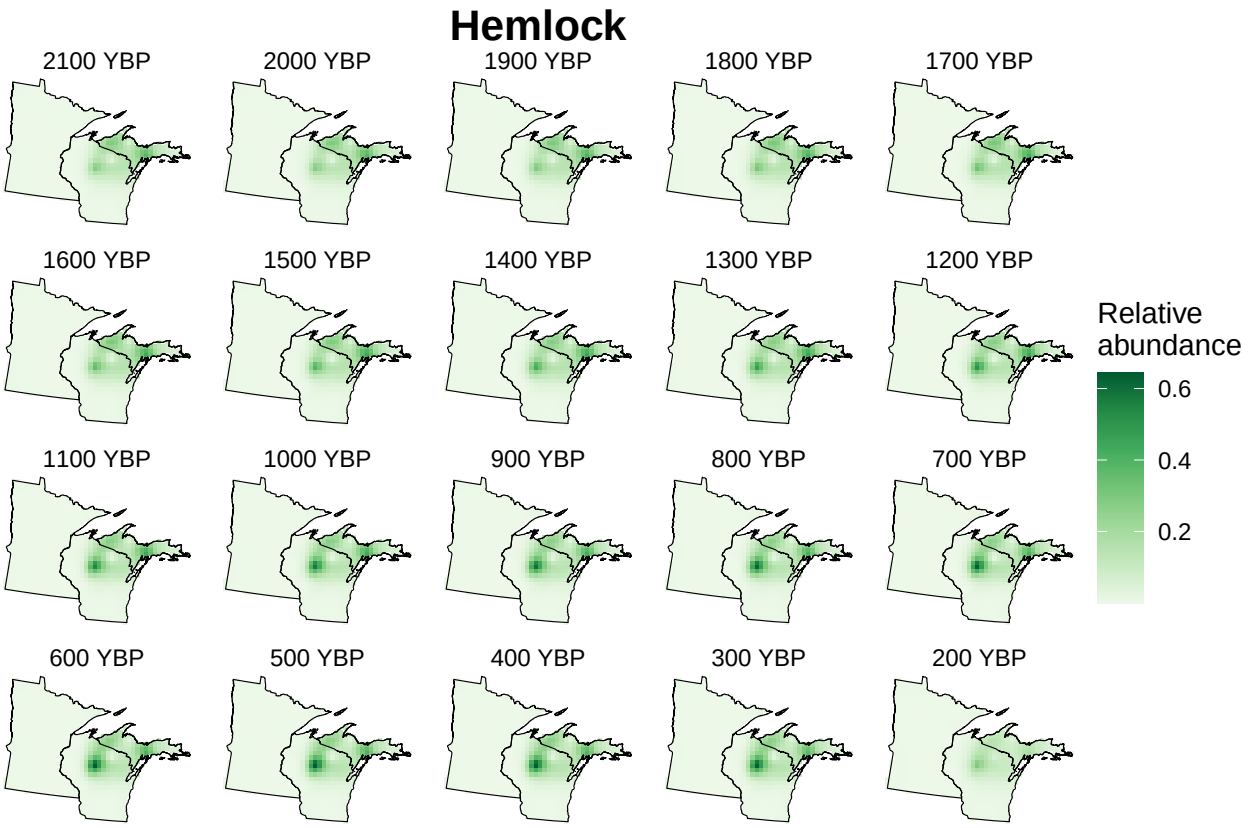
2024-03-29

Purpose

PalEON's previous reconstructions of relative abundance for the last 2,000 years of the pre-Industrial Holocene leveraged the idea that physical processes related to lake sediment settling and ecological processes related to the long lifespans of trees contribute to long memory of relative abundances by borrowing strength over time. This leads to temporally smoothed estimates of relative abundances. Here, we attempt to quantify the autocorrelation of relative abundances so that we can sample a subset of time steps that are relatively independent from one another.

Understanding temporal structure

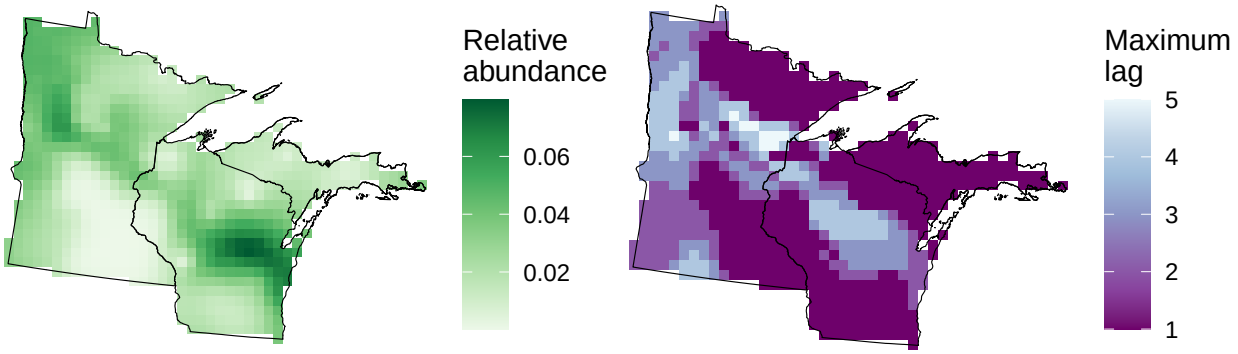
Our data are spatiotemporally smoothed relative abundances of 12 taxa for 2,000 years at 100 year time steps. There are marked changes in relative abundance of some taxa at some locations over this temporal domain, as seen below.



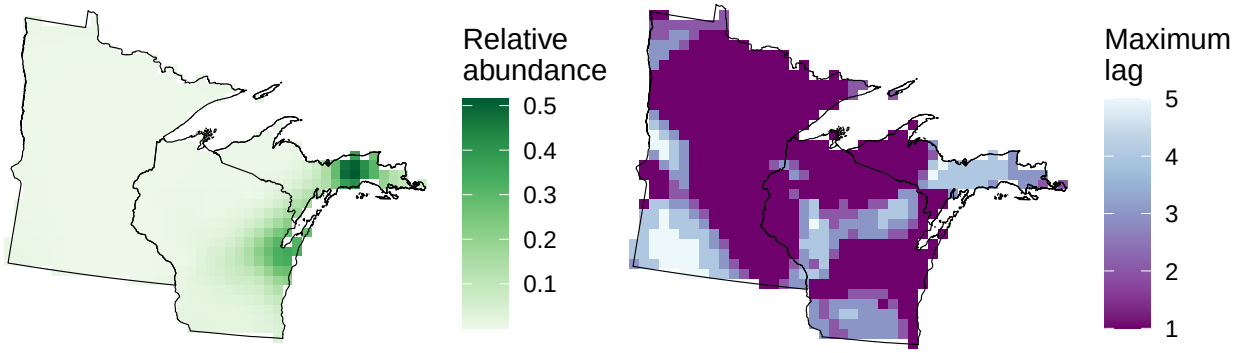
To investigate this temporal structure, I want to start by calculating autocorrelation functions for each taxon at all spatial locations (multivariate autocorrelation).

Here are some plots of the maximum lag that is considered significant at a 95% confidence level according to the autocorrelation function analysis. The left panel in each figure is the relative abundance of the taxon and the right panel is the maximum significant lag. It is important to consider where lags are significant because we expect there to be little change over time in regions where a taxon is never present in great numbers.

