DEPARTMENT OF BIOLOGICAL AND
ENVIRONMENTAL SCIENCES

HOLOCENE VEGETATION CHANGE AND ITS CORRELATES IN FENNOSCANDIA

Raphaela Infanger

Degree project for Master of Science (120 hec) with a major in Biodiversity and Systematics

BIO707, Degree project biology (Master) 60 hec

Second cycle

Semester/year: Autumn 2021 - Spring 2022

Supervisor: Christine D. Bacon, Department of Biological and Environmental Sciences

Examiner: Bengt Oxelman, Department of Biological and Environmental Sciences

(Gribble, 2021)

Table of Contents

Abstract

Climatic and anthropogenic legacy has differently impacted vegetation and its community composition. To determine the extent and origin of changes across Fennoscandia, representative regions were defined that entailed sufficient spatial and temporal data. Fossil pollen grouped into land cover classes represented Fennoscandian vegetation types, and we focused our analyses and comparisons on woodlands versus anthromes (arable lands and pastures). To determine the correlate(s) of vegetation change, we regressed vegetation types against archaeological (summed probability distributions, representing human impact) and climatic data (mean annual temperature) with Granger causality tests. Additionally, multiple change point analyses were conducted on each single dataset in order to reveal the most significant time points of change. We show that humans have been the strongest correlate to changes in vegetation, especially after the onset of farming, while climate was a strong force in the centuries following the glacial retreat. Indications on the timing of establishment and abandonment of land-use practices such as deforestation in the southern and central regions or of animal husbandry in the North have been found and/or recovered based on the discovered change points. Furthermore, we show an increase in populations already at 7000 BP for the Southeast and sings of human induced homogenisation for the midwestern and central regions. These results taken together present a long view of vegetation change in response to both climate and anthropogenic pressures. The strong advantage of interdisciplinary studies like this one is that it enables us to look at a variety of data in a connected way, putting the current state of Fennoscandian biodiversity into context, and allowing for better projections in the future.

Keywords: Fennoscandia, Holocene, human impact, climate impact, vegetation change, fossil pollen

Introduction

The role of biodiversity

Biodiversity expresses itself through a rich variation of genes, species, morphology, and functions (Cardinale et al., 2012). This variation is essential for preserving ecological processes that control the fluxes of energy, nutrients, and organic matter within an environment (Cardinale et al., 2012), and thus makes it the foundation of every healthy forest, pasture, wetland, and crop field. The more diverse a community, the more productive it is (Cardinale et al., 2012). Over the last 50 years, species of earth's flora and fauna have decreased in higher numbers than seen in the past due to demand for ecosystem services, higher human population, and higher average per capita income (Díaz et al., 2019). The current rate of species extinction is reported to be 100 times higher than the background rate and can be considered a mass extinction event (Rounsevell et al., 2020). Assessment of species extinction risks, as well as on climate change or overexploitation of nature, are mainly focused on the past 50 years, because data availability usually only exists for these years (Pereira, Navarro, & Martins, 2012). To understand the current decline in biodiversity and the impact that humans have had on terrestrial landscapes over the last ca. 12,000 years (Ellis et al., 2021), we need to determine past vegetation patterns and set them into context with climate and anthropological events.

Climate in the Holocene

Biodiversity patterns reflect in situ climatic conditions. Some taxa can shift in geographic range rapidly and community composition changes due to climate change, while other taxa lag behind in their adaption, loose access to their habitats, and eventually decrease in population size or go extinct (Nogués-Bravo et al., 2018). Several climatic shifts in the Northern Hemisphere are reported for the Holocene. A short, but significant cooling event called the "8200 yr. event" falls into the generally cool time between 9000 to 8000 BP. This is the last time after the last glacial retreat in Fennoscandia (Norway, Sweden, Finland, and the very north-western part of Russia) at about 9700 BP (Stroeven et al., 2016) when changes in ice sheet extent still drove the cooling temperature (Mayewski et al., 2004). Later climate events were caused by factors like the amount of volcanic aerosols or solar variability (Mayewski et al., 2004). Several drops in temperature that lasted for at least one century (Helama et al., 2021) and variability in humidity (Gunnarson, Borgmark, & Wastegård, 2003) have been reported during the last 10'000 years. Each of these climatic changes contributed to biodiversity and specifically to vegetation patterns.

Main climatic events

Cooling events in northern Europe, additional to the 8200 yr. event, that can be compared to events such as the Little Ice Age happened over timespans from 100 to 800 years starting at ca. 1410 BP, 3620 BP, 5190 BP and 7400 BP (Helama et al., 2021). Several considerable humidity changes that lasted from 100 to 300 years happened over the Holocene in Fennoscandia. Periods of drier conditions started at ca. 6850 BP, 4350 BP, 4050 BP, 3450 BP, 1900 BP, 1550 BP and 600 BP. While periods of wetter conditions started at ca. 5550 BP, 3050 BP, 2050 BP, 1750 BP, 1200 BP and 400 BP (Gunnarson et al., 2003).

Anthropogenic pressure

Throughout the Holocene, climate had a strong impact on vegetation, especially in the Mesolithic, before the onset of farming, when humans lived in hunter-gatherer populations (Kuosmanen et al., 2018). At this time natural vegetation was still dominant in Fennoscandia (O'Dwyer, Marquer, Trondman, & Jönsson, 2021). The onset of agriculture commenced the formation of the cultural landscape where practices included burning, hunting, species propagation, domestication, and/or cultivation (Ellis et al., 2021). Biological and abiotic forces can also create such change, but the strength and therefore the effects of the activities between natural and human forces are different (Berglund, Malmer, & Persson, 1991).

The onset of farming

The onset of farming marked the time where anthropological pressure on the environment grew stronger, particularly in southern areas of Fennoscandia, and in some areas it was even found to be stronger than climatic pressures (Kuosmanen et al., 2018). The earliest signs of agriculture in southern Fennoscandia are found at ca. 6000 BP (Sørensen & Karg, 2014; Berglund et al., 1991). Agriculture arrived at the border from the southern to the central Fennoscandian areas at around 3200–2800 BP (Eriksson & Cousins, 2014). Whereas for northernmost Fennoscandia there is strong indication of permanent cultivation since ca. 2250 – 2500 BP (D'Anjou, Bradley, Balascio, & Finkelstein, 2012; Josefsson, Hörnberg, Liedgren, & Bergman, 2017). For eastern Fennoscandia the earliest signs of agriculture have been found at ca. 4000 – 3500 BP (M. Lahtinen, Oinonen, Tallavaara, Walker, & Rowley-Conwy, 2017), though earliest evidence of cereal cultivation are suggested at ca. 6000 BP (Alenius, Mökkönen, Holmqvist, & Ojala, 2017).

Methodological opportunities

To reveal the causes that shaped past and present biodiversity patterns integrated research is needed (Fordham et al., 2020). The methodological opportunities to do so are wide and there is no standard way in how to explore past vegetation patterns, but interdisciplinary studies open a gateway to the past. For example can a combination of palynological data and archaeological remains give insights in the way and extend areas have been cultivated (Josefsson et al., 2017; Josefsson, Ramqvist, & Hörnberg, 2014), and fossil pollen can be leveraged to understand past vegetation patterns and can be combined with climate data, historical maps, archaeological remains, or charcoal remains to reach a deeper understanding of ecosystem change (Cui et al., 2014; Hannon, Halsall, Molinari, Boyle, & Bradshaw, 2018; O'Dwyer et al., 2021). One way of exploring past vegetation patterns is by focusing on the main changes in vegetation patterns, which can be done by specifically testing for correlations in time series. This can be done with the Granger causality test that infers causal relationships among time series (Shojaie & Fox, 2022). Main changes in correlations can further be investigated with a multiple change point analysis that locates the specific time points at which trends have changed (Lindeløv, 2020). The anthropological impact on these time series can be investigated by the use of Summed Probability Distributions (SPD), where the number of radiocarbon dates that result from archaeological findings act as a demographic proxy (Woodbridge et al., 2021). Fossil pollen data can be treated in different ways depending on the formulation of the hypothesis. Pollen data can be used in a pseudobiomisation approach in which pollen taxa are assigned to different land-cover classes (LCC) (Fyfe, Roberts, & Woodbridge, 2010) and thus function as representatives for the relevant vegetation types that dominate the area of interest. Only through the inclusion of data from different sources, such as pollen, anthropological and climate data and a combination of analysis can we get a clearer picture of how predictive parameters are connected to the dependent.

Aims and approach

Fennoscandia is ecologically interesting because it includes glaciers, arctic tundra and a high amount of woodlands, all of which are highly affected by the warming of the climate (Rantanen et al., 2022). The aim of this study is to determine vegetation change in Fennoscandia throughout the Holocene and determine causal relationships with climate and human impact. We trace the turnover of certain vegetation types and interpret the results with previously published data to assess the anthropological impact on Fennoscandia's landscape from the last glacial retreat at 9700 BP (Stroeven et al., 2016) until the end of the Middle Ages at 400 BP. The study of Fennoscandia includes Norway, Sweden, Finland and Russia's Kola peninsula and Karelia. Fossil plant pollen were grouped into different landcover classes (LCCs) to reconstruct vegetation dynamics through time and space. The LCCs include three types of woodland (deciduous, coniferous and wet woodland), two natural and open habitats (wet meadow and heath) and two anthromes (pasture/meadow and arable lands). Archaeological data was used to reconstruct population changes in the form of summed probability distributions (SPDs) and give insight on the extent to which humans have influenced vegetation change. Lastly, temperature data represents an additional potential cause of vegetation change; climate. Correlation between vegetation types, humans and climate were evaluated through regression analysis in the form of Granger causality analysis. Additional to that, we use change point analyses on the datasets for temporal exploration and timing of the main changes. With this we ask 1) are Holocene changes in Fennoscandian vegetation correlated to increased human footprint, 2) when did these changes occur, and 3) what anthropological events could have caused such changes in vegetation? The hypothesis we test here is that compositional changes in vegetation types in Fennoscandia throughout the Holocene are a response to human population growth. With this, we predict that climate had a stronger impact on vegetation than humans before the onset of farming. Further, we predict that vegetation patterns show a stronger correlation with population changes than with climate changes after the onset of farming.

Future importance

These results, taken together, provide a broad view of the long history of human impact in Fennoscandia, and while coarse, speak to the role of humans in biodiversity loss patterns and anthropological pressure through time in the region. Sampling and combining data from different sources by investigating paleo-archives offers great opportunities to understand biodiversity responses to future climate change and these causes that stand behind it (Nogués-Bravo et al., 2018). By assessing the strength with which each climate and humans impacted on vegetation and further exploration of these changes, we can give a deeper explanation to the way humans have shaped vegetation as we know it today.

Materials and methods

Time and space

Fennoscandia is a transnational area in Northern Europe, extending from 55°N to 72°N and from 4°E to 42°E. The region encompasses the whole area of Norway, Sweden, Finland, and the very north-western part of Russia. The area is defined accordingly to the exposed area of Precambrian rocks that is called the Fennoscandian shield (R. Lahtinen et al., 2011). Thus the definition is based on geological distinctions, which provides a more natural definition than national borders that have been moved several times during the Holocene. Its vegetation zones include (from north to south) the southern arctic, alpine, northern boreal, middle boreal, southern boreal, boreo-nemoral and nemoral zone (Moen & Lillethun, 1999). The area also encompasses the mountain range of the Scandes that runs from southern Norway through the border to Sweden and into northern Norway and brings a high variety in altitude to the area (Eriksson & Cousins, 2014).

The study treats data over the timespan from 9700 BP until the end of the Middle Ages at 400 BP. This covers the Holocene period after the last glacial retreat at 9700 BP (Stroeven et al., 2016) minus the modern time periods that followed the Middle Ages. This timespan was chosen, because it coincides with the beginning of human settlement and the introduction of agricultural practices, which was followed by a series of different economic systems and population growth and movement. Additionally, the fossil flora of the Holocene is similar or identical to modern biota (Jackson & Williams, 2004), which makes it easier to relate to in terms of human impacted change. The detailed timescale we chose to set our results in relation to (Table 1) is based on the most important epochs according to Fennoscandia's history (Josefsson et al., 2014).

Table 1: The timescale of Fennoscandia's important historical epochs through the Holocene (Josefsson et al., 2014).

Archaeological period	Interval
Mesolithic	12000 – 6200 BP
Early Neolithic	6200 – 5300 BP
Middle Neolithic	5300 – 4300 BP
Late Neolithic	4300 – 3700 BP
Early Bronze Age	3700 – 3100 BP
Late Bronze Age	3100 – 2500 BP
Pre-Roman Iron Age	2500 – 2000 BP
Early Roman Iron Age	2000 – 1800 BP
Late Roman Iron Age	1800 – 1600 BP
Migration Period	1600 – 1450 BP
Merovingian Period	1450 – 1200 BP
Viking Period	1200 – 950 BP
Medieval Period	950 – 450 BP

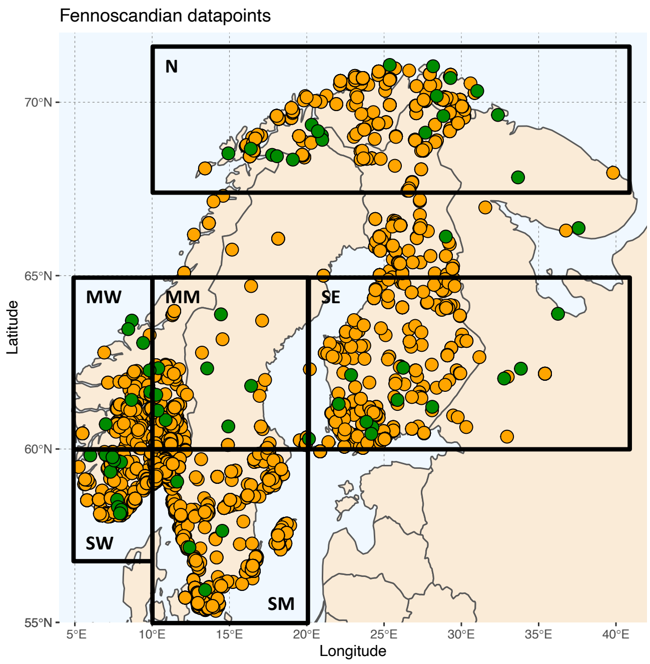

Figure 1: The division of Fennoscandia into the region that has been used in this study. The orange dots mark the locations of the archaeological sites and the green dots the locations of the pollen sites included in the study.

We worked with climate, archaeological and pollen data that was divided into six different areas according to longitude and latitude (Figure 1, Table 2). By taking data from smaller regions into a more broad scaled context information will get lost on the way. Based on the amount of data available, we therefore tried to make the regions that are to be looked at as big as possible whilst loosing as little information as possible. The region ranging from 65°N to 68°N was left out because of the poor spatial coverage of both pollen and archaeological sites. The border between the two western (SW + MW) and the two middle (MM + SM) regions is set to 10°E because this will part the mountainous regions from the low-altitude regions. The border between the two southern (SW + SM) regions to the two middle regions (MW + MM) is set to 60°N because this marks approximately the northern border of the nemoral zone. The south-eastern part includes the whole region east of the Baltic sea between 60 and 65°N. 65°N in the eastern part is approximately the border from the middle boreal to the northern boreal vegetation zone. The northern part, reaching from 68 to 72 °N latitude, was defined, because in this region we have an acceptable spatial coverage of pollen sites. In the region between 65 to 68 °N latitude only two pollen sites were available and thus this region has been excluded from our work. For this and for all subsequent data cleaning and analyses the R statistical platform was used (R Core Team, 2013).

Table 2: *The division of the different Fennoscandian regions in longitude and latitude.*

	longitude (°E)	latitude (°N)
North (N)	4 - 42	68 - 72
Southeast (SE)	21 - 42	60 - 65
Midwest (MW)	4 - 10	61 - 65
Central (MM)	11 - 20	61 - 65
Southwest (SW)	4 - 10	55 - 60
Southcentral (SM)	11 - 20	55 - 60

Pollen data

Pollen data from Fennoscandia was downloaded from the Neotoma Database (Williams et al., 2018) and processed with the R package 'neotoma' (Goring et al., 2015). Fossil pollen data from all 140 registered sites in Fennoscandia was retrieved. The datasets originally come from cores taken from lakes and bogs. From these 140 sites only sites with a robust chronological control and adequate Holocene time resolution were included in the study. Thus, abrupt changes in vegetation will not be related to one dataset suddenly ending. Only the sites fulfilling the following expectations were chosen for further preparation:

- The data has to include pollen data within the time period from 0 to 10'000 BP.
- A pollen sample has to exist for at least every 1000 years through the existing data.
- The data has to include at least 4 samples that are chronologically dated.

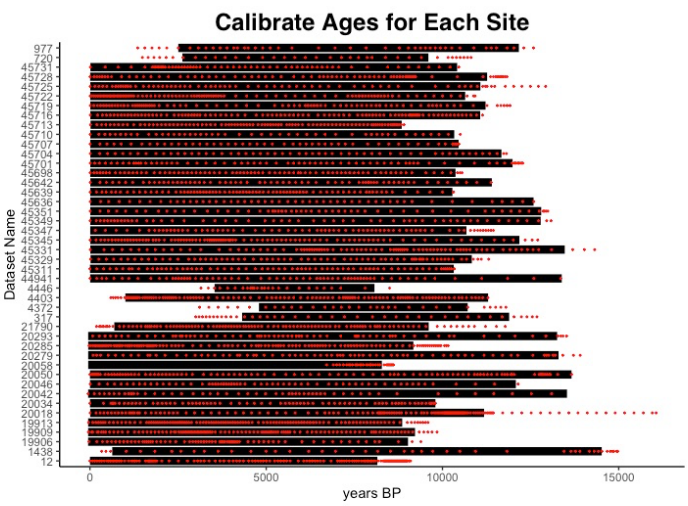

Figure 2: The temporal coverage of the pollen data included in the study.

After the data cleaning 62 sites were left, minus the two sites located between 65° and 68° latitude. Also, the temporal coverage of the data (Figure 2) was examined and assessed as good. For the final 60 sites Age Depth Models were calibrated with the 'Bchron' package (Haslett & Parnell, 2008) using the intcal20 calibration curve (Reimer et al., 2020). For every sample within a site 1000 possible dates were calculated, from which the mean was taken for further analysis. This was done, because calibration methods were updated several times since the first of our pollen samples has been calibrated and therefore the age calibration for all the pollen data needed to be standardised. This specific curve was chosen, because it is intended to specifically calibrate samples from the Northern hemisphere within the timespan from 0 BP – 55000 BP (Reimer et al., 2020).

