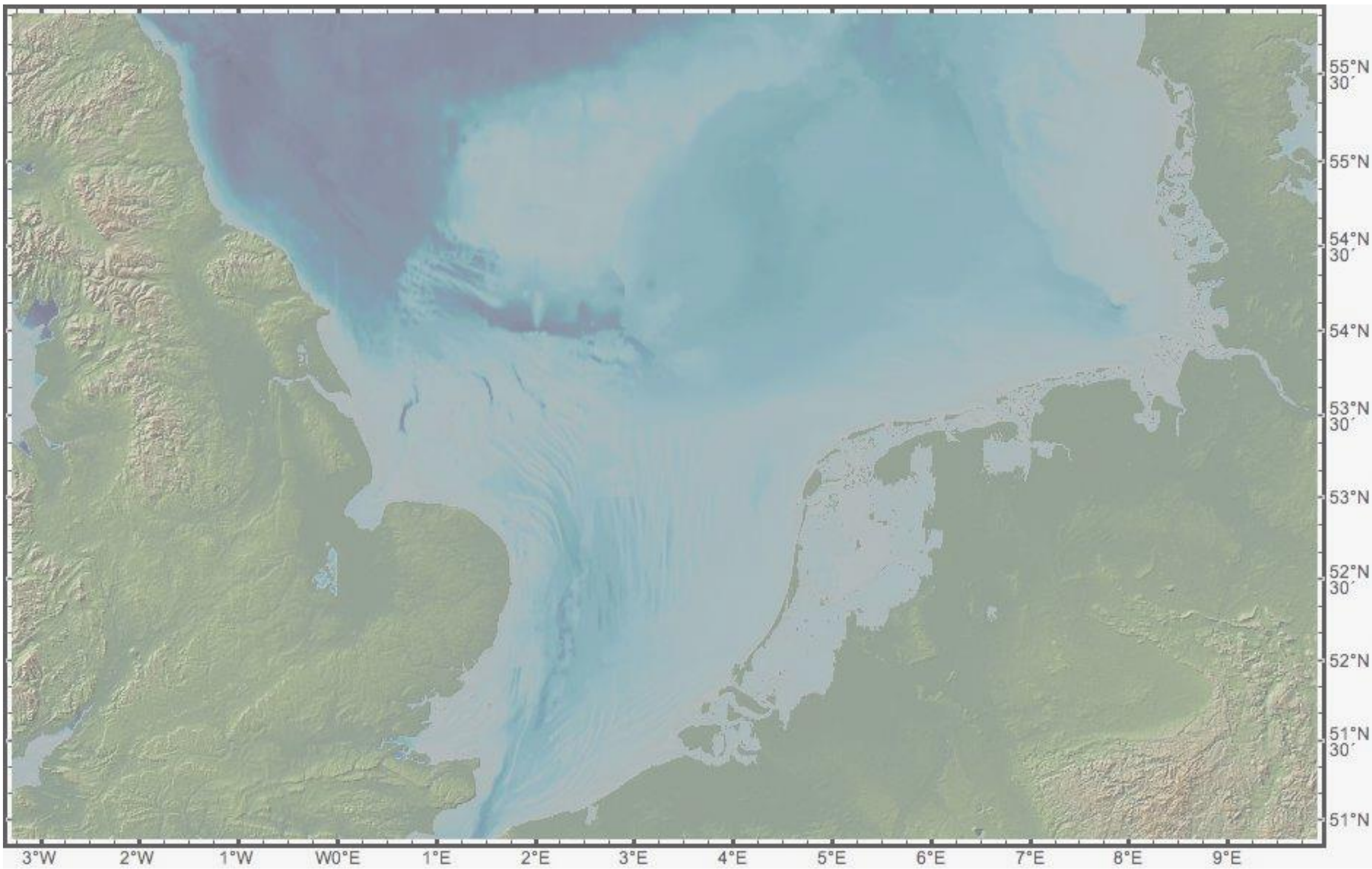



# CLIMATE RECONSTRUCTION FOR THE LATE PLIOCENE – EARLY PLEISTOCENE

## WITH QUANTITATIVE PALEO-CLIMATE METHODS FOR THE NETHERLANDS

LISANNE KROM (4250419)
UNIVERSITY OF UTRECHT
Department of Physical Geographhy

# Climate reconstruction for the Late Pliocene and Early Pleistocene – with quantitative paleo-climate methods for the Netherlands

Lisanne Jeanette Krom
4250419

In partial fulfilment of the degree of Master of Science in the Earth Science

Department of Physical Geography
Utrecht University

Supervisors:
Dr. Timme Donders
Dr. Thomas Giesecke

June 17, 2021

*Image on the front page from: By Marine Geoscience Data System - GeoMapApp, Public Domain, https://commons.wikimedia.org/w/index.php?curid=38675130*

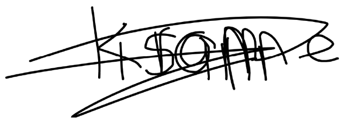
**I declare that:**

1. This is an original report, which is entirely my own work;
2. where I have made use of the ideas of other writers, I have acknowledged the source of all instances,
3. where I have used any diagram or visuals, I have acknowledged the source in all instances,
4. this report has not and will not be submitted elsewhere for academic assessment in any other academic course.

**Student data**:

Name: Lisanne Jeanette Krom
Registration number: 4250419
Data**:** June 17, 2021

Signature:



# Preface and Acknowledge

This guided research about quantitative paleo-climatic reconstruction for the Late Pliocene and Early Pleistocene for the Netherlands has been written as part of the Master Science degree in Earth Science, programme Marine Science at Utrecht University. The research was performed under the supervision of Dr. Timme Donders and Dr. Thomas Giesecke. The aim of this guided research is to reconstruct the climate during the Late Pliocene – Early Pleistocene for the Netherlands based on pollen assemblage data with quantitative paleo-climatic reconstruction methods. Three different fossil data (Noordwijk, Petten and Hank) sets are used to make the reconstruction, whereby the dataset from Whitmore is used as reference data set. The analysing program C2 is used to reconstruct the winter temperature and total precipitation. The technique used for this guided research are transfer functions using Modern Analogue Technique, Weighted Average and Weighted Average Partial Least Squares algorithms. It was a very interesting subject for me because I was unfamiliar with this way of reconstruction. I have learned a lot about reconstruction methods and the way to use them.

First, I would like to thank my supervisor Timme Donders for his great effort and enthusiasm to complete my guided research. I could always text him or we had interesting teams meetings to explore the analysing program and to discuss the data. It was hard for me to work every day on the same desk in the same room as I lived because of the Covid-19 pandemic, however, due to the effort of my friends, roommates, and parents I finished this guided research on time.

## Abstract

This guided research aims to reconstruct the climate during the Late Pliocene / Early Pleistocene for the Netherlands based on pollen assemblage data with quantitative paleo-climatic reconstruction methods. The Late Pliocene and Early Pleistocene are important Epochs because of their characteristics by cyclic growth and decay of terrestrial ice sheets related to the Milankovitch cycles. The Late Pliocene through Early Pleistocene is a frequently targeted interval for paleo-environmental reconstructions because it is considered as an analogue for future climate change. The focus of this guided research is on the Netherlands. Three different terrestrial sections are used from different parts of the Netherlands: Noordwijk, Petten and Hank. The Noordwijk section is from the south-western part of the Netherlands and has a time span from 2.61 – 2.11 Ma years ago, Early Pleistocene. The Petten section has an age range of 1.35 – 2.41 Ma years ago, Early Pleistocene and the Hank section originates from the Late Pliocene (5.45 – 2.48 Ma years ago).

The modern pollen-based quantitative paleo-climatology has undergone the development of a diverse array of statistical techniques to transform fossil assemblage data into past climate estimates. Three main approaches for quantitative reconstruction of past climates from bio-stratigraphical data are known: indicator species approach, assemblage approach and multi-variate transfer function approach. The multi-variate calibration function approach involves underlying statistical models with global estimation of parametric functions for all the taxa present. The analysing program C2 is used to reconstruct the winter temperature and total precipitation. Pollen-climate transfer functions were calculated using WA with inverse deshrinking and 3- component WA-PLS. Calibration set species data values were square-root transformed for WA and WA-PLS regression to reduce noise in the data. Performance statistics were computed for each transfer function, these include the coefficient of determination ($R^2$), Root Mean Square Error of Predication (RMSEP) and maximum bias. Both reconstructions are compared with the percentage deciduous trees in the samples and known data from Dearing Crampton-Flood et al., (2020) for the Late Pliocene and Zagwijn (1992) for Early Pleistocene. From these results, WA seems to fit the best in Late Pliocene data and WA-PLS fits best in the younger data.

WA seems to fit better in the Late Pliocene and WA-PLS fits better for the Early Pleistocene. The problem for WA-PLS is the that it tries to fit all the variation in the reconstruction, for older data this is not the right way because not all pollen taxa do exist today in the study area. Some mistakes inevitable if WA-PLS is used for older data than the Pleistocene. Further research requires more datapoints to make a better connection with other reconstructions.

# Contents

# 1. Introduction

Paleo-ecological and paleo-climatic research is done on almost every region on the world and have produced many fossil pollen records. The fossil pollen records are used to reconstruct past climates for many places on Earth. Most of the paleo-climatic reconstructions are done to reconstruct future climate changes and to indicate potential climate changes in the future.

The aim of this guided research is to reconstruct the climate during the Late Pliocene and Early Pleistocene for the Netherlands based on pollen assemblage data with a quantitative paleo-climatic reconstruction method. The sub-questions for the guided research are:

- Which quantitative paleo-climate technique is the best to reconstruct the climate for the Netherlands during the Late Pliocene and Early Pleistocene?
- Is the dataset of Whitmore et al., (2005) a good dataset to use for a climate reconstruction for the Late Pliocene and Early Pleistocene for the Netherlands?
- Is the summer temperature, reconstructed by Zagwijn (1992) for the Netherlands, in agreement with the reconstructed climate from our reconstruction method?

Three new data-sets with known ages are used for the reconstruction: Hank (Dearing Crampton-Flood et al., 2020) for the Late Pliocene and Noordwijk (Noorbergen et al., 2015) and Petten (Houben, 2019; Krom, 2020) for the Early Pleistocene. The techniques used for this research are transfer functions using Modern Analogue Techniques (MAT), Weighted Average (WA) and Weighted Average Partial Least Squares (WA-PLS) algorithms. MAT, WA and WA-PLS are statistical methods applied on a pollen assemblage data set from the south-eastern part of the United States with 67 different pollen taxa. The climate of the United States have the same climatic characteristics as during the Pliocene and Pleistocene of the Netherlands. The reconstructed summer temperature from the Netherlands for the Early Pleistocene originates from Zagwijn (1992). Zagwijn (1992) made a pollen based July temperature curve with a temperature range between 0 and 20 °C. The temperatures are inferred for the Reuverian (3.1 – 2.5 million years ago (Ma)) and are based on semi-quantitative indicator-taxa approach.

Zagwijn (1992) and Zagwijn (1996) described in his papers about the rarity of the use of terrestrial material instead of sea surface temperatures (SST), most reconstructions are largely based on SST (Dowsett et al., 2010). Most reconstructions are based on the Co-existence Approach (CA) which relies on the presence or absence of a plant taxon within a fossil assemblages. The two techniques WA and WA-PLS are techniques never used for the time interval Late Pliocene and Early Pleistocene, the techniques are based on a dataset from the USA to reconstruct Late Pliocene and Early Pleistocene data from the Netherlands.

# 2. General background information

The Pliocene (5.33 – 2.58 Ma) is an Epoch from the Neogene Period (23.02 – 2.58 Ma) and the Pleistocene (2.58 Ma – 11.7 thousands (ka)) and Epoch from the Quaternary (2.58 to present). Lisiecki and Raymo (2005) described the characteristics of the Late Pliocene and Early Pleistocene, "*one of the characteristics is the cyclic growth of terrestrial ice sheets related to the Milankovitch cycles*". The large terrestrial ice sheets started to develop in the Northern Hemisphere about 2.57 Ma years ago, resulting in multiple glacial – interglacial oscillation driven by variations in orbital insolation as a result of periodic fluctuations in Earth's orbit on Milankovitch timescales of 100 (orbital eccentricity), 41 (Earth's obliquity) and 19 – 23 (precession) ka intervals (Imbrie and Imbrie, 1980; Ruddiman, 1990, 2013; Bradley, 2015). Before about 1.25 Ma, glacial-interglacial oscillations appear to be symmetric with small ice volumes and periodicity of 41 ka years (e.g. Birks, 2019).

The most recent reconstruction for the Pliocene is made by Dearing Crampton-Flood et al., (2020), the reconstruction is based on material from the Southern North Sea Basin from the Hank location. Dearing Crampton-Flood et al., (2020) indicates a 2 to 6 °C warmer SST than present, this makes the Late Pliocene a good interval for prediction of future climate change. During the Late Pliocene and Early Pleistocene the position of the shoreline of the North Sea was positioned more inland with respect to the present-day coastline of the Netherlands (e.g. Westerhoff, 2009; Noorbergen et al., 2015). Ziegler (1990) described "*the Pliocene North Sea had only an opening to the Atlantic in the north, other sides were confined by several landmasses*". Some other connections were confirmed by Gibbard and Lewin (2016), however, only during periods of high sea level. The Southern North Sea basin has a good record of SST evolution and climate change for the north-western European continent because of the high terrestrial input by large European rivers: the Eridanos River and Rhine-Meuse river, described by Gibbard (1988) and confirmed by Dearing Crampton-Flood et al., (2020).

The fossil pollen records are used to reconstruct different climate variables, these reconstructions gave understanding of local climate change over time. Chevalier et al., (2020) described two kinds of dispersal strategies, whereby an-emophilous pollen depends on the wind. The wind mixes the grains in the air from various locations and spread them over large distances. The an-emophilous pollen grains are then further transported by the large European rivers and deposited in the Southern North Sea. So it is possible that not every pollen grain we found in our fossil data-sets originates from that specific location. We found possibly a mix of locally and remotely produced pollen grains in the pollen assemblages, whereby the locally pollen have a higher concentration. The remotely produced pollen have influence on the climate reconstructions and can indicate wrong climate characteristics if they are found in higher concentrations.

## 2.1 Different quantitative reconstruction techniques

Since the start of the 19th century researchers tried to reconstruct the climate of the Late Quaternary. Birks et al., (2010) described the early studies of Blytt (1881), Andersson (1909), Iversen (1944) and Von Post (1946). They used plant macro fossils, animal remains, pollen assemblages and peat stratigraphy to indicate the climate.

Three main approaches for quantitative reconstruction of past climates from bio-stratigraphical data are known (Birks & Birks, 1980; Birks, 1981; 1995; 1998; 2003). These techniques are:

- **Indicator species approach**
  The indicator species approach indicates the presence or absence of one taxa. This technique is the first reconstruction method developed for climate reconstruction. Andersson (1903, 1909, 1910) used this technique to indicate the warm season by the presence of *Corylus avellana*.
- **Assemblage or analogue approach**
  The assemblage or analogue approach indicates the similarity measure between assemblages in all the samples. It considers the fossil assemblages as a whole and the relative abundance of all the different fossil taxa (e.g. Birks & Birks, 1980).
- **Multi-variate calibration function approach**
  The multi-variate calibration function approach involving calibration functions, it uses underlying statistical models with global estimation of parametric functions for all the taxa present (e.g. Birks et al., (2010)).

For this research different multi-variate calibration functions are used, MAT, PCA, WA and WA-PLS. MAT is a calibration method to reconstruct a past environmental parameter, it works by finding modern sites with faunal associations close to the ones in the samples. PCA is used to describe a huge amount of data in smaller components, the principal components. WA and WA-PLS assume a unimodal organism-environmental response model. The difference between both models is the dimensionality, WA uses full dimensionality and WA-PLS a reduced dimensionality.

The co-existence approach, developed by Mosbrugger and Utescher (1997), is the most used method for quantitative paleoclimatic reconstructions. The technique is based on the 'nearest living relative philosophy', the technique corresponds to the 'indicator species approach' of Birks and Birks (1980). The co-existence approach is based on two taxa with a fossil relevance, the best estimate of the mean annual temperature under which the fossil taxa once lived, is the interval where they both lived (Mosbrugger and Utescher, 1997). The co-existence approach is very sensitive and focuses on the extremes in the climate, but it does not show the details. This is a disadvantage, the technique does not use the relative abundance of a fossil taxa, it only uses the presence or absence. For example, the research of Prader et al., (2020), little variation in the temperature is observed while the quantity of specific pollen taxa are constantly changing. The reconstruction is based on the presence of absence of the pollen taxa and not on its quantity. The relative abundance gives additional information to the model, but it is not included.

Mosbrugger and Utescher (1997) did some research about the use of the CA approach, the best results are given for the Oligocene and younger because most of the nearest living relatives are still present. For the Palaeocene and Eocene floras the results are at a lower climatic resolution because the nearest living relatives can only be identified on family level (Mosbrugger and Utescher, 1997). For older material it is difficult to use the technique because there are problems with identifying nearest living relatives.

### 2.1.1 Weighted Average (WA) and Weighted Average Partial Least Squares (WA-PLS)

The assumption of WA and WA-PLS is 'a plant species have their highest abundances in their own climate niches'. The most abundant types are closest to their optimal climatic conditions at the time the pollen assemblages are deposited. The model estimates the climate optima from a calibration set, in this case the dataset from Whitmore et al., (2015), specified on SE USA. Birks et al., (2010) describes the basic idea of WA as "*if a taxon shows an unimodal relationship with a* *particular climate variable x, its abundance will tend to be highest at sites with values of x close to the taxon's environmental optimum or niche peak"*. The taxon's optimum for *x* is the average of all the x values for every site, divided by the taxon's relative abundance.

WA and WA-PLS were first used to reconstruct water chemistry variables in the palaeolimnology (e.g Chevalier et al., 2020). WA has some disadvantages, the modern sample are most of the time unevenly distributed in relation to the environmental gradient of interest, estimation of the optima may be biased (e.g. ter Braak and Looman, 1986). Another weakness is the 'edge-effect', whereby low variables are overestimated, and high values are under-estimated (e.g. Birks et al., 2010; Braak & Juggins, 1993). A solution to reduce the edge-effect is to apply a correction of inferred values referred to as de-shrinking (Birks, 1995). There is estimated how much the range is reduced by modelling a regression model between the reconstructed and the target values from the modern samples. The reconstructions are then rescaled with respect to the range of the target (Nolan et al. 2019).

Ter Braak et al., (1993) developed an improvement on the WA based on partial least squared regression, WA-PLS. WA-PLS is developed to address some of the problems identified in WA, it combines the advantages of WA with the multi-collinearity reduction approach of PLS. WA-PLS suffers less from the edge-effect than WA. It is important that not too many components are selected because this will give the model a poor predictive power. In general, WA-PLS models based on two or three components perform best.

Both methods have been used widely to reconstruct e.g. Late Pleistocene and Holocene precipitation in China (Chen et al., 2015; Li et al., 2017; Liet al., 2012; Lu et al., 2018) and temperature in Siberia (Klemm et al., 2013) and Europe (Finsinger et al., 2010; Lotter et al., 2012; Lotter et al., 2000; Sakari Salonen et al., 2013; Seppä et al., 2009; Veski et al., 2015). WA-PLS performs well in various ecological and climatic conditions (Juggins and Birks, 2012). The models are less sensitive to the size of the studied climate gradient (Chevalier et al., 2020). However, some pollen taxa are multimodal for the variable of interest, mostly by continental or global studies.

### 2.1.2 Modern Analogue Technique (MAT)

MAT is a statistical technique called 'k-nearest-neighbours'. The technique is based on the measure of degree of similarity between fossil pollen assemblages and modern assemblages (Jackson and Williams, 2004). The most similar modern samples, with weights being the inverse of the dissimilarities so that modern samples with the lowest dissimilarity have the greatest weight in the climate reconstruction (e.g. Birks et al., 2010; Overpeck et al., 1985). "*Having found the modern sample(s) that is (are) most similar to the fossil assemblage, the past climate for the fossil sample is inferred to be equivalent to the state of the climate variable(s) of interest for the analogues modern sample(s)*" (Birks et al., 2010).

"*A 'no-analogue' situation occurs when none of the assemblages from the training set resemble the fossil assemblage with sufficient accuracy to be considered an acceptable analogue and can be caused by either a restricted collection of modern samples or by past conditions that have no counterparts in the modern settings*" (Chevalier et al., 2020). Williams and Shuman (2008) did a study about broad-scale cross validation across North-America and comparing modern samples against themselves. They concluded that a small number of closest analogues, from 3 to 7, outperforms the use of only the closest analogue (Williams and Shuman 2008). Including more analogues increases the risk of false positive matches, especially when the calibration dataset encompasses wide spatial areas where the low taxonomic resolution of pollen data can lead to similar pollen compositions to be observed despite experiencing very different climates (Chevalier et al., 2020). William and Shuman (2008) did research about this topic and have proposed that several broadscale North American taxa, like *Pinus, Alnus* or *Tsuga*, could be split into eastern and western pollen taxa to circumvent this issue. When focussing on regional splitting of taxa, a better climatic response is estimated.

The figure below indicates a good example of MAT and is made by Chevalier et al., (2020). The figure is an example of the aridity-humidity gradient, whereby taxon 1 (green) prefers humid conditions and taxon 2 (brown) prefers drier conditions. Three fossil pollen assemblages (diamonds with thick black borders) are compared with the modern assemblages, each had its own humidity value, the coloured circles in the figure. The five closest modern assemblages (white circles) are averaged and estimated the past conditions.