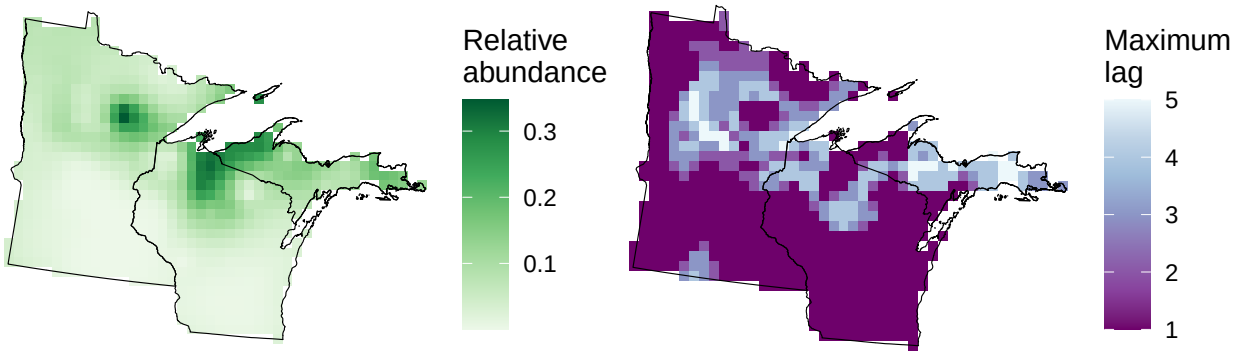
Ash



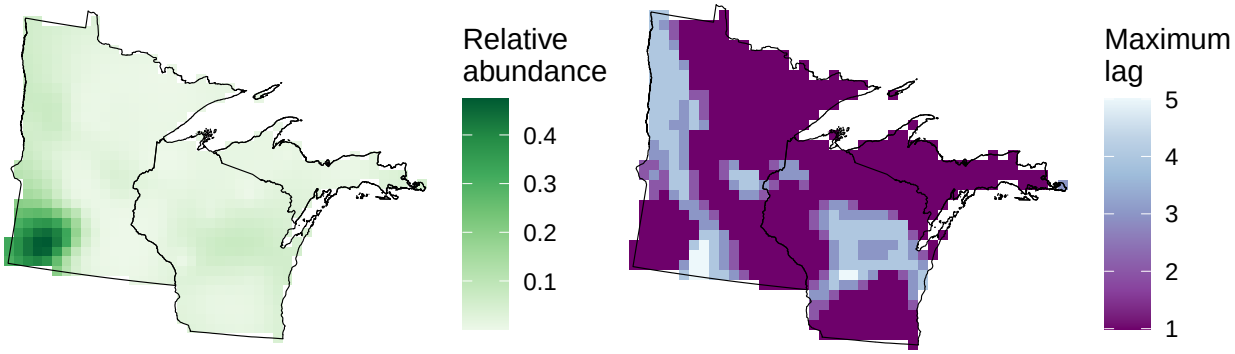
Beech



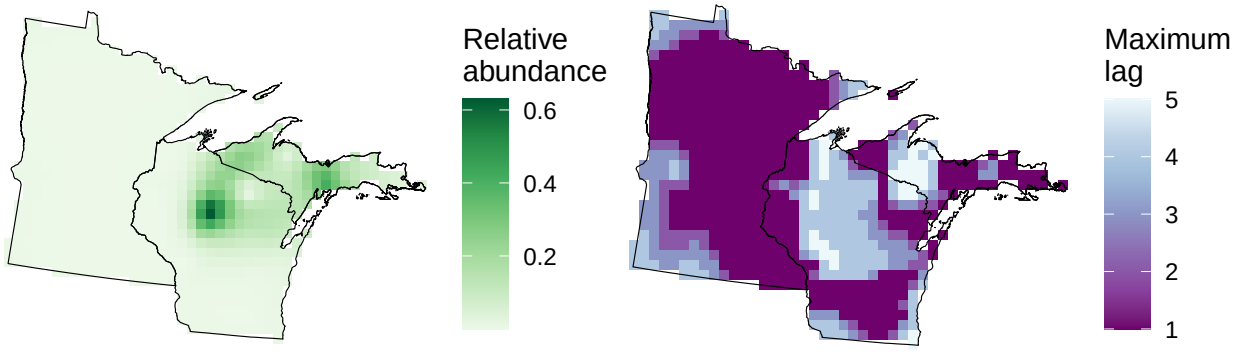
Birch



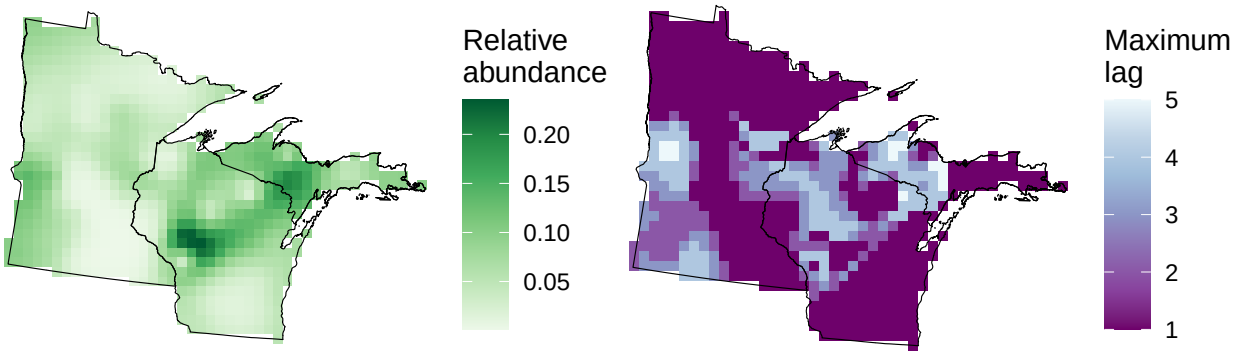
Elm



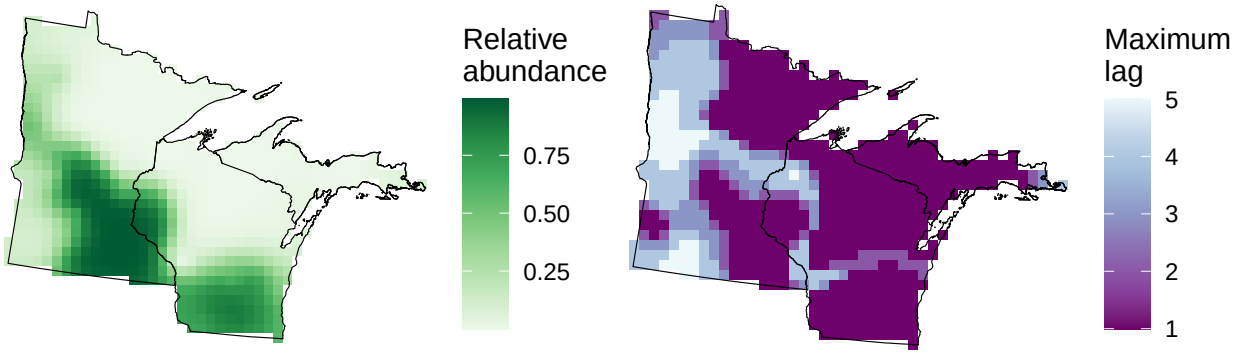
Hemlock



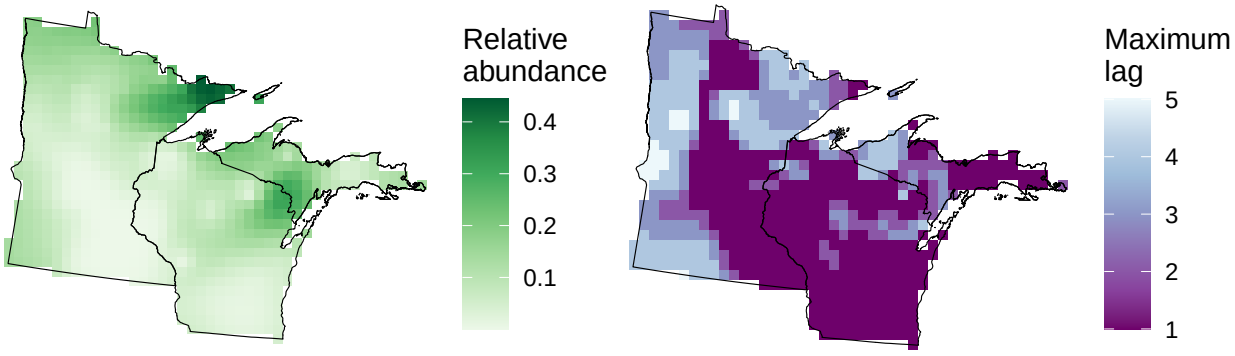
Maple



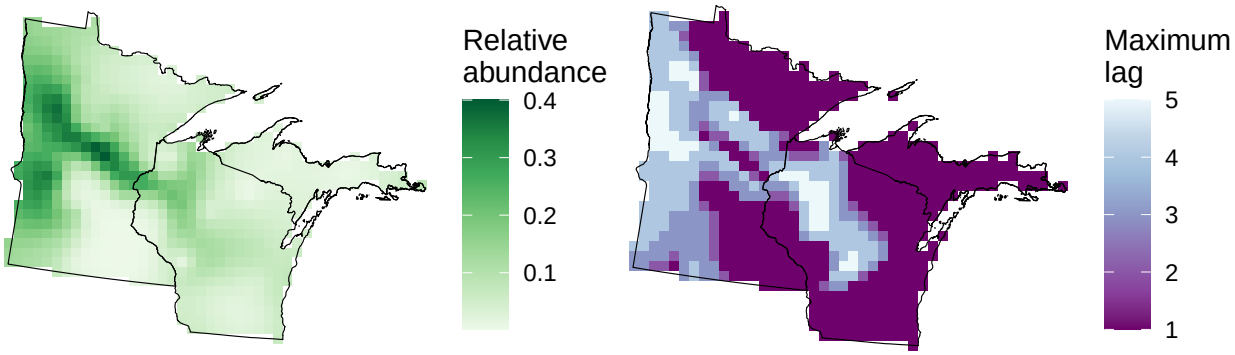
Oak