The taxa names had to be standardised to fit the pollen data of the different authors in one data frame, which left us with 467 taxa. Those taxa are identified either to variety, species, genera or family level. This was done manually and the most specific group available for every taxa was chosen.

Land cover classes (LCC)

The different taxa have been divided into different land cover classes (LCCs). With this method taxa are assigned to a respective LCC according to the indicator species approach (Behre, 1986). We decided to work with LCCs because 1) this approach is not based on species-specific information and can thus be applied to our whole dataset, 2) it is a method of intermediate complexity and thus not too time consuming and 3) the major land-cover changes can be assessed for individual regions (Fyfe et al., 2010). In reality some taxa can be found in different LCCs. This cannot be taken into account by this method. The LCC where the taxa is most frequently found had to be chosen. Three types of woodland were defined (deciduous, coniferous and wet woodland), two natural and open habitats (wet meadow and heath) and two anthromes (pasture/meadow and arable lands). 167 taxa could not be grouped into either of these LCCs because they grow on hills or cliffs, in or around water, are ruderal or frequently occur in more than one of the LCCs. These classes were chosen because they represent the main vegetation patterns of Fennoscandia and can give insight in the early impacts that humans had on this vegetation.

The first round of division into LCCs was made according to the lists found in Fyfe et al. (2010) and Lechterbeck et al. (2014). The remaining taxa were divided based on information retrieved from ecological indicators for Swedish vascular plants (Tyler, Herbertsson, Olofsson, & Olsson, 2021), the online database *Artfakta* (SLU, 2022) and the book *Nordens flora* (Mossberg & Stenberg, 2018). The decision of assigning a taxon to an LCC was made on the considerations that 1) taxa growing on disturbed, non-arable grounds, on exposed places in mountains, hills or cliffs, or in or near waters could not be taken into consideration, 2) taxa that indicate human impact could belong both to pasture or arable lands and were classified as “pasture” when they generally crave low nutrient impact, or as “arable” when they crave high nutrient impact, 3) for taxa where there was only the genus name present in our database but there are several species growing in Fennoscandia, the grouping was based on the most dominant ones. *Nordens flora* and *Artfakta* provide information on the abundance of species. If a genus or a family had to be grouped into a LCC the taxa that represents this taxonomic rank to at least ca. 75% in the North was chosen as representatives, 4) taxa that are cultivated or often grow in or around cultivated areas are grouped under “arable”, and 5) all three sources give information on taxa habitats and indicator potential. If two sources or more identify one habitat as the main habitat of a taxon then this habitat is chosen as LCC. If the sources give more than one habitat as the main living area then the taxon is excluded from the analysis.

Table 3: The LCCs, their summed pollen count numbers and the number of different taxa included into each land cover class.

LCC	Nr. of counts	Nr. of taxa
coniferous woodland	1’424’518	16
deciduous woodland	1’876’111	57
wet woodland / fen carr	1’032’770	27
arable lands	93’538	41
pasture / meadow	582’404	89
wet meadow	262’986	52
heath	374’239	18
Not included	936’115	167
TOTAL	**6’582’681**	**467**

Each of the 60 sites had at least one taxa belonging to every LCC.

Adequacy of the three woodland groupings was tested, because many woodland plants can occur in more than one type of woodland and thus grouping all plants to the right LCC was most challenging for the woodland types. We tested adequacy by running the same analysis on the defined LCCs as for their strongest representatives only. The "coniferous woodland" was tested against *Picea* and *Pinus* only. The "deciduous woodland" was tested against *Betula* only. The "wet woodland" was tested against *Alnus* and *Salix* only. The results looked similar or indistinguishable from the analysis on the complete dataset. The signals for the change point of the LCCs were stronger overall than for the strongest representatives only.

Smoothing

Before the analysis, pollen data was smoothed. This was done, because the pollen data included in this study has been originally sampled by many different researchers, which automatically adds noise to the data. By applying a smoothing factor we try to keep this noise to a minimum. We did this with the inclusion of a simple moving average. Model selection was done by the Bayesian information criterion (BIC). BIC was chosen, because it tends to select models that are more parsimonious than those favoured by the Akaike information criterion (AIC) (Neath & Cavanaugh, 2012), which makes it less sensitive to overfitting. We also compared the model choices of the BIC to the choices of the AIC and they did overlap for every dataset.

Smoothing was done with a simple moving average implemented by the R package 'TTR' (Ulrich, 2021). A moving average is the time series that is constructed by taking the averages of a certain amount values from our past data points (Kulik & Soulier, 2020). The BIC was calculated for every model of every dataset to find out how many time intervals should be used to estimate the next data point (the time periods to move over). We've set a maximum of eight time intervals to make sure that the oldest datapoint in our datasets will still lay at 90000 BP, and because we assumed that a higher value will not improve smoothing, which the BIC results confirmed. Smoothing was then implemented by the R package 'smooth' (Svetunkov, 2017). We decided to select the smoothing window size separately for every site, because of the anticipation that varying data collection procedures at different sites might result in different noise characteristics at different sites.

Archaeological Data

To link vegetation changes to anthropological population increase, we calculated the summed probability distribution (SPD) from archaeological data (E. R. Crema, 2022). Archaeological data comes from the p3k14c database, a synthetic global database of archaeological radiocarbon dates (Bird et al., 2022). All archaeological data from Fennoscandia, based on 94 different kinds of materials, were calibrated with the same method as the pollen data, with the 'Bchron' R package and the intcal20 calibration curve. Missing values throughout time were filled by the use of linear interpolation. There are 94 intervals from the year 400 BP until 9700 BP. The number of interpolations within the datasets range from 0 to 21 and are assumed to have a negligible impact on the result. Data of the most recent years have been excluded, because radiocarbon-based time–frequency is less reliable for recent archaeological samples (E. R. Crema, 2022). We decided to cut our datasets at the time points when human population sizes were assumed to be declining again (Table 4). This has been done because it would be unreasonable to see a steady decline in population size in any of these areas in younger ages.

Table 4: The Fennoscandian regions, their youngest datapoints included, the number of missing data that had to be filled by interpolation, the number of archaeological findings per region and the number of archaeological sites per region.

Region	Youngest age BP	Nr. of interpolations	Findings per region	Sites per region
North	1400	4	1166	163
Southeast	500	9	547	142
Midwest	700	21	677	105
Central	700	12	1285	118
Southwest	700	11	1176	153
Southcentral	1400	0	3587	551

From the resulting datasets summed probability distributions (SPD) have been calculated for every area with the R package ‘rcarbon’ (Enrico R. Crema & Bevan, 2021). The binning parameter within the calculation of the SPD curve was set to 100 years. This parameter represents the length of a time interval within which a local SPD was calculated. Through binning a local SPD was calculated for archaeological sites from which more than one date is reported within the same time interval. The SPDs are summed up and then divided by the number of dates (E. Crema & Bevan, 2021). This process ensures that each site will be equally weighted, independent of the amount of dates recorded (E. R. Crema, 2022). Additionally, the resulting SPDs were compared to the null model, which expects the data to grow exponentially, and negative and positive deviations from this model were calculated. Datasets were cut at 9000 BP to match the pollen datasets.

Climatic Data

The mean annual temperature has been estimated for the timespan from 9000 BP until the youngest ages according to Table 4 for every region individually. To do so version 1.5.1 of PaleoView (Fordham et al., 2017) has been used. PaleoView is a tool for creating continuous climate projections suitable for detecting biotic responses to major climate shifts for the last 21000 years. The modelled climate reconstructions from PaleoView are based on the simulation output from the Community Climate System Model ver. 3 (CCSM3) (Fordham et al., 2017). This model is composed of four separate models that include information on concentrations of greenhouse gases, concentrations of aerosol species (sulfate, dust, carbon, sea-salt, and volcanics), a fixed solar constant and annually cyclic concentrations of ozone. These four models simulate the earth’s atmosphere, ocean, land surface and sea-ice (UCAR, 2020). The main reasons why we decided to use PaleoView is 1) that we will avoid circularity, because the climate estimations are not based on pollen data and 2) that we get individual estimations for every region.

Regression analysis

The first data analyses that have been done were linear regressions in case to determine if there exists a correlation between the different datasets. This will tell us 1) which LCCs are affected by the archaeological data or the climate data and 2) how much these LCCs are affected by the respective predictor. In our case the set predictors were the temperature data and the SPD values. The response data was the LCC count data.

Granger causality analysis

We decided on using Granger causality, because it is a tool specifically created to analyse time series data (Shojaie & Fox, 2022). First, the optimal length of time lags that should be used to explain the next data point in the time series had to be estimated. To do so model selection has been done by cross validation. We chose to use cross validation, because it is a machine learning method, that chooses the model with the best out-of-sample predictive accuracy. It does that by comparing learning algorithms by dividing the data into two segments, where one segment is used to train the model and the other validates the model (Refaeilzadeh, Tang, & Liu, 2018). Autocorrelation has been accounted for and the optimal time lags have been chosen individually for every dataset by using the packages ‘modelr’ (Wickham, 2022) and ‘timetk’ (Dancho & Vaughan, 2022). The maximum lag was set to seven and the number of folds to two for the datasets of the whole Holocene. For the datasets with only the data of before and after the onset of farming the maximum amount of time lags were set to three and the number of folds was set to one. The parameters were adapted for the shorter datasets, because 1) these datasets are too short for a large number of folds and long lags, and 2) because we don’t expect models with many parameters to be statistically supported by such short datasets, which the test runs with more lags also confirmed.

Then, we checked for causality between every predictor and the dependent data, and between both predictors and the dependent data. With the ‘vars’ package (Pfaff, 2008) we constructed a vector autoregressive (VAR) model. We used vector autoregressive models that regress the current value of LCC not just against SPD and climate, but also against the past LCC values. We have to do this because of the existing autocorrelation in the data. We checked if including SPD or climate as predictors in the regression improves the prediction accuracy of the model. Increased predictive power indicates a possible causal relationship between the specific predictor and the dependent variable. This is known as the Granger causality test. We used the multivariate Granger causality test implemented by the 'bruce' package (Bao, 2022). Our Granger causality test compares the VAR model that includes only the LCC with the VAR model that also includes LCC, SPD and climate as predictors and calculates the F-statistic. If significance was found, we used the respective datasets for further analyses in which we calculated the effect sizes. Granger causality analysis has been done on the full datasets plus the datasets consisting only of the data before the onset of farming and the datasets consisting only of the data after the onset of farming.

Commonality analysis is a way to deal with multicollinearity. Because of the fact that our predictors are not independent like in classical regression we have to account for that (Ray-Mukherjee et al., 2014). We used this method to explain how much of the variation in the LCCs is explained by the predictors. This method uses the regression R^2 to explain the unique and common effects of the predictors on the dependent data (Ray-Mukherjee et al., 2014). To do so the following calculations have been made according to Ray-Mukherjee et al. (2014):

- The unique contribution: $U_s = R^2_{l.s.c} - R^2_{l.c}$ / $U_c = R^2_{l.s.c} - R^2_{l.s}$
- The common contribution: $C_{s.c} = R^2_{l.s.c} - U_s - U_c$

The unique contribution of SPD (U_s) is calculated by subtracting the R^2 of the model with only climate as the predictor ($R^2_{l.c}$) from the R^2 of the full model with both climate and SPD as predictors ($R^2_{l.s.c}$). The unique contribution of climate (U_c) is calculated by subtracting the R^2 of the model with only SPD as the predictor ($R^2_{l.s}$) from the R^2 of the full model with both climate and SPD as predictors ($R^2_{l.s.c}$). The common contribution ($C_{s.c}$) is calculated by subtracting the unique contribution of SPD (U_s) and the unique contribution of climate (U_c) from the R^2 of the full model with both climate and SPD as predictors ($R^2_{l.s.c}$).

The northern dataset consisting of the times after the onset of farming was the smallest of our datasets and could not be tested for significances, because it did not include enough datapoints.

Change point analysis

Since we are interested in finding the large trends in vegetation changes instead of smaller fluctuations, we can work with the pollen counts only and do not need to include additional parameters like pollen production rates. Change point analyses were done not only to explain the significant impacts found within the regressions, but also to explore possible links between vegetation types. We decided to use multiple change point analyses (mcp) to show when and where changes in vegetation, climate and human development occurred. Implementation was done through the R package 'mcp' (Lindeløv, 2020). Mcp is a method that calculates multiple change points on an univariate dataset within a Bayesian framework (Lindeløv, 2020). Mcp requests a set number of change points. It divides the dataset into time intervals and uses a separate linear regression model to fit the data within each interval. Since it is a Bayesian method it integrates over many possibilities and averages over them. This leaves us with distributions of where the change points could be. Default priors from the 'mcp' package were used.

We decided on using the following model with adapted amounts of change points and intervals for the different datasets: model = y ~1+x, ~1+x, … . The y-axis contains the information on the count numbers of the respective LCC, SPD or temperature data to be tested, whereas the x-axis gives the according calibrated years BP. The first part of the model (y~1+x) defines the first interval to be shaped according to the movement over time (a slope). The second part defines the change point and the disjoined slope following it (,~1+x). We decided to use this model, because it includes the shape of the curves based on the x-value and because it assumes that change points happened at the time when slopes disjointed from each other. These two assumptions make the model more expressive and will therefore need fewer change points to be explained. This will give us more confidence in finding the most representative changes of the datasets representing the region, and it will be less sensitive to smaller changes that might be based on only a small number of sites, and would therefore not be representative for the whole area. Additionally, the second part of the model (,~1+x) is to be repeated until the number of change points for every LCC in every area is found. We decided to check for individual numbers of change points based on the assumption that vegetation change is different respective to the geographical area and the biological differences between LCCs. The decision of the number of change points has been made as followed: At first, the dataset with the respective LCC and the calibrated years BP has been tested for five change points with three chains, an iteration value of 3000 and a burn-in value of 50% of the iteration value (=1500). The two parameter values were gradually increased until the Rhat values for all the ages, intervals and change points have arrived at 1.0, which indicates convergence of the MCMC method. The following order of adaptions has been followed:

1. adapt = 1500 iter = 3000
2. adapt = 3000 iter = 4500
3. adapt = 6000 iter = 4500
4. adapt = 12000 iter = 6000
5. adapt = 20000 iter = 10000

Then, the Bayes factors for the hypothesis that the calculated mean of every changepoint is the actual value of the change point has been tested. The amount of change points that show a Bayes factor higher than 10, which strongly supports our hypothesis on the estimation of the change point's location (Andraszewicz et al., 2015), has been used for further calculation. The whole procedure has been repeated until all the change points showed a Bayes value higher than 10 and a Rhat value of 1.0. The model has been defined based on this number of change points. Change points have also been calculated for the SPD and temperature data.

Results

Our methodological approach applied to three datasets, land-cover class (LCC, representing vegetation types), Summed Probability Distributions (SPD, a proxy for human population size) and temperature (representing climate and its change through time). This approach produced two kinds of results. First, the effect sizes from the Granger causality test indicated which LCCs were statistically correlated with which predictor (Table 5). Second, the change points from the mcp analysis identified significant changes through time within each dataset.

Table 5: The commonality coefficients for the data that showed significant correlations are present for the three different datasets (the whole Holocene, the time before the onset of farming, the time after the onset of farming). The unique and common contributions of the predictors that showed significance are given. P-values are showed as followed: <0.1=.(low), <0.05=(moderate), <0.01=**(high), <0.001=***(very high).*

			coniferous woodland	deciduous woodland	wet woodland	wet meadow	pasture	arable land	heath
NORTH	Holocene (9000 – 1400 BP)	SPD							
		Climate							
		Both							
	Before (9000 – 2600 BP)	SPD							
		Climate							7.17% *
		Both							-3.87% .
	After (2500 – 1400 BP)	SPD	x	x	x	x	x	x	x
		Climate	x	x	x	x	x	x	x
		Both	x	x	x	x	x	x	x

SOUTHEAST	**Holocene (9000 – 500 BP)**	SPD	16.31% **				8.98% **	9.46% .	18.06% ***
		Climate	27.88% ***	16.45% .					
		Both	-7.77% ***				2.19% **		-0.18% ***
	Before (9000 – 4100 BP)	SPD							
		Climate		6.15% .					
		Both							
	After (4000 – 500 BP)	SPD					9.37% .		9.37% **
		Climate	27.00% *						
		Both	5.80 .						-0.80% **
MIDWEST	**Holocene (9000 – 700 BP)**	SPD					3.67% .		
		Climate			8.98% *			4.77% *	2.54% .
		Both			0.25% *				
	Before (9000 – 3300 BP)	SPD				14.87% *			
		Climate			6.15% .				10.22% *
		Both							9.08% *
	After (3200 – 700 BP)	SPD		15.45% .	24.30% *			13.82% .	
		Climate				16.01% *			
		Both			0.93% .	7.89% *			
CENTRAL	**Holocene (9000 – 700 BP)**	SPD		18.61% **	11.50% .	4.12% .	22.68% ***	19.39% **	22.37% **
		Climate							
		Both		-5.73% .		2.16% .	1.14% ***	2.53% **	4.78 *
	Before (9000 - 3300 BP)	SPD	8.77% *						
		Climate	10.86% *					5.40% .	
		Both	-7.81% *		6.55% *				
	After (3200 – 700 BP)	SPD					24.91% *	40.25% *	25.48% *
		Climate	22.65% *						
		Both	1.23% .				1.28% .	3.15% *	-0.23% .
SOUTHWEST	**Holocene (9000 – 700 BP)**	SPD	4.11% .	8.66% **	4.40% .			17.43% **	
		Climate	3.66% .	10.02% **	4.01% .				
		Both		-6.78% **	-2.42% .		4.27% .	0.99% *	
	Before (9000 – 6100 BP)	SPD							
		Climate							
		Both	11.80% *	12.64% .					
	After (6000 – 700 BP)	SPD		7.86% *	11.34% *	5.73% .		3.38% *	5.48% .
		Climate				7.28% *			5.10% .
		Both		0.17% .	-0.19% *	-0.95% *			-0.60% .
SOUTHMID	**Holocene (9000 – 1400 BP)**	SPD	22.47% *	7.58% *		7.81% *		23.88% **	
		Climate							
		Both		0.29% *		-0.08% *		-1.31% *	
	Before (9000 – 6100 BP)	SPD				41.13% *			
		Climate			20.40% .				
		Both					4.94% .		
	After (6000 – 1400 BP)	SPD	32.09 % **						
		Climate	15.25% *		6.68% .		14.66% *		
		Both	-7.88% **						

All LCCs

The number of correlations of every LCC for each time frame, plus the total correlations overall are shown in Table 6. We find that the coniferous woodlands showed the most correlations to temperature after the onset of farming and show the most correlation to temperature over all the different LCCs. But in every region the coniferous woodlands (the same applies to wet meadows) are correlated to temperature, they also showed a significant common correlation with both SPD and temperature taken together. The deciduous and wet woodlands, arable lands and heaths showed more correlation to temperature before, and more to SPD, after the onset of farming. The wet meadows were correlated to SPD across the Holocene, and the times before and after the onset of farming, but they were only correlated to temperature after the onset of farming. The pastures had few correlations before the onset of farming (a single correlation to both predictors) and arable lands had no correlation to temperature after the onset of farming and no correlation to SPD before the onset of farming. The common contributions lay between -7.88% and 9.88%. The unique contributions are higher than the common contributions. The unique contributions of climate lay between 3.66% and 27.88%. The unique contributions of SPD are higher than the ones for climate and lay between 3.38% and 41.13%. The mean unique contribution of SPD over the whole Holocene lays at 13.24%, while the mean unique contribution of SPD for the time after the onset of farming lays at 16.35%.