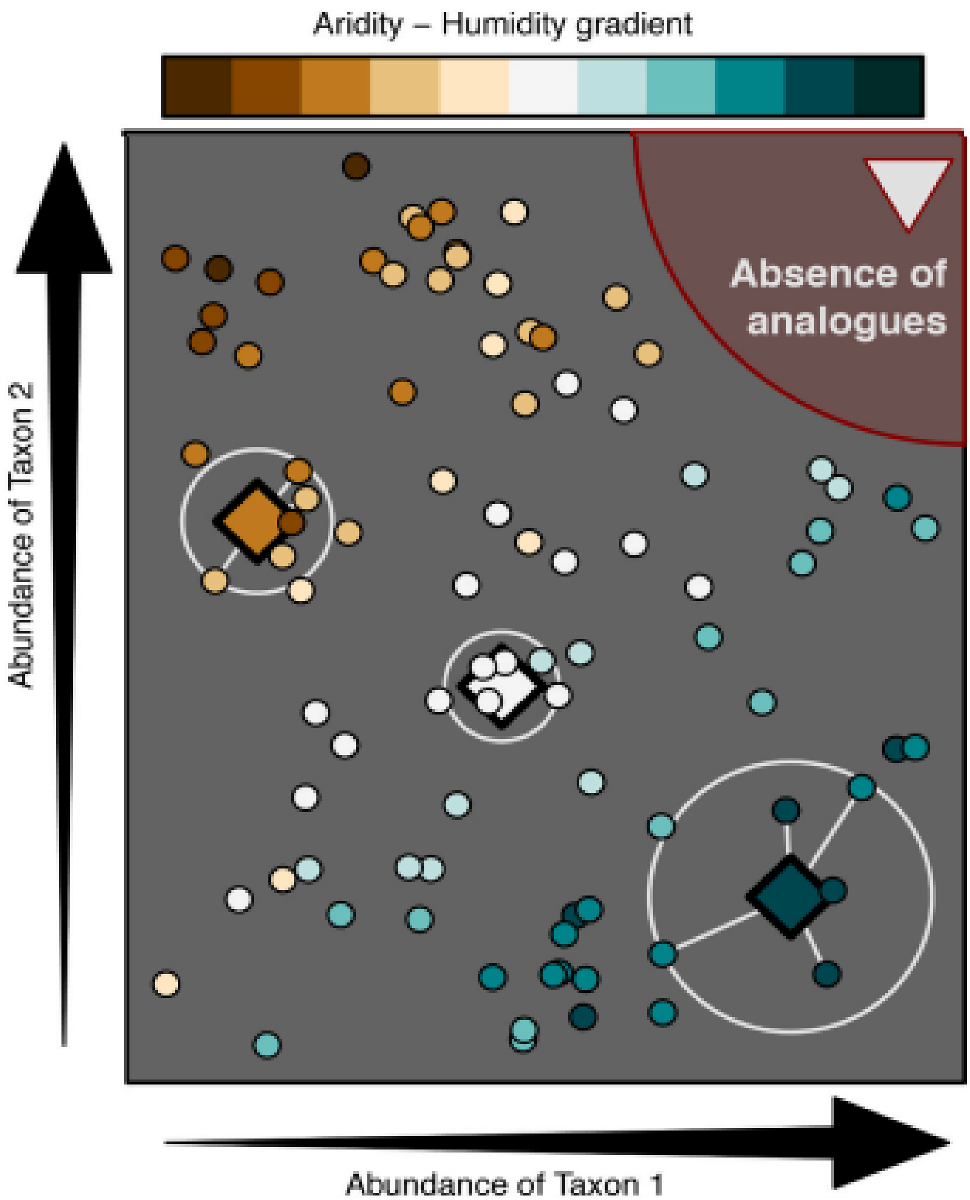



*Figure 1 Conceptual diagram describing MAT. Taxon 1 prefers humid climates (green) and taxon 2 prefers drier conditions (brown). Three fossil pollen assemblages (diamonds) are compared to a modern set, each one associated with a humidity value (coloured circles). The five closest analogues (white circle) are retained and averaged to estimate past conditions. (Chevalier et al., 2020)*

### 2.1.3 Principal Component Analysis and Bootstrapping

PCA is a multi-variate analyse method to describe a huge amount of data into a smaller one that still contains most of the information in the large dataset. The purpose of the technique is to find a new variable in the dataset with a maximal variance. Normally, reducing the number of variables of a dataset decreases the accuracy of a dataset, however, it makes the data simpler. The variables in the dataset are called the 'principal components' and indicate the maximum variation in the data. This technique makes it easier to analyse the dataset due to the smaller datasets that are created with the same amount of information.

The figure below is made by Chevalier et al., (2020) and represents a conceptual diagram of WA-PLS. Three different taxa are used, the black part indicates all the taxa and are not considered in this example. The graph on the right indicates the ecological gradient of the taxon, the modern assemblages. Climate values are extracted from each modern sample, one taxon occurs in and weighted by the corresponding pollen percentages to derive the climate optimum of the taxon. the small circles on the right side. For more information about WA and WA-PLS, there is referred to the paper of Chevalier et al., (2020).

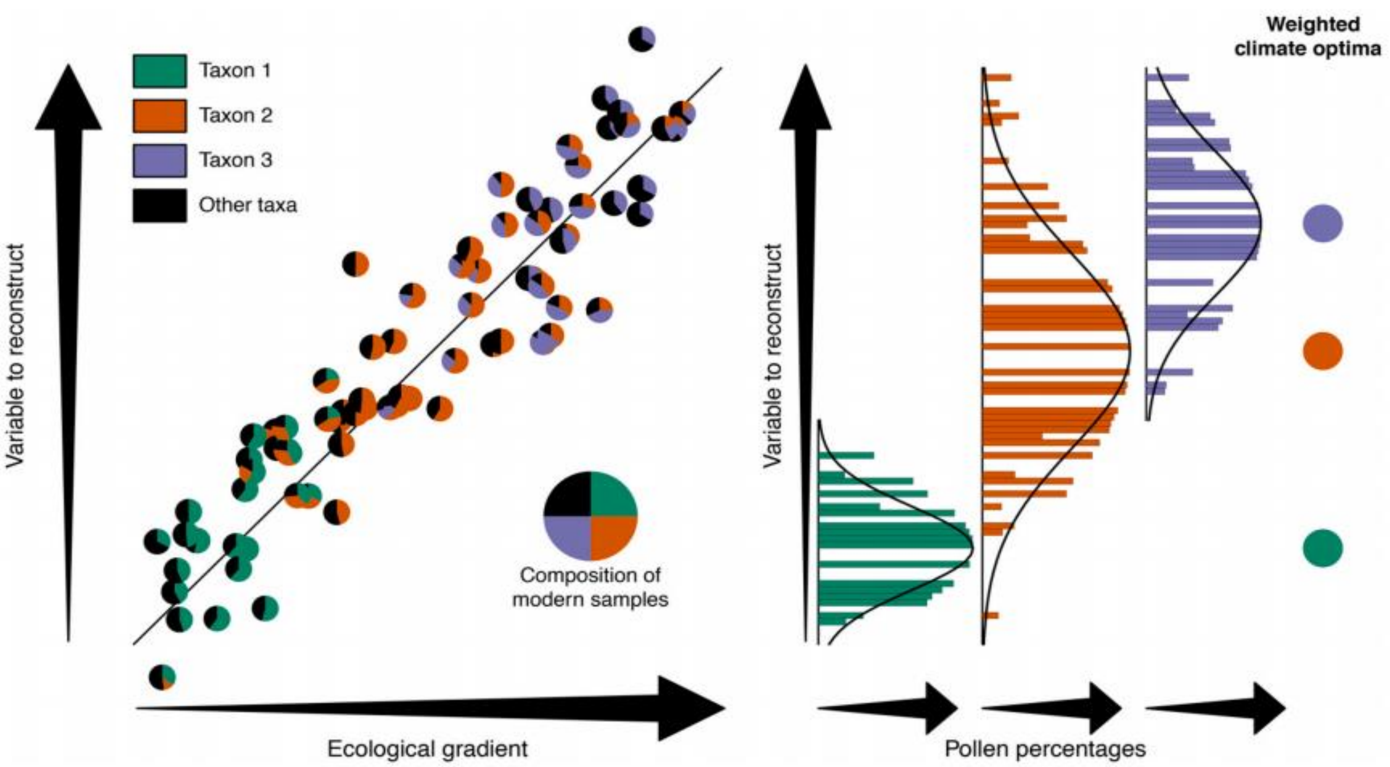



*Figure 2 Conceptual diagram describing WA-PLS. The black part indicates all the taxa which are not considered in the example. On the right, the ecological gradient is showed, the modern assemblage. Climate values are extracted from each modern sample to derive the climate optimum of the taxa, the small circles on the right (Chevalier et al., 2020).*

Bootstrapping is used as an approach for each model. Chevalier et al., (202) explained bootstrapping as "*if we call n the number of samples composing the modern data set, the bootstrap approach consists of randomly sampling n samples with replacement to create a new dataset of the same size*". The unselected samples become part of the test data set, however, some samples are selected multiple times randomly. For each model run, a specific bootstrap set is chosen and used for the specified reconstruction.

## 2.2 Modern dataset

The dataset contain 4634 sites from across North America and contain four types of information for each record: (1) pollen counts for 134 pollen taxa that are either found across North America or are regionally important, (2) site identification, geographic coordinates , source, depositional environment and auxiliary identification codes, (3) environmental data, including elevation-corrected climatic data based on the Climatic Research Unit gridded climatology (New et al., 2002) with improved lapse rate correction and (4) information about natural and potential vegetation derived from cartographic sources (Whitmore et al., 2005).

Most of the pollen data are from lake sediments (2261 sites), however mostly in the south-eastern part of the USA, samples are taken from moss pollsters (860 sites). Most of the lake sediments are from Canada, while samples from the south-eastern US, northern Pacific Coast and the Central Plains are from moss pollsters and terrestrial/soil samples from the southwestern part of the US. The original data for the modern pollen database came from core-top samples originating from many sites where fossil data were analysed during the past 40 years and local and regional surface sample datasets accumulated expressly for calibration studies (Whitmore et al., 2005). The pollen are processed following a standard procedure with a small amount of sediment and some series of acid and bases to remove the sedimentary material, the pollen grains remain intact (e.g. Whitmore et al., 2005). The classification of the pollen is based on taxonomic and morphological hierarchies defined in the NAPD (Grimm, 2000a).

A three-step procedure was used to indicate the climate and bio-climate variables. First, the local lapse rate for the long-term monthly mean temperature, precipitation and possible sunshine data were calculated (e.g Whitmore et al., 2005). Second, the lapse rate at the surrounding 10-min grid points were used to adjust the sunshine data values at those points to the elevation of the target point (e.g. Whitmore et al., 2005). Third, the values were interpolated at the location of the target point using geographically weighted bilinear interpolation (e.g. Whitmore et al., 2005). Whitmore et al., (2005) defined the bioclimatic variables as *"calculation from the basic climatic variables including growing-degree days with a 0 and 5 °C, a chilling period, mean temperature of the coldest and warmest month of the year and the January/annual and July/annual precipitation ratios".*

# 3. Material and methods

## 3.1 Location of the modern and fossil datasets

Three different fossil datasets from the Netherlands: Noordwijk, Petten and Hank are used for the climate reconstruction. Svenning (2003) described Pliocene and Pleistocene extinctions in Europe and the correspondence in cool-temperate tree genera for North-America and Europe. Some species occurring during Late Pliocene and Early Pleistocene in Europe, are still present in North America (Florida region) and are now extinct in Europe (e.g. Svenning, 2003). The modern dataset from Whitmore et al., (2015) originates from North America and Greenland and is limited to the SE USA region for this guided research. The Florida region is chosen because of the climate characteristics, the climate conditions during the Late Pliocene and Early Pleistocene for the Netherlands were almost the same as the current climate conditions in Florida. The winter temperature is used for the reconstruction because the summer temperatures in Florida are likely higher comparison with the climate conditions in Netherlands during the Late Pliocene and Early Pleistocene. Also, winter temperatures are more limiting to warm temperate to subtropical vegetation, so it is more likely to reflect realistic trends by using summer temperatures.

### 3.1.1 Location Noordwijk

Noordwijk is in the southwestern part of the Netherlands (Figure 3) in the province of South-Holland. The environment around Noordwijk is created during one of the last ice periods due to the interaction of the sea and the Old Rhine river. The dataset of Noordwijk was drilled under authority of the Geological Survey of the Netherlands (Rijksgeologische Dienst: RGD) using air lifting drill methods (e.g. Noorbergen, 2015). The 73 samples were taken at a depth from 454.1 – 52.5 m and have an age from 2.61 – 2.11 Ma years ago and contain 45 different pollen types. The borehole consists of different Formations, mainly the lower Maassluis, Westkappele Ground and Upper Red Grag Formation. For more information about the Noordwijk core there is referred to Noorbergen (2015).

### 3.1.2 Location Petten

Petten is located at the North Sea in the northwestern part of the Netherlands (Figure 3) in the province North-Holland. Two different boreholes were drilled in Petten, borehole 1 (BH1) is used for the reconstruction. The top of the borehole is at an elevation of 3.26 m above Normaal Amsterdams Peil (NAP). Down to 101,50 m depth, sonic drilling was performed. Below that, wireline rotary coring was performed. TNO (2019) partially describes the core and said that the core was recovered in unconsolidated, dominantly sandy Pleistocene sequences. The borehole was taken at a depth from 383,60 – 196.85 m depth, has 71 samples and contain 33 different pollen species. The samples of Petten have an age range of 2.41 – 1.35 Ma. For more information about the Petten core there is referred to TNO (2019) and the thesis from Krom (2019).

### 3.1.3 Location Hank

Hank is in the southern part of the Netherlands (Figure 3), within the current Rhine-Meuse-Scheldt delta. The site is located within the Ruhr Valley Rift, a region that experienced relatively high tectonic subsidence during the Late Cenozoic (e.g. van Balen et al., 2000; Dearing Crampton-Flood et al., 2020). The borehole was drilled by air-lifting technology to a base of 404 m in 2001. The Geological Survey of the Netherlands (TNO) and Dearing Crampton-Flood et al., (2020) described the lithology of the core, the base consists of the upper part of the marine Breda Formation, followed by sandy and clay-rich marine belonging to the Oosterhout Formation. Sometimes the marine sediments contain shell fragments and crags. The dataset contain 27 samples and 43 different pollen species. The age ranges from 2.48 – 5.42 Ma years ago. For more information about the Hank core there is referred to Dearing Crampton-Flood et al., (2020).


*Figure 3 Location of the fossil datasets, from north to south: Pettten, Noordwijk and Hank*

## 3.2 Creating datasets and models in C2

The combination of the pollen species in the dataset from Whitmore et al., (2015) presents a specific climatic pattern, the characteristics of each sample are presented in the environmental datasheet. As described in paragraph 2.3, the south-eastern part of the United States is used for this guided research. The latitude range was the limited factor to choose the samples from the dataset of Whitmore et al., (2015). The latitude range is 25° - 50° N, the value is determined because of the sample amount and the pollen species occurring in between this region. The bi-modal distribution of the pollen is important to keep in mind, mainly for *Pinus*. This genus has different species, occurring in warmer (lowland) and colder (mountains) environments. The gradient is used to include both *Pinus*, however, it could gave problems for other species like the Cupressaceae and Taxodiaceae group. The longitude ranges from -52.6° - 99.3° E.

The three fossil datasets contain pollen taxa which do not exist anymore or do not occur in the modern dataset because of climate limits. The modern and fossil dataset must be at the same taxonomical resolution otherwise the program cannot link them with each other. The merged dataset can be found in Appendix K, the fossil pollen species have been combined with other pollen species from the same family to make a new dataset. *Alisma, Armeria, Azolla* (Mosquito

Fern), *Brassicaceae., Dipsacaceae, Engelhardia, Eucommia, Isoëtes, Lythrum, Myriophyllum, Parthenociccus, Pilularia, Platycarya, Potamogeton, Reveesia, Sapotaceae, Sciadopitys, Symplocos, Trudopollis, Typhaceae* and *Vitis* have no taxonomic representation in the modern pollen dataset and are not used for the reconstruction. *Carpinus* and *Ostrya* are merged into *Ostrycar*, the other fossil pollen species with another 'modern' name are mentioned in Table 1. The decision to call them by this specific family name usually stems from the family branch they origin from. Usually, the fossil pollen is a clade of the same family as the modern variant.

*Table 1 Family names of the fossil ones*

| **Fossil name** | **Family name** | **Merged with** |
|---|---|---|
| *Asteraceaea* | Asteraceae | Aster |
| *Calluna* | Ericaceae | Ericaceaea |
| *Cedrus* | Pinaceae | Larix |
| *C. Tubuliflorae* | Asteraceae | Aster |
| *C. Liguliflorae* | Asteraceae | Aster |
| *Dryopteris* | Dryopteridaceae | Polypodiaceae |
| *Filipendula* | Rosaceae | Rosaceae |
| *Gramineae* | Poaceae | Poaceae |
| *Illex* | Aquifoliaceae | Aquifoliaceae |
| *Juniperus* | Cupressaceae | Cupressaceae |
| *Lonicera* | Caprifoliaceae | Caprifoliaceae |
| *Nymphaea* | Nymphaeaceae | Liliaceae |
| *Rhus* | Anacardiaceae | Anacardiaceae |
| *Sequoia* | Cupressaceae | Taxodium |
| *Spergula* | Caryophyllaceae | Caryophyllaceae |
| *Succisa* | Caprifoliaceae | Caprifoliaceae |
| *Thelypteris* | Thelypteridaceae | Polypodiaceae |
| *Umbelliferaea* | Apiaceae | Apiaceae |
| *Viscum* | Santalaceae | Arceuthobium |

The fossil datasets have been merged into one big dataset with the data from Noordwijk (1 – 73), Petten (74 – 145) and Hank (146 – 172). Appendix K shows the combination of the part of the dataset from Whitmore et al., (2015) and the fossil datasets. The first three columns show the pollen species from the fossil datasets and the fourth the names used in the dataset. Thereafter, a new pollen sum and pollen percentages are calculated.

Two multi-variate functions (PCA and Correspondence Analysis (CA)) are used to indicate the correspondence between the data from Whitmore et al., (2015) and the fossil data is. After analysing the first results, *Pinus,* and *Pinus haploxylon*-type are eliminated from the data because of the outside position in the graphs. *Pinus* has a bimodal distribution, it has two different peaks, one in the cold part of the gradient and one in the warm part of the gradient. The bi-modal distribution caused that *Pinus* is not a good pollen species to use as an indicator for climate reconstruction. The first PCA is done with all the data to confirm the outliers of the CA graph, the second one is done without the sample outliers and *Pinus* and *Pinus haploxylon*-type. Second, MAT is used to observe the reliability (test of 'closeness') of the fossil samples to the reference data, the values are used to describe the reliability of the reconstructions. The distance to the closest modern analogue is used to describe the reliability, the 'goodness of fit' measured by the $1^{st}$ and $5^{th}$ percentile with the reference sample, calculated as squared residual distance, are considered very good and good analogues, respectively (Jackson and Williams, 2004). One percentile indicates that 99% of the data has a close analogue in the reference data. The percentiles are showed in Table 4. The last step is the calculation of the pollen-climate transfer function using WA and WA-PLS with inverse de-shrinking for WA and 3rd component for WA-PLS. Prentice (1980) suggest to square root transform species data for WA and WA-PLS regression to reduce noise in the data. Two parameters are used to reconstruct the climate: the winter temperature and the total precipitation. The summer climate in Florida has more tropical characteristics than during the Late Pliocene and Early Pleistocene in the Netherlands. The same reason could be given for the choice to reconstruct the total precipitation, during the tropical warm summers a lot of water evaporates. Also, Florida has a kind of monsoon period with a lot of precipitation, probably different in comparison with the Pliocene in Europe. The calculation for the mean winter temperature and total precipitation of WA and WA-PLS transfer function was performed with C2 version 1.7 software (Juggins, 2016) for the period 5.45 – 1.34 Ma years ago. Both methods were used to indicate which method is better to use for this time period. Most papers use the co-existence approach for this time period, this guided research will introduce transfer-functions like WA and WA-PLS to this time period. The performance statistics were computed for each transfer functions, these include the coefficient of determination ($R^2$), $R^2$ bootstrapping and Root Mean Square Error or Prediction (RMSEP). These statistics are shown in Table 1 and Table 2.

For some samples, the age is interpolated because the age was unknown, the average of the sum of the sample above and below is calculated to determine the age. The reconstructed data is discussed, and the minima and maxima temperature and precipitation are explained by the combination of pollen found in the samples. The results from the WA and WA-PLS reconstruction are compared with each other to decide which method is better to use for this period. WA is the simplest approach with the least assumptions, it assumes an unimodal relationship between species and environments. This relationship is tested by plotting the gradient of species against the temperature (Appendix A). WA-PLS used residual variance to increase the model fit. Also, the percentage of deciduous trees is determined to indicate potential climate cycles.

# 4. Results

Appendix A, B, C and D show the pollen diagrams of Whitmore et al., (2015), Noordwijk, Petten and Hank, respectively. The dominant pollen taxa (Appendix E) in Noordwijk and Petten are *Alnus, Artemisia, Betula, Corylus, Cyperaceae, Ericaceae, Picea, Pinus, Poaceae, Polypodium, Quercus* and *Sphagnum*. For Petten *Alnus, Cheno/Am, Corylus, Ericaceae, Picea, Pinus,* Poaceae, *Polypodium, Quercus* and *Sphagnum*. The dataset of Hank has five dominant pollen species: *Abies, Picea, Pinus, Polypodium* and *Taxodium*. The three datasets have a dominance in the bisaccate pollen, *Pinus* is already removed from the figure. *Pinus* has also a high dominance in the modern dataset in comparison to other pollen species. One explanation for the high percentage of *Pinus* pollen is the easy way of dispersion of the pollen. Bisaccate pollen are easily transported by the wind and are most sensitive to differential transport processes (e.g. Donders et al., 2018). Because of this effect, it cannot be said with certainty of the *Pinus* pollen, found in the fossil and modern datasets, are regionally deposited.

## 4.1 General overview of the data

Appendix E shows the dominant pollen species with a timescale (2.56 – 2.12 Ma) on the y-axis for Noordwijk. The first observation is the occurrence of *Polypodium* until an age of 2.287 Ma years ago and the dominance of *Ericaceae* for the whole section. *Poaceae* has its highest dominance during 2.38 – 2.29 Ma. Another observation is the decrease of *Polypodium* and *Poaceae* at almost the same moment, also *Picea* has a remarkable decrease and do not fluctuate that much anymore. *Betula* has one peak at an age of 2.29 Ma. *Sphagnum* and *Alnus* fluctuate for the whole section between 5 and 20%. *Aster* has a peak at an age of 2.55 Ma.

The Petten dataset has an approximate time range from 2.41 – 1.35 Ma years ago. The same decrease in *Poaceae* and *Polypodium* is recognised at an age of 2.287 Ma years ago. *Polypodium* has a new high peak at an age of 1.47 Ma years ago. For Noordwijk, *Ericaceae* is present for the whole section, the section of Petten indicates the decrease in *Ericaceae* from 1.67 Ma years ago. *Quercus* becomes more important at an age of 1.97 Ma years ago.

The time range of the dataset from Hank is from 5.43 – 2.48 Ma years ago. Only four pollen species are dominant for this dataset, these pollen species are *Abies, Picea, Polypodium* and *Taxodium*. *Taxodium* is the most dominant pollen species in this dataset, during the highest dominance periods the other dominant pollen species are almost zero. *Polypodium* is dominant for the period 2.98 – 2.5 Ma years ago.