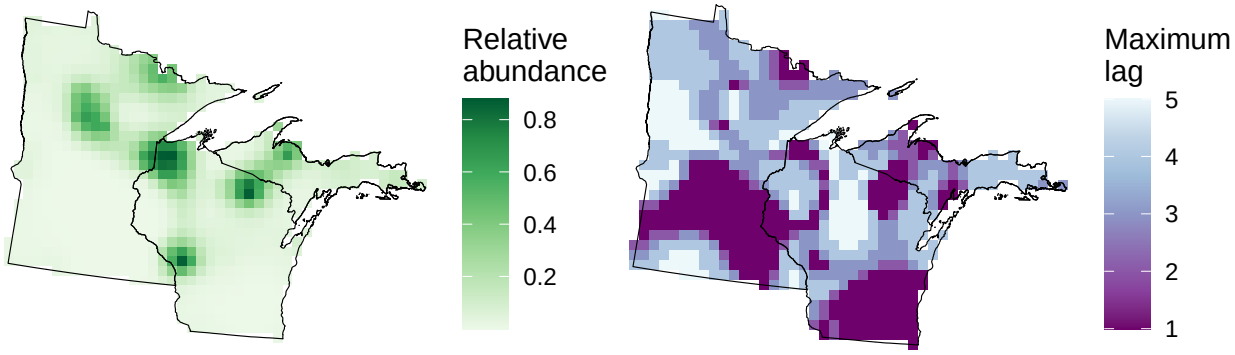
Other conifer



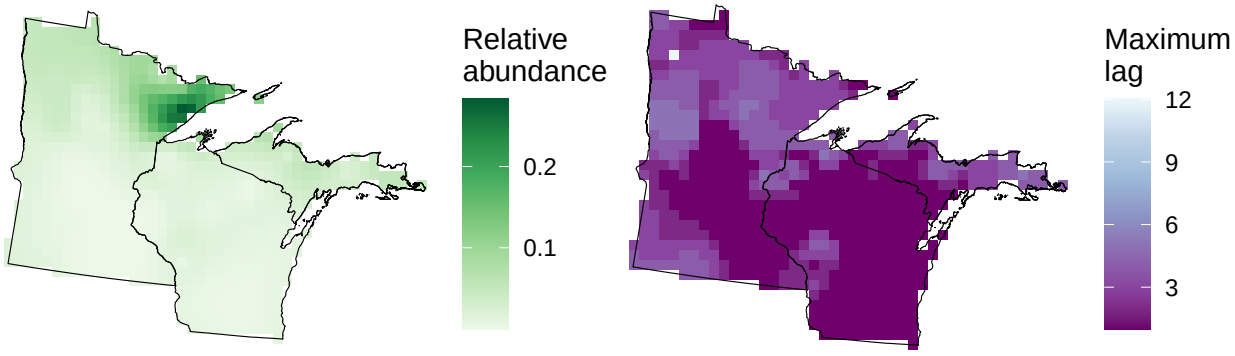
Other hardwood

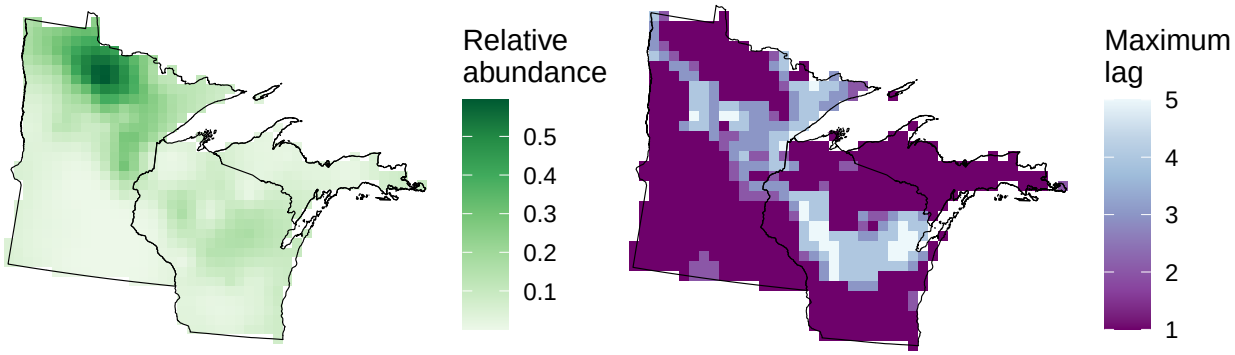


Pine



Spruce



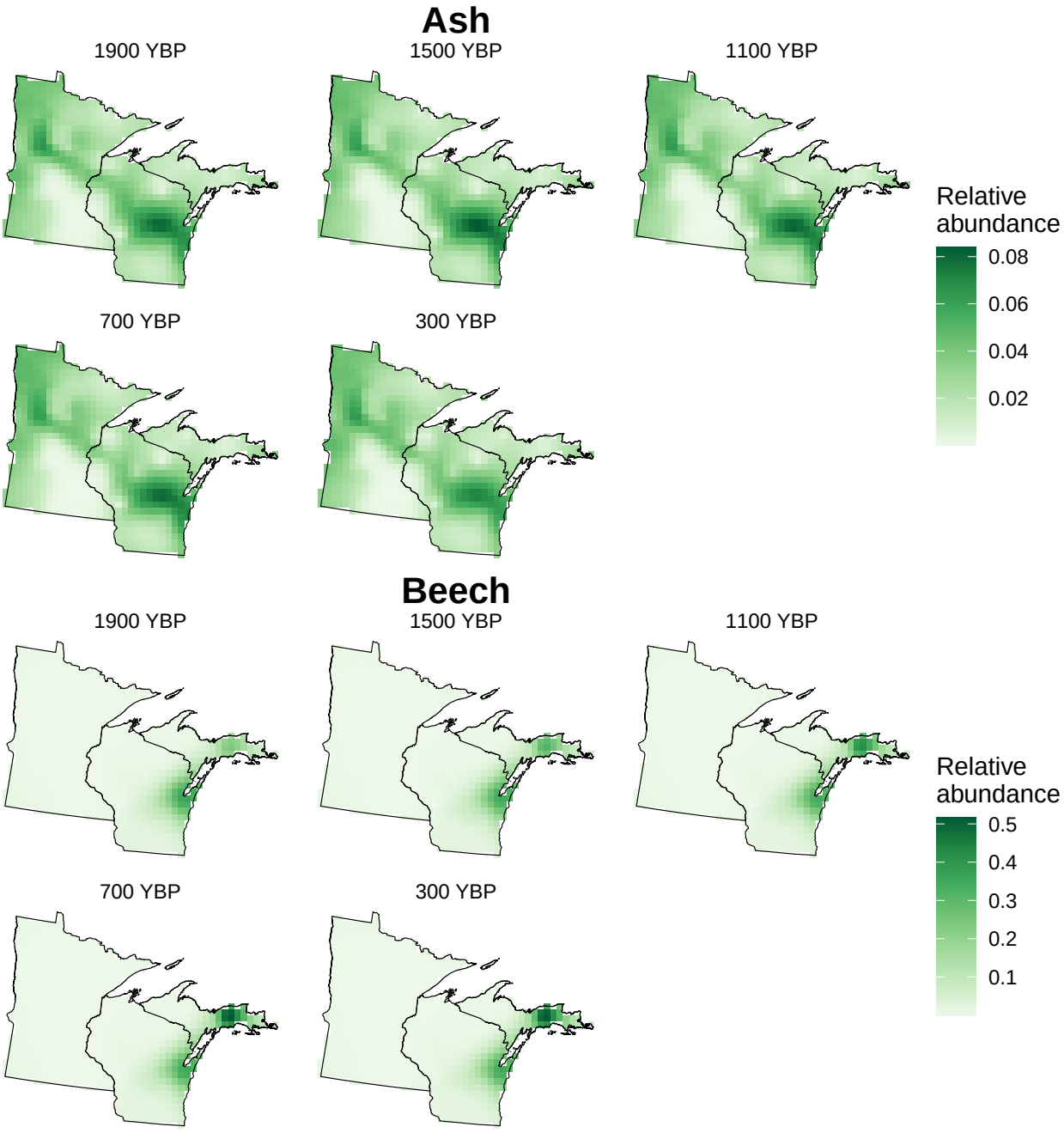


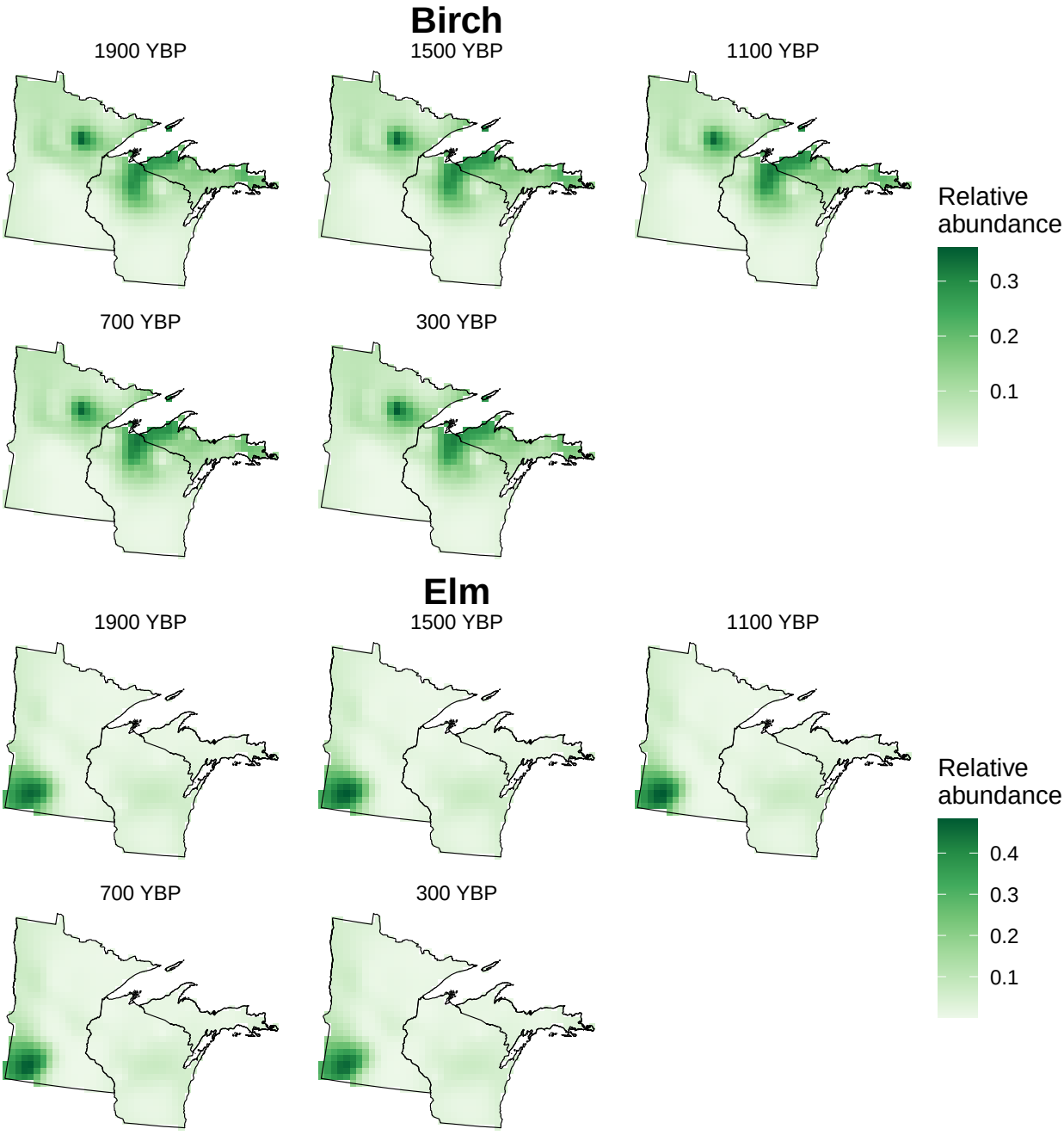
Sub-sampling in time

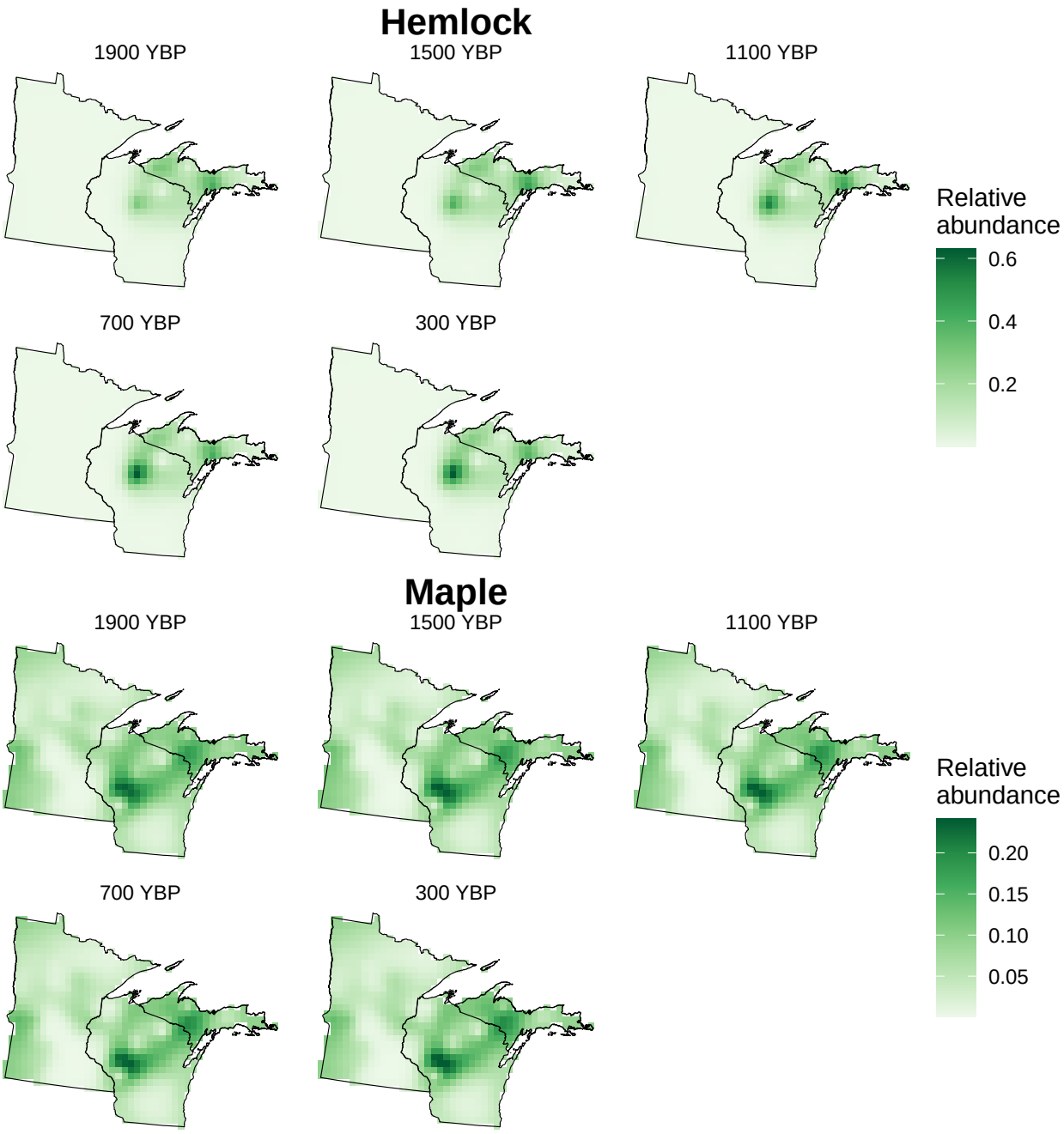
To me, this seems to indicate that we should sample every 4-5 time steps. This is what those options would look like. Remember that we will likely not have the first couple of time steps because we don't have climate reconstructions for those time steps and we also don't want to inculde the final time step, 200 YBP because we can see the effects of industrialization/logging.