Table 6: The amount of significant correlations for the different datasets.

			coniferous woodland	deciduous woodland	wet woodland	wet meadow	pasture	arable land	heath	Total correlations
CORRELATIONS	The whole Datasets	SPD	3	3	2	2	3	4	2	19
		Climate	2	2	2	0	0	1	1	8
		Both	1	3	2	2	3	3	2	16
	Times before the onset of farming	SPD	1	0	0	2	0	0	0	3
		Climate	1	1	2	0	0	1	2	7
		Both	2	1	1	0	1	0	2	7
	Times after the onset of farming	SPD	1	2	2	1	2	3	3	14
		Climate	3	0	1	2	1	0	1	8
		Both	3	1	2	2	1	1	3	13

Arable lands correlated most with SPD (seven times), and were correlated twice with temperature. Wet meadows correlated six times with SPD, and twice with temperature. Pastures correlated five times with SPD and a single time with temperature. The heaths were the vegetation type that correlated the least with SPD, and they only did so after the onset of farming. The deciduous woodlands showed to be the woodland type that correlated to temperature most often. When comparing the overall correlations to each other we find that generally SPD was more often correlated to the LCCs after the onset of farming than before the onset of farming and also more often than to climate. Climate was more often correlated to the LCCs before the onset of farming than the SPDs, though not more often than after the onset of farming.

North

The heaths from before the onset of farming were the only LCC that showed a significant correlation with temperature as its predictor (for this and all correlation coefficients in the results, see Table 4; Figure 3). No correlation was found between any of the LCCs with SPD as their predictor.

The change points found individually for every northern dataset are visualized in Figure 3. The changepoints of the correlated dataset are visualized in the Appendix in Figure 11. At 7930 BP temperature rose significantly as did the occurrence of wet woodlands (7960 BP). The heaths show a significant peak at 7680 BP when also temperature peaks. All the LCCs decline after 7000 BP and incline again before 6000 BP. By this incline of LCCs also temperature inclines and reaches its highest point during the Holocene. The coniferous woodlands, wet woodlands and the wet meadows all decline in a phase when also temperature has a steady low over some centuries (between ca. 5500 and 5000 BP). The SPD curve indicates a higher population level between ca. 4000 and 3500 BP, which also ends with a significant incline in human population size at 3450 BP. The only other changes found at 4000 BP are the weak increase in both pastures and heaths. The northern region had a significant increase in SPD 100 years after the onset of farming (2500 BP). Despite a temporal lag (ca. 650 yr.), a significant rise in arable lands was also detected after the onset of farming. Also the coniferous woodland showed a significant rise at the same time as the rise in SPD (and the onset of farming).

Southeast

The coniferous woodlands showed a correlation with temperature of a higher significance ($p<0.001$) for the whole Holocene than for the time after the onset of farming (4000 BP; $p<0.5$). The correlation between temperature and coniferous woodlands also showed a higher significance and a higher unique contribution (U_c=27.88%) than for the correlation with SPD (U_s=16.31%) over the whole Holocene. The deciduous woodlands only showed weakly significant correlations with temperature over the whole Holocene and before the onset of farming. The two anthromes and the heaths show correlations to the SPD over the whole Holocene, while pastures and heaths even showed correlation for the time after the onset of farming. The unique contributions generally were lower for the correlated anthromes than for the correlated woodlands.

The change points found individually for every south-eastern dataset are visualized in Figure 4. The change points of the correlated datasets are visualized in the Appendix in Figure 12. In the Southeast the rise in temperature at 6000 BP is the only dated significance found in temperature. Right after this peak, when temperature is declining again a slight decline in coniferous, deciduous and wet woodlands and wet meadows is visible in the mcp results. Also, it is first between 5000 and 4000 BP, when temperatures rose again, that we see significant increases in every of the three woodlands and the wet meadows. No significant changes accompany the dated onset of farming at 4000 BP. The short peak in SPDs right after 3500 BP is accompanied by a short rise in pastures. The significant rise in SPDs at 2950 BP does not overlap with any other significant or mentionable change. Right after a short lasting low in temperature at about 2100 BP both the coniferous (significantly at 2049 BP) and the deciduous woodlands start to decline. Pastures, arable lands and heaths did first show significant rises or mentionable variations after the onset of farming. The rise in south-eastern SPDs at 1380 BP overlaps with the significant rises in pastures (1350 BP), heaths (1250 BP), deciduous woodlands (1270 BP) and wet meadows (1250 BP) (**Error! Reference source not found.**). Whilst later around 1000 BP SPDs started to rise again and overlap with the significant rise in arable lands at 950 BP and a decrease in pastures at the same time.

Midwest

Five of the seven LCCs did show correlations with the SPDs, whilst four out of these five showed an additional correlation to climate. Generally, the unique contributions of SPDs (mean contribution = 14.42%) are higher than those of climate (mean contribution = 8.11%). Though the significance of all the correlations in this region lay between weak ($p<0.1$) and moderate ($p<0.5$).

The change points found individually for every midwestern dataset are visualized in Figure 5. The change points of the correlated datasets are visualized in in the Appendix in Figure 13 and Figure 14. The peak in temperature at ca. 6000 BP goes along with the significant rise (6450 BP) and fall (5740 BP) of deciduous woodlands and the significant decrease (6010 BP) in wet woodlands. Around 4000 BP we have a weak rise in SPD, a clear peak in temperature, significant rises in coniferous (3940 BP), deciduous (4070 BP) and wet woodlands (4020 BP), wet meadows (4120

BP) and heaths (3980 BP), a significant decline in arable lands (3930 BP) and a weak peak in pastures. At 2850 BP, 350 years after the dated onset of farming (3200 BP), there is a significant increase in SPDs. Around the same time at 3000 BP we see a peak in temperature, significant declines in coniferous (2900 BP), deciduous (2950 BP) and wet woodlands (3000 BP) and wet meadows (3030 BP). Heaths start to decrease at 3000 BP as well, though show first a significance in decrease at around 3980 BP, whilst arable lands start to steadily increase around the dated onset of farming. After 1000 BP we see steady rises in SPDs (750 + 850 BP), whilst pastures start to decrease again (930 BP) after reaching its highest peak, and the open landscapes heaths, arable lands and wet meadows peak in count numbers, whilst the woodlands show relatively low count numbers.

Central

All the six LCCs correlate with the SPDs, whilst only two LCCs (coniferous woodland and arable land) correlate with temperature. The correlations between SPDs and the anthromes and the heaths showed high to very high significances and also the unique contributions of SPDs on vegetation lay between 19.39% and 22.68% for the whole Holocene, whilst the moderate significances after the onset of farming showed 24.91% to 40.25% unique contribution of SPDs. The significances for the Central region were the strongest.

The change points found individually for every central dataset are visualized in Figure 6. The change points of the correlated datasets are visualized in in the Appendix in Figure 15 and Figure 16. After 8000 BP, when temperature rose from a low point, every one of the LCCs showed a weak to strong decline in pollen counts. The deciduous woodlands did even show a significant incline before the decline at 8180 BP, whilst pastures and heaths showed a significant decline at 7950 BP and 7880 BP. After 6000 BP, when temperature significantly increases (5930 BP) and peaks, also arable lands, coniferous woodlands and wet meadows do significantly increase at 5850 BP, 5930 BP and 5920 BP. Here the significant change points before the onset of farming (3200 BP) of arable lands do overlap with the significant change point of climate at 6000 BP and both show significant inclines at 3780 BP and 3890 BP, which coincides again with the decline of temperature after it peaked. Around 4000 BP we see even significant inclines in deciduous woodlands (3860 BP), wet woodlands (3840 BP) and pastures (4110 BP). After the onset of farming at 3200 BP, we see a clear general trend of increasing woodlands and decreasing open habitats that coincide with the growth of population. We see overlaps in change points through this time at around 2500 BP, when SPDs start to increase (2610 BP) and deciduous woodlands significantly decrease (2400 BP), and after 2000 BP, when SPDs start to grow even steadier and pastures (1950 BP) and arable lands (1880 BP) significantly increase.

Southwest

All the LCCs except of pastures did significantly correlate to the SPDs over the whole Holocene and/or over the time after the onset of farming (6000 BP), whilst 5 of them (all except arable lands) did also correlate to temperature over the whole Holocene or after the onset of farming. The mean unique contributions of SPDs over the whole Holocene on the three woodland types is 5.72%. The mean unique contributions of climate over the whole Holocene on the three woodland types is 5.90%. SPDs show additional correlations to two woodland types (deciduous and wet) after the onset of farming.

The change points found individually for every southwestern dataset are visualized in Figure 7. The change points of the correlated datasets are visualized in in the Appendix in Figure 17 and Figure 18. Right after the onset of farming at 6000 BP we see a significant rise (5950 BP) and peak in temperature that is only followed by the significant decrease in wet woodlands at 5520 BP. After 4000 BP there are significant rises in coniferous (3620 BP) and wet woodlands (3510 BP). There is even a rise in deciduous woodlands visible. The significant rise in SPDs at about 2300 BP draws with it non-significant, but visible rises in wet meadows, arable lands and heaths, all of which are also significantly correlated to SPDs after the onset of farming. At 870 BP the SPDs started to significantly rise. This rise falls together with the significant increase in wet meadows (1030 BP), pastures (1050 BP), arable lands (1150 BP) and heaths (980 BP). At this time, even the three woodlands show increases.

Southcentral

The coniferous woodlands, the deciduous woodland, the wet meadows and the arable lands did correlate to SPDs over the whole Holocene. Interestingly did three LCCs (coniferous woodlands, wet woodlands and pastures) correlate to temperature after the onset of farming (6000 BP).

The change points found individually for every southcentral dataset are visualized in Figure 8. The change points of the correlated datasets are visualized in in the Appendix in Figure 19 and Figure 20. The onset of farming (6000 BP) was followed by a significant rise in SPDs (5800 BP). We see a significant increase in wet woodland at 5050 BP that coincides with an increase in heaths right before 5000 BP. They both start do decrease again 300 to 500 years later. A significant increase in pastures was first found at 4050 BP. Arable lands show a small rise shortly before 4000 BP, but then show a significant decrease at 3670 BP, which coincides with the increase in SPDs (3510 BP) and is shortly followed by a significant increase in wet meadows at 3360 BP. Deciduous woodlands, pastures and wet meadows that all start to significantly increase between 3300 BP and 4100 BP, continue to do so until the end of the Holocene.

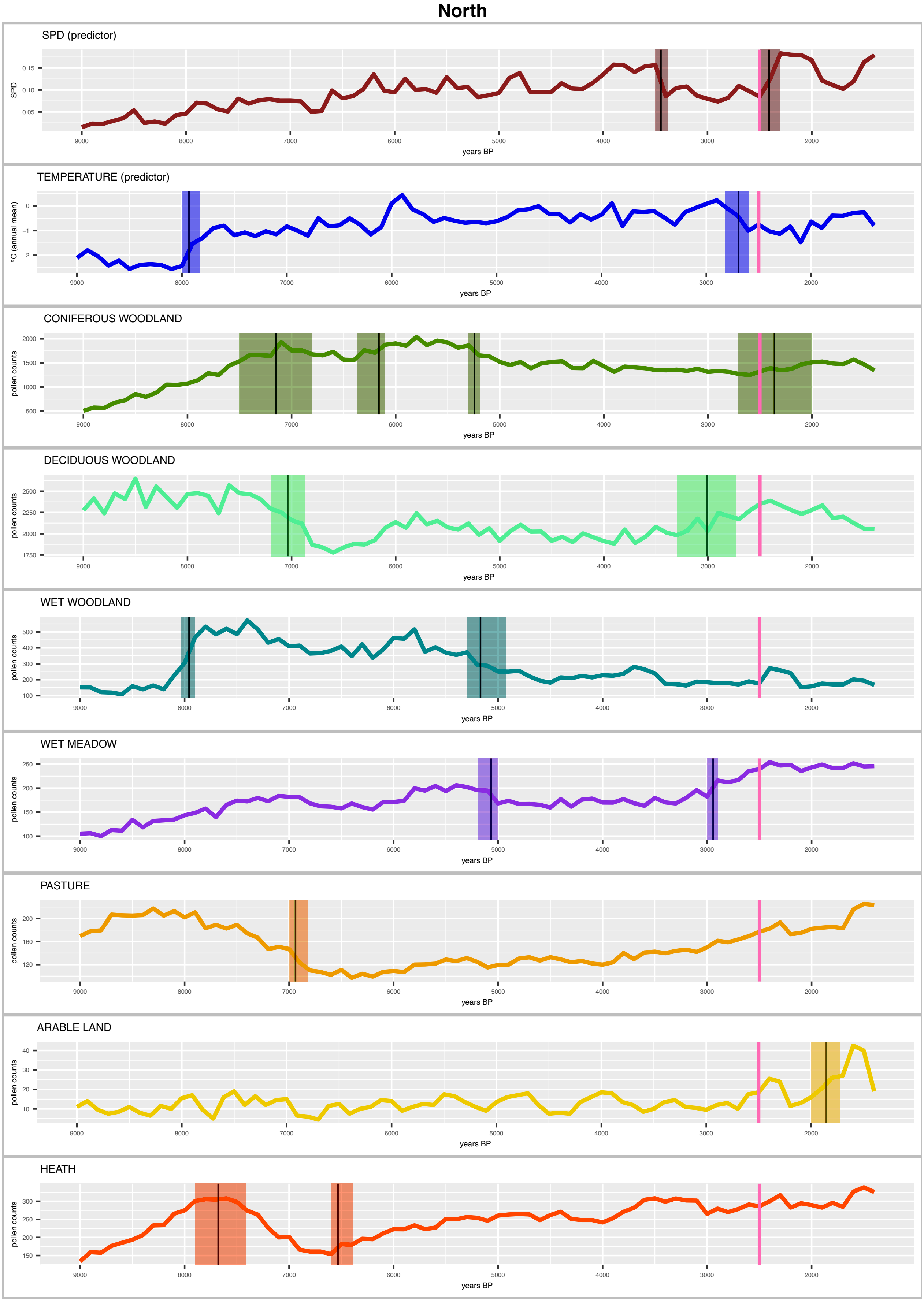

Figure 3: All change points found for the datasets from the North over the years from 9000 BP to 1400 BP. The pink vertical line indicates the onset of farming at 2500 BP.

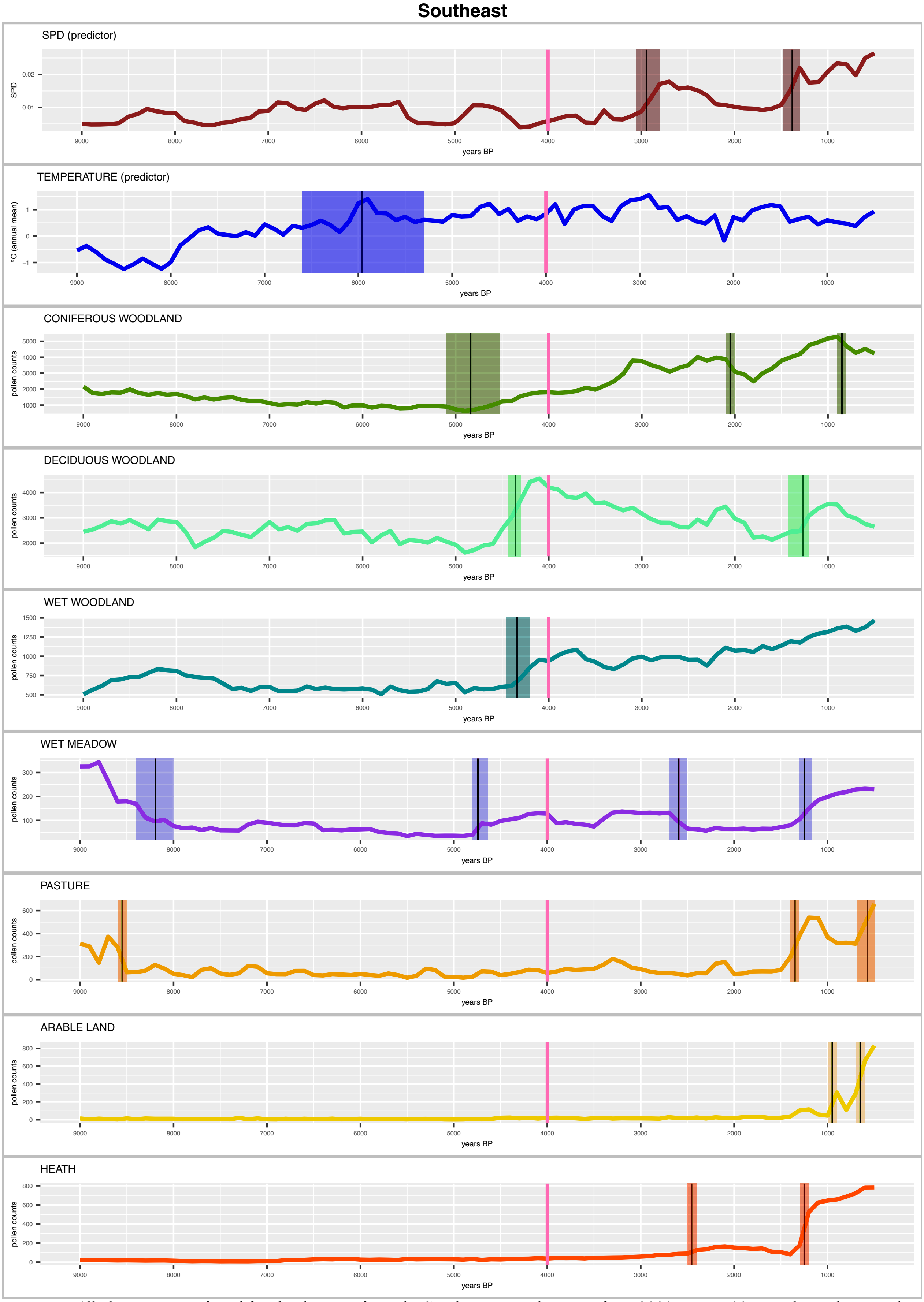

Figure 4: All change points found for the datasets from the Southeast over the years from 9000 BP to 500 BP. The pink vertical line indicates the onset of farming at 4000 BP.