### 4.1.1 Ordination results

The idea of PCA is to reduce the number of variables of the dataset, the new eigenvectors preserving as much information as possible. Appendix A shows the PCA without the removal of *Pinus* and *Pinus Haplox,* these figures show that the PCA's have very long gradients. Assumed is the linear relationship between the environment and the in- or decrease of species (Appendix A). Figure 4 shows the PCA for the modern and fossil data without *Pinus* and *Pinus haploxlyon*-type. The first subdivision spread the data into *Cheno/Am, Polypodium, Ericaceae, Picea* and *Abies* and *Betula, Tsuga, Ulmus, Quercus* and *Poaceae*. The lower part is clearly divided into *Quercus* and *Betula*, while for the upper part the division is less clear, but it is made between *Cheno/Am* and *Picea*. The fossil datapoints are more centred in the middle and upper part, whereby the Petten samples are spread over the whole upper part and Hank and Noordwijk almost only in the middle. A lot of *Polypodium* and *Ericaceae* are observed in these samples.

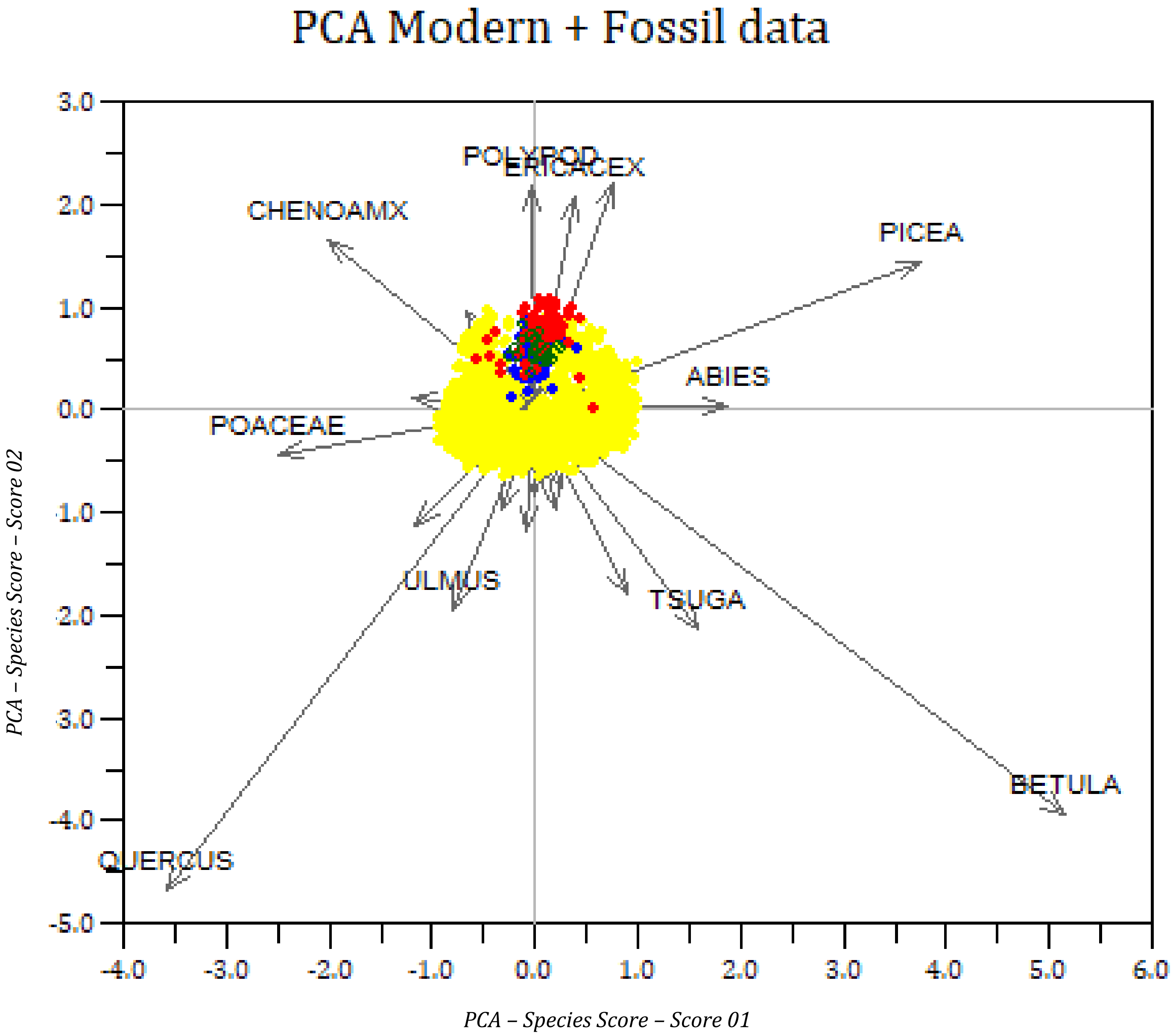



*Figure 4 Principal component analysis for the modern and fossil data. Yellow dots = Florida, Blue dots = Noordwijk, Red dots = Petten and Green dots = Hank*

Figure 5 represents the CA results for the modern and fossil data without *Pinus* and *Pinus haplox*. The red triangles indicate the optimum for the pollen taxa. *Arecaceae, Taxodium, Malvaceae, Caryophyllaceae, Ranuncul* and *Sphagnum* optima plot are distinct from the other taxa, whereby *Quercus, Betula, Cheno/Am* and *Picea* went to another corner. Samples further from this optimum point contain less of these species. The blue dots (Noordwijk) have the highest correspondence with the Florida dataset, green (Hank) seems to have the least resemblance to the modern dataset. The samples with the least resemblance are Sample 88, 118, 160, 163 and 169. The samples are outliers in this analyse because those samples are from the southern part of Florida and contain a lot of *Taxodium*. Those samples are relative warm in comparison to the other samples, maybe too warm for the Late Pliocene and Early Pleistocene for the Netherlands.

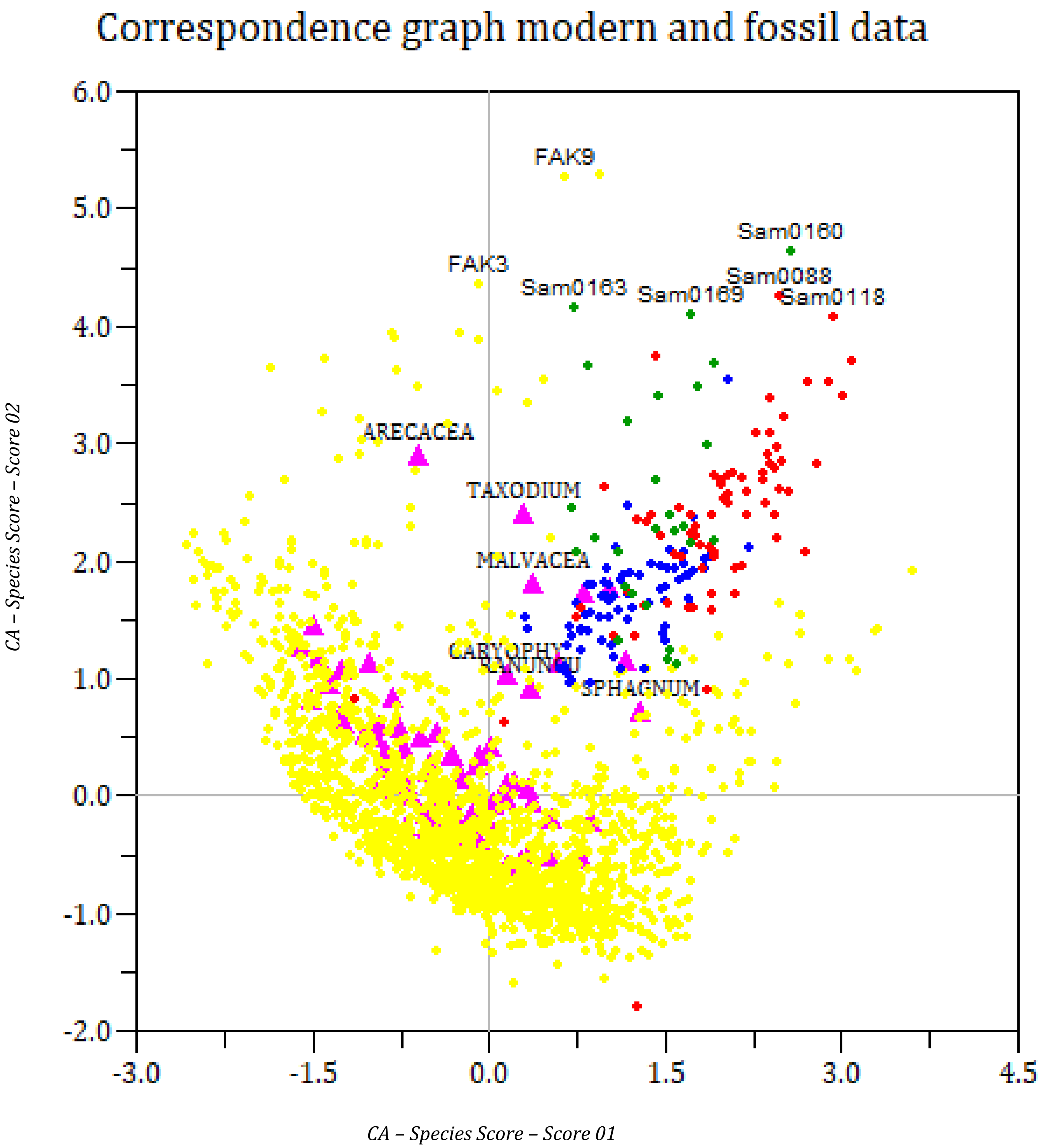



*Figure 5 Correspondence graph for the modern and fossil data. Yellow dots = Florida, Blue dots =Noordwijk, Red dots = Petten and Green dots = Hank*

## 4.2 Climate reconstruction

The performance of WA and WA-PLS are presented in Table 2 and Table 3 and the 'observed vs predicted' is presented in Appendix G and H, the performance of the MAT are presented in Table 4. Table 2 are the first results without removing any sample or taxa, Table 3 are without sample GPD1649, GPD1950, GPD1951 and GPD 1952 and the taxa *Pinus* and *Pinus haploxyon*. The samples have been chosen because of the outlying positions in the observation graphs (Appendix G). The inverse de-shrinking method is chosen for WA in stead of the classic method because the classic methods have a negative change in comparison to the inverse de-shrinking method. For the WA-PLS method both is chosen for component 3 because the %change to component 4 is minimal and to avoid overfitting in the model. Component 3 has a significant change to component 2 and a minimal difference with component 4. A small difference is observed between Table 1 and Table 2, only for the winter temperature reconstruction of WA is no rise in the bootstrapped $R^2$ value seen. All the others have a small rise in the value, so it seems that the reconstruction is a little bit better now.

*Table 2 Performance WA and WA-PLS*

| ***WA Total precipitation*** | | | | ***WAPLS Total precipitation*** | | | |
|---|---|---|---|---|---|---|---|
| *Code* | *Boot $R^2$* | *%Change* | *RMSEP* | *Code* | *Boot $R^2$* | *%Change* | *RMSEP* |
| **WA_INV** | 0.512816 | | 180.037 | **Comp. 1** | 0.51262 | | 180.176 |
| **WA_Cla** | 0.513366 | -34.9606 | 242.979 | **Comp. 2** | 0.5956 | 8.70763 | 164.487 |
| **WATOL_INV** | 0.402529 | | 200.244 | **Comp. 3** | 0.642429 | 5.43333 | 155.55 |
| **WATOL_Cla** | 0.403677 | -55.291 | 310.961 | **Comp. 4** | 0.668901 | 1.81805 | 152.722 |
| | | | | **Comp. 5** | 0.661611 | -0.0215982 | 152.755 |
| ***WA Winter temperature*** | | | | ***WAPLS Winter temperature*** | | | |
| **WA_INV** | 0.714195 | | 3.83462 | **Comp. 1** | 0.713716 | | 3.84124 |
| **WA_Cla** | 0.714653 | -16.7989 | 4.47879 | **Comp. 2** | 0.807457 | 17.2186 | 3.173888 |
| **WATOL_INV** | 0.741713 | | 3.70423 | **Comp. 3** | 0.829968 | 5.51694 | 2.94848 |
| **WATOL_Cla** | 0.742278 | -14.6081 | 4.23534 | **Comp. 4** | 0.837909 | 1.7796 | 2.87854 |
| | | | | **Comp. 5** | 0.838871 | -0.0718837 | 2.8701 |

*Table 3 Performance WA and WA-PLS after removing samples and Pinus and Pinus haploxyon*

| ***WA Total precipitation*** | | | | ***WAPLS Total precipitation*** | | | |
|---|---|---|---|---|---|---|---|
| *Code* | *Boot $R^2$* | *%Change* | *RMSEP* | *Code* | *Boot $R^2$* | *%Change* | *RMSEP* |
| **WA_INV** | 0.516889 | | 179.384 | **Comp. 1** | 0.505986 | | 179.558 |
| **WA_Cla** | 0.51772 | -35.8083 | 243.618 | **Comp. 2** | 0.580981 | 7.71401 | 166.643 |
| **WATOL_INV** | 0.41889 | | 195.82 | **Comp. 3** | 0.634105 | 5.97508 | 156.393 |
| **WATOL_Cla** | 0.419841 | -51.5904 | 296.845 | **Comp. 4** | 0.652612 | 2.07448 | 152.999 |
| | | | | **Comp. 5** | 0.655908 | 0.0725082 | 153.246 |
| ***WA Winter temperature*** | | | | ***WAPLS Winter temperature*** | | | |
| **WA_INV** | 0.729023 | | 3.73047 | **Comp. 1** | 0.728387 | | 3.73866 |
| **WA_Cla** | 0.729342 | -15.1814 | 4.29717 | **Comp. 2** | 0.807716 | 15.2154 | 3.16981 |
| **WATOL_INV** | 0.741262 | | 3.371282 | **Comp. 3** | 0.8323 | 5.92074 | 2.98347 |
| **WATOL_Cla** | 0.741684 | -14.3765 | 4.2466 | **Comp. 4** | 0.839519 | 1.63764 | 2.93669 |
| | | | | **Comp. 5** | 0.841089 | 0.106758 | 2.93166 |

*Table 4 MAT percentile values for squared residual distance values before and after removing samples and Pinus / Pinus haploxyon*

| **Percentile** | **Value** | **Value after removing samples** |
|---|---|---|
| 1 | 9.72514 | 9.71247 |
| 2 | 11.2313 | 11.2218 |
| 5 | 13.675 | 13.6659 |
| 10 | 16.0908 | 16.0823 |
| 20 | 19.4073 | 19.4015 |

Table 4 shows the values for the MAT, those values indicate the goodness of fit. Fossil samples with a distance to the closest analogue larger than the 2nd and 5th percentile of the distances of the modern samples in the calibration dataset were classified as having 'no-close' and 'no-good' analogue, respectively (following e.g. Gouw-Bouman et al., 2019). The 20th percentile of MAT is presented in Figure 7 together with the reconstruction for the winter temperature and total precipitation. The age range for this graph is from 1.34 – 5.34 Ma years ago and include the Late Pliocene (cream colour) and the Early Pleistocene (blue/grey colour). The three different data sets are indicated by blue, red and green boxes, Noordwijk, Petten and Hank, respectively. The data points are marked by red circles and the errors bars are the small blue lines. The minimum and maximum temperature/precipitation is the same for the WA and WA-PLS to make a good comparison between both reconstruction methods. The percentage trees are the deciduous trees (Figure 8), the 100% sum is a calculation of deciduous trees, bisaccates, ferns, grasses, and other pollen types.

The resulting temperature range for the winter temperature is between -16 °C and 16 °C and the precipitation between 600 and 2100 mm per year. The minimum and maximum temperature lies between -12.47 °C and 13.77 °C for WA and -14.11 °C and 13.39 °C for WA-PLS. These optimum are reached at an age of 1.338 and 3.79 Ma years ago for WA and 1.338 and 3.217 Ma years ago for WA-PLS. The minimum and maximum precipitation lies between 924.4 and 1918 mm per year at an age of 1.34 and 3.79 Ma years ago for WA and 675.8 and 1529 mm per year, 2.372 and 1.597 Ma years ago for WA-PLS. During the Late Pliocene, the average temperature of this period is 4.12 °C/0.91 °C while the average temperature of the Early Pleistocene is below zero, -3.67 °C/-4.05 °C for the WA/WA-PLS reconstruction. A clear trend to a cooler world is observed at the transition to the Early Pleistocene for both reconstructions.

## 4.3 Optimum in the winter temperature and total precipitation

For the first 0.2 Ma year (2.38 – 2.58 Ma), the temperature is relatively low between -8.693 °C and 0 degrees for WA and -9.58 °C and 2.07 °C for WA-PLS. Also the percentage trees is low between 0-40%. However, the WA, WA-PLS and the percentage trees follow a similar pattern, when the percentage increases the temperature do the same. The next 0.2 Ma year (2.38 – 2.18 Ma) has a temperature range between -6.64 °C and 0.45 °C for WA and -10.31 °C and -0.31 °C for Wa-PLS. The percentage trees is between 2 and 87%, more difference than 0.2 Ma earlier.

### 4.3.1 Late Pliocene reconstruction

First, the general trend in the Late Pliocene is discussed, some differences between the WA and WA-PLS reconstruction are discussed. The highest temperature for WA is 13.77 °C, sample 162 (3.79 Ma) with a high concentration of *Pinus* and *Taxodium*. The highest temperature for WA-PLS is 13.39 °C, sample 160 (3.22 Ma), this sample has a peak abundance of *Osmunda* (86%), while a low concentration of *Taxodium* is found. Probably this is a sample from a sea-level low-stand (Dearing Crampton-Flood et al., 2020) and is an outlier. In general, the samples with higher temperatures contain high concentrations of *Taxodium, Quercus* and *Polypodium*. The lowest temperature for WA is -1.26 °C, sample 157 (2.60 Ma), this sample contain a lot of bisaccates (*Abies, Pinus* and *Pinus haploxyon*) and some *Tsuga* and *Taxodium* (~5%). Sample 167 (4.75 Ma) is the coldest sample for the WA-PLS reconstruction with a temperature of -5.42 °C. This sample consists of some *Quercus* (3.2%), *Sphagnum* (2.4%) and *Taxodium* (85). The colder samples also contain a high abundance of *Abies*, while the warmer samples contain only *Pinus* and *Pinus haploxyon*. Other samples with relative high temperatures are 163, 169 and 172 (respectively 3.97; 4.97 and 5.43 Ma), these samples contain a higher concentration of *Alnus*, *Quercus* and *Sphagnum*. The relative colder samples are 155, 161 and 171 (respectively 2.90; 3.24 and 5.17 Ma), these samples contain a higher concentration in *Ericaceae, Betula, Poaceae, Polypodium* and *Ulmus*.

The wettest sample for the WA reconstruction is 162 (3.79 Ma), also the warmest sample for the Late Pliocene reconstructed by WA. The same phenomenon is observed for WA-PLS, the same sample caused the wettest conditions (160, 3.22 Ma). The driest condition for WA is caused by sample 155 (2.94 Ma), this sample has a high diversity, including the bisaccates (41%), *Taxodium* (4%), *Ulmus* (3%) and *Polypodium* (21%). The driest sample for WA-PLS reconstruction is the same as the coldest conditions, sample 167. Other relative wet samples for both methods are 158, 163, 169 and 171 with ages 3.05, 3.97, 4.97 and 5.17 Ma years ago, respectively. These samples contain, except the bisaccates and the species mentioned above, *Alnus, Ericaceae, Osmunda* and *Poaceae*. Other relative dry samples are 156 and 165 (2.93 and 4.37 Ma years ago), with except for the bisaccates and the species mentioned above, only some *Ericaceae*.

### 4.3.2 Early Pleistocene reconstruction

The Early Pleistocene has more samples than the Late Pliocene, so more variation in the temperature and precipitation is observed. Only 7 samples are above zero degrees for the WA and 8 for the WA-PLS reconstruction. For WA, the highest temperature is observed for sample 97 (2.36 Ma) with 5.39 °C. The sample contains a high amount of *Polypodium* (50%) and *Betula Osmunda* and *Cheno/Am*, all 13%. The highest temperature for WA-PLS method is also observed in sample 97, 5.29 °C. All the other samples with a temperature above zero, for both WA and WA-PLS, contain a high amount of *Polypodium* and do not contain as many bisaccates as during the Late Pliocene. Other pollen species with relative high concentrations in the warmer samples are *Alnus, Poaceae, Ericaceae, Sphagnum, Cyperaceae* and *Tsuga.* The lowest temperature for WA method is caused by sample 74 (2.29 Ma), this sample contain 50% *Betula* and 50% *Pinus.* Another sample with a relative low temperature is sample 63 (-10.03 °C; 2.23 Ma), this sample also contain a high amount of *Pinus* (36%) and *Ericaceae* (22%). For WA-PLS, again sample 74 caused the lowest temperature (-14.11 °C), other cold samples are number 85 and 142 (2.32 and 2.57 Ma). Number 85 contain 50% *Polypodium* and 30% *Ericaceae*, number 142 also contains a lot of *Pinus* (29%), *Polypodium* (18%) and some *Picea* (22%). Other pollen species which are associated to colder conditions are *Alnus, Cyperaceae, Poaceae* and *Quercus*.

The precipitation range is almost the same for the Early Pleistocene as during the Late Pliocene. The wettest sample for the WA method is number 97 (2.36 Ma) with 1433 mm per year. This sample is again the same sample as the warmest sample for the WA method, the same phenomenon as during the Late Pliocene. For WA-PLS, sample 88 is the wettest sample with 1529 mm per year at an age of 2.34 Ma years ago. This sample is not the same as the wettest sample for the WA-PLS method. However, this sample also contains a high amount of *Polypodium* (86%), other pollen species in the sample are *Ericaceae* (5%), *Poaceae* (2%) and *Alnus* (2%). Other wet sample contain lot of *Ericaceae*, *Pinus* and *Sphagnum.* The driest sample for the WA-method is sample 89 (2.34 Ma), this sample is very close to the wettest sample of the

WA-method. The most important pollen species for this sample is *Poaceae* (92%). The WA approach produces in general warmer and wetter reconstruction than WA-PLS. The WA-PLS method reconstructs an overall drier environment than the WA-method, with values around the 700 mm per year in comparison to WA with ~950 mm per year. The driest sample for WA-PLS is 32 (1.84074 Ma), this sample contain *Pinus* (26%), *Poaceae* (18%) and *Artemisia* (15%). Other pollen species which cause dry conditions are *Betula, Cheno/Am, Cyperaceae* and *Quercus*.