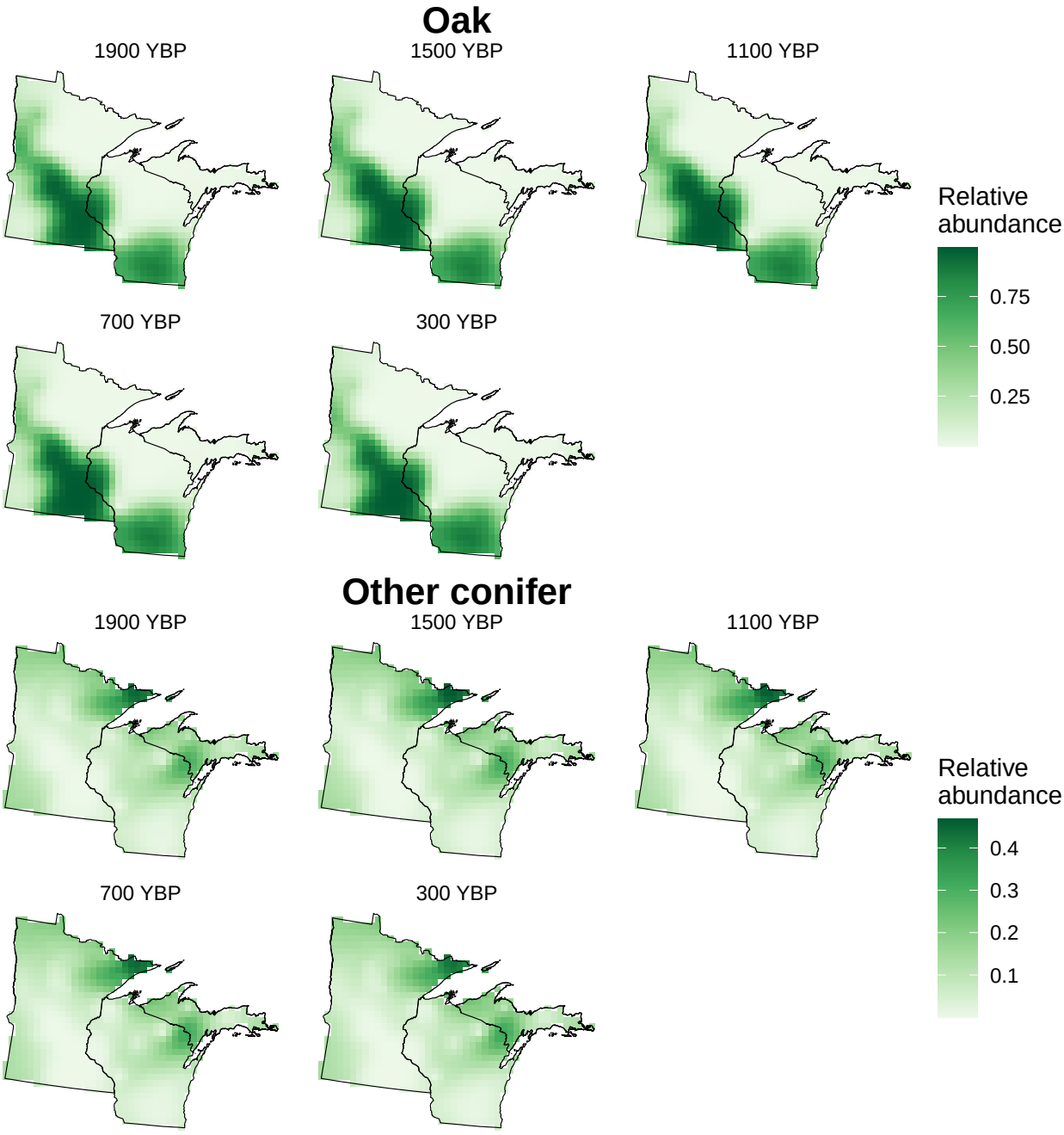
Sampling every 4 time steps

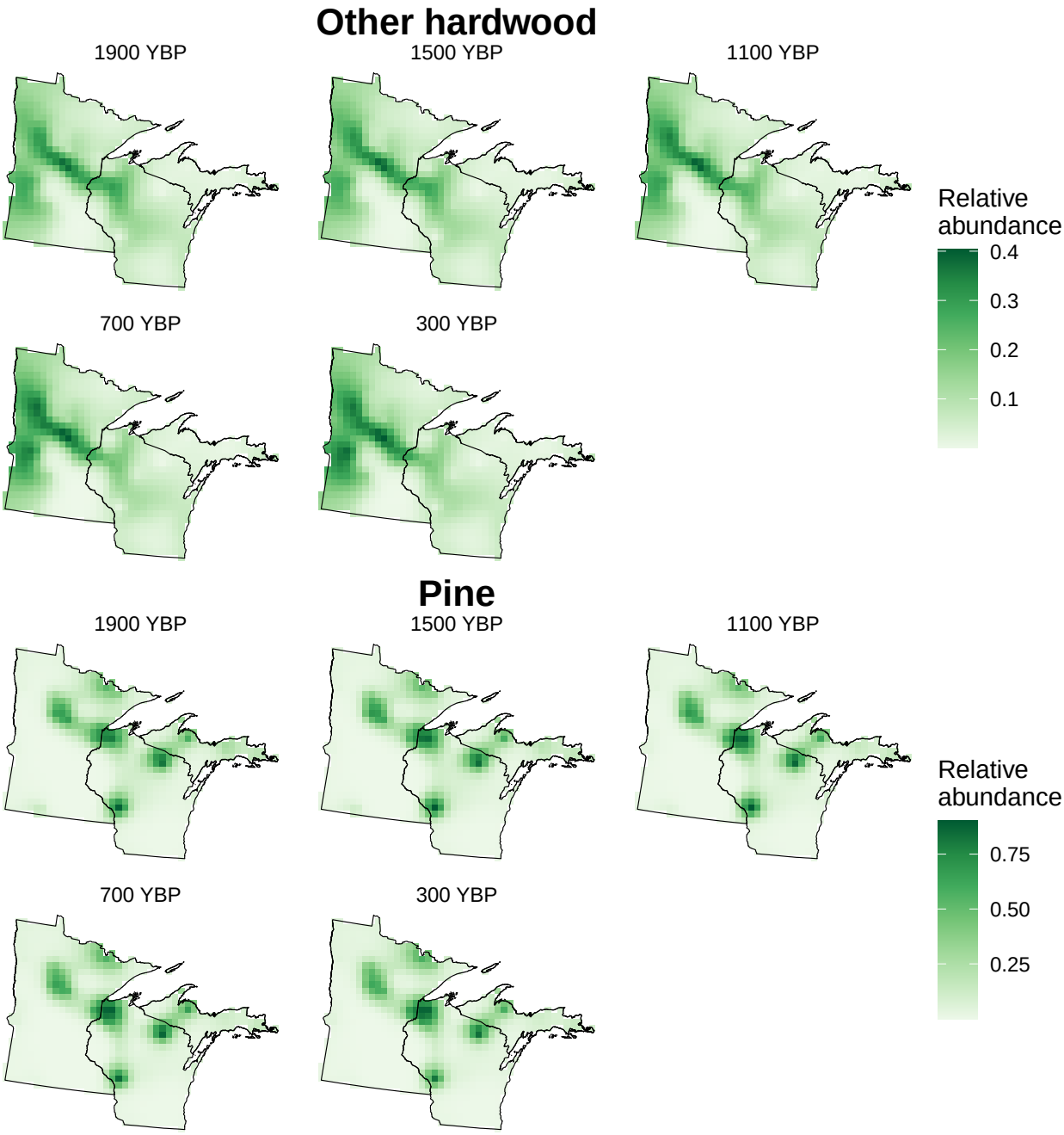
```
# Sample times 4 apart
keep_times <- seq(from = 19, to = 3, by = -4)
# Order of facets
keep_times_order <- c('1900 YBP', '1500 YBP', '1100 YBP',
                      '700 YBP', '300 YBP')

# Plot
ash_melt |>
  dplyr::filter(time %in% keep_times) |>
  dplyr::mutate(time = as.character(time),
                time = dplyr::if_else(time == '19', '1900 YBP', time),
                time = dplyr::if_else(time == '15', '1500 YBP', time),
                time = dplyr::if_else(time == '11', '1100 YBP', time),
                time = dplyr::if_else(time == '7', '700 YBP', time),
                time = dplyr::if_else(time == '3', '300 YBP', time)) |>
  ggplot2::ggplot() +
  ggplot2::geom_raster(ggplot2::aes(x = x, y = y, fill = ash)) +
  ggplot2::geom_sf(data = states, color = 'black', fill = NA) +
  ggplot2::facet_wrap(~factor(time, levels = keep_times_order)) +
  ggplot2::scale_fill_distiller(palette = 'Greens', direction = 1,
                                na.value = 'white', name = 'Relative\nabundance') +
  ggplot2::ggtitle('Ash') +
  ggplot2::theme_void() +
  ggplot2::theme(plot.title = ggplot2::element_text(size = 16, face = 'bold', hjust = 0.5))
```

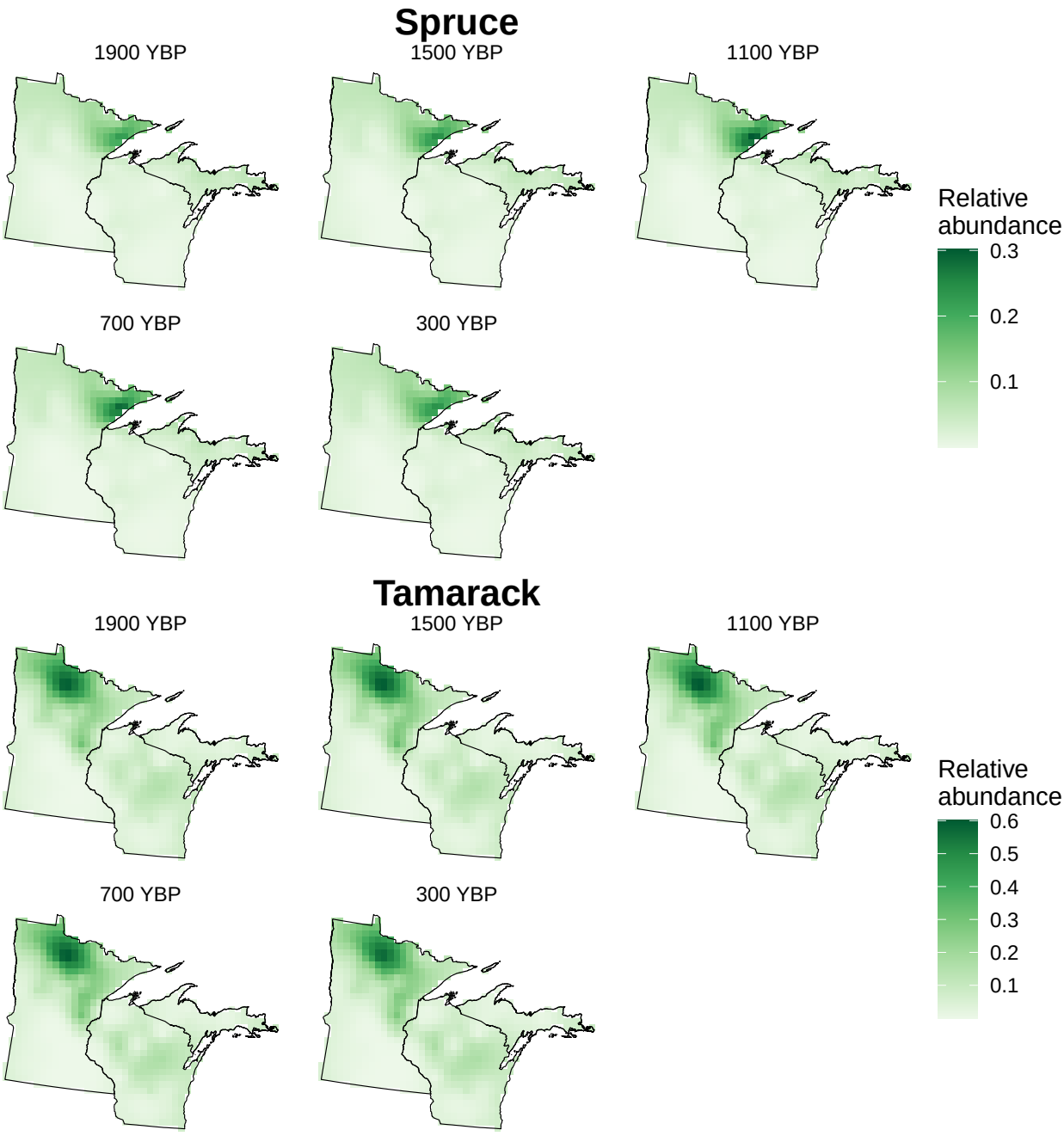

Ash
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.08
0.06
0.04
0.02
Beech
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.5
0.4
0.3
0.2
0.1