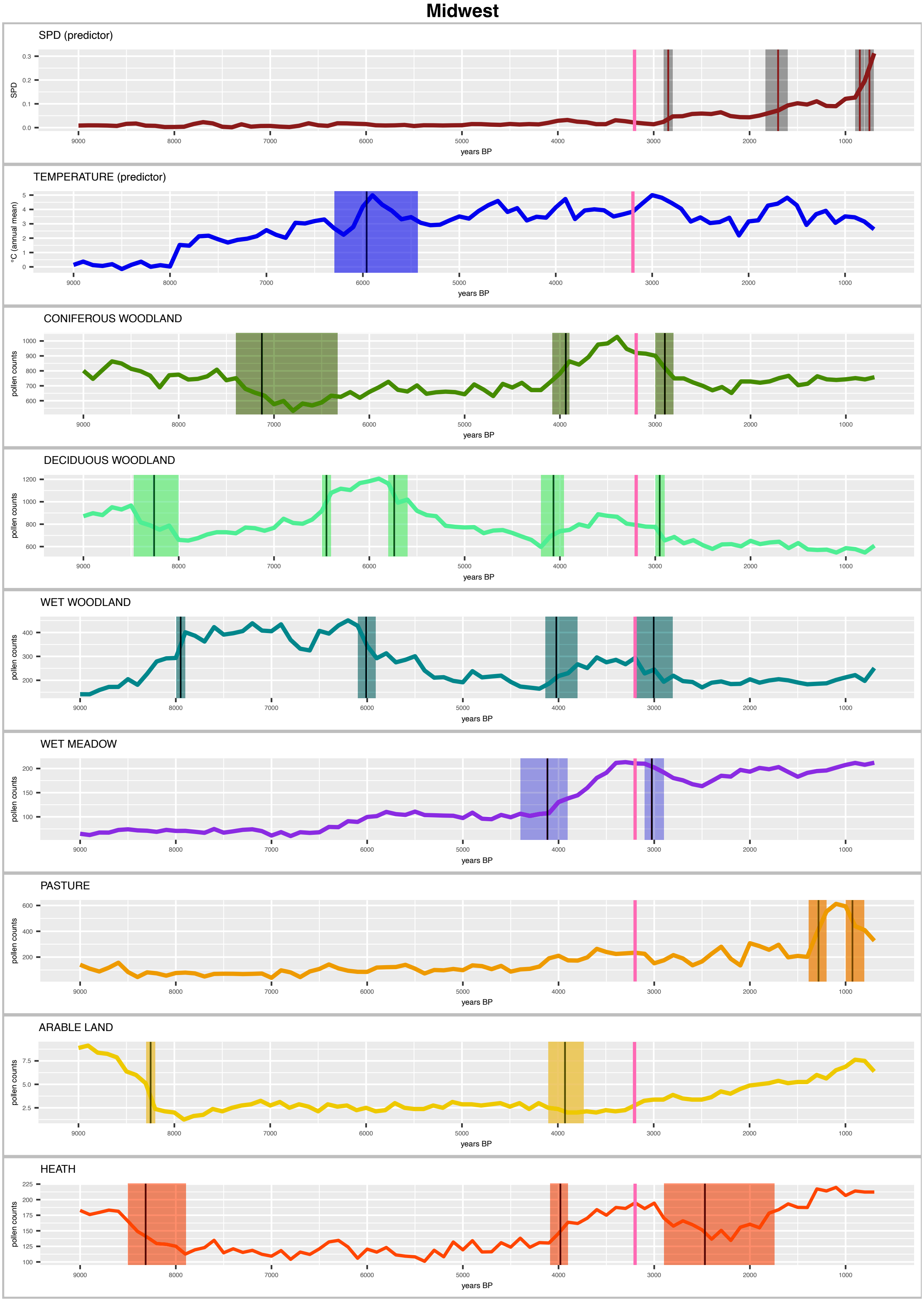

Figure 5: All significant change points found for the datasets of the Midwest over the years from 9000 BP to 700 BP. The pink vertical line indicates the onset of farming at 3200 BP.

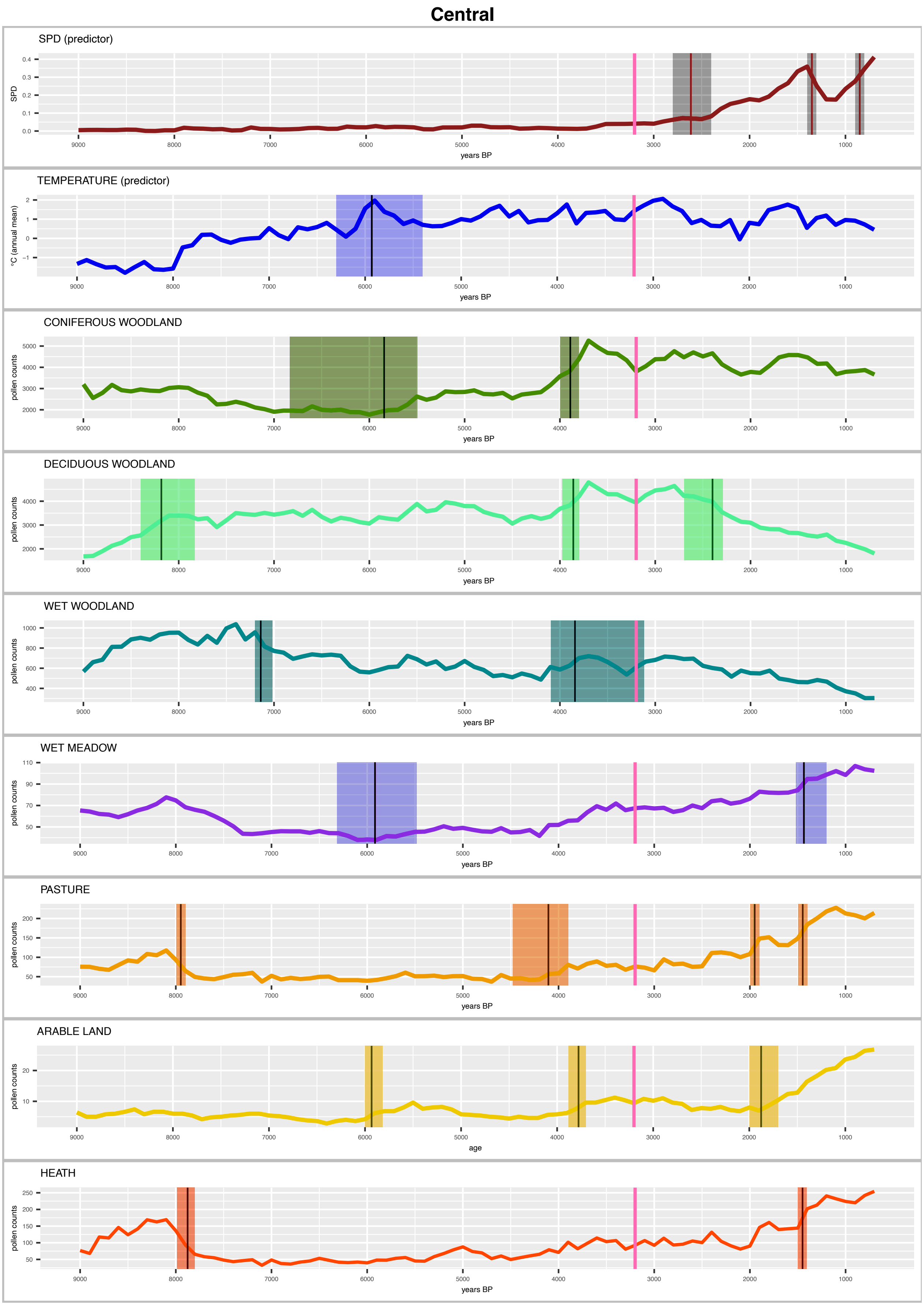

Figure 6: All significant change points found for the datasets of the central region over the years from 9000 BP to 700 BP. The pink vertical line indicates the onset of farming at 3200 BP.

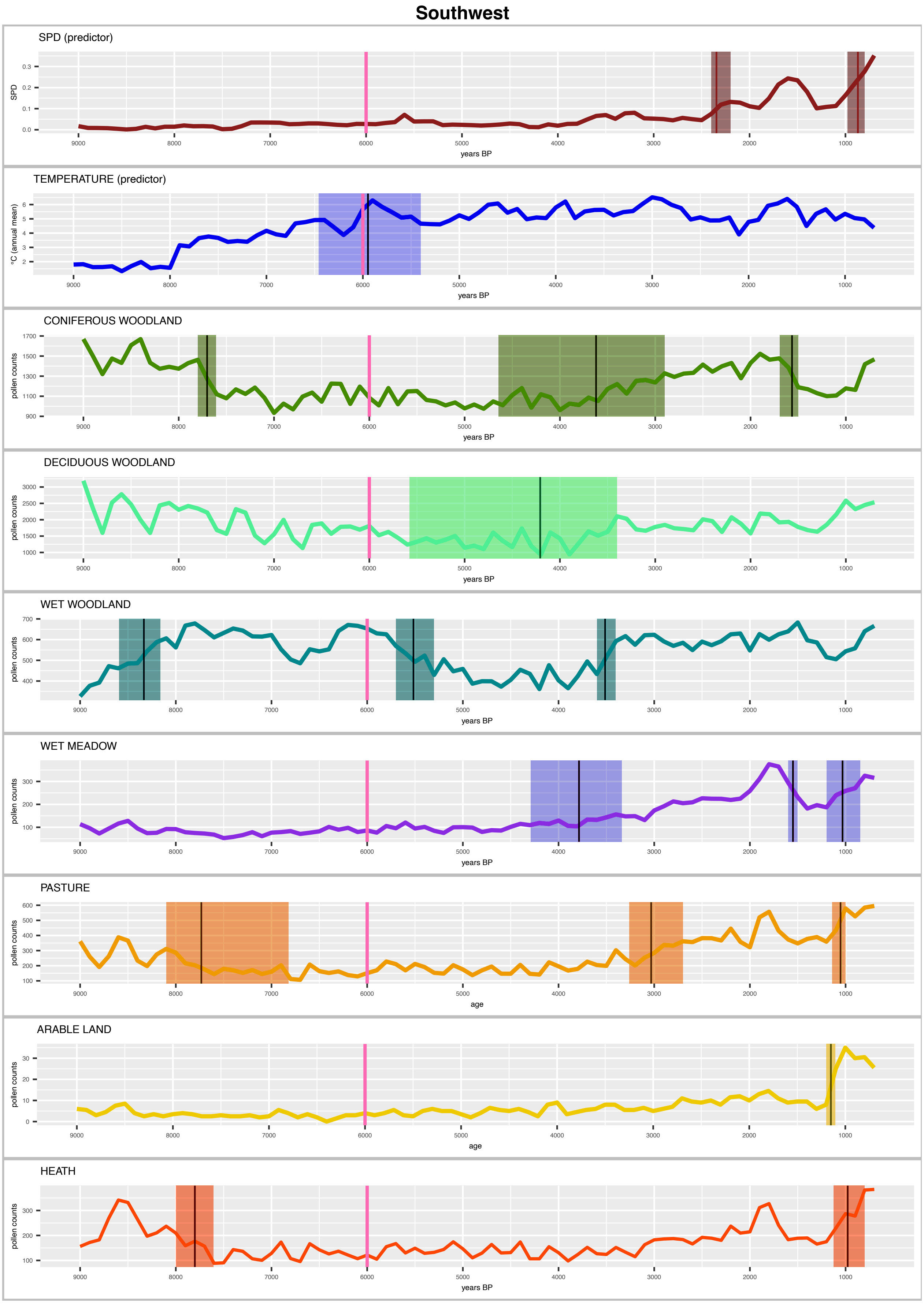

Figure 7: All significant change points found for the datasets of the Southwest over the years from 9000 BP to 700 BP. The pink vertical line indicates the onset of farming at 6000 BP.

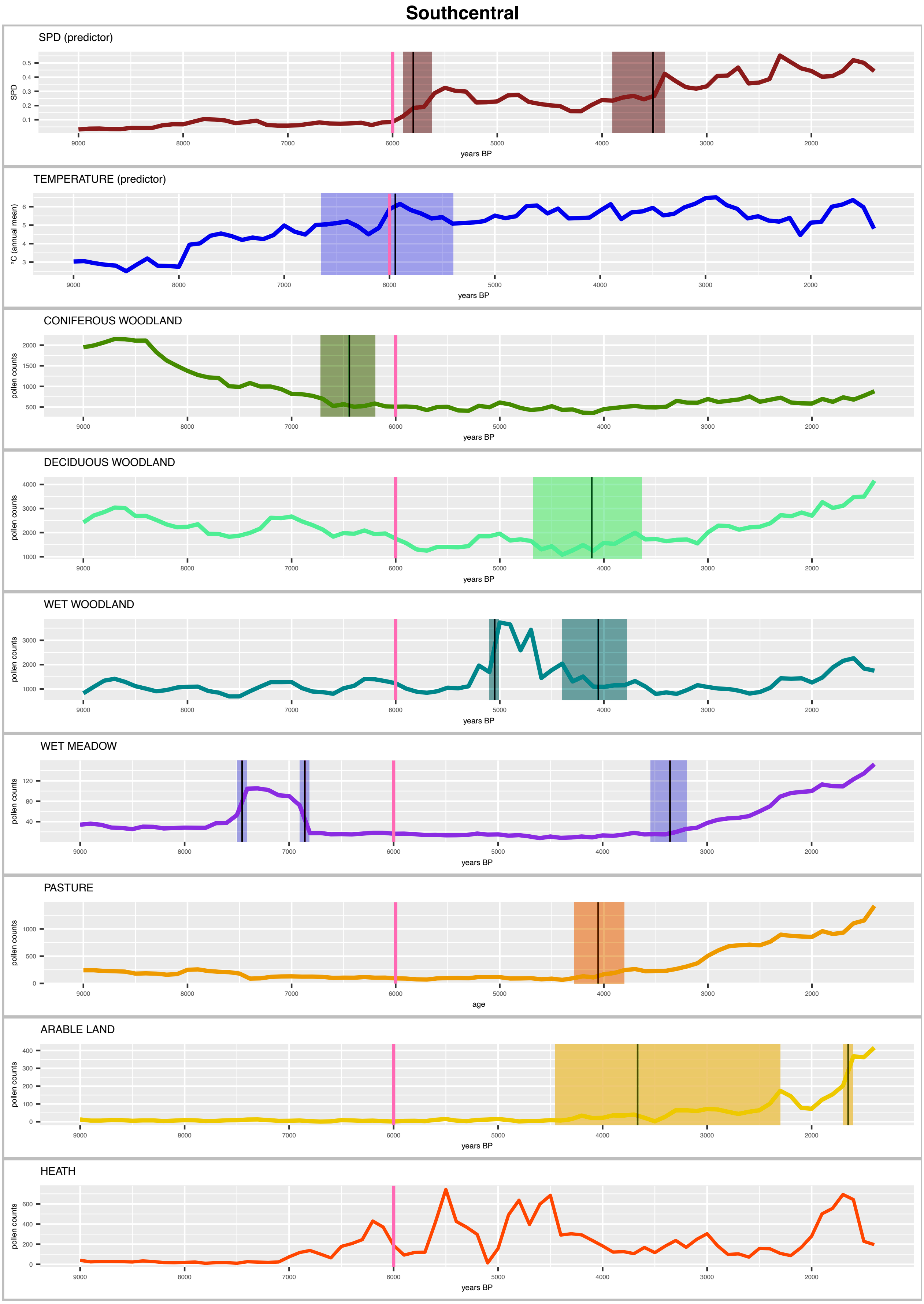

Figure 8: All significant change points found for the datasets of the Southcentral over the years from 9000 BP to 1400 BP. The pink vertical line indicates the onset of farming at 6000 BP.

SPDs

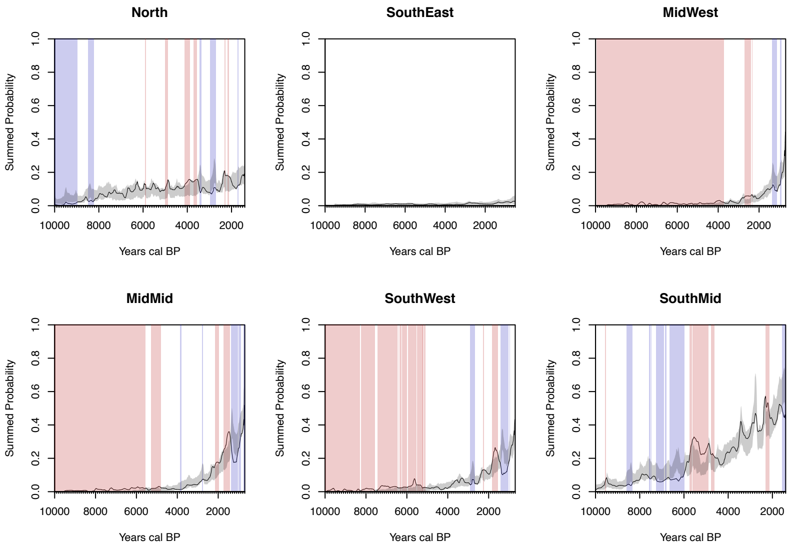

Figure 9: The calculated SPDs and their null models from every of the six regions.