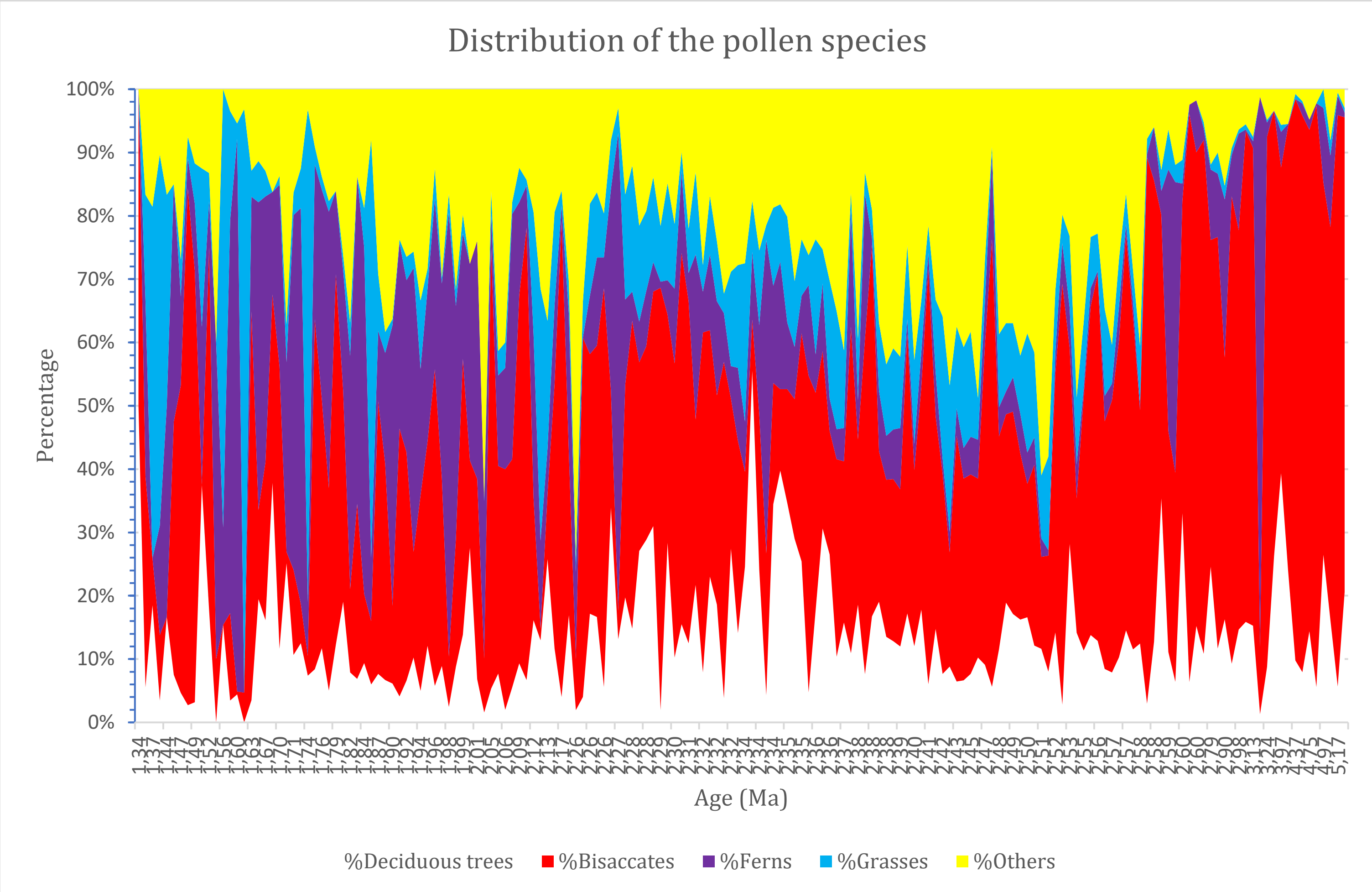



*Figure 6 Distribution of the pollen species with Age (Ma). Deciduous trees are Acer, Alnus, Anacardia, Arecaceae, Betula. Carpinus, Carya, Celtis, Corylus, Fagus, Fraxinus, Juglans, Liquidambar, Nyssa, Quercus, Salix, Taxodium, Salix and Ulmus. The bisaccates are Abies, Picea, Pinus, Pinus haploxyon and Tsuga. The ferns are Osmunda and Polypodium. The grasses are Poaceae.*

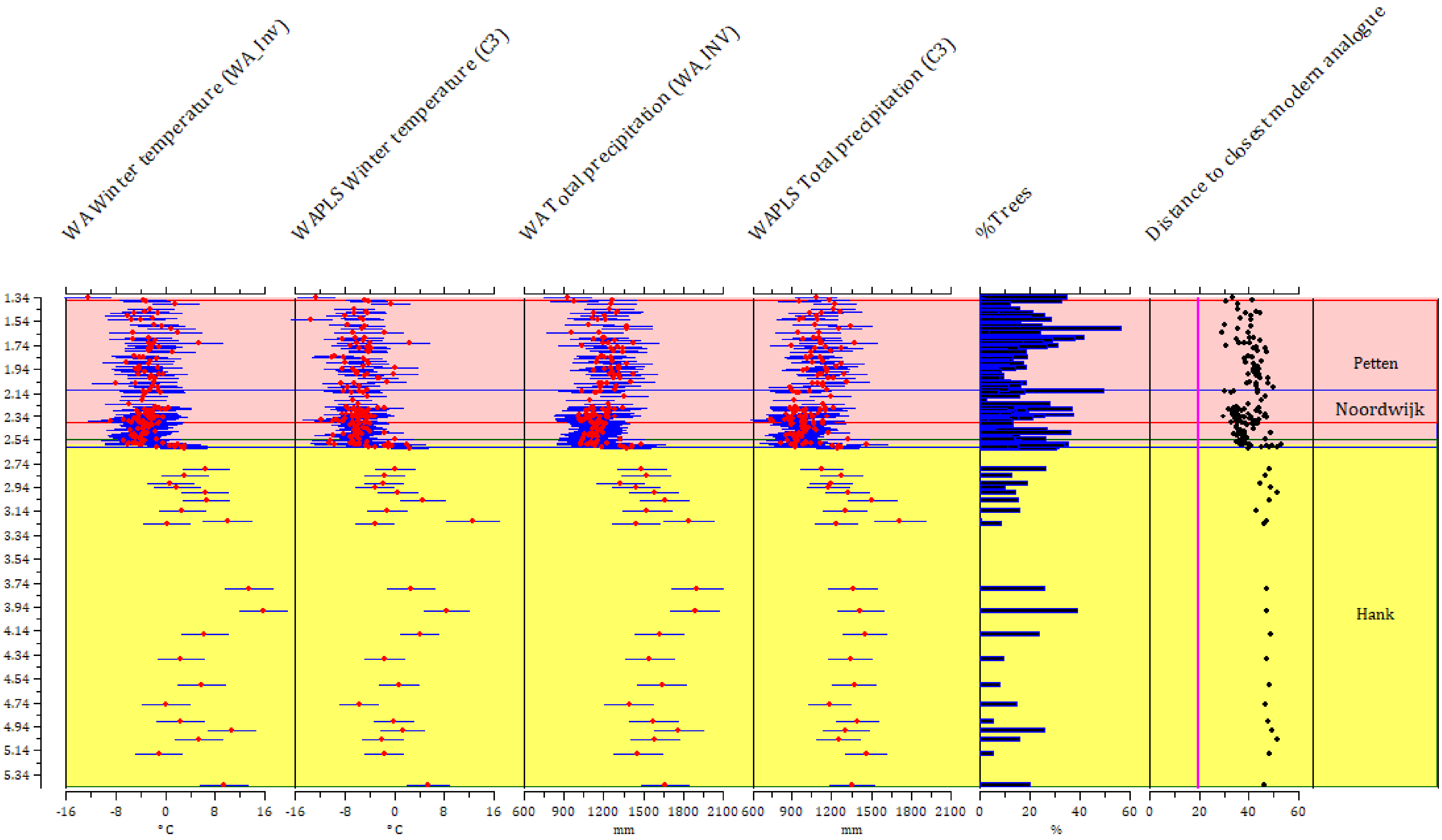



*Figure 7 Winter temperature and Total precipitation reconstruction for the period 1.34 – 5.45 Ma years ago. The yellow colour is the Late-Pliocene and the pink colour the Early-Pleistocene. Two different reconstruction methods are used: WA and WA-PLS. On the right side the percentage trees (without bisaccates) and the distances to the closest analogue are showed. The red dots are the data points, and the blue lines are the error bars. The blue box indicates the samples of Noordwijk, the red box of Petten and the green box of Hank. The pink line in the 'distance of closest analogue' indicates the 20th percentile line.*

# 5. Discussion

## 5.1 How accurate are the climate reconstructions?

It is important to know if the reference dataset is a valid dataset to use for this kind of climate reconstructions and period for the Netherlands. Two options are used to check whether the reconstruction made is correct. The first option is to compare the reconstruction with the percentage of deciduous trees, the section option is to check our reconstructions with known data from other reconstructions from the same period.

### 5.1.1 Comparison with percentage deciduous trees

From Marshak, 2010, pp 795, *"Deciduous forest are common in warmer climates and coniferous forests are more common in colder environments*". Bisaccates are more dominant during colder climates and grass pollen and deciduous forests are more dominant during warmer climate periods (e.g. Marshak, 2010, pp 795 – 795). However, during the Pliocene bisaccates are more common in warmer areas due to well developed soil. The precipitation graph should follow roughly the same curve as the temperature graph because during warmer periods more precipitation is expected. Warm air contain more water vapor, the warm air rises and cools. As a result, the water vapor condenses and the small droplets form large clouds which cause the precipitation.

The percentage deciduous trees (Appendix M) is showed in Figure 7, these trees are: *Acer, Alnus, Anacardia, Arecaceae, Betula, Carpinus, Carya, Celtis, Corylus, Fagus, Fraxinus, Juglans, Liquidambar, Nyssa, Quercus, Salix, Taxodium, Tilia* and *Ulmus*. Figure 6 shows the distribution of the pollen taxa with age (Ma). The Pliocene had a higher percentage of deciduous trees, bisaccates, ferns and grasses than the Pleistocene, a new indication of a warmer period during the Pliocene. The figure shows the trend into a colder period whereby the 'others' increase. The WA reconstruction (winter temperature) for the Late Pliocene follows the pattern of the extreme in the %trees graph. The precipitation graph follows roughly the same pattern as the temperature graph. The WA method fits better in the %trees graph than the WA-PLS method for both reconstructions. The peak, observed around 3.217 Ma, is not observed in the %trees graph. This peak is caused by a high amount of *Osmunda*. *Osmunda* is an indication of wet and warm conditions during the Late Pliocene. *Osmunda* is a terrestrial plant mainly found in temperate or sub-tropical regions, this indicates a relative more wat and warmer conditions during this period with less deciduous trees.

The Pleistocene has more samples than the Pliocene, because of the high number of samples, the Pleistocene is subdivided into different parts to observe the data closer. In general, the WA-PLS reconstruction follows roughly the same pattern as the %trees graph. Two peaks are observed closer because the peaks are relatively high in AP% in comparison with the reconstruction temperature. The peaks are caused by sample 23 and 26, these two sample contain *Alnus* (11/16 %), *Betula* (9/8 %), *Poaceae* (8/9%), *Polypodium* (9/16 5) and *Quercus* (25/- 5).

The period (2.18 – 1.98 Ma) has a temperature range between -7.929 °C and -0.755 °C for WA and -6.813 °C and 3.056 °C for WA-PLS. The tree percentage varies between 5 and 87% with two peaks which are not observed in the WA and WA-PLS reconstruction. Also in the precipitation graphs, these increases in %trees are not observed. The two peaks are caused by sample 3 and 131 and contain mainly ferns, like *Sphagnum* (18/5 %) and *Polypodium* (5/13 5) and grasses like *Poaceae* (20/5 %). The deciduous trees found in these samples are *Alnus, Betula* and *Corylus*. For the period 1.98 – 1.78 Ma, one outlier is observed caused by sample 103. This sample contain mainly ferns like *Polypodium* and *Sphagnum*, also some *Betula* is observed. The period, 1.78 – 1.58 Ma, follows roughly the same pattern as the reconstructed data. One data point has a relative high percentage of deciduous trees and a very low temperature of -8 °C. This datapoint is sample 93 and has a high percentage of ferns (*Osmunda, Polypodium*) and some deciduous trees (*Alnus, Betula, Quercus*). The youngest part of the Pleistocene has one peak with high percentages of deciduous trees and a relative low temperature/percentage. This is caused by sample 83 with *Alnus* (25%), *Quercus* (12,5%), *Ericaceae* (12,5%), *Poaceae* (25%) and *Polypodium* (25%). In general, the WA-PLS temperature range is higher than the WA reconstruction, however, the same mismatches with the %deciduous trees are observed. The most common pollen taxa are *Polypodium*, *Poaceae* and *Sphagnum* which are found in the outlying samples. The combination of these pollen taxa causes probably a higher temperature than the %deciduous trees predicts.

### 5.1.2 Comparison with other datasets

The reconstructed data from Hank is compared with a temperature reconstruction from Dearing Crampton-Flood et al., (2020). Dearing Crampton Flood et al., (2020) reconstructed the mean temperature by using three independent organic temperature proxies for sea surface and terrestrial temperatures based on different lipid biomarkers. A multi-proxy record of Pliocene climate change is made based on the sedimentary record for the borehole of Hank. This borehole is the same as used in this guided research.

The reconstructed data from this guided research follows roughly the same pattern as the reconstruction from Dearing Crampton-Flood et al., (2020), however, some differences are observed (Figure 9). The differences are observed between 2.58 – 2.93 Ma and 3.32 – 4.37 Ma. An important note to this comparison is the amount of datapoints, the reconstruction from Dearing Crampton-Flood et al., (2020) contain more datapoints than this research had used. More datapoints makes the reconstruction more precise and accurate. Some increase of decrease in temperature are not registered in our reconstructions because the datapoints are not present. Because of the lack of datapoints, the differences described below can be explained. For the period 2.58 – 2.98 Ma, the reconstructed GDGT data from Dearing Crampton-Flood et al., (2020) varies between 4.4 – 7.9 °C and the reconstructed data in this guided research fluctuates between 11.07 – 16.84 °C. During the other period (3.32 – 4.37 Ma), the reconstruction of Dearing Crampton-Flood et al., (2020) fluctuates between 3.91 – 12.4 °C and a variety of 10.15 – 24.95°C for this reconstruction. It is hard to conclude from this comparison which transfer function is better to use because both, WA and WA-PLS, following the same pattern. WA seems to fit a little better than WA-PLS.

Another dataset which used to compare the reconstructed data is the data from Zagwijn (1992). Zagwijn (1992) made the best guess for the Late Pliocene and Early Pleistocene for the Netherlands, however, his data is not validated, and the age model is old. Zagwijn (1992) used the minimal summer temperature following the method of Iversen (1944) (Figure 8). His guess for the mean temperature in July varies from 0 to 20 °C and during the Late Pliocene the

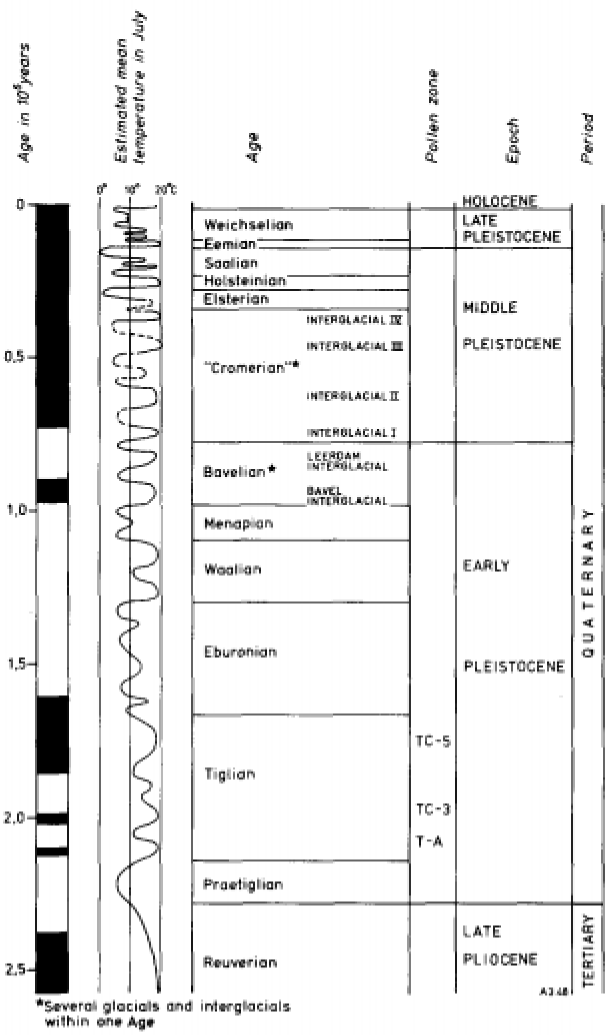



*Figure 8 Temperature reconstruction from Zagwijn (1992)*

summer temperature does not drop below the 10 °C. Zagwijn' temperature guess for the Early Pleistocene is warmer during the transition to the Early Early Pleistocene and the same during the Early Pleistocene. The data from Zagwijn (1992) indicates that our temperature reconstruction is in the right temperature range.

Comparing our reconstructed data with other datasets indicate that the temperatures are roughly in the same temperature ranges. For further indication it is good to compare the data with more datasets and to use more data points for a better comparison.

### 5.1.3 The differences between WA and WA-PLS

WA-PLS is a modification of WA, however, despite the modification WA is the most used method for paleo-environmental reconstructions. WA-PLS wants to fit all the variation into the reconstruction, while WA just showed the reconstruction without fitting it all. The WA-PLS method indicates a higher temperature for both Late Pleistocene and Early Pleistocene. The WA indicates more precipitation for the Late Pliocene and less precipitation for the Early Pleistocene in comparison with the WA-PLS method. Two methods are used to indicate which method is better to use for this timespan and region. The first method is the percentage deciduous trees in the samples. The WA seems to fit better in the older data from the Late Pliocene, while the younger data better fits in the WA-PLS reconstruction.

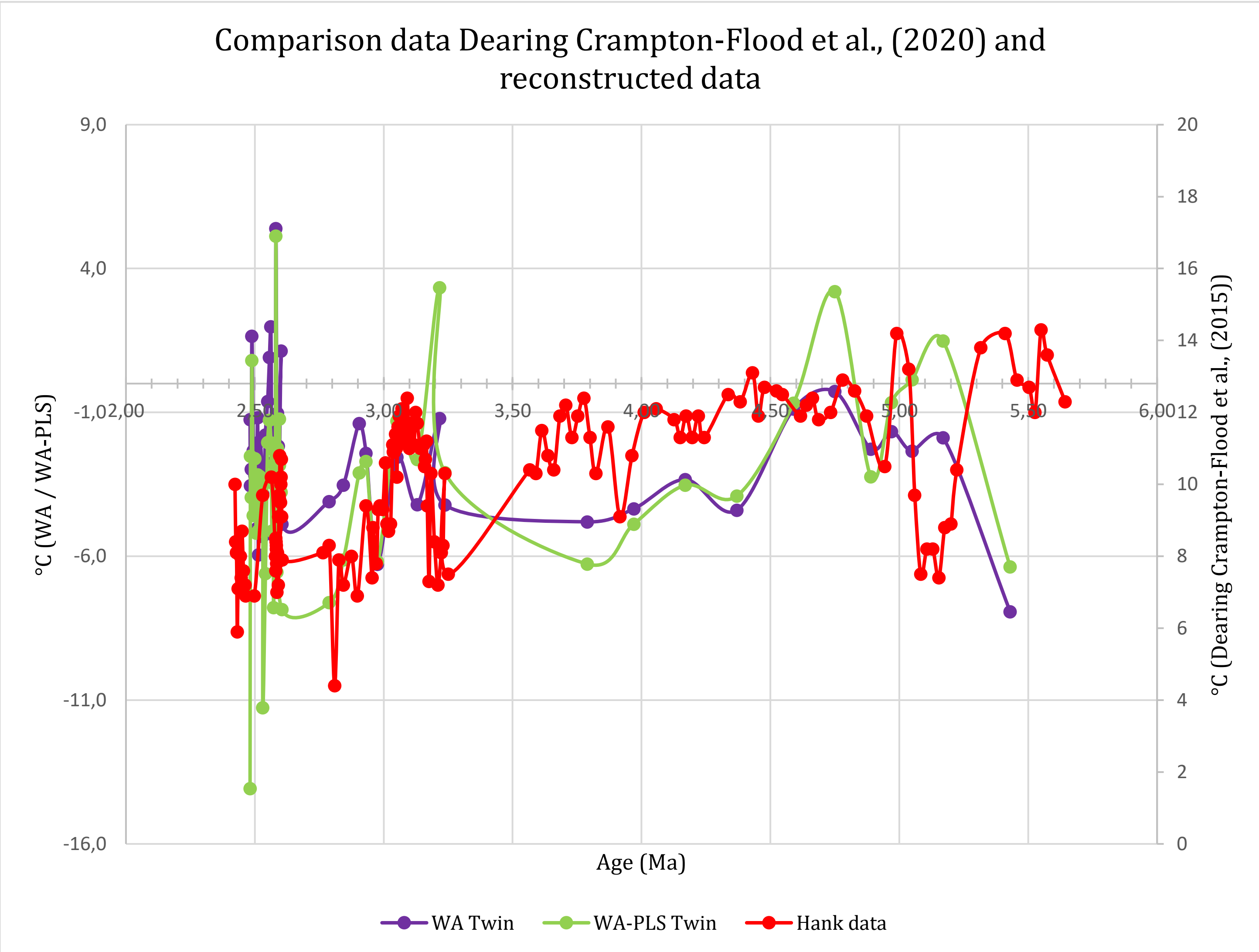



*Figure 9 Comparison reconstructed WA and WA-PLS method with data from Dearing Crampton-Flood et al., (2020). The winter temperature is purple (WA) and green (WA-PLS). The data from Dearing Crampton-Flood et al., (2020) is red.*

An observation was the existence of the ferns during the periods of high percentage of deciduous trees with colder temperatures. The ferns seems to create a colder temperature with an average precipitation pattern. Especially *Polypodium*, a spore, has a high occurrence during the colder temperature periods with a high occurrence of deciduous trees. This observation was also done in Chapter 4 Results, whereby the colder temperatures were characterised by the occurrence of *Polypodium* for both Late Pliocene and Early Pleistocene. A suggestion should be to do a run without the ferns for future research to indicate whether the ferns actually bias the temperature reconstructions.