Birch
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.3
0.2
0.1
Elm
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.4
0.3
0.2
0.1


Hemlock
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.6
0.4
0.2
Maple
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.20
0.15
0.10
0.05


Oak
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.75
0.50
0.25
Other conifer
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.4
0.3
0.2
0.1

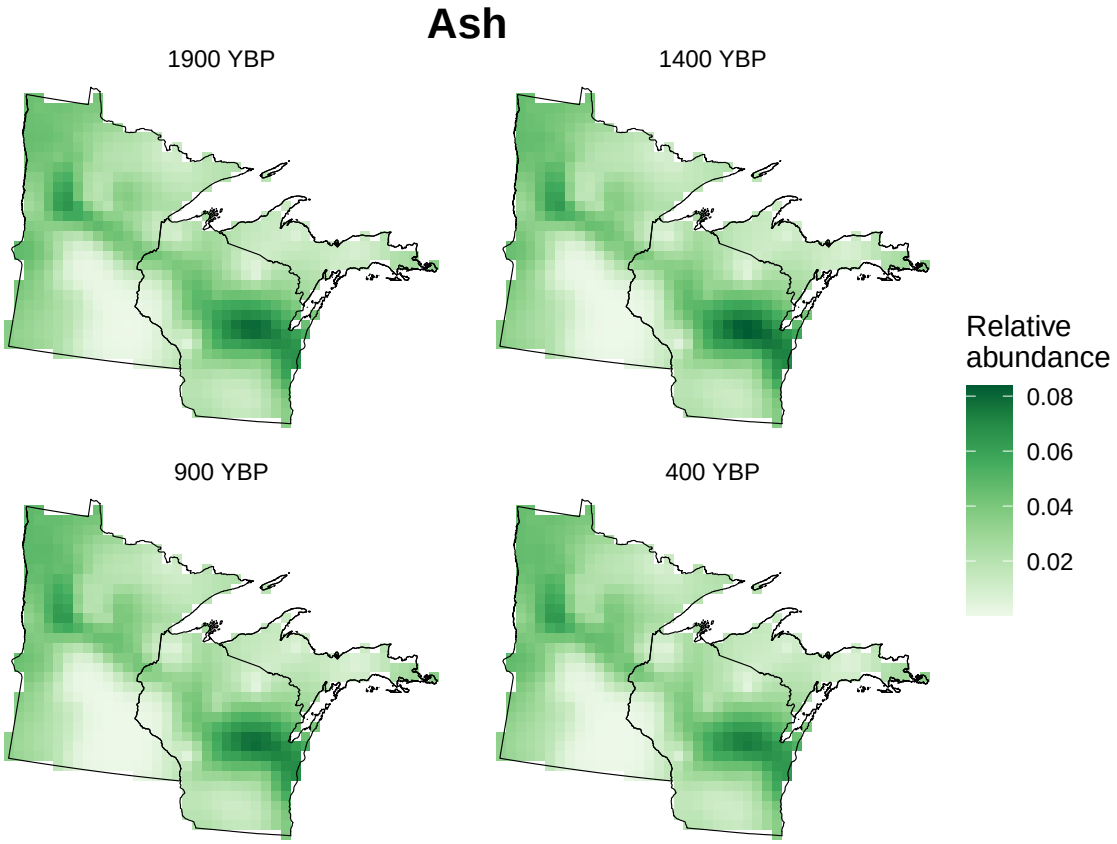

Other hardwood
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.4
0.3
0.2
0.1
Pine
1900 YBP
1500 YBP
1100 YBP
700 YBP
300 YBP
Relative abundance
0.75
0.50
0.25