The positive and negative deviations of the calculated SPDs from the exponential null models are visible in Figure 9. The North shows many negative and positive deviations through the whole Holocene and it does not show a clear exponential increase. Whereas the south-eastern curve does not show any deviations from the null model and shows only very low SPDs. The two central (MidWest + MidMid) and the southern (SouthWest + SouthMid) datasets show an exponential increase in SPD values.

Discussion

Vegetation changes and its correlates

This is the first time that correlations between vegetational changes, human population change and climate, and the strength of these correlates' impacts have been modelled and measured on nearly whole Fennoscandia over most of the Holocene. To test 1) if compositional changes in vegetation types in Fennoscandia throughout the Holocene are a response to human population growth, 2) that climate had a stronger impact on vegetation before the onset of farming and 3) that population and vegetation patterns show a stronger correlation after the onset of farming, the correlations between vegetation types, human population growth and climate have been evaluated, and the most significant time points of change have been revealed.

Based on the rather low common contributions and the fact that common contributions most often occurred in datasets that correlate with high significances, we can assume that in these cases the additional predictor acted as a complementary predictor and not as a central correlate. Thus, the individual correlates show to have important individual impacts on vegetation. By further investigation of these correlates, it became clear that human population growth was the main correlate to vegetation change throughout the Holocene, and had an especially high impact, based on the amount of correlations, change points and the effect size, after the onset of farming. Climate did not show any correlation trends, though showed a higher amount of change points, that party overlapped with vegetation changes, before the onset of farming. Our results on the general impacts of human population change and climate on Fennoscandian vegetation lay in accordance to findings of another study on Fennoscandia (Kuosmanen et al., 2018).

The fact that we have found more correlations between vegetation patterns and human population sizes over the whole Holocene shows that humans have affected landscape already before the onset of farming and supports the findings on the long impact humans have had on vegetation (Ellis et al., 2021).

Climate (mean annual temperature)

Climate seemed to have impacted northern vegetation in the years between 8000 BP and 5000 BP, when in- and decreases of woodlands, wet meadows, pastures and heaths follow the main trends of the warming and cooling of temperatures in this phase. Especially the warming after 8000 BP and the cooling phase in 5200 BP can be mirrored in the vegetation data of the North. Also in the southwestern and central region we see clear influences of climate on coniferous woodlands, pastures and heaths around 8000 BP. Later, at 6000 BP in the central region, vegetation starts to rise around the same time temperature peaks. This is especially well visible in the increase of coniferous woodlands that increase to do so into a humid phase around 5500 BP (Gunnarson et al., 2003). The steep increases in woodlands in the central and midwestern regions at around 4000 BP could be related to the ending of a dry phase at this time (Gunnarson et al., 2003). In the Southeast are correlations to increases in woodlands (especially deciduous) between 5000 and 4000 BP also more likely to be correlated to increasing temperatures and drier climate than to a decrease in human population size. The increases in vegetation after 4000 BP in the central and midwestern region could be correlated to either humans or climate, this is hard to distinguish due to weak and overlapping correlations. Though, they probably happened due to the ending of a dry period and an increase in humidity (Gunnarson et al., 2003). Also, the decrease of southern coniferous woodlands started because of the warming temperature in the early Holocene (9000 – 8000 BP).

Overall, climate seems to have impact northern vegetation patterns the most, whilst having least over regional influences in the southern regions, where human population sizes started to grow already in the Neolithic.

Human population size (SPD)

Human population size didn't seem to have much impact in the North. This is also a region that has a long history of foraging (Sjögren & Damm, 2019) and has been much less intensively used by people (Eriksson & Cousins, 2014). Also, due to pastoral and arable patterns it seems that pastoral practices including animal husbandry are prevalent on cultivation in the North. Most likely, agriculture was not a system in the North until after 2000 BP due to the cold climatic conditions and due to the non-linear increase of arable pollen counts and population growth. Whereas the first increase in human population in the Southeast is already dated to about 7000 BP. This is congruent with the first big spread of Neolithic communities from south-eastern Europe (Alenius, Mökkönen, & Lahelma, 2013), which seems to be the earliest rise in population throughout Fennoscandia. This early rise in population has also been reconstructed in Tallavaara's (2010) SPD curve of south-eastern Fennoscandia. Pastoral activities started to notably increase with the increase in population after the Migration Period, to which time also the onset of cultivation through slash-and-burn cultures have been dated (Eriksson & Cousins, 2014). Those cultures have also affected landscape openness, which is seen in increasing pastures, wet meadows and heaths.

Signs for the arrival of farming in the Midwest become clear in the Late Bronze Age at around 3000 BP when decreases in woodlands coincide with a constant increase in arable lands and SPDs. The beginning of the Medieval Period shows clear signs of human induced homogenisation, when landscapes have clearly opened up in relation to the ongoing rise in human population in this region.

In the central region, by the arrival of the Pre-Roman Iron Age we see an opening of the landscape and thus homogenisation coinciding to human population growth. This trend intensifies in the Late Roman Iron Age, which marks a time of intensified agricultural development in the form of dairy production and hay-making (Eriksson & Cousins, 2014). The high contributions of SPD on the trend in pastures, arable lands and heaths is well visible and continuing onto the Middle Ages. Interestingly, in the central region we do not see declines in pastures or arable lands during the Migration and Viking Periods, when population sizes are low. Something we can observe in the two southern regions.

The climate cooling after 6000 BP happened at the same time as the start of human-induced opening of the landscape in the southern regions (O'Dwyer et al., 2021), which is well visible in the southwest and can also be seen weakly in the wet and deciduous woodlands of the southcentral region. The correlations of the southcentral region would suggest that this long-lasting opening of the vegetation was human induced as it also coincides with rising human populations whilst the southwestern region shows correlations to both climate and humans, especially for the woodlands.

In both southern areas landscape started to open up in the Bronze Age. In the southcentral region this starts already in the early Bronze Age due to agricultural use and deforestation (Eriksson & Cousins, 2014; O'Dwyer et al., 2021), where an increase in human populations coincides with the decrease in deciduous and wet woodlands and the continuous growth of wet meadows and pastures. The deforestation phase during the transition from the Late Bronze Age to the Early Roman Iron Age (Eriksson & Cousins, 2014) is especially well visible in the southwest. Whilst in the southcentral region an increase of human population size and arable lands after the onset of the Pre Roman Iron Age could be due to the introduction of stabling of livestock (Eriksson & Cousins, 2014). The southwestern decrease in population in the beginning of the Migration period, when open land was abandoned (O'Dwyer et al., 2021), can be related to the decrease of pastures, arable lands, heaths and wet meadows. Only after the end of the Viking period and with the beginning of the Middle Ages population size start to rise again sharply.

The overall southern and central correlation numbers and effect sizes support the assumption that the southern and central regions have mainly been impacted by humans.

The summed probability distributions (SPD)

The SPD curve of the North does not show an exponential increase and therefore deviates from the null model in both positive and negative ways. To better understand which deviations might be realistic and which not we compared our SPD results to an SPD curve based on Norway's north that lays in the region of our northern Fennoscandian region (Jørgensen, 2020). The results lay to a high extend in accordance to each other. Also, the SPD curve of the Southeast was too low to be properly compared to the null model, probably due to a combination of the size of the area and the comparably low amount of archaeological findings. Therefore we did compare it to the SPD curves by Tallavaara (2010). We compared our curve mainly to the southern, but also considered the central curve of Tallavaara (2010). We found a lot of congruence between the curves until 2000 BP.

The issues around SPDs

The problems around SPDs are 1) its non-random and systematic nature of chronological uncertainty, 2) the sampling errors, and 3) the wide range of possible population curves (E. R. Crema, 2022). They are non-random, because already the archaeological sites are chosen for some specific reasons. We expect to find certain artefacts from a certain time at these places. Another issue is that the "summed" probabilities that describe if population increased or decreased at certain times do not actually relate to the same event, because they are based on any kind of archaeological findings that can be associated to anthropological activities. Therefore it is unclear how the probabilities itself can be interpreted (E. R. Crema, 2022). The events that the dates in this study are associated with can only be defined as "anthropic", because they are not limited to specific findings, but instead every finding connected to human presence is included. These problems have to be considered when interpreting the data. Alternative ways to estimate historical human population changes could be implemented by using relative chronology based on pottery or similar historical sources, though here data availability can be limiting for over regional studies. Data availability was the main reason why we decided to work with SPDs, despite the insecurities around the approach.

The methodology

Granger causality tests and change point analyses are methods that are constructed to deal with time series data. The fact that our results manage to recover over regional trends in vegetation, climate and human demography through time, reassures the effectiveness of these methods and their ability to reflect over regional patterns.

Improvements

Covariance-stationary has not been checked for, because we didn't have any reason to believe that our data was not covariance-stationary. An improvement could be to carefully check for that, and to adjust for if needed. Also, we did not explicitly model the dependence between humans and climate. Network Granger causality is a method, that would have been able to do so (Shojaie & Fox, 2022), though we couldn't find a way to implement network Granger causality in R. Implementing this method, could further improve the results. Neither could the boot strapping of the commonality coefficients be implement in R. Boot strapping these coefficients would bring a higher reliability to the results. Another issue concerning the commonality coefficients was that when predictors are collinear, there is some shared information between them. When that shared information helps to predict the response, the common contribution is positive. When shared information is unrelated to the response, cancelling this information out by choosing the right regression coefficients helps to predict the response and the common contribution is negative. Negative common effects suggest that it is the part of human activity that is not dependent on climate that is driving vegetation changes (Prunier, Colyn, Legendre, Nimon, & Flamand, 2015), which is hard to interpret and is therefore not done in this study. The significant common contributions have only been looked at, to know if significant amounts of vegetation change have been explained by both predictors together. We examined this closer, because we wanted to know to what extend humans might have been driven by climate, which high common contributions could have suggested.

Conclusion

Humans have been the strongest cause of vegetation changes over the Holocene, and especially after the onset of farming, while climate was a strong force in the centuries following the glacial retreat. Indications on the establishment and abandonment of different land-use practices have been found and/or recovered. This includes signs of the use of slash-and-burn systems in the Southeast, animal husbandry in the central and northern regions and deforestation in the southwest and southcentral regions. We see signs of human induced homogenisation of the landscape in the Midwest (Medieval Period) and in the central region (Pre-Roman Iron Age).

We showed that it is possible to examine over regional patterns by combining local and regional data from different sources. Through the implementation of methods that specifically aim to analyse time series data we can correctly account for existing autocorrelation. Also, multiple change point analysis proved to be a useful tool in depicting the main changes in vegetation based on fossil pollen data. Altogether, we present a long view of vegetation change in response to both climate and anthropogenic pressures.

The strong advantage of interdisciplinary studies like his one is that it enables us to look at a variety of data in a connected way, that puts the current state of Fennoscandian biodiversity into context, which allows for better projections in the future.

Future work

The vegetational patterns can become even clearer through the extension of the dataset with archaeobotanical data. This could give clearer insights into the ratio of the assemblage of different species that had effects on humans and vice versa. With an extended dataset even simple presence and absence of species could be investigated. Also, there are climate patterns visible which can only be further investigated by expanding the climate variable with additional parameters like humidity, sedimentary stable oxygen isotope records, tree line shifts or changes in lake levels. Finally, SPD distributions could be additionally supported by proxies such as palaeoecological estimates of openness or through relative chronologies based on pottery or similar historical sources.

Acknowledgments

First and foremost, I have to thank my research supervisor Christine D. Bacon. With her dedicated involvement she improved this work and kept me motivated through the process of writing this thesis. Special thanks go to Jakub Truszkowski, who was an immense support during the development of the method. I am also grateful to Claudia Speciale, who assisted me with the historical interpretation of the results. I thank Bengt Oxelman, who is the examiner of this work. Final thanks go to the people at the University of Gothenburg, especially to the Gothenburg Global Biodiversity Centre, who constantly supported me with feedback on the improvement of this work.

References

Alenius, T., Mökkönen, T., Holmqvist, E., & Ojala, A. (2017). Neolithic land use in the northern Boreal zone: high-resolution multiproxy analyses from Lake Huhdasjärvi, south-eastern Finland. *Vegetation History and Archaeobotany*, *26*(5), 469–486. https://doi.org/10.1007/s00334-017-0606-2

Alenius, T., Mökkönen, T., & Lahelma, A. (2013). Early Farming in the Northern Boreal Zone: Reassessing the History of Land Use in Southeastern Finland through High-Resolution Pollen Analysis. *Geoarchaeology*, *28*(1), 1–24. https://doi.org/10.1002/gea.21428

Andraszewicz, S., Scheibehenne, B., Rieskamp, J., Grasman, R., Verhagen, J., & Wagenmakers, E. J. (2015). An Introduction to Bayesian Hypothesis Testing for Management Research. *Journal of Management*, *41*(2), 521–543. https://doi.org/10.1177/0149206314560412

Bao, H. W. S. (2022). Bruce R: Broadly useful convenient and efficient R functions. *R Package Version 0.8.9*, *2*.

Behre, K.-E. (1986). Anthropogenic Indicators in Pollen Diagrams. *The Journal of Ecology*, *76*(3), 903. https://doi.org/10.2307/2260581

Berglund, B. E., Malmer, N., & Persson, T. (1991). Landscape-Ecological Aspects of Long-Term Changes in the Ystad Area Author (s): Björn E . Berglund , Nils Malmer and Thomas Persson Published by : Oikos Editorial Office in the Ystad area, (41), 405–424.

Bird, D., Miranda, L., Vander Linden, M., Robinson, E., Bocinsky, R. K., Nicholson, C., … Freeman, J. (2022). P3K14C, a Synthetic Global Database of Archaeological Radiocarbon Dates. *Scientific Data*, *9*(1), 1–19. https://doi.org/10.1038/s41597-022-01118-7

Cardinale, B. J., Duffy, J. E., Gonzalez, A., Hooper, D. U., Perrings, C., Venail, P., … Grace, J. B. (2012). Biodiversity loss and its impacton humanity, 0–9. https://doi.org/10.1038/nature11148

Crema, E., & Bevan, A. (2021). Analysing radiocarbon dates using the rcarbon package.

Crema, E. R. (2022). *Statistical Inference of Prehistoric Demography from Frequency Distributions of Radiocarbon Dates: A Review and a Guide for the Perplexed*. *Journal of Archaeological Method and Theory*. Springer US. https://doi.org/10.1007/s10816-022-09559-5

Crema, Enrico R., & Bevan, A. (2021). Inference from large sets of radiocarbon dates: Software and methods. *Radiocarbon*, *63*(1), 23–39. https://doi.org/10.1017/RDC.2020.95

Cui, Q. Y., Gaillard, M. J., Lemdahl, G., Stenberg, L., Sugita, S., & Zernova, G. (2014). Historical land-use and landscape change in southern Sweden and implications for present and future biodiversity. *Ecology and Evolution*, *4*(18), 3555–3570. https://doi.org/10.1002/ece3.1198

D'Anjou, R. M., Bradley, R. S., Balascio, N. L., & Finkelstein, D. B. (2012). Climate impacts on human settlement and agricultural activities in northern Norway revealed through sediment biogeochemistry. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(50), 20332–20337. https://doi.org/10.1073/pnas.1212730109

Dancho, M., & Vaughan, D. (2022). Package ' timetk ' R topics documented :

Díaz, S., Settele, J., Brondízio, E. S., Ngo, H. T., Agard, J., Arneth, A., … Zayas, C. N. (2019). Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science*, *366*(6471). https://doi.org/10.1126/science.aax3100

Ellis, E. C., Gauthier, N., Goldewijk, K. K., Bird, R. B., Boivin, N., Díaz, S., … Watson, J. E. M.