The second method is the comparison with known data from Dearing Crampton-Flood et al., (2020). It is hard to indicate which method is better because of the lack of datapoints for our reconstruction. For the older data, the WA seems more likely because the extremes are more visible in Figure 9 than by the WA-PLS method. Mainly during the period 3.50 – 4.50 Ma, a low resolution have been reconstructed for this guided research and it is hard to make any conclusions.

## 5.2 How accurate is the data from Whitmore et al., (2015)?

Sample GPD1649, GPD1950, GPD1941 and GPD1952 were removed from the dataset because of the outlying position in the PCA and CA graph. GPD1649 contain a high percentage of *Pinus, Quercus* and *Arecaceae*. GPD1950 contain a high percentage of *Abies, Betula, Pinus, Polypodium* and *Quercus* and sample GPD1952 *Betula* and *Quercus*. The removed samples have a latitude between 29° and 35°. *Quercus* is a dominant pollen species in between these latitudes. Also sample FAK3 and FAK 9 are removed, these two samples contain a high amount of *Taxodium* and *Polypodium*. These two samples are relative warm, known by the high percentage of *Taxodium*. This could be a reason why these two samples were outliers in the PCA and CA. As indicated by Table 2 and Table 3, not much changed after removing these six samples. Removing these samples makes the reconstruction a little bit precise but do not change enough to see large changes.

A specific part of the dataset of Whitmore et al., (2015) was used for this reconstruction. The limit of 51° N is made because of the temperature range from the pollen species. Going further north than 51° N, the climate is more similar to montane and continental Europa instead of the lowland sea climate in the Netherlands during the Late Pliocene and Early Pleistocene. The real problem is the occurrence of similar types like *Juniperus* in cool areas that (within the Cupressaceae type) had a broader/warm distribution in the Pliocene in Europe. Also some cold pollen species do not occur in the Florida region, like *Abies* and *Picea*, now relatively a warm and cold pollen species. In the dataset of Hank they occur at the same time, an indication that some plants changed their climatic characteristics which makes it hard to compare the fossil datasets with the modern ones. Because these two species do not overlap in the modern dataset (Appendix A). Other pollen taxa which have more correspondence in the fossil dataset instead of the modern one were Arecaceae, Caryophyllaceae, Malvaceae, *Ranunculus, Sphagnum* and *Taxodium*, a non-analogue condition. These are the taxa that characterize the right/upper side of the CA plot.

The occurrence of *Betula* and *Quercus* at the same time occurs only in the upper part (Early Pleistocene), during the Late Pliocene they barely occur together. One of the options is the occurrence of different *Quercus* species.  Another observation in the pollen species was the occurrence of the bisaccates. After analysing the PCA and the CA, *Pinus,* and *Pinus haploxyon*-type are removed from the dataset because they were outliers in the graphs. It was suspected that this pollen taxon was located nearby the other bisaccates. Another reason to remove *Pinus* was because of the dominant occurrence in the samples and the strong bi-modal distribution, this could give a disturbance in the reconstruction has too much emphasis on the reconstruction.

During the Late Pliocene and Early Pleistocene, *Pinus* occurs in a different environment than present. In the Florida data, *Pinus* only has a decrease at a latitude of ~38 °. During the Early-Pleistocene, *Pinus, Picea* and *Taxodiaceae* pollen grains are dominant with *Fagus, Quercus, Zelkova* and *Ulmus* pollen grains (Fujiki and Ozawa, 2008). In the Early Pleistocene, pollen grains (*Liquidambar, Tsuga, Pterocarya* and *Carya*) disappeared (e.g. Svenning, 2003). *Pinus, Picea* and *Taxodiaceae* pollen grains were still dominant and *Tsuga* pollen increased (Fujiki and Ozawa, 2008; Svenning, 2003). The climate became colder from the Late Pliocene to the Early Pleistocene and *Pinus, Tsuga* and *Abies* grew in the highland areas at altitudes from 1500 to 3000 m above sea level (Fujiki and Ozawa, 2008). In Florida, the *Pinus* occur in warm lowlands, while during the Late Pliocene and Early Pleistocene *Pinus* occur in the colder highlands in Europe. *Pinus haploxyon* occurs in the warmer coastal lowlands during the Pliocene. Another aspect was the differences between the occurrence of *Picea, Pinus* and *Abies*. During the Pliocene *Picea* and Abies had a high occurrence in overall, currently *Picea* and *Abies* occurring in the higher colder areas. Pinus has a higher abundance during warmer temperatures and wetter conditions while *Abies* abundance results in colder temperatures due to its present distribution in that habitat. In general, the bisaccates had different environments during the Pliocene and the Pleistocene for Europe and America. This makes it hard to make a good reconstruction with these taxa in the reconstruction, for future research it seems good to indicate which kind of the bisaccates occur in the different areas and to specify their climate characteristics.

Another striking aspect is the occurrence of *Osmunda*, *Osmunda* is an indicator of warm and wet conditions during the Late Pliocene. The Whitmore data contain two samples with a high amount of *Osmunda*, these samples are GPD165 and GPD591. The average temperature for this sample is 16.23 °C with a total precipitation of 1105 mm/year. The average winter temperature is 6.7 °C degrees, in comparison to other samples with less *Osmunda*, this temperature is low. GPD591 has an average temperature of 3.34 °C degrees and a winter temperature of -3.8 °C degrees.

For future research, the dataset of Whitmore et al., (2015) is a good dataset to use, however, it is important to know the climatic characteristics from the dominant pollen species for now and during the Late Pliocene and Early Pleistocene.

# 6. Conclusions

The aim of this guided research was to reconstruct the climate during the Late Pliocene and Early Pleistocene for the Netherlands based on pollen assemblages data with quantitative paleo-climatic reconstruction methods. The reconstruction for this timespan is based on three different terrestrial datasets from the Netherlands. The datasets are in Noordwijk, Petten and Hank. Two different pollen-climate transfer functions were calculated using two way Weighted Averaging (WA) with inverse deshrinking and 3- component Weighted Averaging Partial Least Squares (WA-PLS) regression. Calibration set species data values were square-root transformed for WA and WA-PLS regression to reduce noise in the data.

Two different climate parameters are reconstructed: winter temperature and total precipitation. *Pinus* and *Pinus haploxyon* have a very high presence in all the data samples, so these pollen species are removed from the reconstruction, substantiate by the PCA and CA. *Pinus* and *Pinus haploxyon* were both outliers in these graphs.

Two methods were used to indicate which technique is better to use for climate reconstruction for this time span. The first method was to compare the reconstruction with the percentage deciduous trees and the second one was to compare the reconstruction with known data. Concluded from the first method, the WA method is better to use with older data (Late Pliocene), WA-PLS can be used for younger data. WA-PLS wants to fit all the variation in the reconstruction, while this is for older data not the right way to reconstruct. As we see in our data, some pollen species do not exist anymore or have another climate characteristic in comparison with their present pollen species. If WA-PLS tries to fit all the variation in the reconstruction, some mistakes are going to be made in the reconstruction.

The other method was to compare our data with known data from Dearing Crampton-Flood et al., (2020) and to indicate if the same trends from Zagwijn (1992) are observed for the Pleistocene. Again, WA seems to fit better in the reconstruction, however, it is hard to make a good conclusion because of the differences in datapoints. For a long time, Zagwijn (1992) has made the best estimate for the temperature range in the Netherlands for the Late Pliocene and Early Pleistocene. Zagwijn (1992) made a summer temperature reconstruction, while our data is a winter temperature reconstruction. The data of Zagwijn (1992) is compared with the average temperature reconstruction. The Late Pliocene data seems right, the temperature is in the same range as Zagwijn (1992) reconstructed. The transition to the Early Pleistocene seems too cold, however, the reconstruction from Zagwijn (1992) is older, however, not wrong, and is not dated accurately.

For further research, it would be better to use more fossil datapoints for a better comparison with other data. It is hard to indicate if the data is right because, especially for the Late Pliocene, not enough data points are available. Also, another dataset could be used as calibration set, for example for the South-Eastern part of Asia. It is also good to know the difference between the climate characteristics of some pollen species now and during the Late Pliocene and Early Pleistocene. Also, the influence of the Ferns on the climate was a striking element found in this research. Another method which could be used for the reconstruction is CREST, a method from Chevalier. The CREST method is related to a Bayesian approach combines presence-only occurrence data and modern climatology's to estimate the conditional response of a given taxon to a variable of interest.

# 7. References

**Avizinis, J., Web III, T.** (1985) *The computer file of modern pollen and climatic data at Brown University*. Department of Geological Sciences, Brown University, Providence.

**Bartlein, P.J., Harrison, S.P., Brewer, S.C., Connor, S., Davis, B.A.S., Gajewski, K., Guiot, J., Harrison-Prentice, T.I. Henderson, A., Peyron, O., Prentice, I.C., Scholze, M., Seppä, H., Shuman, B.N., Sugita, S., Thompson, R.S., Viau, A.E., Williams, J.W., Wu, H.** (2011) *Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis*. Clim. Dyn. 37, 775-802.

**Birks, H.J.B., Birks, H.H.** (1980) *Quaternary paleoecology*. London (Reprinted 2004 by Blackburn Press, New Jersey): Edward Arnold.

**Birks, H.J.B.** (1981) *The use of pollen analysis in the reconstruction of past climates: a review.* In Wigley TML, Ingram MJ, Farmer, G. editors: Climate and history. Cambridge: Cambridge University Press; p 111 – 138.

**Birks, H.J.B.** (1995) *Quantitative palaeoenvironmental reconstructions*. Stat. Model. Quat. Sci. data. Tech. Guid. 5, 161-254.

**Birks, H.J.B.** (1998) *Numerical tools in palaeolimnology-progress, potentialities, and problems*. J. paleolimnol. 20, 307-332.

**Birks, H.J.B, Heiri, O., Seppã, H., Bjuna, A.E.** (2010) *Strengths and weakness of quantitative climate reconstructions based on Late-Quaternary biological responses*. Open Ecol. J., 3, 68-110.

**Birks, H.J.B** (2019) *Contributions of Quaternary botany to modern ecology and biogeography,* Plant ecology & Diversity, 12:3-4, 189 – 385, DOI: 10.1080/17550874.2019.1646831.

**Birks, H.H..** (2003) *The importance of plant macrofossils in the reconstruction of Lateglacial vegetation and climate: examples from Scotland, western Norway and Minnesota, USA*. Quat Sci Rev. 22: 453-473.

**Bradley, R.S.** (2015) *Paleoclimatology: Reconstructing climates of Quaternary*, p. 675.

**Bradshaw, R.H., Webb III, T.** (1985) *Relationship between contemporary pollen and vegetation data from Wisonsin and Michigan, USA*. Ecology 69, 941 – 955.

**Brewer, S.C., Guiot, J., Barboni, D.** (2007) *Pollen methods and studies, use of pollen as climate proxies*. Encyclopedia of Quaternary Science, Elsevier, Oxford, pp. 2497 – 2508.

**Chalk, T.B., Capron, E., Drew, M., Panagiotopoulos, K.** (2017) *Interglacials of the 41 ka world and the Mid-Pleistocene transition*. PAGES Mag. 25(3): 155.

**Cheddadi, E.** (2002) European Pollen database. IGBP PAGES/World Data Centra for Paleoclimatology. NOAA/NGDC Paleoclimatology Program, Boulder, Co, USA.

**Chen, F., Xu, Q., Chen, J., Birks, H.J.B., Liu, J., Zhang, S., Jin, L., An, C., Telford, R.J., Cao X., Wang Z., Zhang, X., Selvaraj, K., Lu. H., Li, Y., Zheng, Z., Wang, H., Zhou, A., Dong, G., Zhang, J., Huang, X., Bloemendal, J., Rao, Z.** (2015) *East Asian summer monsoon precipitation variability since the last deglaciation*. Sci. Rep. 5, 1 – 11.

**Chevalier, M., Davis, B.A.S., Heiri, O., Seppä, H., Chase, B.M., Gajewski, K., Lacourse, T., Telford, R.J., Finsinger, W., Guiot, J., Kühl, N., Maezumi, S.Y., Tipton, J.R., Carter, V.A., Brussel, T., Phelps, L.N., Dawson, A., Zanon, M., Vallé, F., Nolan, C., Mauri, A., de Vernal, A., Izumi, K., Holmström, L., Marsicek, J., Goring, S., Sommer, P.S., Chaput, M., Kypriyanov, D.**

(2020) *Pollen-based climate reconstruction techniques for late Quaternary studies*. Earth-Science Reviews 103384.

**Cole, H.** (1969) *Objective reconstruction of the paleo-climate record, through application of eigenvectors of present-day pollen spectra and climate to the late Quaternary pollen stratigraphy*. Ph.D. Dissertation, University of Wisconsin, Madison.

**Dahl, E.** (1998) *The phytogeography of Northern Europe (British Isles, Fennoscandia and ajacent areas)*. Cambridge: Cambridge University Press.

**Davis, R.B., Webb III, T.** (1975) *The contemporary distribution of pollen in eastern North America: a comparison with the vegetation*. Quaternary Research 5, 395 – 434.

**Dearing Crampton-Flood, E., Peterse, F., Munsterman, D., Sinninghe Damsté, J.S**. (2018) *Using tetraether lipids archived in North Sea Basin sediments to extract North Western European Pliocene continental air temperatures*, Earth Planet. Sc. Lett. 490, 193 – 205.

**Dearing Crampton – Flood, E., Noorbergen, L.J., Smits, D., Christine Boschman, R., Donders, T.H., Munsterman, D.K., ten Veen, J., Peterse, F., Lourens L., Sinninghe Damsté, J.S.** (2020) *A new age model for the Pliocene of the southern North Sea basin: a multi-proxy climate reconstruction*. Clim. Past, 16, 523 – 541.

**Delcourt, P.A., Delcourt, H.R., Davidson, J.L.,** (1983). *Mapping and calibration of modern pollen-vegetation relationships in the Southeastern United States*. Review of Palaeobotany and Palynology 39, 1 – 45.

**Donders, T.H., van Helmon, N.A.G.M., Verreussel, R., Munsterman, D., ten Veen, J., Speijer, R.P., Weijers, J.W.H., Sangiorgi, F., Peterse, F., Reichart, G., Sinninghe Damsté, J.S., Lourens, L., Kuhlmann, G., Brinkhuis, H.** (2018) *Land-sea coupling of early Pleistocene glacial cycles in the Southern North Sea exhibit dominant Northern Hemisphere forcing*. Climate of the Past, 14. 397 – 411.

**Dowsett, H.J., Robinson, M.M., Haywood, A.M., Hill, D.J., Dolan, A.M., Stoll, D.K., Chan, W.L., Abe-Ouchi, A., Chandler, M.A., Rosenbloom, N.A., Otto-Bliesner, B.L.** (2012) *Assessing confidence in Pliocene sea surface temperature to evaluate predictive models*. Nat. Clim. Change 2, 365-371.

**Duval, M., Arnold, L.J., Pares, J.M., Hoffman, D.L.** (2015) *The Jaramillo Subchron and the Early Middle Pleistocene Transition in continental records from a multidisciplinary perspective*. Quatern Int. 289: 1 – 6.

**Finsinger, W., Colombaroli, D., De Beaulieu, J.-L., Valsecchi, V., Vannière B., Vescovi, E., Chapron, E., Lotter, A.F., Magny, M., Tinner, W.** (2010) *Early to mid-Holocene climate change at Lago dell'Accesa (Central Italy): climate signal or anthropogenic bias?* J. Quat. Sci. 25, 1239 – 1247.

**Gibbard, P.L.** (1988) *The history of the great northwest European rivers during the past three million years*. Philos. T. Roy. Soc. Lond. B. 318, 559-602.

**Gibbard, P.L., Lewin, J.** (2016) *Filling the North Sea Basin: Cenozoic sediments sources and river styles* (André Dumont medallist lecture 2014), Geol. Belg., 10, 201 – 217.

**Grimm, E.** (2000a) *North American Pollen Database*. IGBP PAGES/World Data Centre for Paleoclimatology. NOAA/NGDC Paleoclimatology Program, Boulder, CO, USA.

**Grimm, E**. (2000b) *Global Pollen Database*. IGBP PAGES/World Data Centre for Paleoclimatology. NOAA/NGDC Paleoclimatology Program, Boulder. CO. USA.

**Head, M.J., Gibbard, P.L.,** (2015) *Early-Middle Pleistocene transition: Linking terrestrial and Marine realms*. Quatern Int. 289: 7 – 46.

**Huizer, J., Weerts, H.J.T.** (2003) *Beschrijving lithostratigrafische eenheid: de Formatie van Maasluis.* Nederlands Instituut voor Toegepaste Geowetenschappen TNO.

**Imbrie J., Kipp, N.G.** (1971) *A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core. Late Cenozoic Glacial Ages*. 3, 71-181.

**Imbrie, J., van Donk, J., Kipp, N.G.** (1973) *Paleo-climatic investigation of a Late Pleistocene Caribbean Deep-Sea core: comparison of isotopic and faunal methods*. Quat. Res. 3, 10-38.

**Imbrie, J., Imbrie, J.Z.** (1980) *Modelling the climatic response to orbital variations*. Science 29, Vol 207, 943 – 953.

**Iversen, J.** (1944) *Visum, Hedera and Illex as climate indicators. A contribution to the study of the post-glacial temperature climate*. Geologiska Föreningens I Stockholm Förhandlingar. 66(3):463-483.

**Jackson, S.T., Williams, J.W**. (2004) *Modern analogues in Quaternary paleo-ecology: Here today, gone yesterday, gone tomorrow?* Annual Review of Earth and Planetary Science, 32. 496 – 537.

**James, F.C. Johnston, R.F., Wamer, N.O., Niemi, G.J., Boecklen, W.J.** (1984) *The Grinnellian niche of the wood thrush*. Am Nat. 124; 17 – 30.

**Klemm, J., Herzschuh, U., Pisaric, M.F.J., Telford, R.J., Heim, B., Pestryakova, L.A.** (2013) *A pollen-climate transfer function from the tundra and taiga vegetation in Arctic Siberia and its applicability to a Holocene Record*. Palaeogeogr. Palaeoclimatol. Palaeoecol. 386, 702 – 713.

**Li, J., Dodson, J., Yan, H., Cheng, B., Zhang, X., Xu, Q., Ni, J., Lu, F.** (2017) *Quantitative precipitation estimates for the northeastern Qinghai-Tibetan Plateau over the last 18000 years.* J. Geophys. Res. 122, 592 – 602.

**Lisiecki, L.E., Raymo, M.E.** (2005) *A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}O$ records*. Paleo-oceanography, Vol. 20, pp 1 – 17.

**Litt, T., Ohlwein, C., Neumann, F.H., Hense, A., Stein, M.** (2012) *Holocene climate variability in the Levant from the Dead Sea pollen record*. Quat. Sci. Rev. 49, 95 – 105.

**Lotter, A.F., Birks, H.J.B., Eicher, U., Hofmann, W., Schwander, J., Wick, L.** (2000) *Younger Dryas and Allerod summer temperature at Gerzensee (Switzerland) inferred from fossil pollen and cladoceran assemblages*. Palaeogeogr. Palaeoclimatol. Palaeoecol. 159, 349 – 361.

**Lotter, A.F., Heri, I., Brooks, S., van Leeuwen, J.F.N., Eicher, U., Ammann, B.** (2012) *Rapid summer temperature changes during termination 1a: high-resolution multi-proxy climate reconstruction from Gerzensee (Switserland)*. Quat. Sci. Rev. 36, 103 – 113.

**Lu, F., Ma, C., Zhu, C., Lu, H., Zhang, X., Huang, K., Guo, T., Li, K., Li, L., Li, B., Zhang, W.** (2018) *Variabilitiy of East Asian summer monsoon precipitation during the Holocene and possible forcing mechanisms.* Clim. Dyn. 0, 1 – 21.

**Marshak, S.** (2012) *Earth portrait of a planet (Fourth edition)*. W.W Northon & Company.

**Maslin, M.A., Brierley, C.M.** (2015) *The role of orbital forcing in the Early Middle Pleistocene Transition*. Quatern Int. 289: 47 – 55.

**Mosbrugger, V., Utescher, T**. (1997) *The Co-existence approach – a method for quantitative reconstructions of Tertiary palaeoclimate data using plant fossils*. Palaeogeography, Palaeoclimatology, Palaeoecology, 134. 61 – 86.

**New, M., Lister, D.H., Hulme, M., Makin, I.** (2002) *A high-resolution data set of surface climate over global land areas*. Clim. Res. 21, 1-25.

**Nolan, C., Tipton, J.R., Booth, R.K., Hooten, M.B., Jackson, S.T.** (2019) *Comparing and improving methods for reconstructing peatland water-table depth from testate amoebae*. The Holocene 39, 1350 – 1361.

**Noorbergen, L.J., Lourens, L.J., Munsterman, D.K., Verreussel, R.M.C.H**. (2015) *Stable isotope stratigraphy of the early Quaternary of borehole Noordwijk, southern North Sea*.

**Overpeck, J.T., Webb, T., Prentice, I.C. (**1985) *Quantitative interpretation of fossil pollen spectra – dissimilarity coefficients and the method of modern analogs*. Quat Res. 23: 87 – 108.

**Prader, S., Kotthoff, U., Greenwood, D.R., McCarthy, P.M.G., Schmiedl, G., Donders, T.H.** (2020) *New Jerney's paleoflora and eastern North American climate through Paleogene-Neogene warm phases*. Review of Palaeobotany and Palynology, Vol. 297, 104224.

**Prell, W.L.** (1985) *The stability of low-latitude sea-surface temperatures: an evaluation of the CLIMAP reconstruction with emphasis on the positive SST anomalies*. Technical Report TE025. Department of Energy Washington DC, pp 1-60.

**Richard, P.,** (1976) *Relations entre la vegetation actuelle et le spectre pollinique au Québec*. Naturaliste Canadien 103, 53 – 66.

**Ruddiman, W.F.** (1990) *Changes in climate and biota on geologic time scale*. Tree, Vol. 5, no. 9, pp 285 – 288.

**Ruddiman, W.F.** (2013). *Earth's climate: past and future*. 3rd ed. New York: WH Freeman.

**Sakari Salonen, J., Helmens, K.F., Seppä, H., Birks, H.J.B.** (2013) *Pollen-based climate reconstructions over long glacial-interglacial timescales: Methodological tests based on the Holocene and MIS 5d-c deposits at Sokli, northern Finland*. J. Quat. Sci 28, 271 – 282.