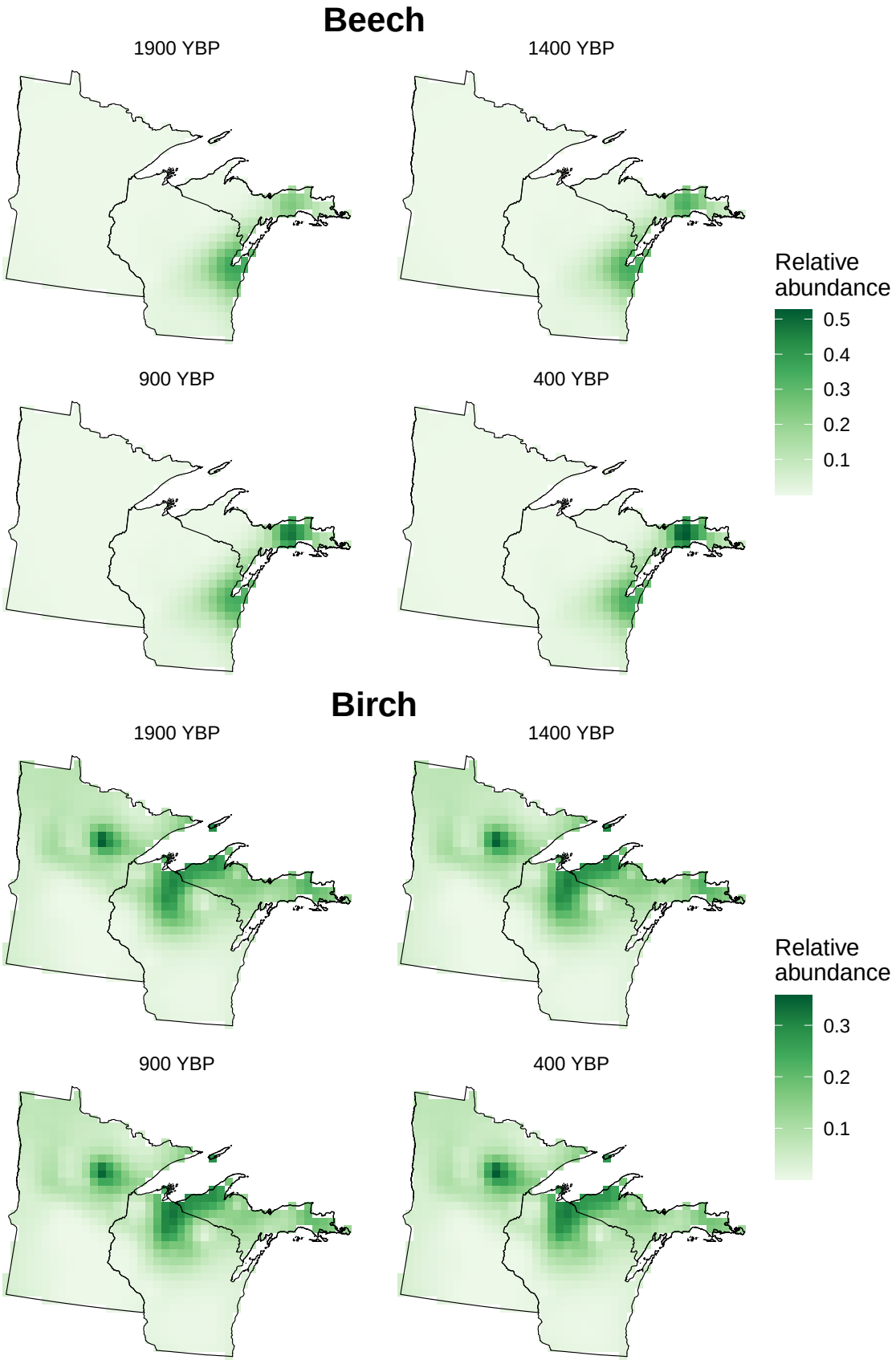


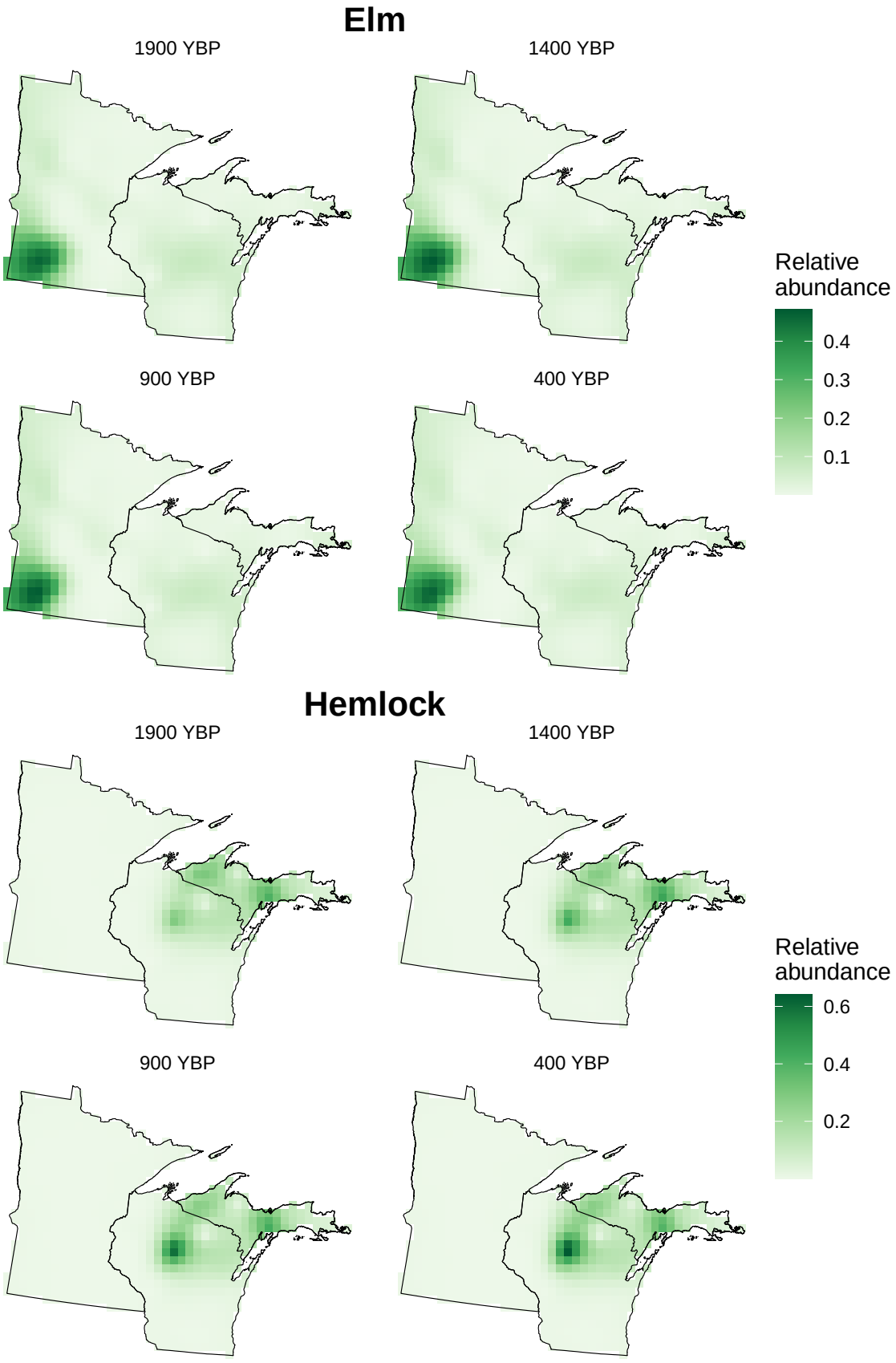
Sampling every 5 time steps

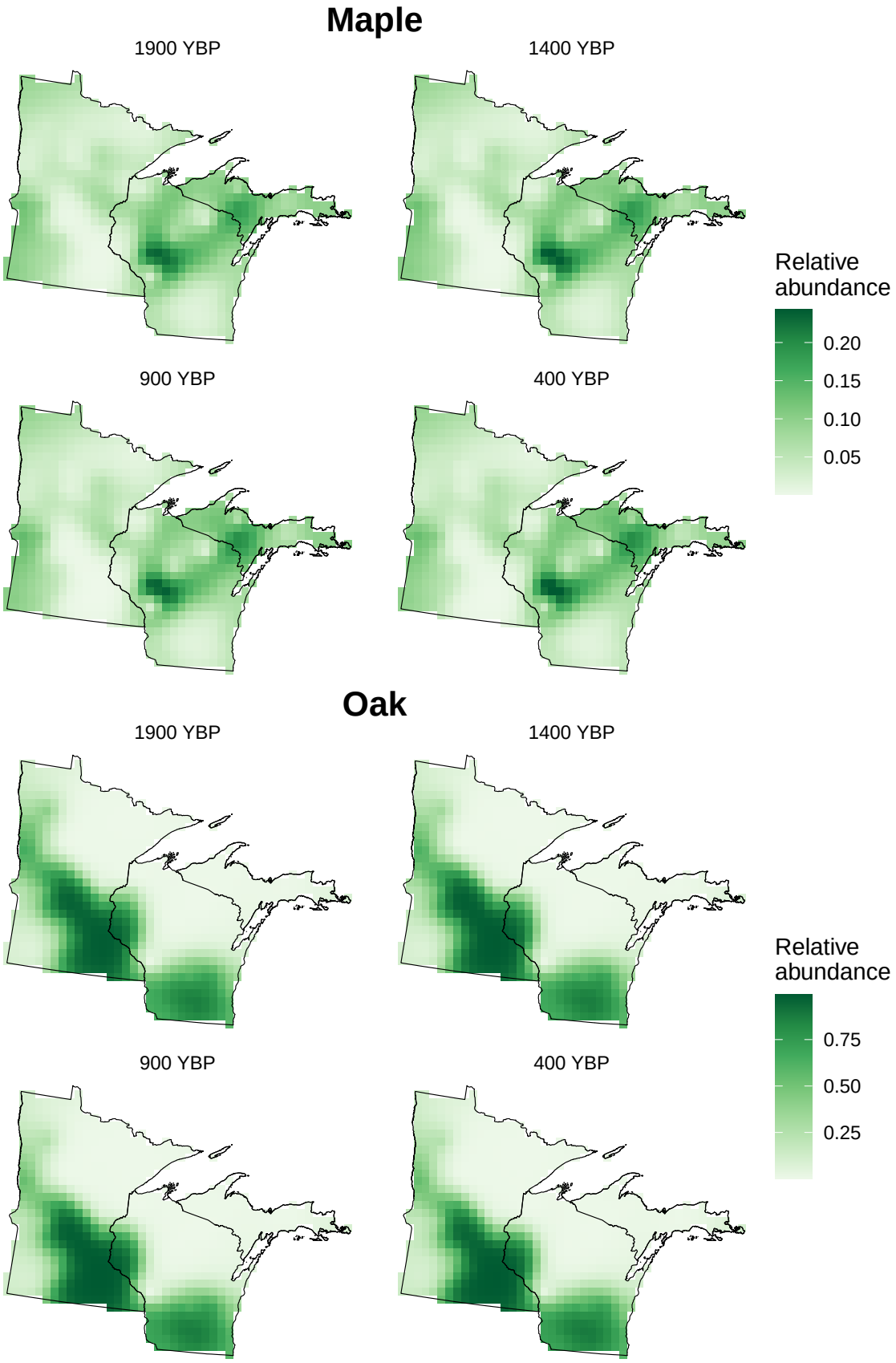
```
keep_times <- seq(from = 19, to = 3, by = -5)
keep_times_order <- c('1900 YBP', '1400 YBP',
                      '900 YBP', '400 YBP')

ash_melt |>
  dplyr::filter(time %in% keep_times) |>
  dplyr::mutate(time = as.character(time),
                time = dplyr::if_else(time == '19', '1900 YBP', time),
                time = dplyr::if_else(time == '14', '1400 YBP', time),
```

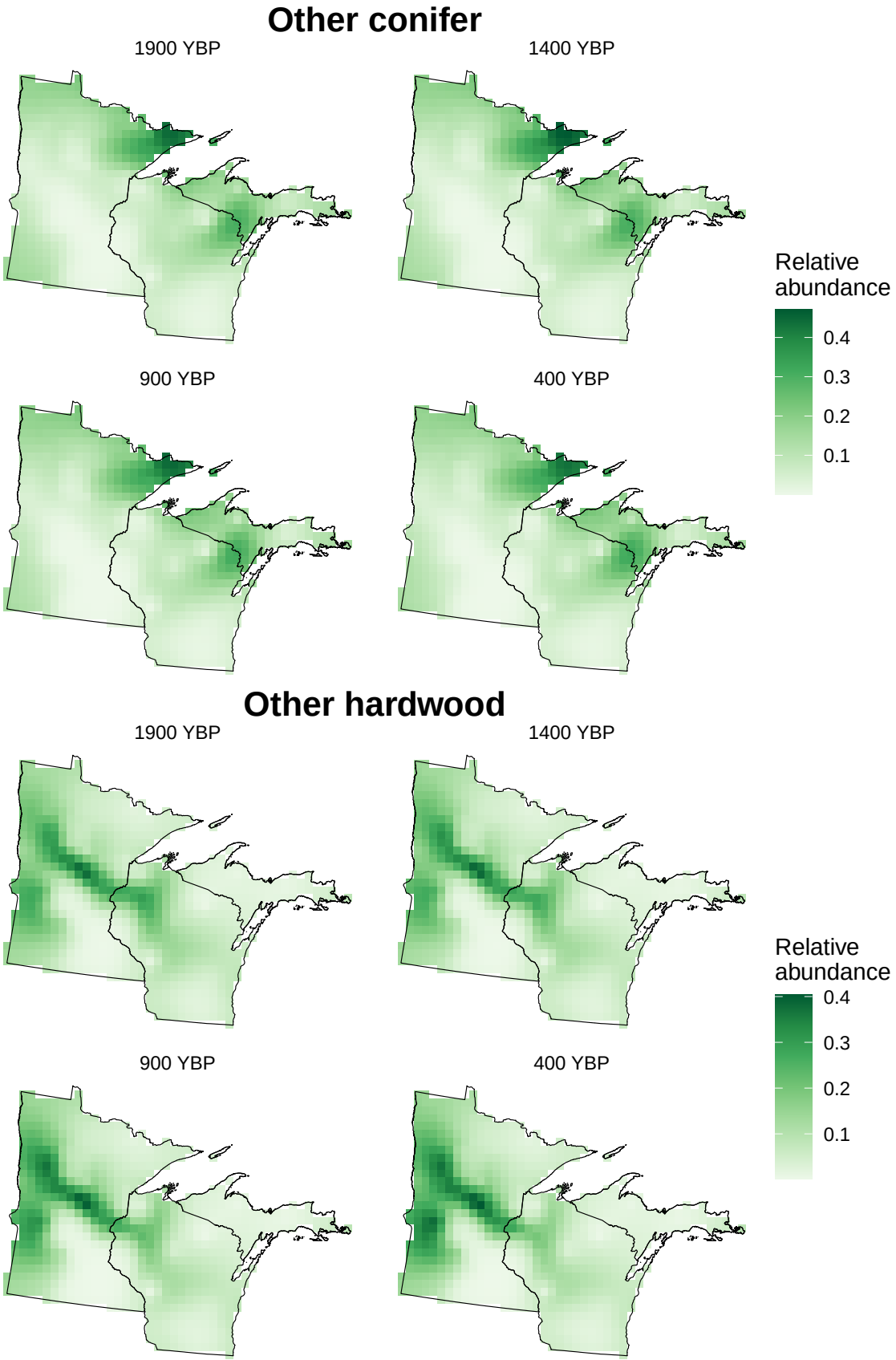
```
              time = dplyr::if_else(time == '9', '900 YBP', time),
              time = dplyr::if_else(time == '4', '400 YBP', time)) |>
ggplot2::ggplot() +
ggplot2::geom_raster(ggplot2::aes(x = x, y = y, fill = ash)) +
ggplot2::geom_sf(data = states, color = 'black', fill = NA) +
ggplot2::facet_wrap(~factor(time, levels = keep_times_order)) +
ggplot2::scale_fill_distiller(palette = 'Greens', direction = 1,
                              na.value = 'white', name = 'Relative\nabundance') +
ggplot2::ggtitle('Ash') +
ggplot2::theme_void() +
ggplot2::theme(plot.title = ggplot2::element_text(size = 16, face = 'bold', hjust = 0.5))
```