(2021). People have shaped most of terrestrial nature for at least 12,000 years. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(17), 1–8. https://doi.org/10.1073/pnas.2023483118

Eriksson, O., & Cousins, S. A. O. (2014). Historical landscape perspectives on grasslands in Sweden and the Baltic Region. *Land*, *3*(1), 300–321. https://doi.org/10.3390/land3010300

Fordham, D. A., Jackson, S. T., Brown, S. C., Huntley, B., Brook, B. W., Dahl-Jensen, D., … Nogues-Bravo, D. (2020). Using paleo-archives to safeguard biodiversity under climate change. *Science*, *369*(6507), eabc5654. https://doi.org/10.1126/science.abc5654

Fordham, D. A., Saltré, F., Haythorne, S., Wigley, T. M. L., Otto-Bliesner, B. L., Chan, K. C., & Brook, B. W. (2017). PaleoView: a tool for generating continuous climate projections spanning the last 21 000 years at regional and global scales. *Ecography*, *40*(11), 1348–1358. https://doi.org/10.1111/ecog.03031

Fyfe, R., Roberts, N., & Woodbridge, J. (2010). A pollen-based pseudobiomisation approach to anthropogenic land-cover change. *Holocene*, *20*(7), 1165–1171. https://doi.org/10.1177/0959683610369509

Goring, S., Dawson, A., Simpson, G. L., Ram, K., Graham, R. W., Grimm, E. C., & Williams, J. W. (2015). Neotoma: A programmatic interface to the neotoma paleoecological database. *Open Quaternary*, *1*, 1–17. https://doi.org/10.5334/oq.ab

Gribble, R. (2021). *Human impact on environment* [Painting]. Retrieved 09/13/2022 from https://www.richellegribble.com/the-nomadic-artist/2019/2/6/human-impact-on-environment-in-the-middle-of-nowhere

Gunnarson, B. E., Borgmark, A., & Wastegård, S. (2003). Holocene humidity fluctuations in Sweden inferred from dendrochronology and peat stratigraphy. *Boreas*, *32*(2), 347–360. https://doi.org/10.1080/03009480310001641

Hannon, G. E., Halsall, K., Molinari, C., Boyle, J., & Bradshaw, R. H. W. (2018). The reconstruction of past forest dynamics over the last 13,500 years in SW Sweden. *Holocene*, *28*(11), 1791–1800. https://doi.org/10.1177/0959683618788669

Haslett, J., & Parnell, A. (2008). A simple monotone process with application to radiocarbon-dated depth chronologies. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *57*(4), 399–418. https://doi.org/10.1111/j.1467-9876.2008.00623.x

Helama, S., Stoffel, M., Hall, R. J., Jones, P. D., Arppe, L., Matskovsky, V. V., … Oinonen, M. (2021). Recurrent transitions to Little Ice Age-like climatic regimes over the Holocene. *Climate Dynamics*, *56*(11–12), 3817–3833. https://doi.org/10.1007/s00382-021-05669-0

Jackson, S. T., & Williams, J. W. (2004). Modern analogs in quaternary paleoecology: Here today, gone yesterday, gone tomorrow? *Annual Review of Earth and Planetary Sciences*, *32*, 495–537. https://doi.org/10.1146/annurev.earth.32.101802.120435

Jørgensen, E. K. (2020). The palaeodemographic and environmental dynamics of prehistoric Arctic Norway: An overview of human-climate covariation. *Quaternary International*, *549*(November 2017), 36–51. https://doi.org/10.1016/j.quaint.2018.05.014

Josefsson, T., Hörnberg, G., Liedgren, L., & Bergman, I. (2017). Cereal cultivation from the Iron Age to historical times: evidence from inland and coastal settlements in northernmost Fennoscandia. *Vegetation History and Archaeobotany*, *26*(3), 259–276. https://doi.org/10.1007/s00334-016-0586-7

Josefsson, T., Ramqvist, P. H., & Hörnberg, G. (2014). The history of early cereal cultivation in northernmost Fennoscandia as indicated by palynological research. *Vegetation History and*

Archaeobotany, *23*(6), 821–840. https://doi.org/10.1007/s00334-014-0446-2

Kulik, R., & Soulier, P. (2020). Moving averages. *Springer Series in Operations Research and Financial Engineering*, 425–452. https://doi.org/10.1007/978-1-0716-0737-4_15

Kuosmanen, N., Marquer, L., Tallavaara, M., Molinari, C., Zhang, Y., Alenius, T., … Seppä, H. (2018). The role of climate, forest fires and human population size in Holocene vegetation dynamics in Fennoscandia. *Journal of Vegetation Science*, *29*(3), 382–392. https://doi.org/10.1111/jvs.12601

Lahtinen, M., Oinonen, M., Tallavaara, M., Walker, J. W. P., & Rowley-Conwy, P. (2017). The advance of cultivation at its northern European limit: Process or event? *Holocene*, *27*(3), 427–438. https://doi.org/10.1177/0959683616660164

Lahtinen, R., Hölttä, P., Kontinen, A., Niiranen, T., Nironen, M., Saalmann, K., & Sorjonen-Ward, P. (2011). Tectonic and metallogenic evolution of the Fennoscandian Shield: Key questions with emphasis on Finland. *Special Paper of the Geological Survey of Finland*, *2011*(49), 23–33.

Lechterbeck, J., Edinborough, K., Kerig, T., Fyfe, R., Roberts, N., & Shennan, S. (2014). Is Neolithic land use correlated with demography? An evaluation of pollen-derived land cover and radiocarbon-inferred demographic change from Central Europe. *Holocene*, *24*(10), 1297–1307. https://doi.org/10.1177/0959683614540952

Lindeløv, J. K. (2020). mcp: An R Package for Regression With Multiple Change Points. *Journal of Statistical Software*. Retrieved from https://osf.io/fzqxv/

Mayewski, P. A., Rohling, E. E., Stager, J. C., Karlén, W., Maasch, K. A., Meeker, L. D., … Steig, E. J. (2004). Holocene climate variability. *Quaternary Research*, *62*(3), 243–255. https://doi.org/10.1016/j.yqres.2004.07.001

Moen, A., & Lillethun, A. (1999). *National atlas of Norway: vegetation*. Norwegian mapping authority.

Mossberg, B., & Stenberg, L. (2018). *Nordens flora*. Bonnier fakta.

Neath, A. A., & Cavanaugh, J. E. (2012). The Bayesian information criterion: Background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, *4*(2), 199–203. https://doi.org/10.1002/wics.199

Nogués-Bravo, D., Rodríguez-Sánchez, F., Orsini, L., de Boer, E., Jansson, R., Morlon, H., … Jackson, S. T. (2018). Cracking the Code of Biodiversity Responses to Past Climate Change. *Trends in Ecology and Evolution*, *33*(10), 765–776. https://doi.org/10.1016/j.tree.2018.07.005

O'Dwyer, R., Marquer, L., Trondman, A. K., & Jönsson, A. M. (2021). Spatially Continuous Land-Cover Reconstructions Through the Holocene in Southern Sweden. *Ecosystems*, *24*(6), 1450–1467. https://doi.org/10.1007/s10021-020-00594-5

Pereira, H. M., Navarro, L. M., & Martins, I. S. (2012). Global biodiversity change: The Bad, the good, and the unknown. *Annual Review of Environment and Resources*, *37*(October), 25–50. https://doi.org/10.1146/annurev-environ-042911-093511

Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*, *27*(4 SE-Articles), 1–32. https://doi.org/10.18637/jss.v027.i04

Prunier, J. G., Colyn, M., Legendre, X., Nimon, K. F., & Flamand, M. C. (2015). Multicollinearity in spatial genetics: Separating the wheat from the chaff using commonality analyses. *Molecular Ecology*, *24*(2), 263–283. https://doi.org/10.1111/mec.13029

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., … Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth and Environment*, *3*(1), 1–10. https://doi.org/10.1038/s43247-022-00498-3

Ray-Mukherjee, J., Nimon, K., Mukherjee, S., Morris, D. W., Slotow, R., & Hamer, M. (2014). Using commonality analysis in multiple regressions: A tool to decompose regression effects in the face of multicollinearity. *Methods in Ecology and Evolution*, *5*(4), 320–328. https://doi.org/10.1111/2041-210X.12166

Refaeilzadeh, P., Tang, L., & Liu, H. (2018). Cross-Validation BT - Encyclopedia of Database Systems. In L. Liu & M. T. Özsu (Eds.) (pp. 677–684). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4614-8265-9_565

Reimer, P. J., Austin, W. E. N., Bard, E., Bayliss, A., Blackwell, P. G., Bronk Ramsey, C., … Talamo, S. (2020). The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0-55 cal kBP). *Radiocarbon*, *62*(4), 725–757. https://doi.org/10.1017/RDC.2020.41

Rounsevell, M. D. A., Harfoot, M., Harrison, P. A., Newbold, T., Gregory, R. D., & Mace, G. M. (2020). A biodiversity target based on species extinctions. *Science*, *368*(6496), 1193–1195. https://doi.org/10.1126/science.aba6592

Shojaie, A., & Fox, E. B. (2022). Granger Causality: A Review and Recent Advances. *Annual Review of Statistics and Its Application*, *9*(1969), 289–319. https://doi.org/10.1146/annurev-statistics-040120-010930

Sjögren, P., & Damm, C. (2019). Holocene vegetation change in northernmost Fennoscandia and the impact on prehistoric foragers 12 000–2000 cal. a BP–A review. *Boreas*, *48*(1), 20–35.

SLU. (2022). *Artbestämning - Artfakta från SLU Artdatabanken*. Retrieved 09/13/2022 from https://artfakta.se/artbestamning

Sørensen, L., & Karg, S. (2014). The expansion of agrarian societies towards the north - new evidence for agriculture during the Mesolithic/Neolithic transition in Southern Scandinavia. *Journal of Archaeological Science*, *51*, 98–114. https://doi.org/10.1016/j.jas.2012.08.042

Stroeven, A. P., Hättestrand, C., Kleman, J., Heyman, J., Fabel, D., Fredin, O., … Jansson, K. N. (2016). Deglaciation of Fennoscandia. *Quaternary Science Reviews*, *147*, 91–121. https://doi.org/10.1016/j.quascirev.2015.09.016

Svetunkov, I. (2017). Statistical models underlying functions of "smooth" package for R. *Working Paper of Department of Management Science*, 1–52. Retrieved from http://www.lums.lancs.ac.uk

Team, R. C. (2013). R: A language and environment for statistical computing.

Tyler, T., Herbertsson, L., Olofsson, J., & Olsson, P. A. (2021). Ecological indicator and traits values for Swedish vascular plants. *Ecological Indicators*, *120*(August 2020), 106923. https://doi.org/10.1016/j.ecolind.2020.106923

UCAR. (2020). *CCSM3.0 IPCC vs. non-IPCC runs*. Retrieved 09/13/2022 from https://www.cesm.ucar.edu/models/ccsm3.0/ipcc_definition.html

Ulrich, M. J. (2021). Package ' TTR .'

Wickham, H. (2022). Package ' modelr .'

Williams, J. W., Grimm, E. C., Blois, J. L., Charles, D. F., Davis, E. B., Goring, S. J., … Takahara, H. (2018). The Neotoma Paleoecology Database, a multiproxy, international, community-curated data resource. *Quaternary Research (United States)*, *89*(1), 156–177. https://doi.org/10.1017/qua.2017.105

Woodbridge, J., Fyfe, R., Smith, D., Pelling, R., de Vareilles, A., Batchelor, R., … Davies, A. L. (2021). What drives biodiversity patterns? Using long-term multidisciplinary data to discern centennial-scale change. *Journal of Ecology*, *109*(3), 1396–1410. https://doi.org/10.1111/1365-2745.13565

Appendix

Popular Science Summary

Popular science summary
Holocene vegetation change and its correlates in Fennoscandia
Raphaela Infanger

Early humans left more traces than we might be aware of

Our time travel through Fennoscandia starts after the retreat of the last big glaciers at about 9700 BP and lasts till the end of the Medieval times at about 500 BP. This timespan includes the first human settlements, the development of a range of different working tools, the arrival of agriculture and cattle and many different cultural epochs. All these events are known to have influenced vegetational changes in different Fennoscandian areas. But we want to look at the big picture; the whole area over nearly 10000 years' time.

> **Fennoscandia**, consisting of Norway, Sweden, Finland and a small eastern part of Russia, is distinct in its geological and anthropological history.
> Its vegetation zones include (from north to south) the southern arctic, alpine, northern boreal, middle boreal, southern boreal, boreo-nemoral and nemoral zone

When we discuss modern changes, we talk about the environment that surrounds us; its biodiversity loss, how it looked to us some years ago and how it does now. But what was before the time that we remember? How did the places we know look like many thousand years ago? How do we know that "wild" and "natural" environments really never have been impacted and changed by humans? I guess we don't and never quite will know, but the growing amount of knowledge and methods allows us to gain some more insight in the past every day. To know if, when and how humans have impacted changes in vegetation is essential to be able to assess these questions on modern human impacts.

Fossil pollen data, stemming from lakes and bogs, show us what type of vegetation was dominant at the different time epochs and how this dominance changed. If we see changes in vegetation structures, let's say that coniferous forests suddenly became much less but meadows became more abundant, then we will check for the cause of these changes. For this, we reconstruct human population sizes with data on archaeological remains, which will give us an idea of the anthropological pressure at different times. Finally, climate data can tell us if there were fluctuations in temperature that could have been the cause for shifts in vegetation. With this knowledge we hope to be able to tell to what amount humans have influenced changes in vegetation and also what humanitarian actions could lay behind it, as for example certain agricultural practices, the introduction of cattle, forestry practices, and so on.

BACK TO THE FUTURE: In the end, it is crucial to have knowledge on the impact humans and climate have had in the past and how this impact and the practices behind it have expressed themselves on landscapes today. We not only show that humans in Fennoscandia have been the main drivers of vegetation change for a long time, but also that these impacts grew with the onset of farming. We also saw that climate was a strong force in the centuries following the glacial retreat, especially in the times before the onset of farming. Finally, indications on the establishment and abandonment of different land-use practices have been found and/or recovered, which helps us to better understand the history that shaped Fennoscandia's landscape and can help us to better take care of this diversity in the future.

Figure 10: Popular Science Summary of this work

The correlated datasets

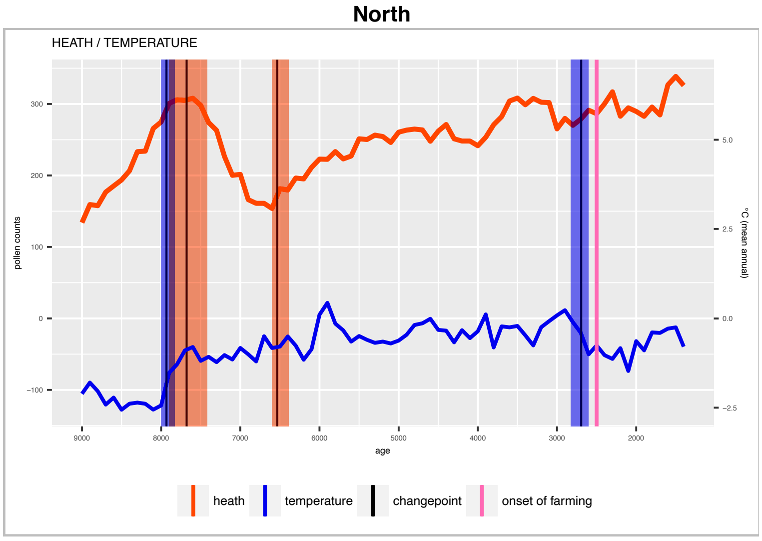

Figure 11: The changepoints of the significant correlation between the heaths and temperature that has been found for the North.

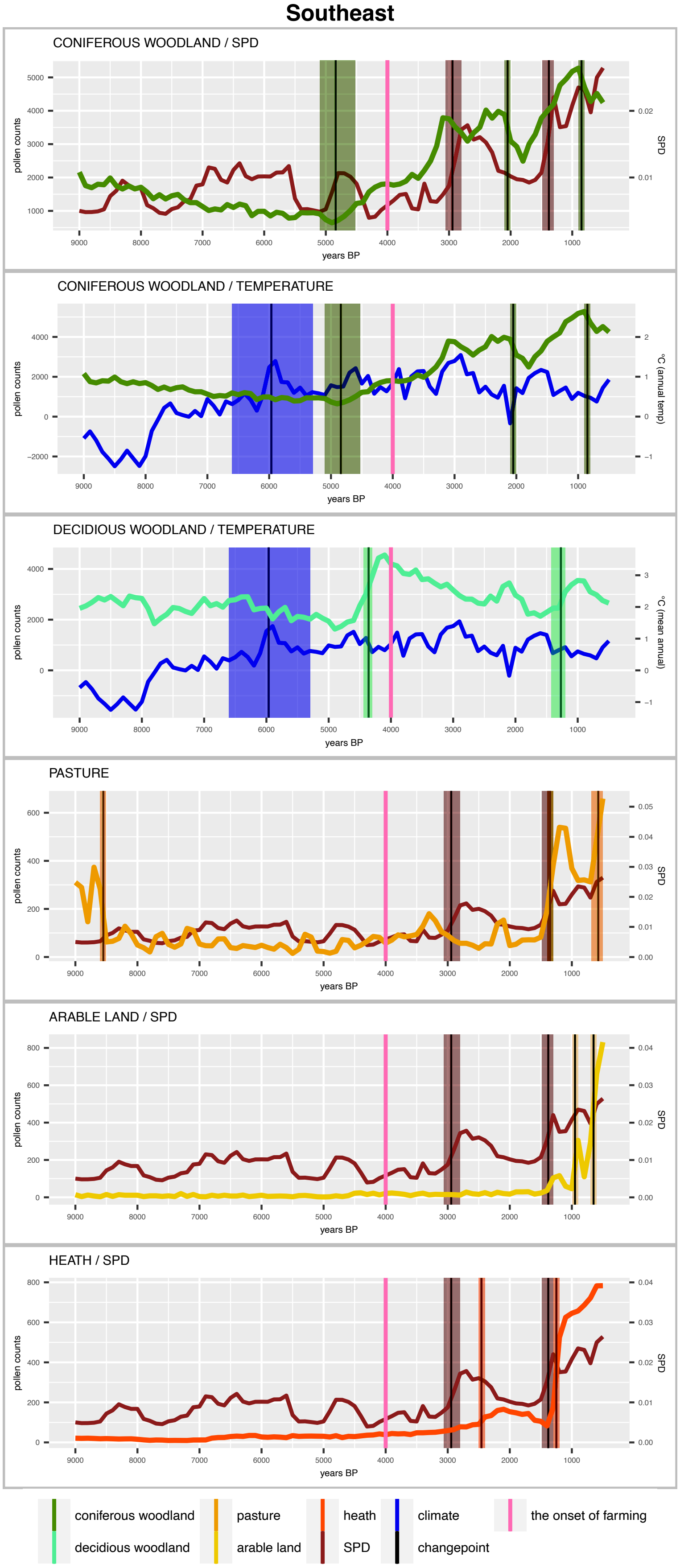

Figure 12: The aligned change points of the south-eastern datasets that correlated with each other with both temperature and SPD as predictors.