**Seppä, H., Bjune, A.E., Telford, R.J., Birks, H.J.B., Veski, S.** (2009) *Last nine-thousand years of temperature variability in Northern Europe*. Clim. Past 5, 523 – 535.

**Simpson, G.L.** (2007) *Analogue methods in paleo-ecology: using the analogue package*. Journal of Statistical Software, 22, 1-29.

**Simpson, G.L.** (2007) *Analogue methods in paleo-climatology.* In: *Tracking environmental change using lake sediments*. Volume 5: Data handling and Numerical Techniques (Eds. Birks, H.J.B., Lotter, A.F., Juggings, S., Smol, J.P). Springer Dordrecht (in press).

**Svenning, J.** (2003) *Deterministic Plio-Pleistocene extinctions in the European cool temperate tree flora*. Ecology Letters, 6. 646 – 653.

**ter Braak, C.J.F., Juggings, S**. (1993) *Weighted Averaging Partial Least Squares Regression (WA-PLS) – an improved method for reconstruction environmental variables from species assemblages*. Hydrobiologia, 269, 485 – 502.

**ter Braak, C.J.F., Looman, C.W.N.** (1986) *Weighted averaging, logistic regression and the Gaussian response model.* Vegetatio 65, 3-11.

**Van Balen, R.T., Houtgast, R.F., van der Wateren, F.M., Vandenberghe, J., Bogaart, P.W**. *Sediment budget and tectonic evolution of the Meuse etachment in the Ardennes and the Roer Valley Rift System*. Global Planet. Change. 27, 113 – 129.

**Veski, S., Seppä, H., Stancikaite, M., Zernitskaya, V.P., Reitalu, T., Gryguc, G., Heinsalu, A., Stivrins, N., Amon, L., Vassiljev, J., Heiri, O.** (2015) *Quantitative summer and winter temperature reconstructions from pollen and chironomid data between 15 and 8 ka BP in the Baltic-Belarus area*. Quat. Int. 388, 4 -11.

**Voosen, P.** (2018) *Sticky glaciers slowed tempo of ice ages*. Science, 361(6404):739.

**Webb III, T, Bryson, R.A.** (1972) *Late- and Postglacial Climate change in the Northern Midwest, USA: quantitative estimates derived from fossil pollen spectra by multi-variate statistical analysis*. Quat. Res. 2, 70-115.

**Webb III, T., McAndrews, J.H.** (1976) *Corresponding patterns of contemporary pollen and vegetation in Central North America*. Geological Society of America Memoir 145, 267 – 299.

**Westerhoff, W.E.** (2009) *Stratigraphy and sedimentary evolution: The Lower Rhine-Meuse system during the Late Pliocene and Early Pleistocene (southern North Sea Basin)*, PhD Thesis, Vrije Universiteit Amsterdam.

**Westerhoff, W.E., Donders, T.H., Trabucho Alexandre, N., Busschers, F.S.** (2020) *Early Pleistocene Tiglian Sites in the Netherlands: A revised view on the significance for Quaternary stratigraphy*. Quaternary Science Reviews 242.

**Whitmore, J., Gajewski, K., Sawada, M., Williams, J.W., Shuman, B., Bartlein, P.J., Minckley, T., Viau, A.E., Webb III, T., Shafer, S., Anderson, P., Brubaker, L.** (2005) *Modern pollen data from North America and Greenland for multi-scale paleo-environmental applications*. Quaternary Science Reviews 24, 1828 – 1848.

**Williams, J.W., Shuman, B.N.** (2008) *Obtaining accurate and precise environmental reconstructions from the modern analogue technique and North America surface pollen dataset*. Quat. Sci. Rev. 27, 669 – 687.

**Zagwijn, W.H.** (1992) *The beginning of the ice age in Europe and its major subdivions*. Quaternary Science Reviews, Vol. 11. Pp 583 – 591.

**Ziegler, P.A.** (1990) *Geological Atlas of Westen and Central Europe*, 2nd edn, Geological Society Publishing House, 239, pp.

# List with figures

**Figures:**

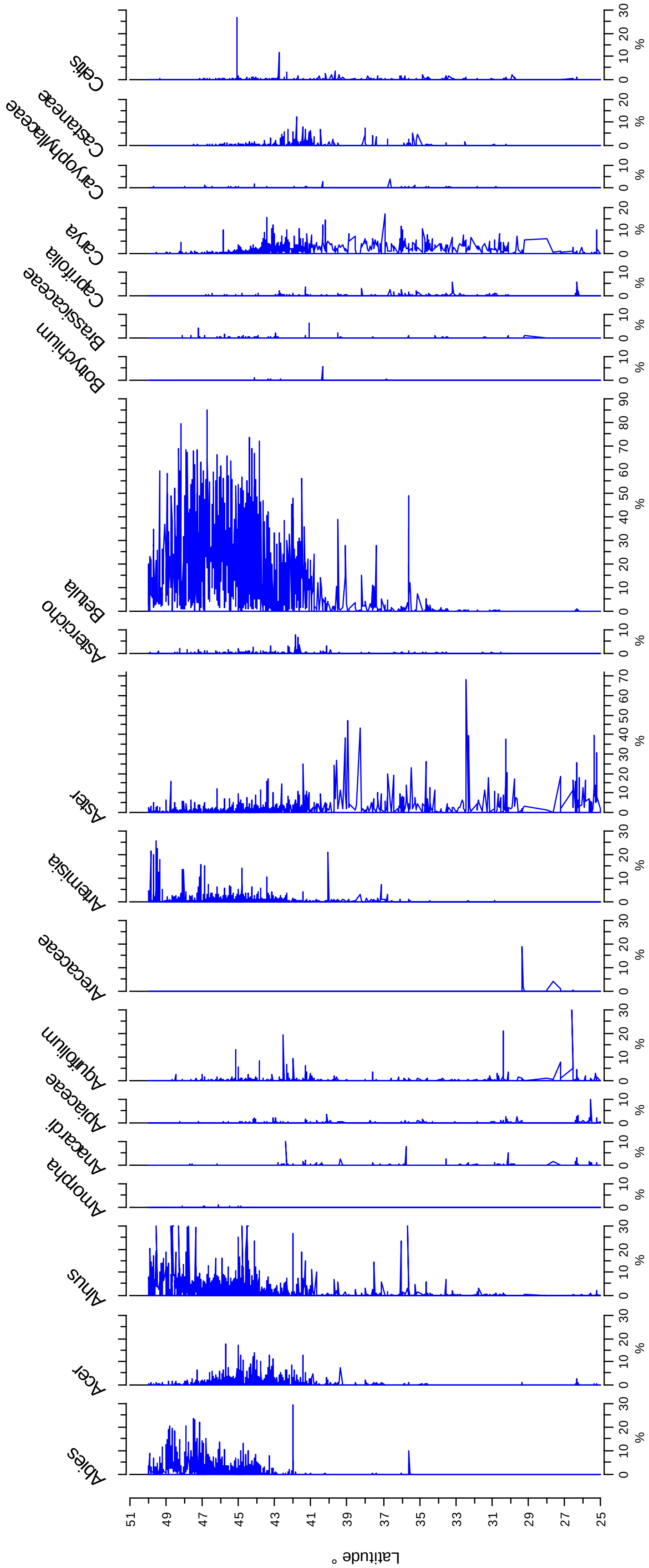


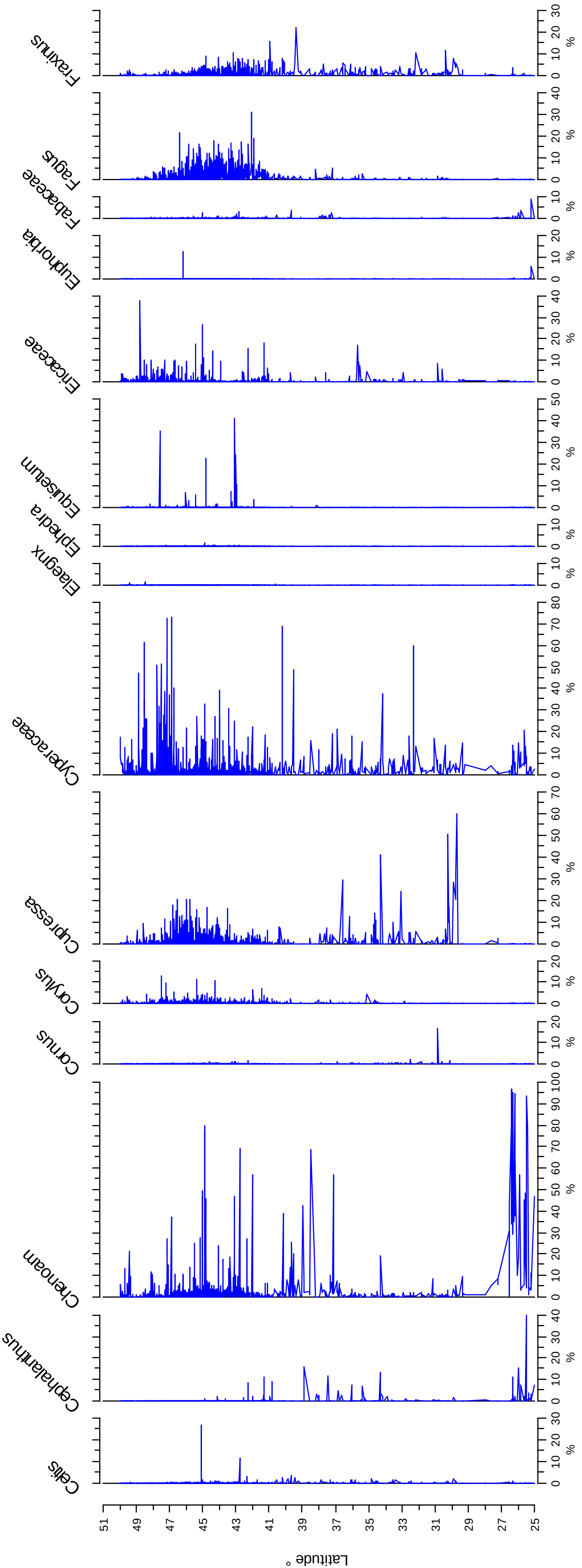
**Tables:**

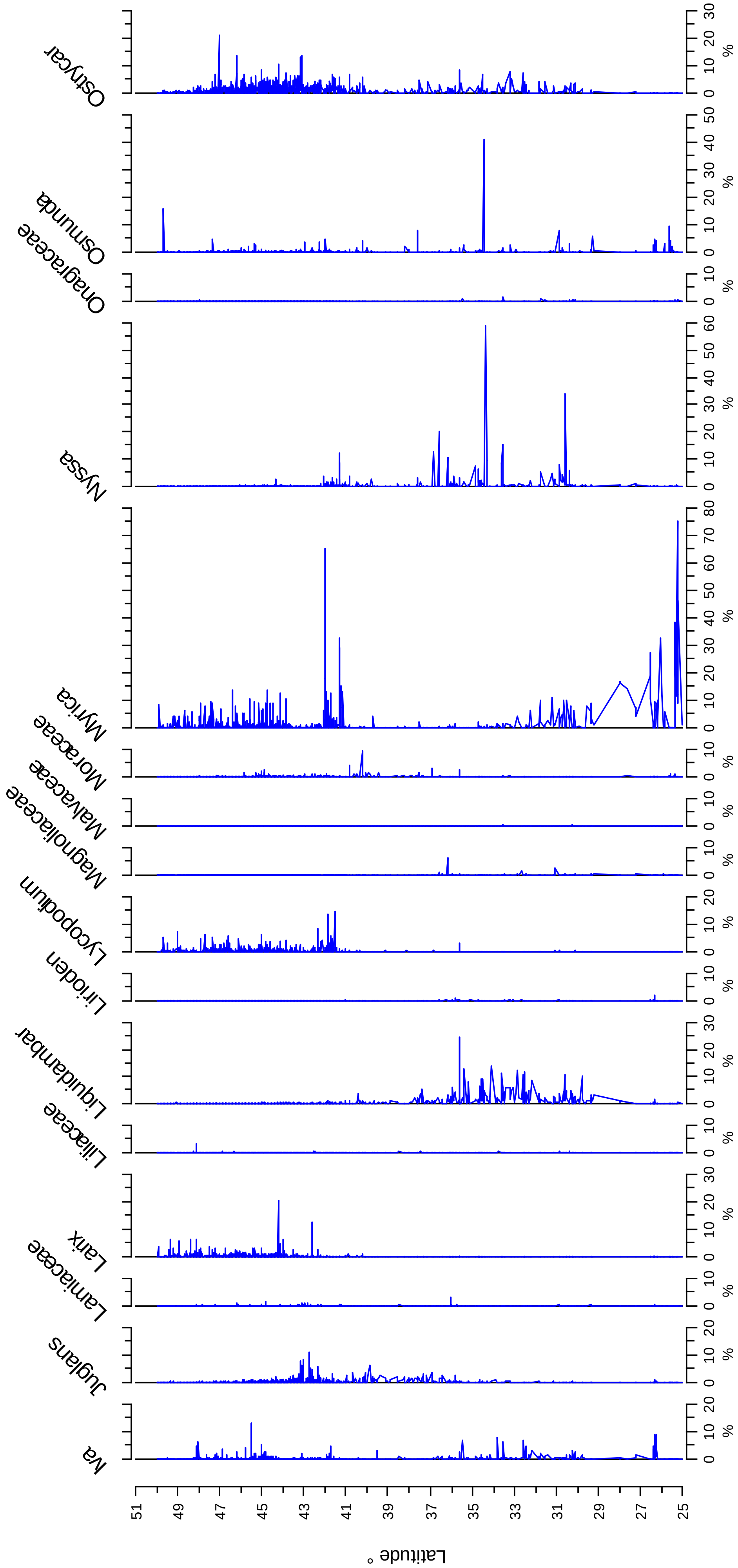
# Appendices

A. Pollen diagram modern dataset Florida
B. Pollen diagram fossil dataset Noordwijk
C. Pollen diagram fossil dataset Petten
D. Pollen diagram fossil dataset Hank
E. Dominant pollen species for fossil datasets with age
F. Principal Component Analyse Modern and Fossil data
G. Correspondence graph with all the data
H. Observation graphs WA and WA-PLS
I. Observation graphs WA and WA-PLS with removed samples
J. First climate reconstruction
K. Calibration sample number – age
L. Fossil pollen species combined with modern pollen species
M. Percentage trees & bisaccates

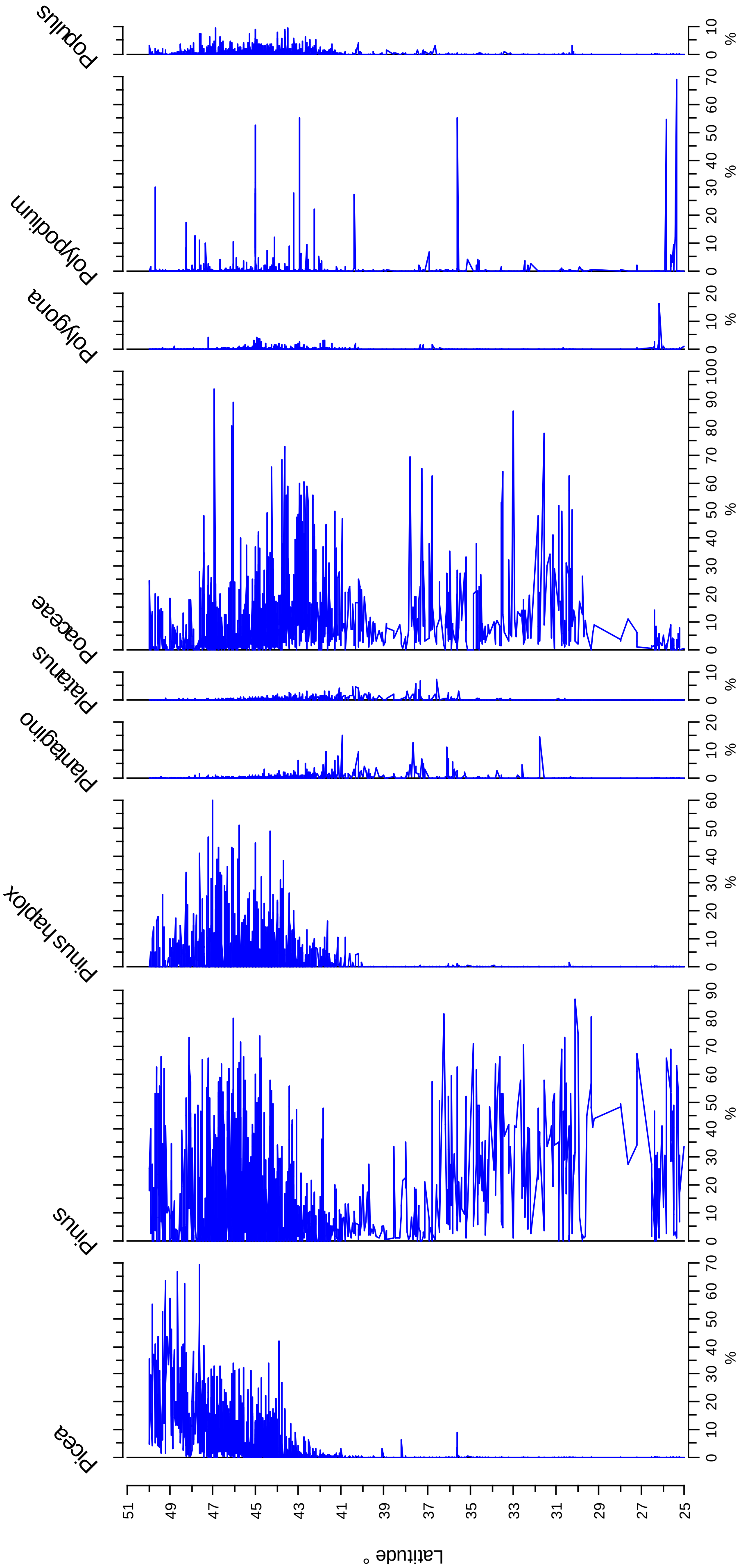
# A. Pollen diagram modern data set Florida