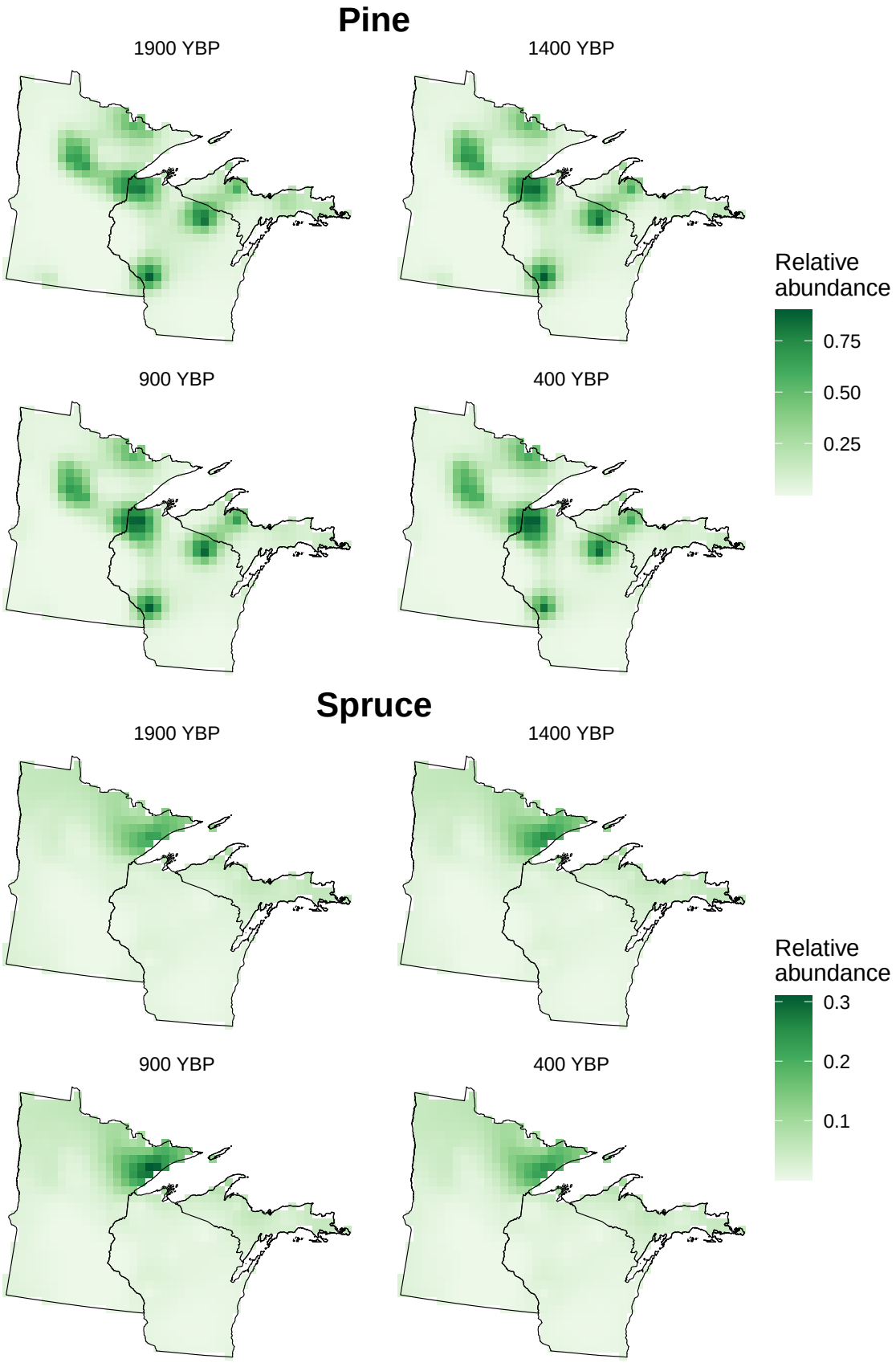


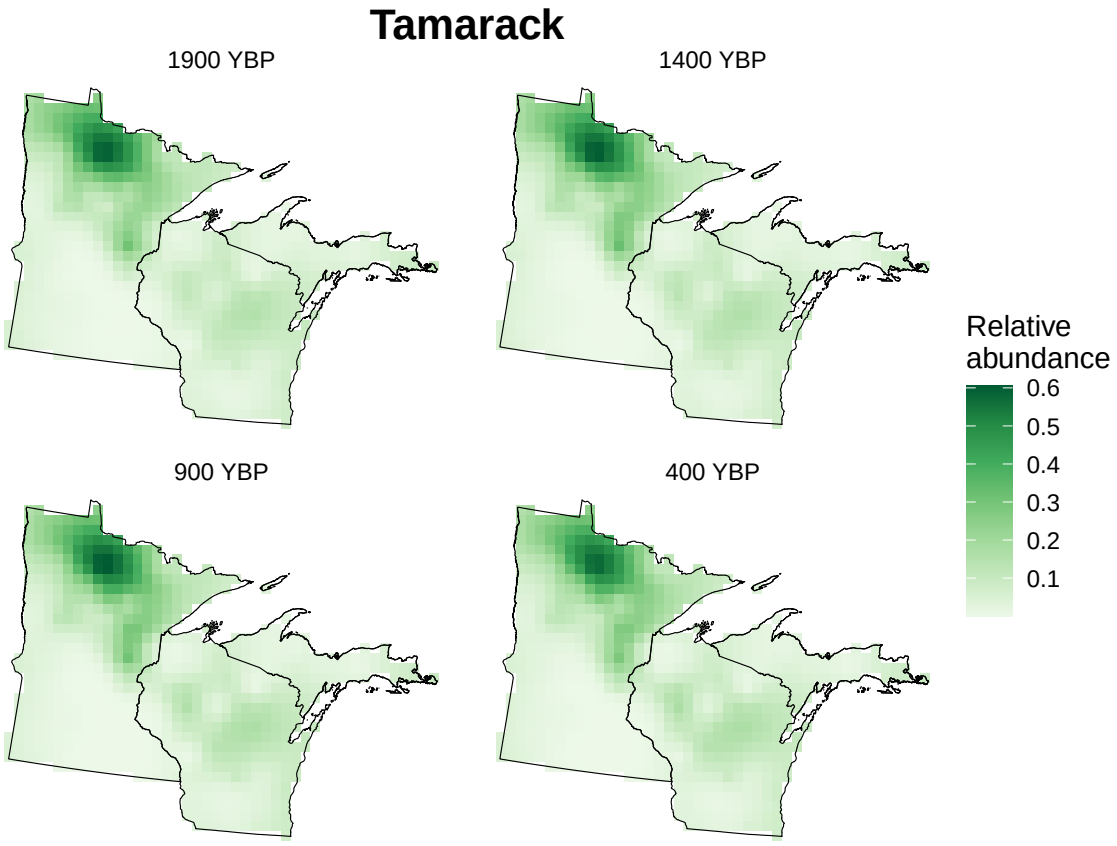

Beech
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.5
0.4
0.3
0.2
0.1
Birch
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.3
0.2
0.1


Elm
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.4
0.3
0.2
0.1
Hemlock
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.6
0.4
0.2


Maple
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.20
0.15
0.10
0.05
Oak
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.75
0.50
0.25


Other conifer
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.4
0.3
0.2
0.1
Other hardwood
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.4
0.3
0.2
0.1


Pine
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.75
0.50
0.25
Spruce
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.3
0.2
0.1


Tamarack
1900 YBP
1400 YBP
900 YBP
400 YBP
Relative abundance
0.6
0.5
0.4
0.3
0.2
0.1