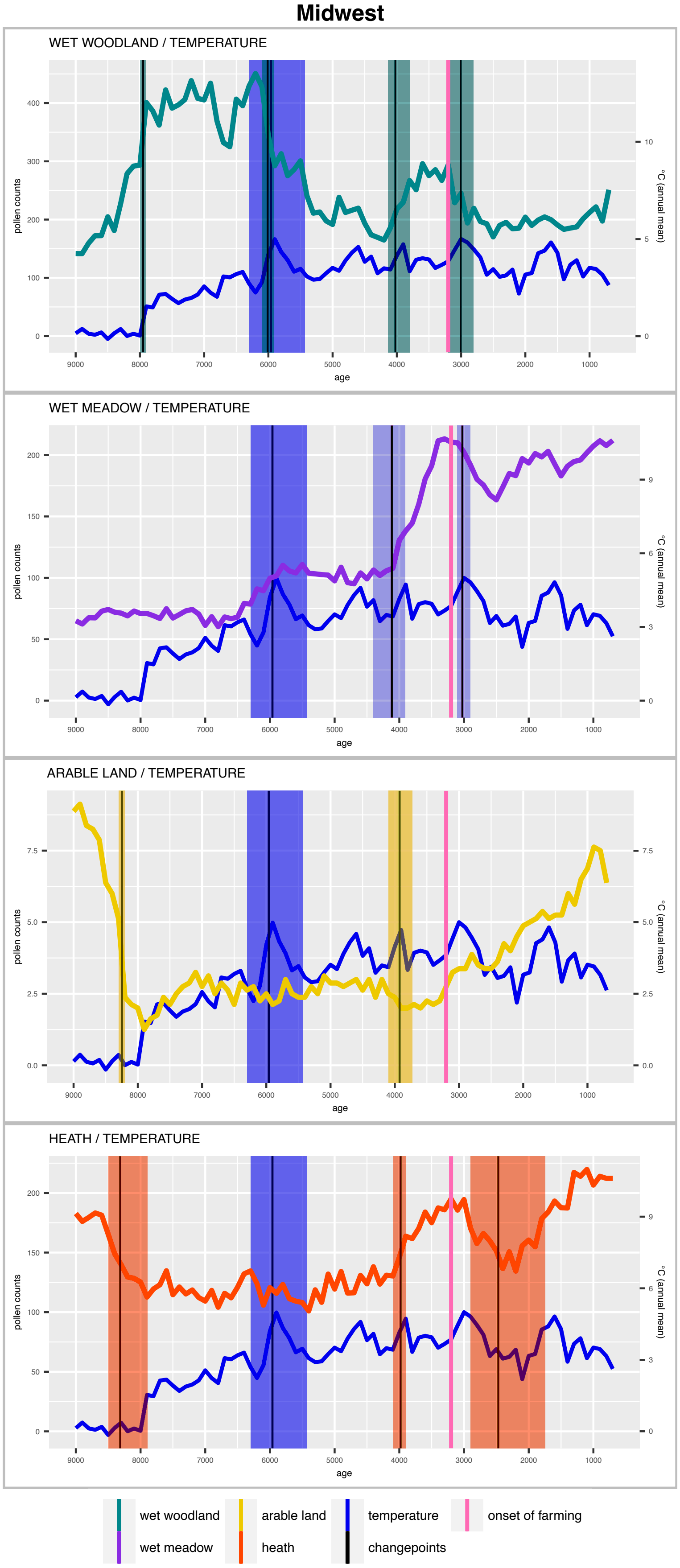

Figure 13: The aligned change points of the midwestern datasets that correlated with each other with temperature as a predictor.

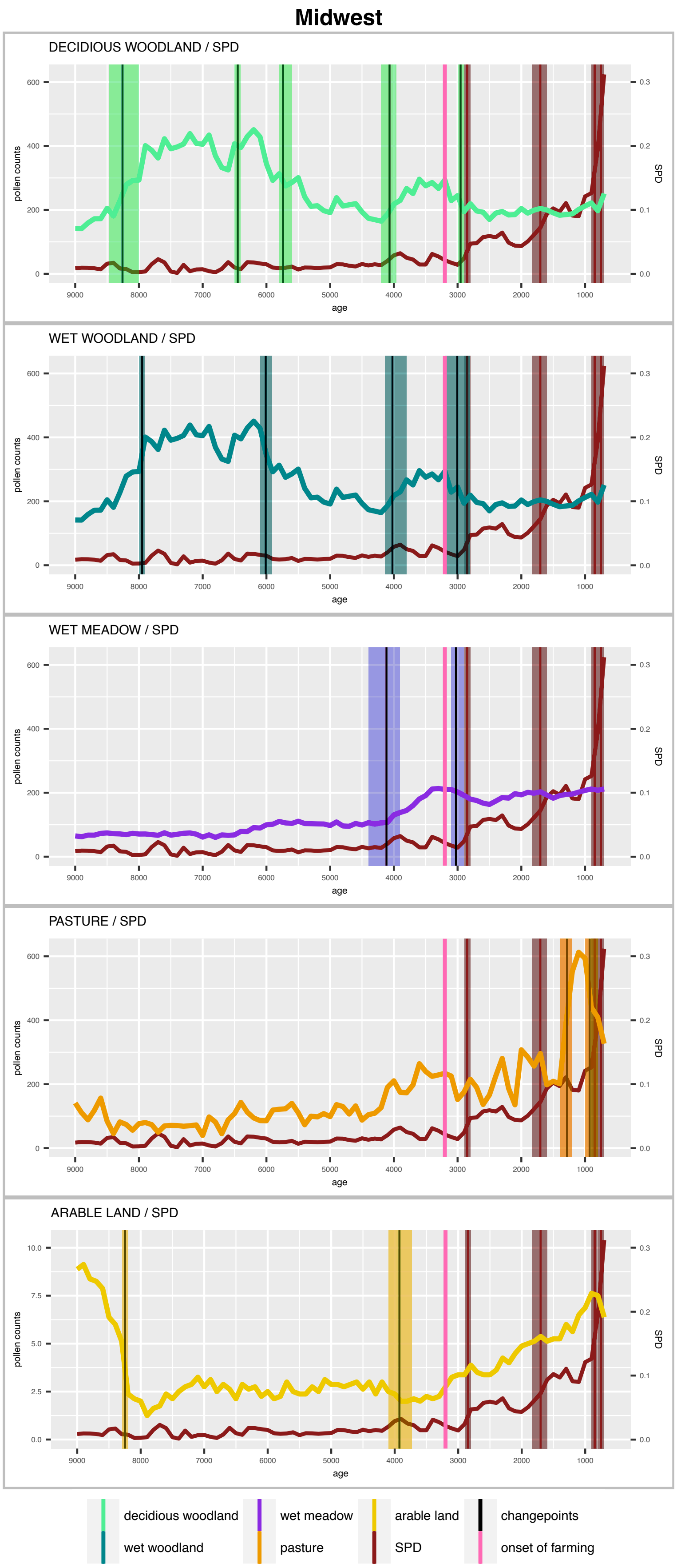

Figure 14: The aligned change points of the midwestern datasets that correlated with each other with SPD as a predictor.

Figure 15: The aligned change points of the central Fennoscandian datasets that correlated with each other with temperature as a predictor.

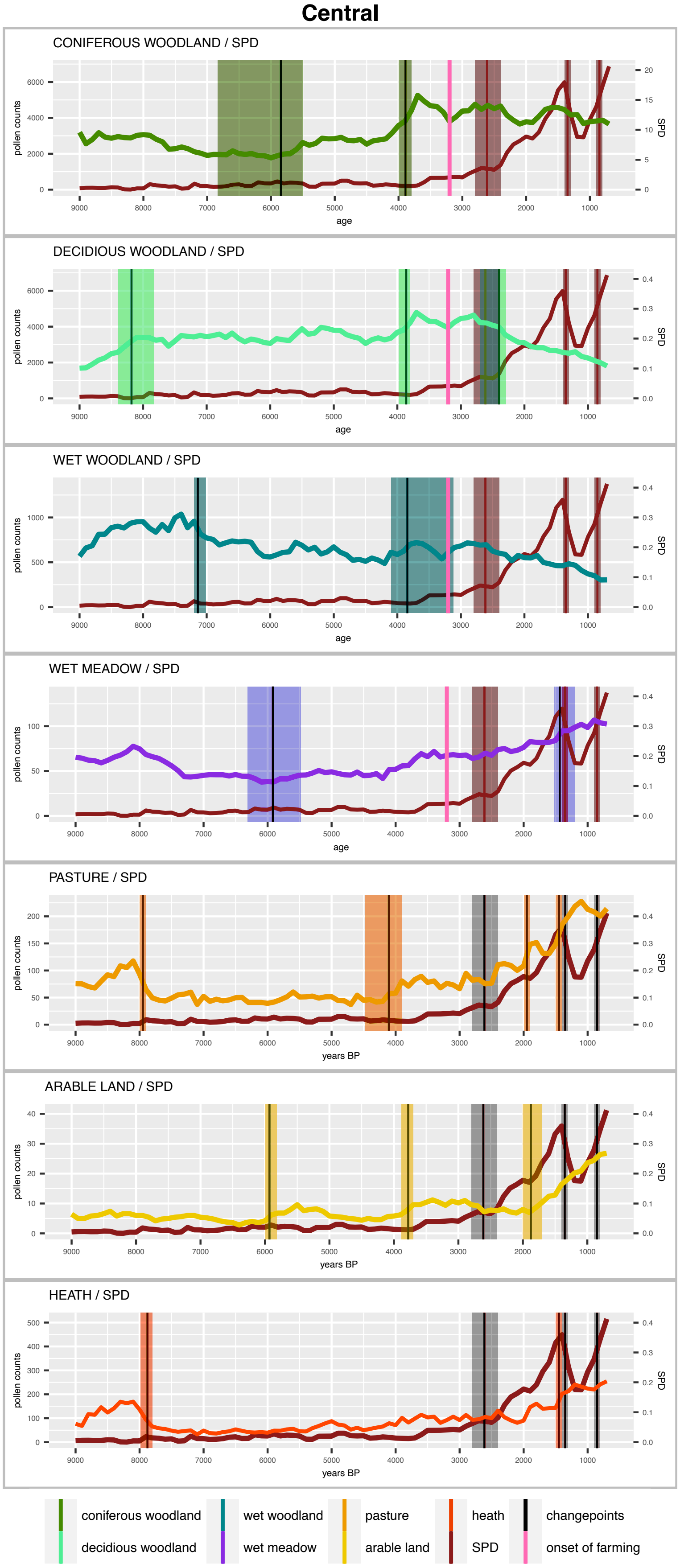

Figure 16: The aligned change points of the central Fennoscandian datasets that correlated with each other with SPD as a predictor.

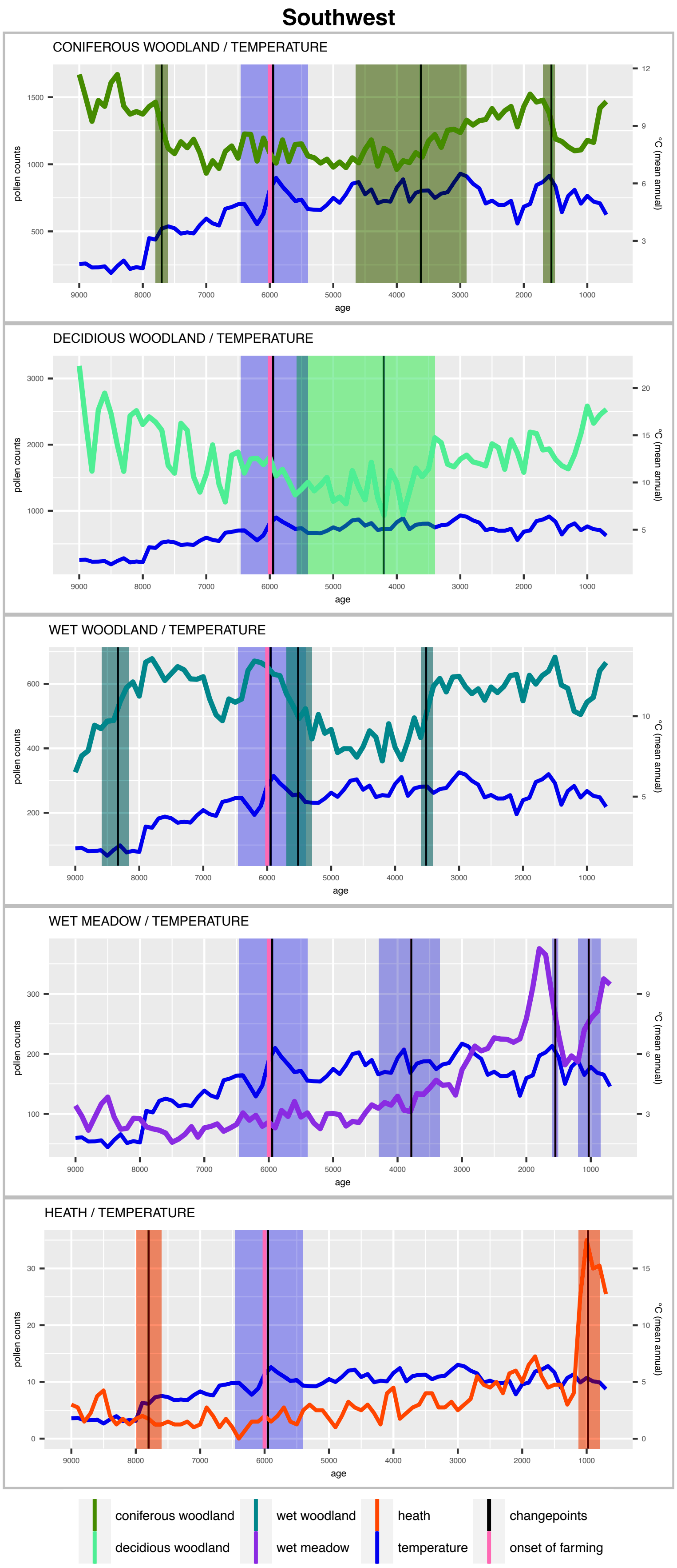

Figure 17: The aligned change points of the southwestern datasets that correlated with each other with temperature as a predictor.

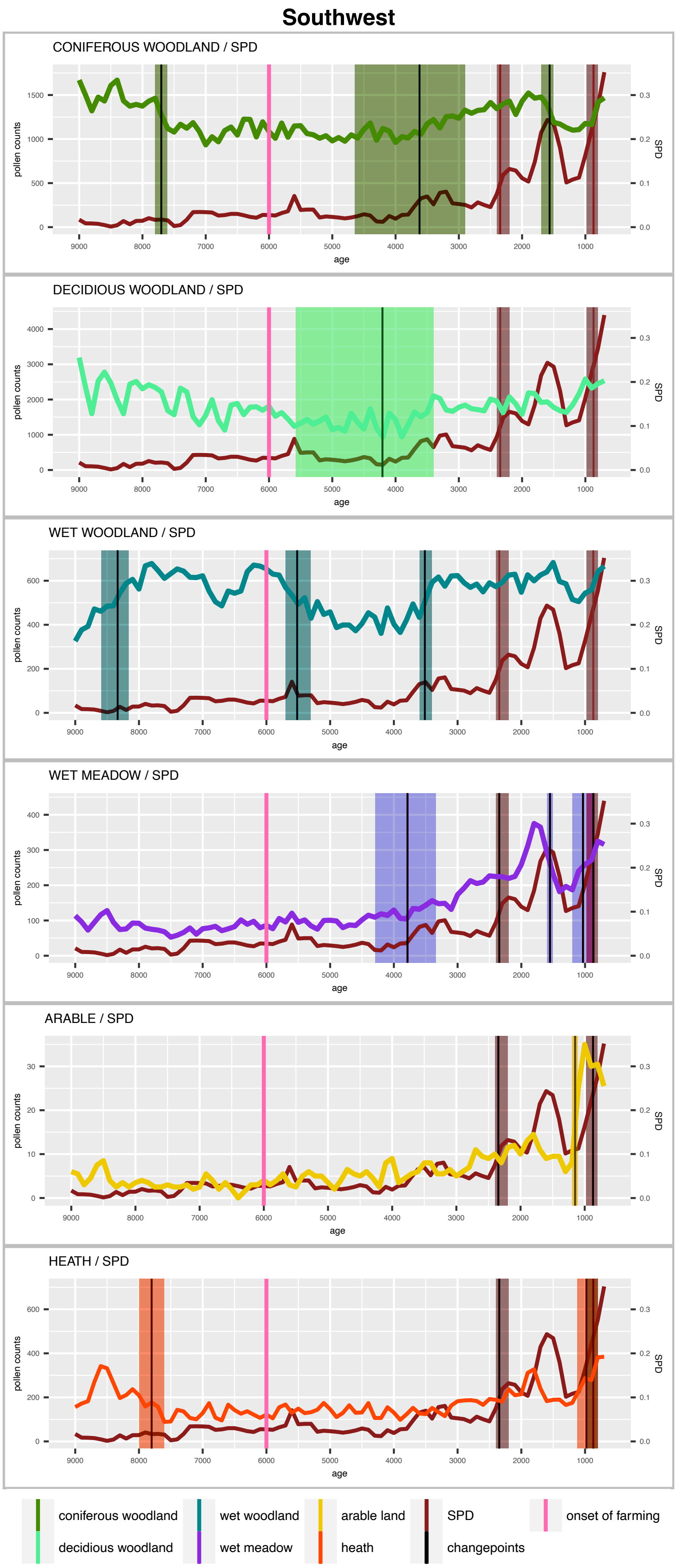

Figure 18: The aligned change points of the southwestern datasets that correlated with each other with SPD as a predictor.

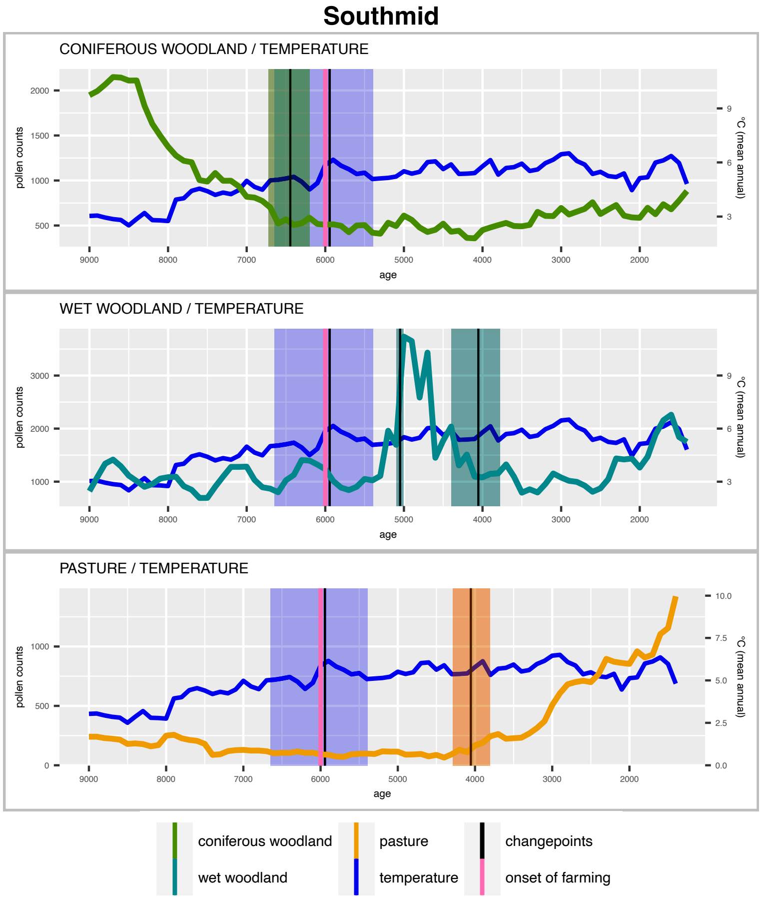

Figure 19: The aligned change points of the southcentral datasets that correlated with each other with temperature as a predictor.