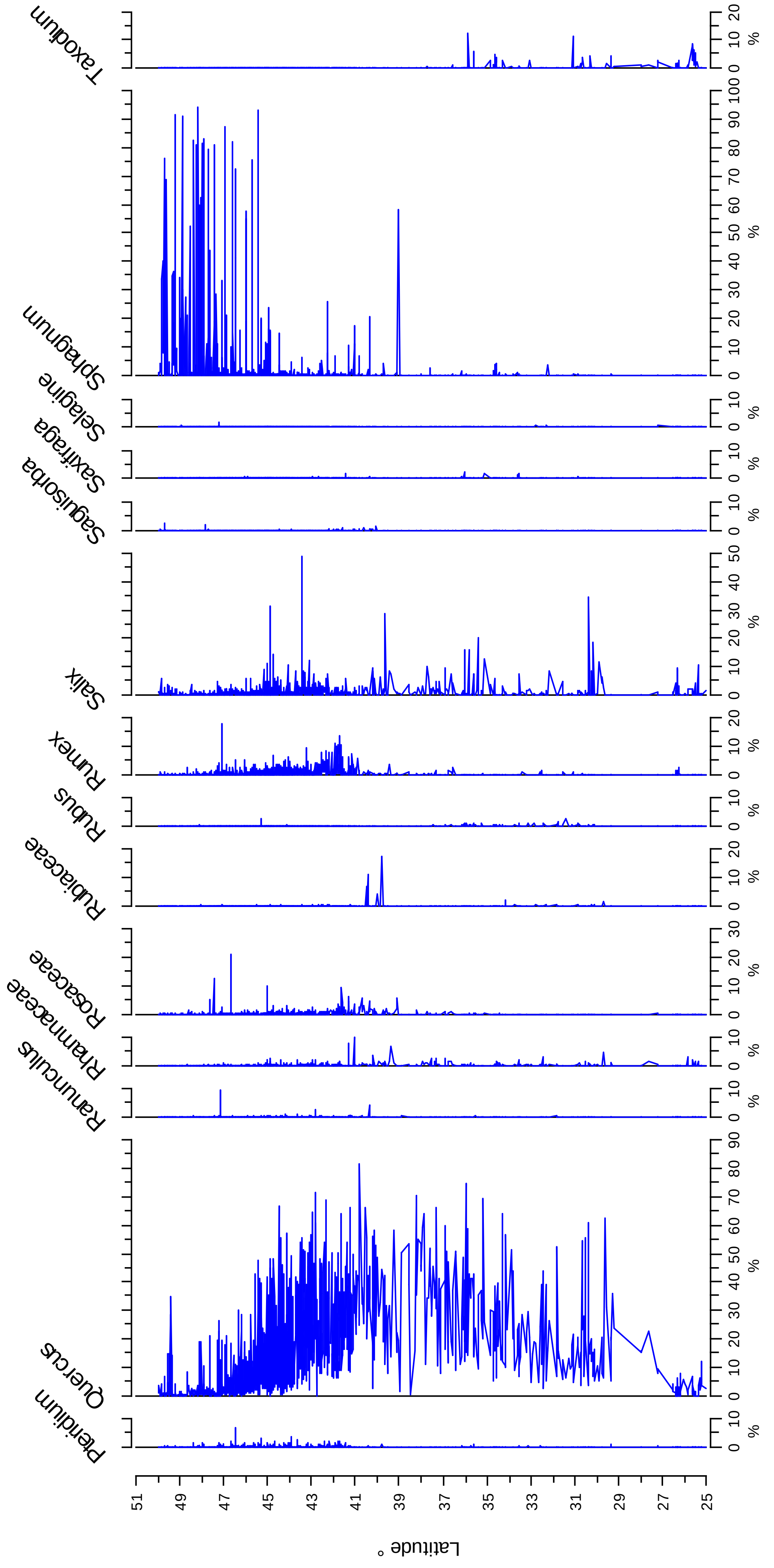


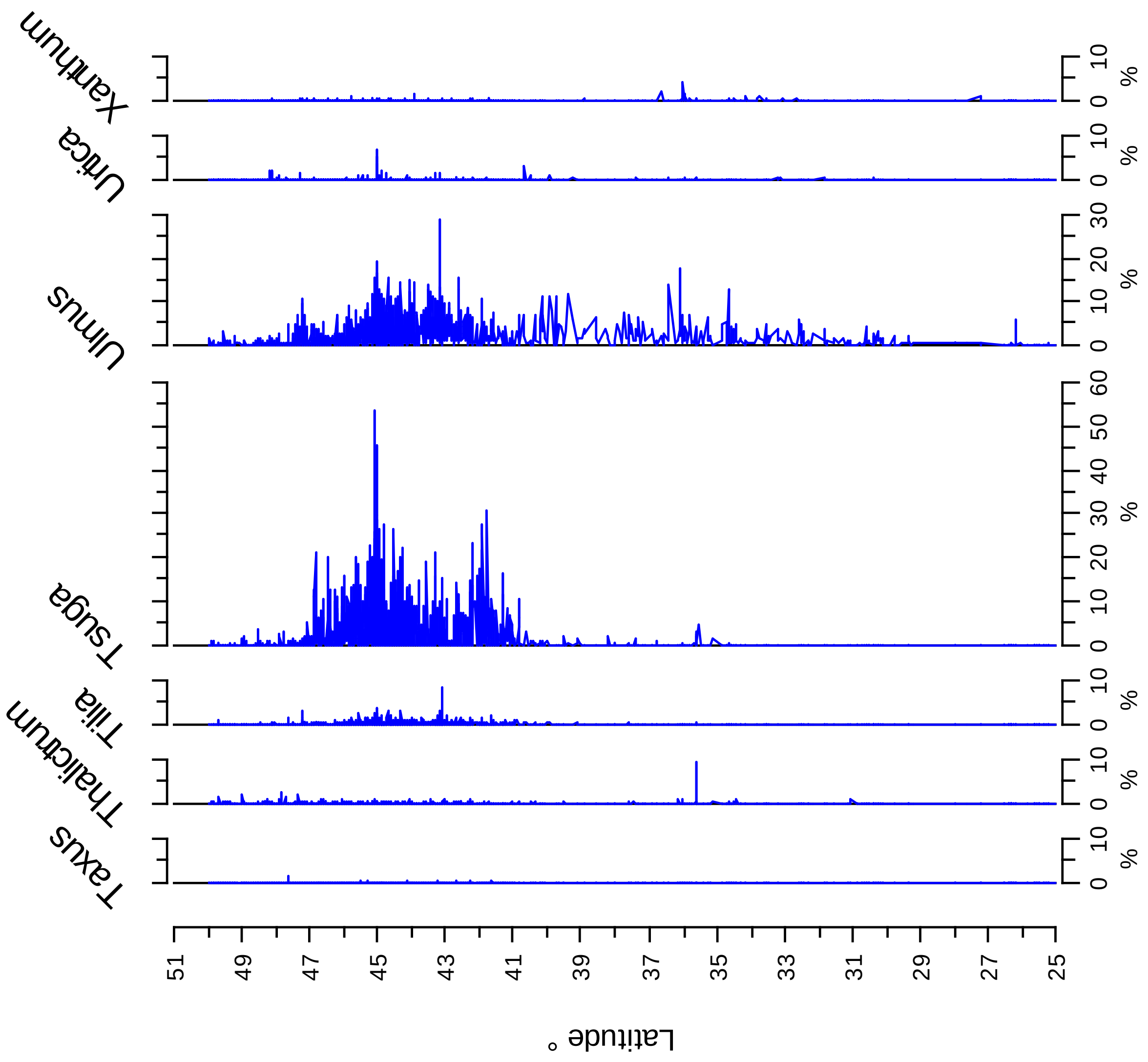

Celtis
Cephalanthus
Chenoam
Cornus
Corylus
Cupressa
Cyperaceae
Elaegnx
Ephedra
Equisetum
Ericaceae
Euphorbia
Fabaceae
Fagus
Fraxinus
%
Latitude °



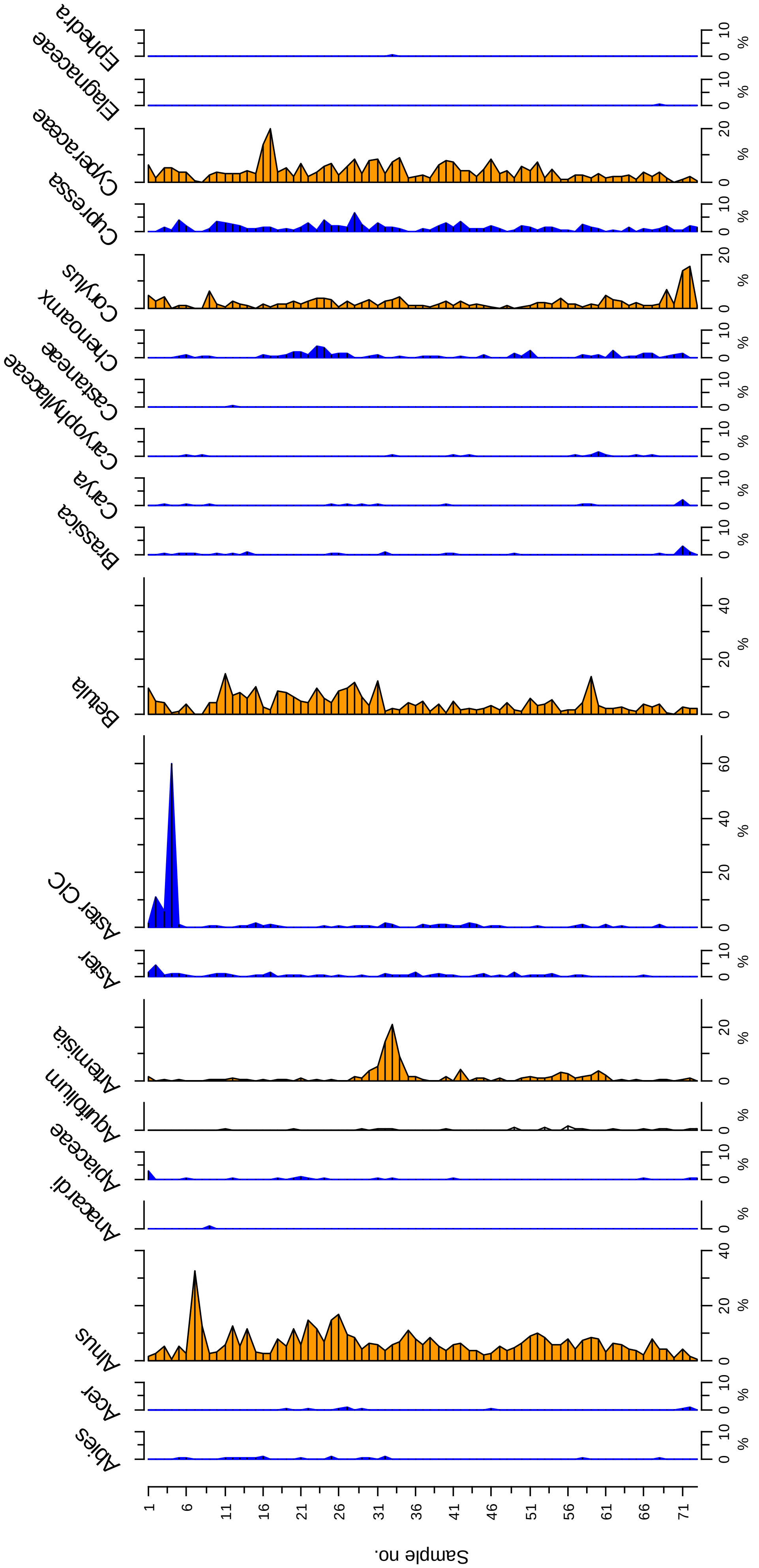
Ostycar
Osmunda
Onagraceae
Nyssa
Myrica
Moraceae
Malvaceae
Magnoliaceae
Lycopodium
Liquidambar
Liliaceae
Larix
Lamiaceae
Juglans
Latitude °
%

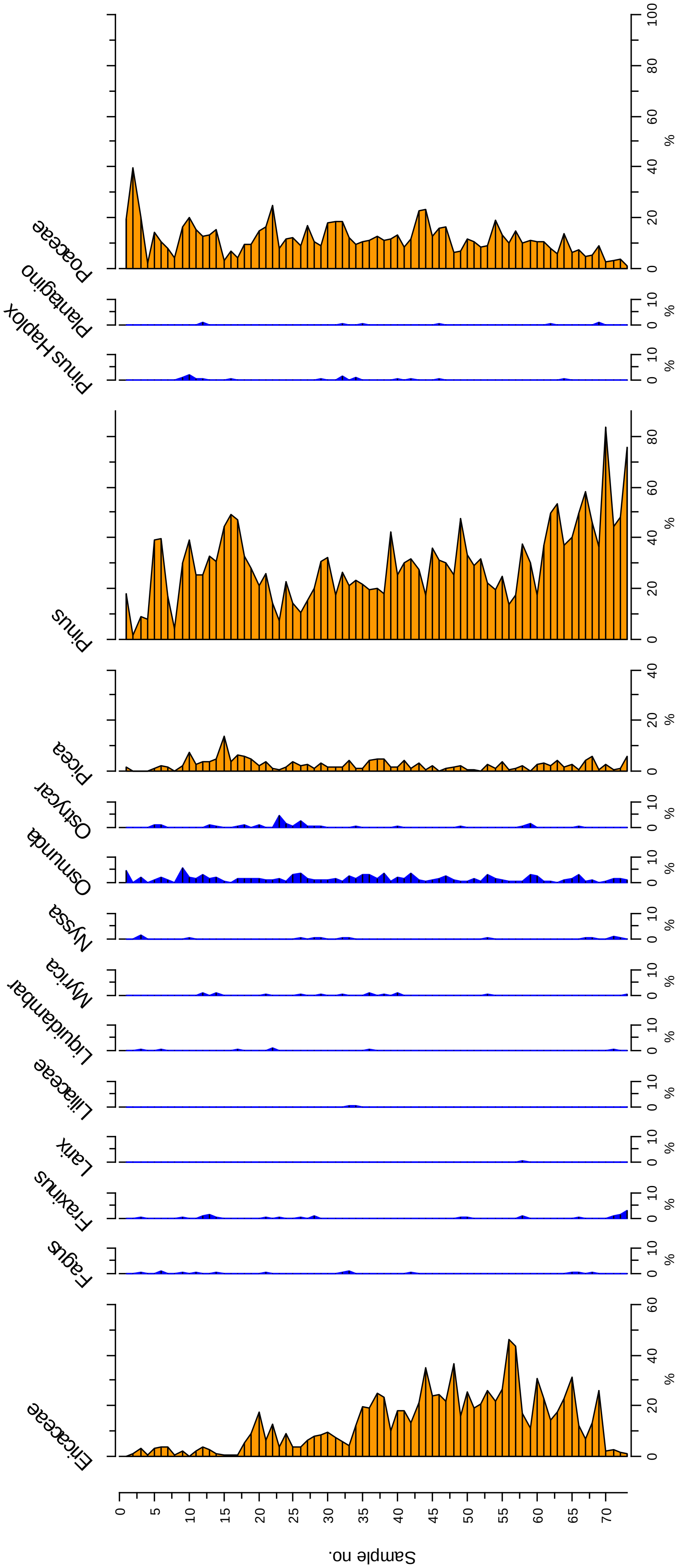

Picea
Pinus
Pinus haplox
Plantagino
Platanus
Poaceae
Polygona
Polypodium
Populus
%
Latitude °

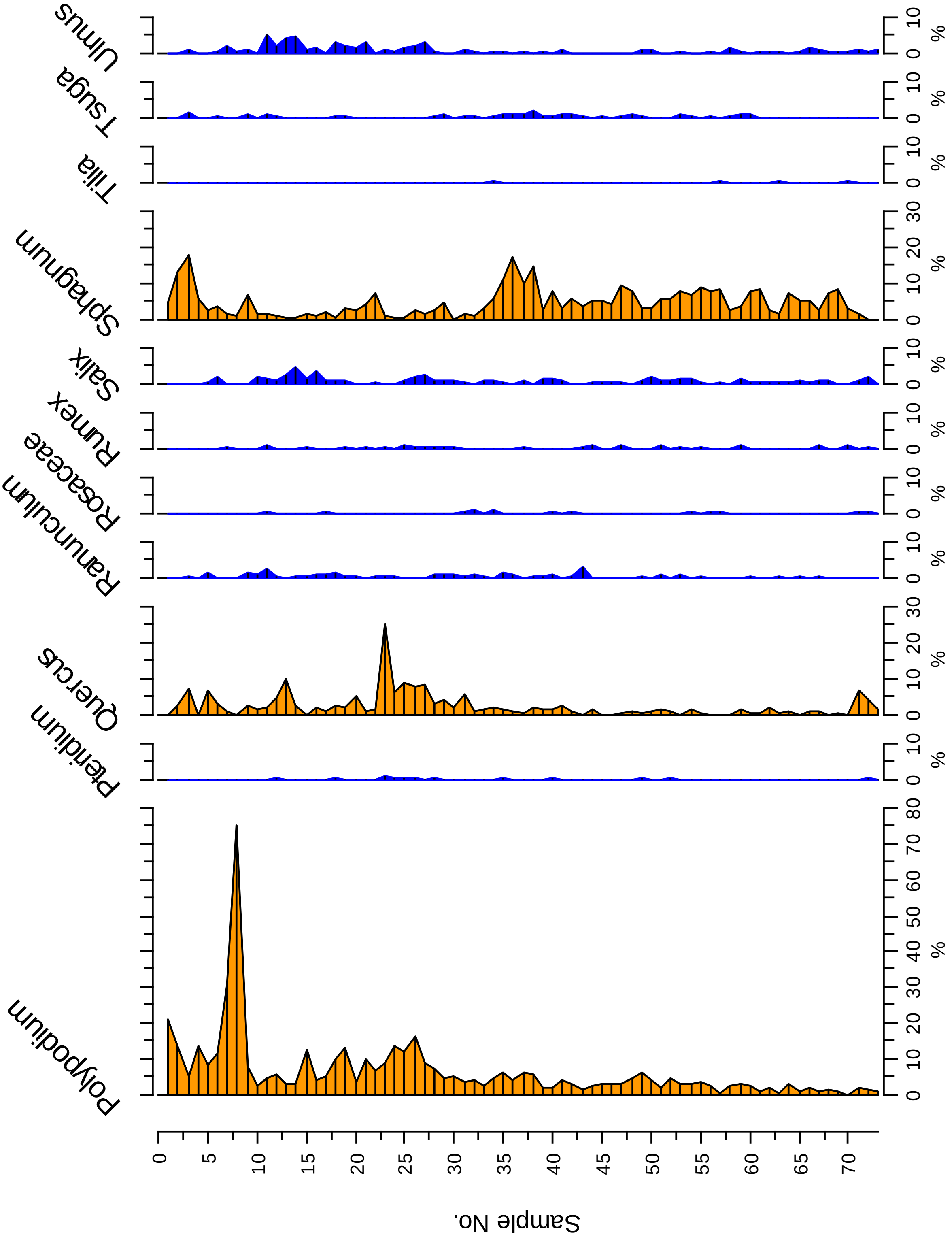

Pteridium
Quercus
Ranunculus
Rhamnaceae
Rosaceae
Rubiaceae
Rubus
Rumex
Salix
Saguisorba
Saxifraga
Selagine
Sphagnum
Taxodium
Latitude °


Xanthium
Urtica
Ulmus
Tsuga
Tilia
Thalictrum
Taxus
%
0
10
20
30
40
50
60
51
49
47
45
43
41
39
37
35
33
31
29
27
25
Latitude °

B. Pollen diagram fossil dataset Noordwijk

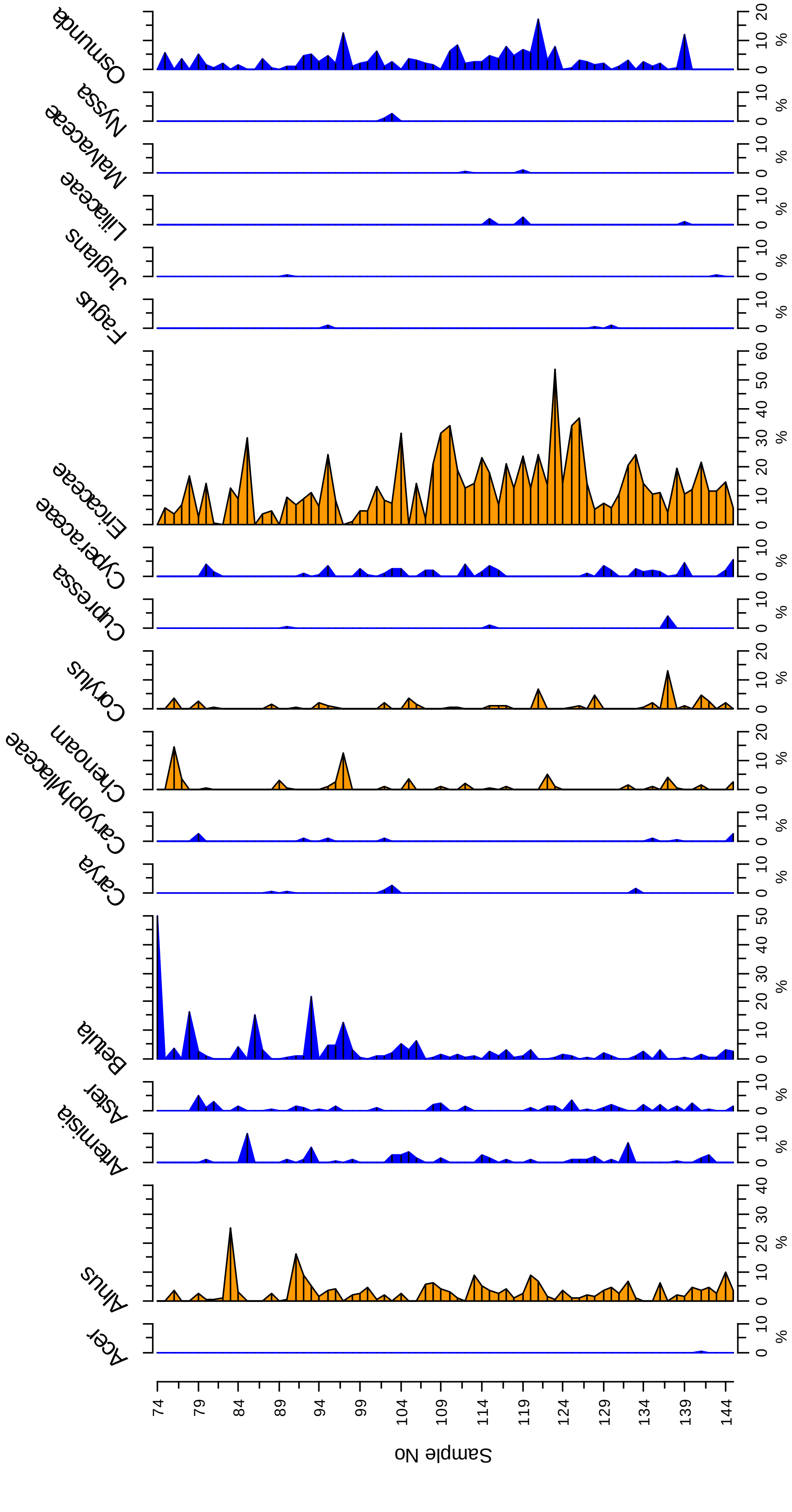


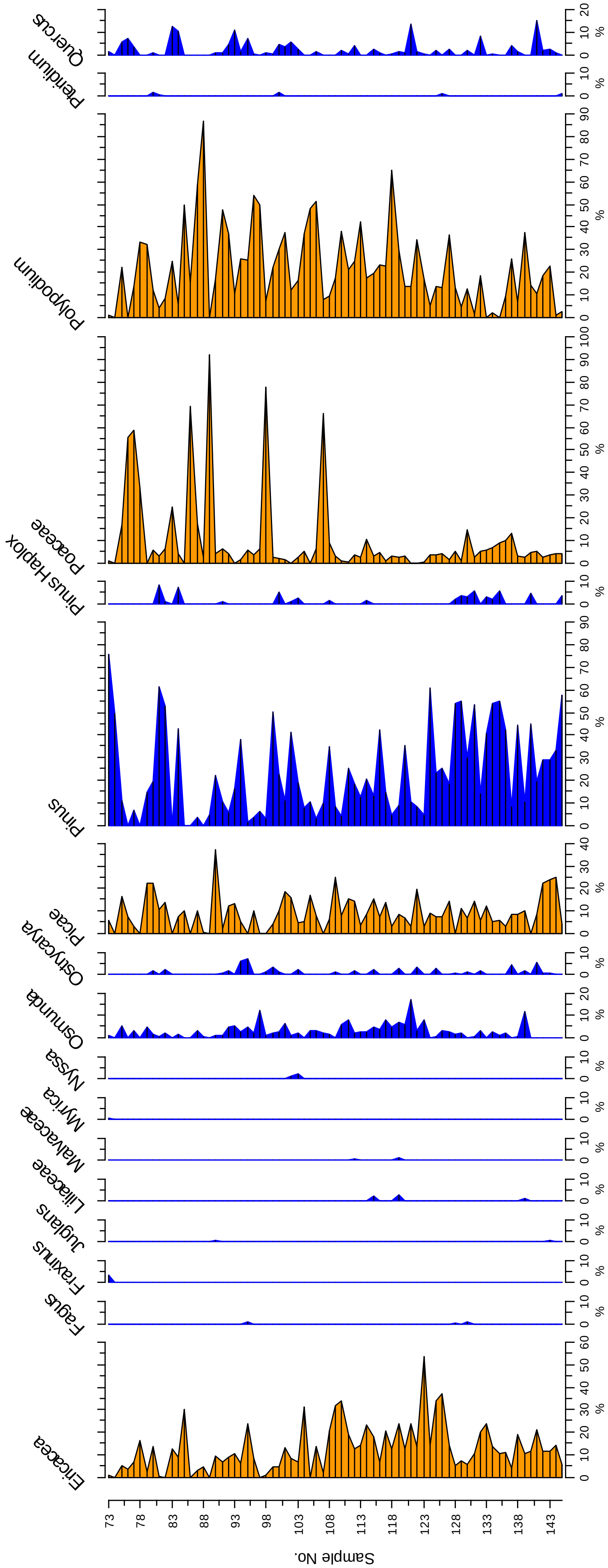

Poaceae
Plantagino
Pinus Haplox
Pinus
Picea
Ostrycar
Osmunda
Nyssa
Myrica
Liquidambar
Liliaceae
Larix
Fraxinus
Fagus
Ericaceae
Sample no.
%