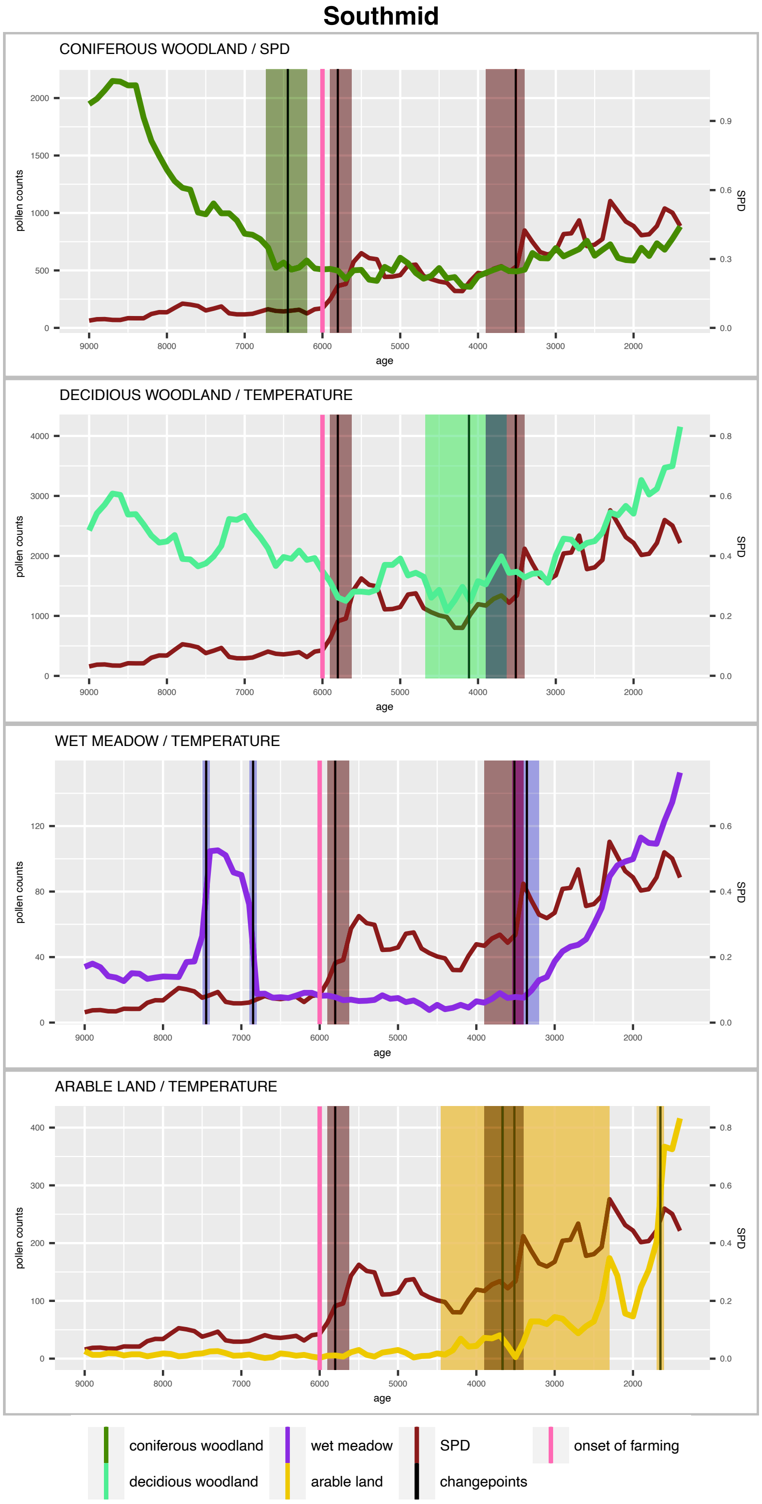

Figure 20: The aligned change points of the southcentral datasets that correlated with each other with SPD as a predictor.

Data availability

The change points

Table 7: The significant change points (Mean age) are given for every dataset of every region, including the lower boundary (Min age) and the upper boundary (Max age) of their confidence intervals.

Region	Dataset	Mean age	Min age	Max age
North	SPD	2410	2309	2499
		3447	3384	3500
	Climate	2694	2600	2824
		7932	7825	7999
	coniferous woodland	2361	2000	2706
		5243	5185	5300
		6161	6100	6370
		7148	6800	7506
	deciduous woodland	3007	2735	3299
		7037	6867	7199
	wet woodland	5169	4922	5299
		7957	7900	8034
	wet meadow	2942	2898	2999
		5067	5000	5192
	pasture	6939	6821	6999
	arable land	1855	1725	1999
	heath	6533	6385	6599
		7678	7414	7899
Southeast	SPD	1380	1300	1484
		2944	2800	3059
	Climate	5964	5294	6601
	coniferous woodland	849	802	899
		2049	2002	2098
		4840	4524	5100
	deciduous woodland	1270	1200	1428
		4358	4300	4438
	wet woodland	4339	4200	4455
	wet meadow	1248	1166	1299
		2594	2500	2693
		4742	4632	4800
		8193	8001	8399
	pasture	574	500	681
		1350	1304	1399
		8549	8504	8599

	arable land	650	603	698
		949	902	996
	heath	1249	1202	1297
		2456	2400	2505
Midwest	SPD	750	705	799
		850	805	899
		1702	1600	1833
		2848	2800	2895
	Climate	5961	5431	6296
	coniferous woodland	2899	2810	2999
		3939	3900	4077
		7127	6334	7399
	deciduous woodland	2953	2900	2999
		4068	3958	4199
		5739	5598	5799
		6450	6403	6498
		8258	8000	8472
	wet woodland	3005	2806	3199
		4023	3803	4139
		6001	5913	6099
		7949	7903	7998
	wet meadow	3025	2900	3102
		4116	3906	4399
	pasture	928	804	999
		1283	1200	1385
	arable land	3925	3730	4099
		8248	8200	8294
	heath	2470	1745	2899
		3981	3900	4087
		8315	7894	8499
Central	SPD	851	804	899
		1350	1304	1399
		2613	2399	2800
	Climate	5931	5401	6299
	coniferous woodland	3891	3800	3997
		5845	5494	6835
	deciduous woodland	2399	2290	2697
		3860	3799	3976
		8182	7831	8399

	wet woodland	3842	3116	4095
		7139	7015	7199
	wet meadow	1436	1200	1518
		5918	5484	6315
	pasture	1448	1401	1497
		1950	1900	1996
		4106	3900	4479
		7948	7900	7996
	arable land	1877	1700	1999
		3777	3700	3882
		5931	5816	5999
	heath	1450	1404	1499
		7877	7800	7987
Southwest	SPD	871	800	978
		2346	2200	2400
	Climate	5947	5400	6459
	coniferous woodland	1565	1500	1695
		3621	2901	4643
		7702	7609	7799
	deciduous woodland	4207	3400	5576
	wet woodland	3513	3405	3599
		5517	5302	5698
		8334	8161	8590
	wet meadow	1032	849	1199
		1550	1503	1599
		3786	3339	4290
	pasture	1053	999	1139
		3032	2700	3259
		7733	6824	8099
	arable land	1150	1104	1199
	heath	978	800	1125
		7801	7605	7999
Southcentral	SPD	3514	3400	3900
		5802	5623	5900
	Climate	5944	5394	6650
	coniferous woodland	6445	6197	6724
	deciduous woodland	4116	3636	4678

	wet woodland	4053	3778	4398
		5049	3778	4398
	wet meadow	3355	3196	3540
		6850	6805	6899
		7449	7401	7496
	pasture	4054	3803	4284
	arable land	1650	1601	1696
		3665	2300	4453
	heath	x	x	x

The pollen sites included in the study

Table 8: The pollen sites with their longitudes and latitudes for the North that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
20	27.67406	69.12326
317	28.407205	70.18109
720	31.025865	70.325835
4257	20.97475	69.04527
4286	21	68.916667
4372	20.3166667	69.35
4468	28.83694	69.603
20034	20.716667	69.166667
20279	25.36871	71.074485
20285	28.169065	71.038055
20293	29.295155	70.69954
44941	16.3814	68.660445
45311	19.1	68.345556
45636	14.93089	68.52691
45639	17.75935	68.48221
45642	18.071667	68.444167
24757	32.366667	69.633333
4169	33.666667	67.833333

Table 9: The pollen sites with their longitudes and latitudes for the Southeast that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
4092	24.25122	60.62404
4133	25.86881	61.41513
4156	20.14137	60.29171
4168	23.85616	60.78455
4259	26.23393	62.34206
4393	24.17711	60.438365

4420	22.066667	61.3
4472	28.116667	61.216667
4539	22.866667	62.133333
4543	36.25	63.9
4017	33.85	62.316667
3928	32.766667	62.033333

Table 10: The pollen sites with their longitudes and latitudes for the Midwest that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
977	8.689085	63.702905
20042	7	60.716667
20046	8.666667	61.416667
45331	8.454444	63.4575
45345	9.416944	63.059722
45347	9.833333	62.266667
45349	9.876667	61.649167

Table 11: The pollen sites with their longitudes and latitudes for the central region that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
19906	13.557865	62.324285
19909	16.40472	61.82125
19913	14.45	63.883333
20018	10.398245	62.331455
21790	14.92652	60.65189
45351	10.267778	61.5625
45698	10.874495	60.83513
45701	10.35931	61.108565

Table 12: The pollen sites with their longitudes and latitudes for the Southwest that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
20050	6	59.816667
45704	7.98669	59.626125
45707	7.305195	59.34377
45710	7.540515	59.67017
45713	7.252055	59.79602
45716	6.992657	59.8420545
45719	7.734035	58.537409
45722	7.78555	58.3260095
45725	8.004555	58.244145
45728	7.913611	58.14445
45731	7.434405	59.764875

Table 13: The pollen sites with their longitudes and latitudes for the Southcentral that have been retrieved from the Neotoma Database. site=site number registered in Neotoma.

site	longitude	latitude
12	13.42774	55.934475
1438	14.53458	57.64055
4403	12.41426	57.16313
45329	11.617778	59.059722

The taxa included in every LCC

Table 14: All the taxa that have been grouped to the LCCs "coniferous woodland", "deciduous woodland", "wet woodland", "wet meadow", "pasture", "arable land" and "heath" are listed here.

coniferous woodland	deciduous woodland	wet woodland	wet meadow	pasture	arable land	heath
Abies	*Acer*	*Alnus*	*Angelica*	*Achillea*	*Arctium*	*Arctous.alpina*
Arctostaphylos	*Adoxa.moschatellina*	*Alnus.glutinosa*	*Caltha*	*Alchemilla*	*Artemisia*	*Calluna*
Arctostaphylos.uva.ursi	*Anemone*	*Alnus.incana*	*Caltha.palustris*	*Anthemis*	*Artemisia.norvegica*	*Calluna.vulgaris*
Huperzia	*Anemone.nemorosa*	*Alnus.viridis*	*Carex*	*Apiaceae*	*Asteroideae*	*Cassiope*
Huperzia.selago	*Betula*	*Athyrium.alpestre*	*Cladium.mariscus*	*Arenaria*	*Avena.Triticum*	*Diapensia*
Linnaea.borealis	*Betula.Corylus.Myrica*	*Athyrium.filix.femina*	*Cyperaceae*	*Armeria*	*Cannabaceae*	*Diapensia.lapponica*
Lycopodium	*Betula.fruticosa*	*Blechnum.spicant*	*Cyperaceae.1*	*Armeria.Limonium*	*Cannabis*	*Dryas.octopetala*
Lycopodium.annotinum	*Betula.humilis*	*Chrysosplenium*	*Drosera*	*Aster*	*Cannabis.sativa*	*Empetrum*
Melampyrum	*Betula.nana*	*Frangula*	*Drosera.anglica*	*Asteraceae*	*Capsella*	*Empetrum.nigrum*
Picea	*Betula.pendula*	*Frangula.alnus*	*Drosera.intermedia*	*Bistorta*	*Caryophyllaceae*	*Erica.arborea*
Picea.abies	*Betula.pubescens*	*Lysimachia.vulgaris*	*Drosera.rotundifolia*	*Bistorta.officinalis.B..vivipara*	*Centaurea.cyanus*	*Erica.cinerea.E..tetralix*
Pinus	*Carpinus*	*Matteuccia*	*Drosera.rotundifolia.D..anglica*	*Bistorta.vivipara*	*Cerealia*	*Ericaceae*
Pinus.sylvestris	*Carpinus.betulus*	*Matteuccia.struthiopteris*	*Droseraceae*	*Botrychium*	*Fagopyrum*	*Ericales*
Taxus	*Convallaria*	*Phegopteris.connectilis*	*Equisetum*	*Botrychium.lunaria*	*Fagopyrum.esculentum*	*Jasione*
Taxus.baccata	*Convallaria.majalis*	*Rhododendron.tomentosum*	*Eriophorum*	*Brassica*	*Galeopsis*	*Jasione.montana*
Vaccinium	*Cornus*	*Rubus*	*Filipendula*	*Brassicaceae*	*Hordeum*	*Juniperus*
	Cornus.mas.C..suecica	*Rubus.chamaemorus*	*Juncaceae*	*Campanula*	*Humulus*	*Juniperus.communis*

	Cornus.sanguinea	*Rubus.fruticosus*	*Lentibulariaceae*	*Campanulaceae*	*Humulus.Cannabis*	*Lotus.corniculatus*
	Cornus.suecica	*Rubus.idaeus*	*Lotus*	*Caprifoliaceae*	*Humulus.lupulus*	
	Corylus	*Rubus.saxatilis*	*Lotus.pedunculatus*	*Centaurea.jacea*	*Juglans*	
	Corylus.avellana	*Salix*	*Menyanthes.trifoliata*	*Centaurea.nigra*	*Juglans.regia*	
	Corylus.Myrica	*Salix.herbacea*	*Micranthes.stellaris*	*Centaurea.scabiosa*	*Linum.usitatissimum*	
	Fagus	*Sambucus*	*Narthecium.ossifragum*	*Cerastium*	*Lycopsis.arvensis*	
	Fagus.sylvatica	*Sambucus.nigra.S..racemosa*	*Ophioglossum*	*Cerastium.fontanum*	*Papaver.radicatum*	
	Fraxinus	*Saussurea*	*Ophioglossum.vulgatum*	*Cichorieae*	*Papaver.rhoeas*	
	Fraxinus.excelsior	*Saussurea.nuda*	*Oxyria*	*Cichorioideae*	*Persicaria.maculosa*	
	Hedera	*Sphagnum*	*Oxyria.digyna*	*Cirsium*	*Plantago.major*	
	Hedera.helix		*Parnassia*	*Cirsium.Carduus*	*Plantago.major.P..media*	
	Ilex		*Parnassia.palustris*	*Dianthus*	*Polygonum.aviculare*	
	Larix		*Pedicularis*	*Dipsacoideae*	*Rumex*	
	Lonicera.periclymenum		*Pedicularis.palustris*	*Euphrasia*	*Rumex.longifolius*	
	Mercurialis		*Peucedanum*	*Fabaceae*	*Rumex.Oxyria*	
	Mercurialis.perennis		*Pinguicula*	*Galium*	*Rumex.Oxyria.digyna*	
	Micranthes.nivalis		*Plantago.maritima*	*Gentiana*	*Scleranthus.annuus*	
	Micranthes.nivalis.M..tenuis		*Rhynchospora*	*Gentiana.campestris*	*Secale*	
	Polygonatum		*Rhynchospora.alba*	*Gentiana.nivalis*	*Secale.cereale*	
	Polygonatum.verticillatum		*Scheuchzeria*	*Gentiana.pneumonanthe*	*Sinapis*	
	Populus		*Scheuchzeria.palustris*	*Gentiana.purpurea*	*Spergula*	
	Populus.tremula		*Schoenus*	*Gentianaceae*	*Spergula.arvensis*	
	Prunus		*Selaginella*	*Gentianella.campestris*	*Spergula.Spergularia*	
	Prunus.padus		*Selaginella.selaginoides*	*Gentianella.detonsa*	*Triticum*	

Quercus	*Silene.dioica*	*Geraniaceae*
Rhamnus	*Silene.flos.cuculi*	*Geranium*
Rhamnus.cathartica	*Teucrium*	*Geranium.sylvaticum*
Sanicula	*Thelypteris.palustris*	*Geum*
Sorbus	*Tofieldia*	*Geum.rivale*
Sorbus.aria	*Tofieldia.pusilla*	*Helianthemum*
Sorbus.aucuparia	*Trollius*	*Hypericum*
Stachys.sylvatica	*Trollius.europaeus*	*Hypericum.calycinum*
Tilia	*Valeriana*	*Hypericum.perforatum*
Tilia.cordata	*Valeriana.excelsa*	*Hypericum.pulchrum*
Ulmus	*Valeriana.officinalis*	*Knautia.arvensis*
Ulmus.glabra		*Lathyrus*
Viburnum		*Onobrychis*
Viburnum.opulus		*Pimpinella*
Viscum		*Pimpinella.saxifraga*
Viscum.album		*Plantago*
		Plantago.lanceolata
		Plantago.media
		Poaceae
		Polygala
		Primula
		Primula.farinosa
		Ranunculus
		Ranunculus.acris
		Rhinanthus
		Rubiaceae
		Rumex.acetosa
		Rumex.acetosa.R..acetosella
		Rumex.acetosella

	Sanguisorba
	Sanguisorba.officinalis
	Serratula
	Silene.vulgaris
	Stellaria
	Succisa
	Succisa.pratensis
	Taraxacum
	Thalictrum
	Trifolium
	Trifolium.pratense
	Trifolium.repens
	Urtica
	Veronica
	Vicia
	Vicia.cracca
	Vicia.Lathyrus
	Vicia.sylvatica
	Viola.palustris