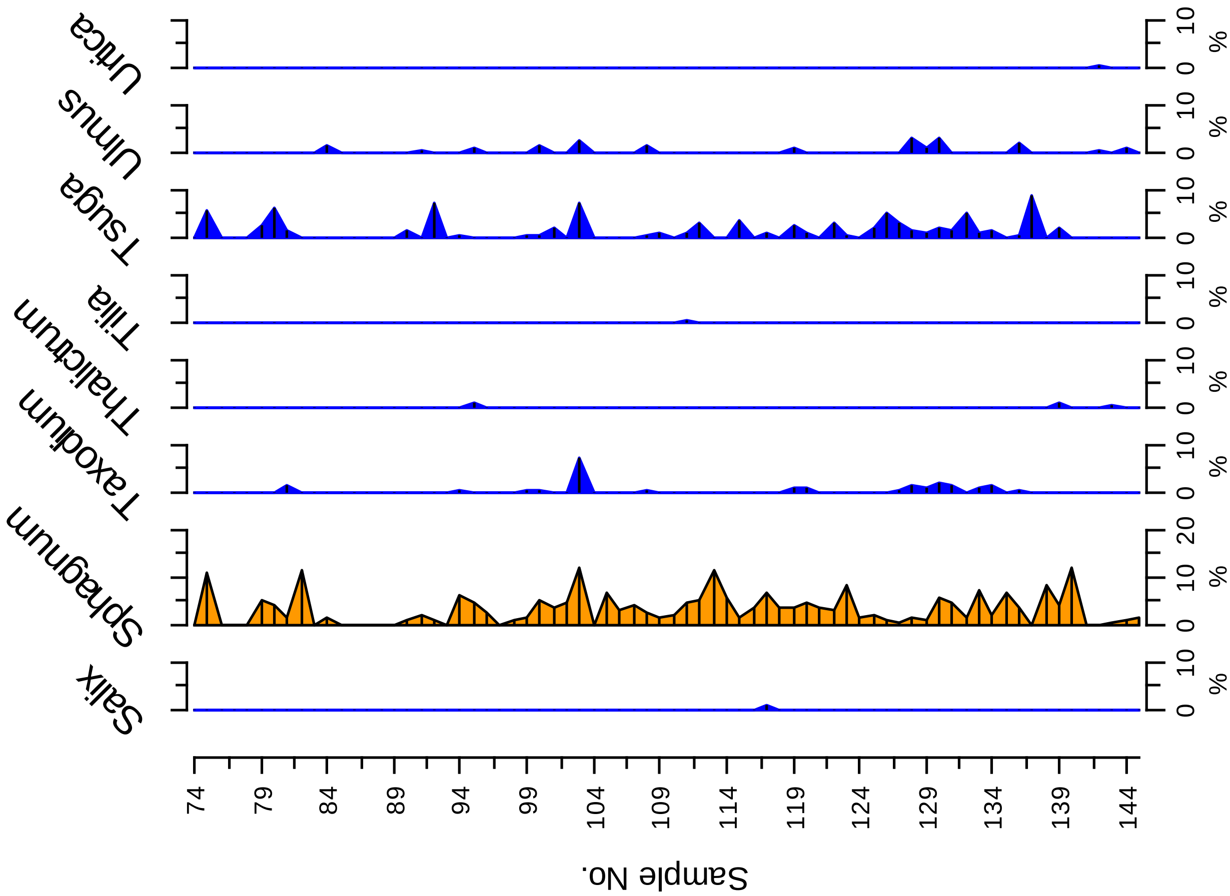


Ulmus
Tsuga
Tilia
Sphagnum
Salix
Rumex
Rosaceae
Ranunculum
Quercus
Pteridium
Polypodium
%
Sample No.

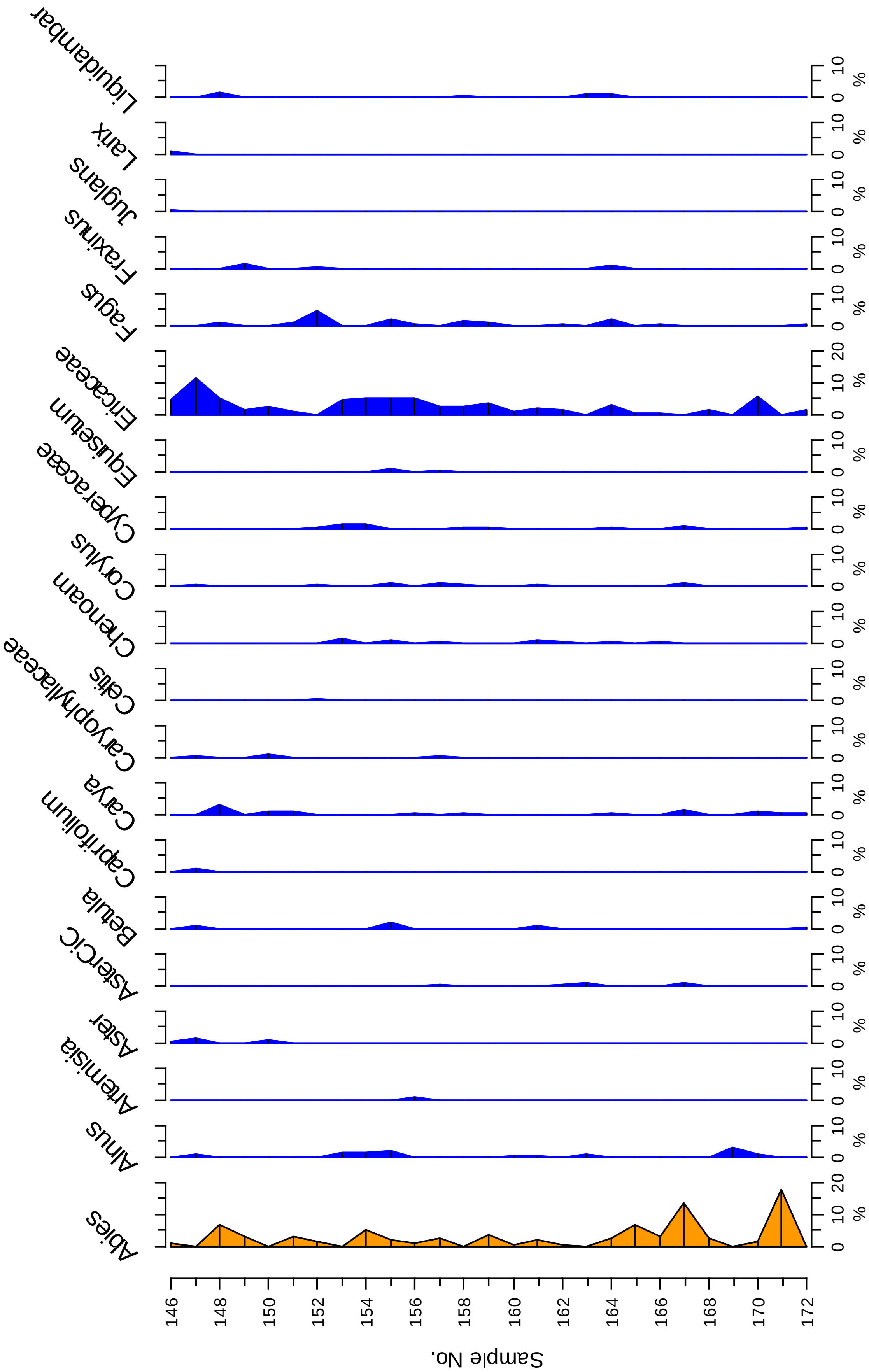
## C. Pollen diagram Petten

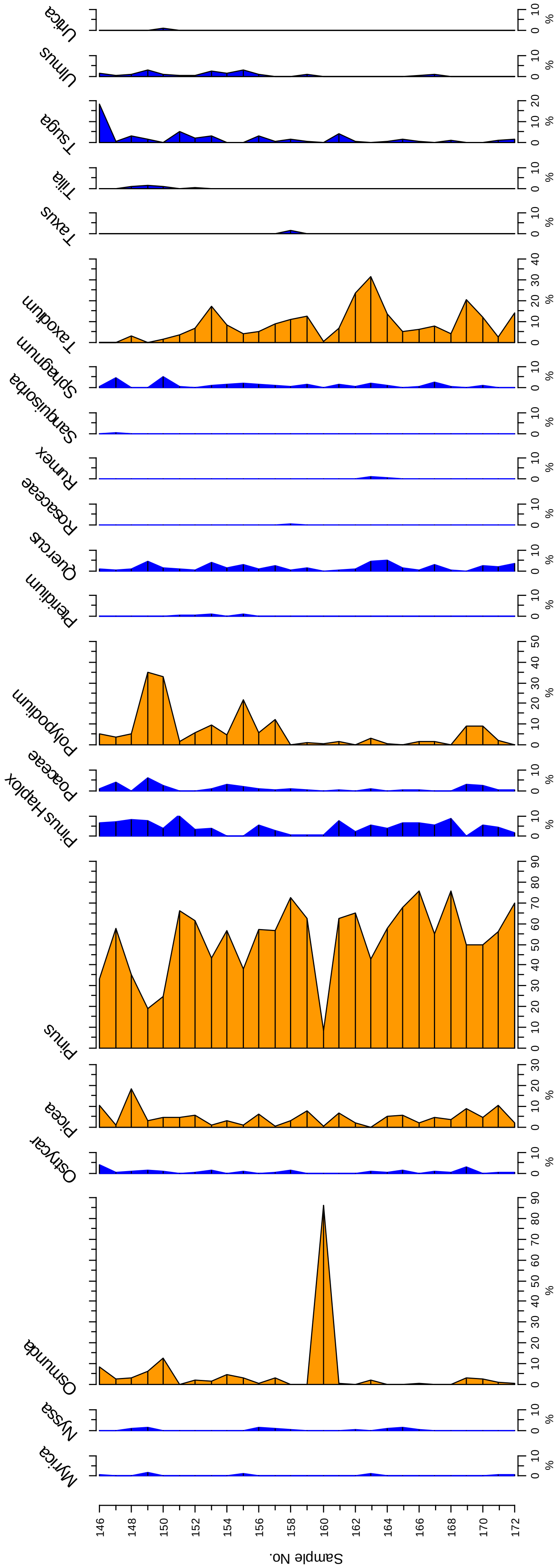



Quercus
Pteridium
Polypodium
Poaceae
Pinus Haplox
Pinus
Picae
Ostrycarya
Osmunda
Nyssa
Myrica
Malvaceae
Liliaceae
Juglans
Fraxinus
Fagus
Ericacea
Sample No.


Urtica
Ulmus
Tsuga
Tilia
Thalictrum
Taxodium
Sphagnum
Salix
0
10
20
%
74
79
84
89
94
99
104
109
114
119
124
129
134
139
144
Sample No.

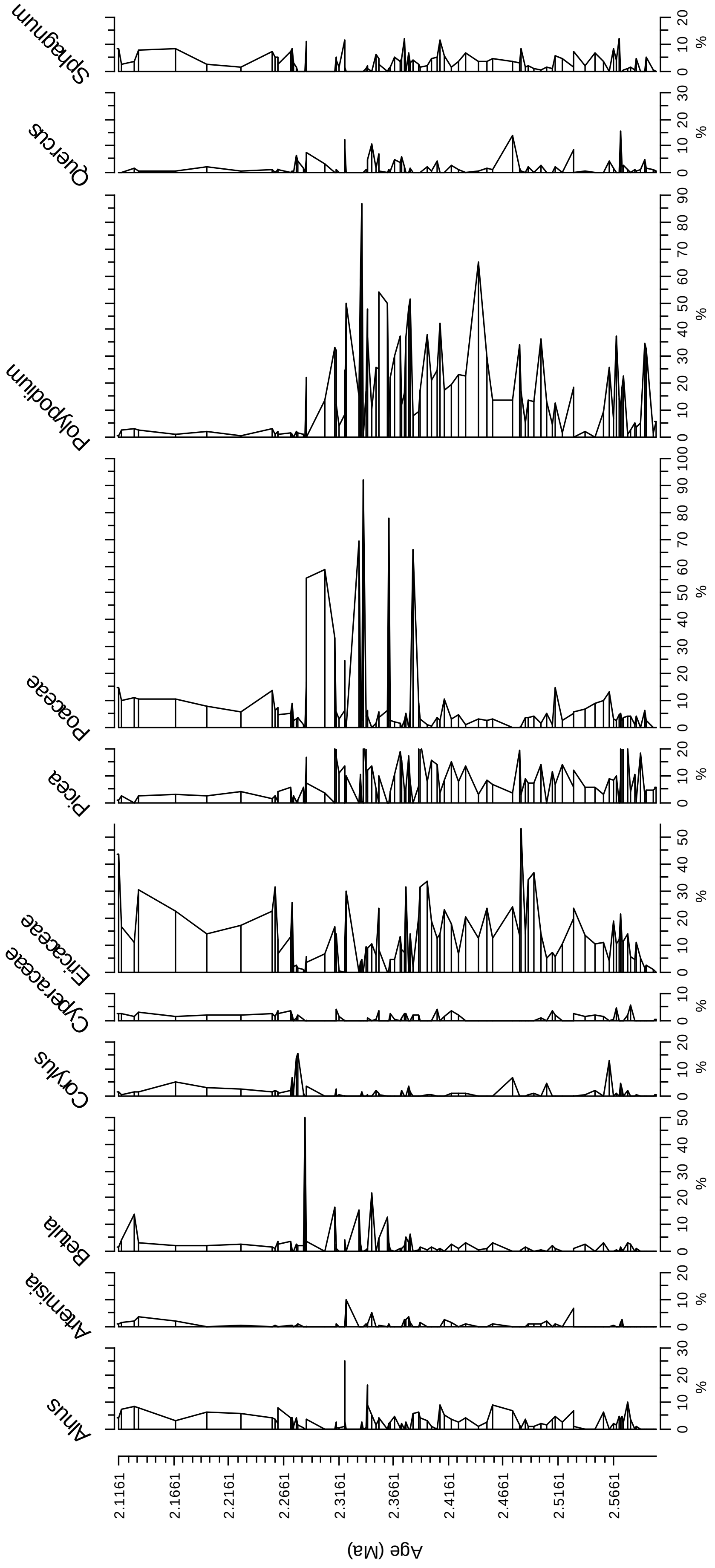
## D. Pollen diagram Hank



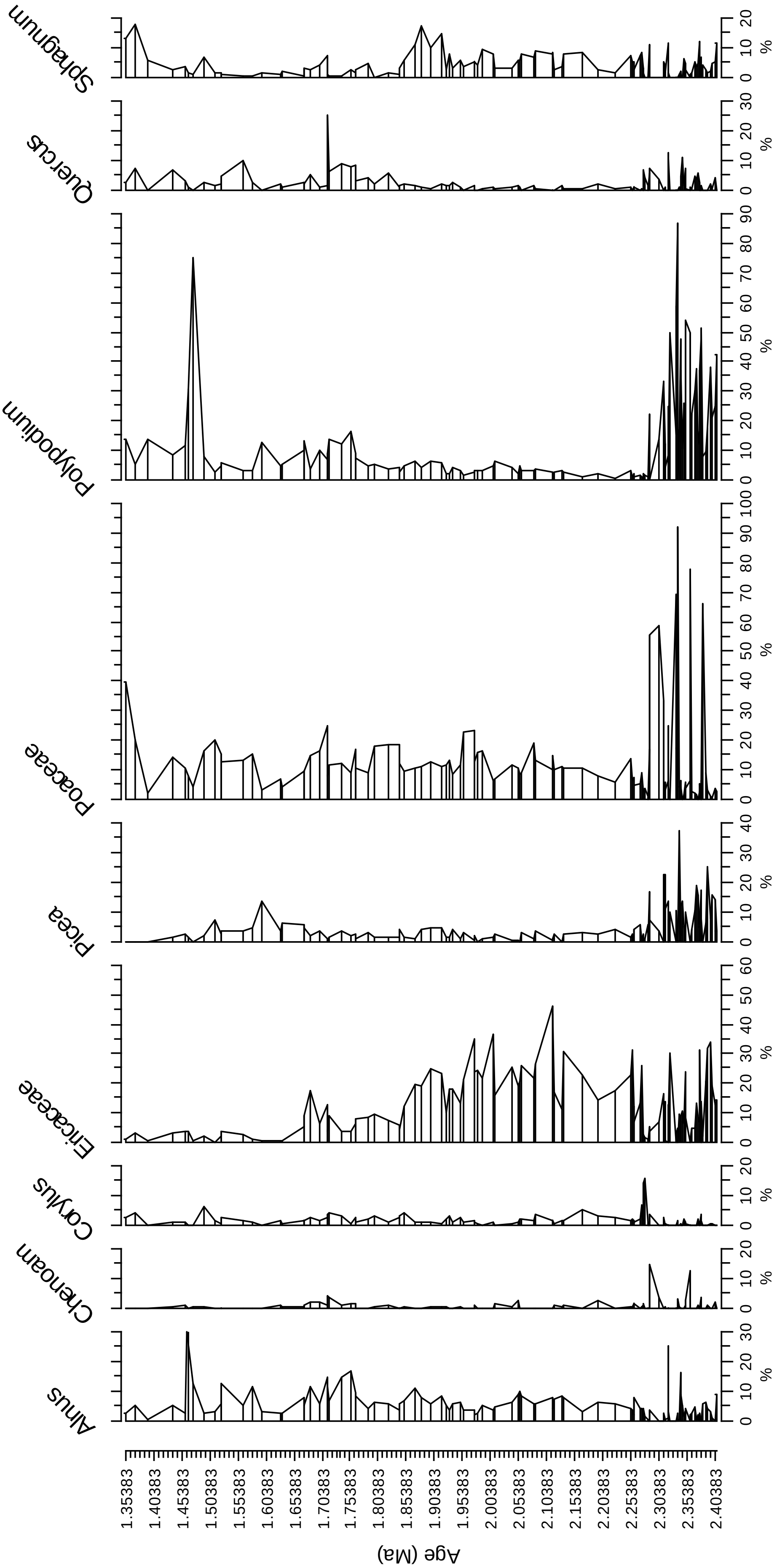

Urtica
Ulmus
Tsuga
Tilia
Taxus
Taxodium
Sphagnum
Sanguisorba
Rumex
Rosaceae
Quercus
Pteridium
Polypodium
Poaceae
Pinus Haplox
Pinus
Picea
Ostrycar
Osmunda
Nyssa
Myrica
Sample No.

# E. Dominant pollen species for fossil data sets with age

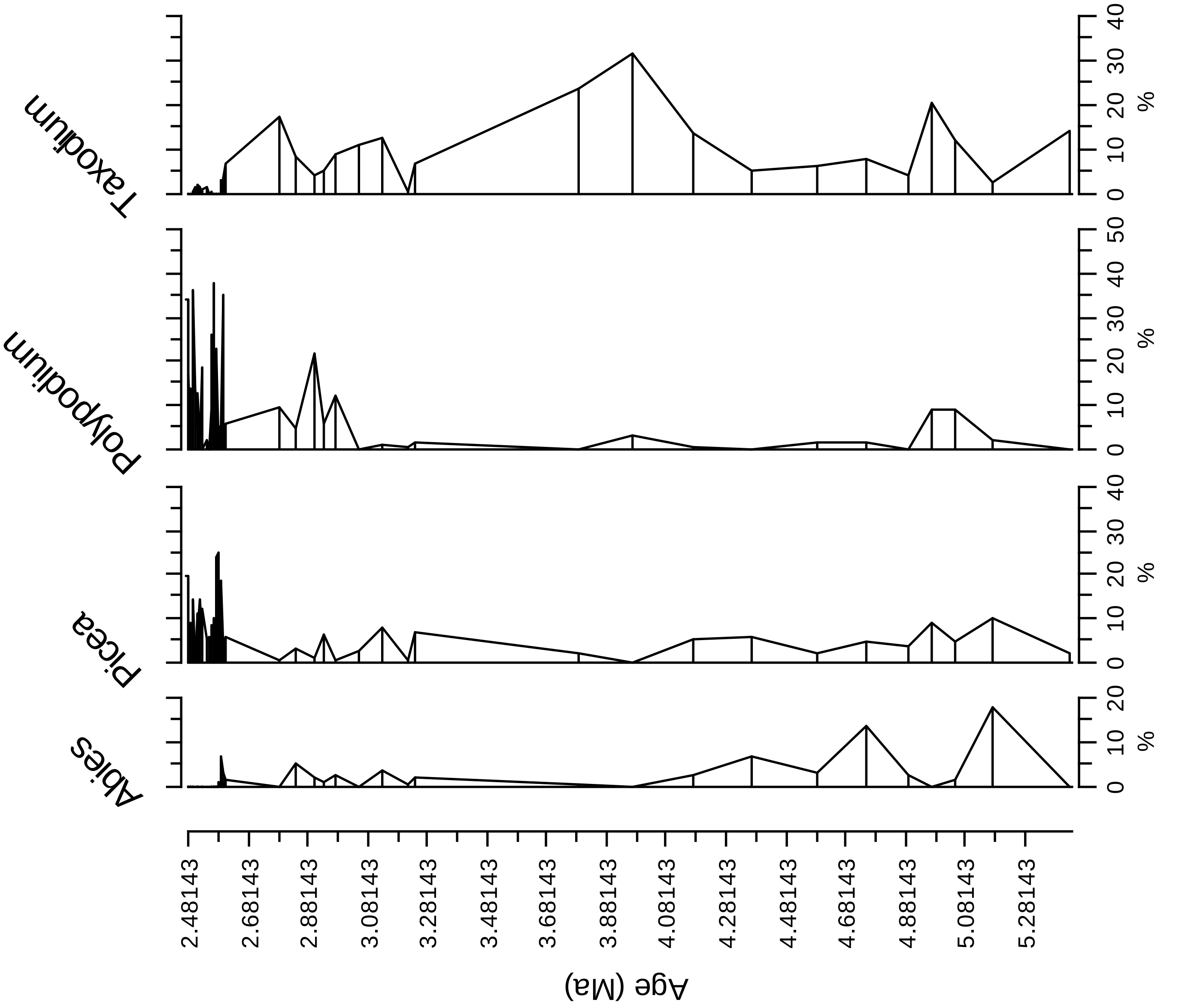
**Noordwijk**



**Petten**


Sphagnum
Quercus
Polypodium
Poaceae
Picea
Ericaceae
Corylus
Chenoam
Alnus
%
Age (Ma)
1.35383
1.40383
1.45383
1.50383
1.55383
1.60383
1.65383
1.70383
1.75383
1.80383
1.85383
1.90383
1.95383
2.00383
2.05383
2.10383
2.15383
2.20383
2.25383
2.30383
2.35383
2.40383

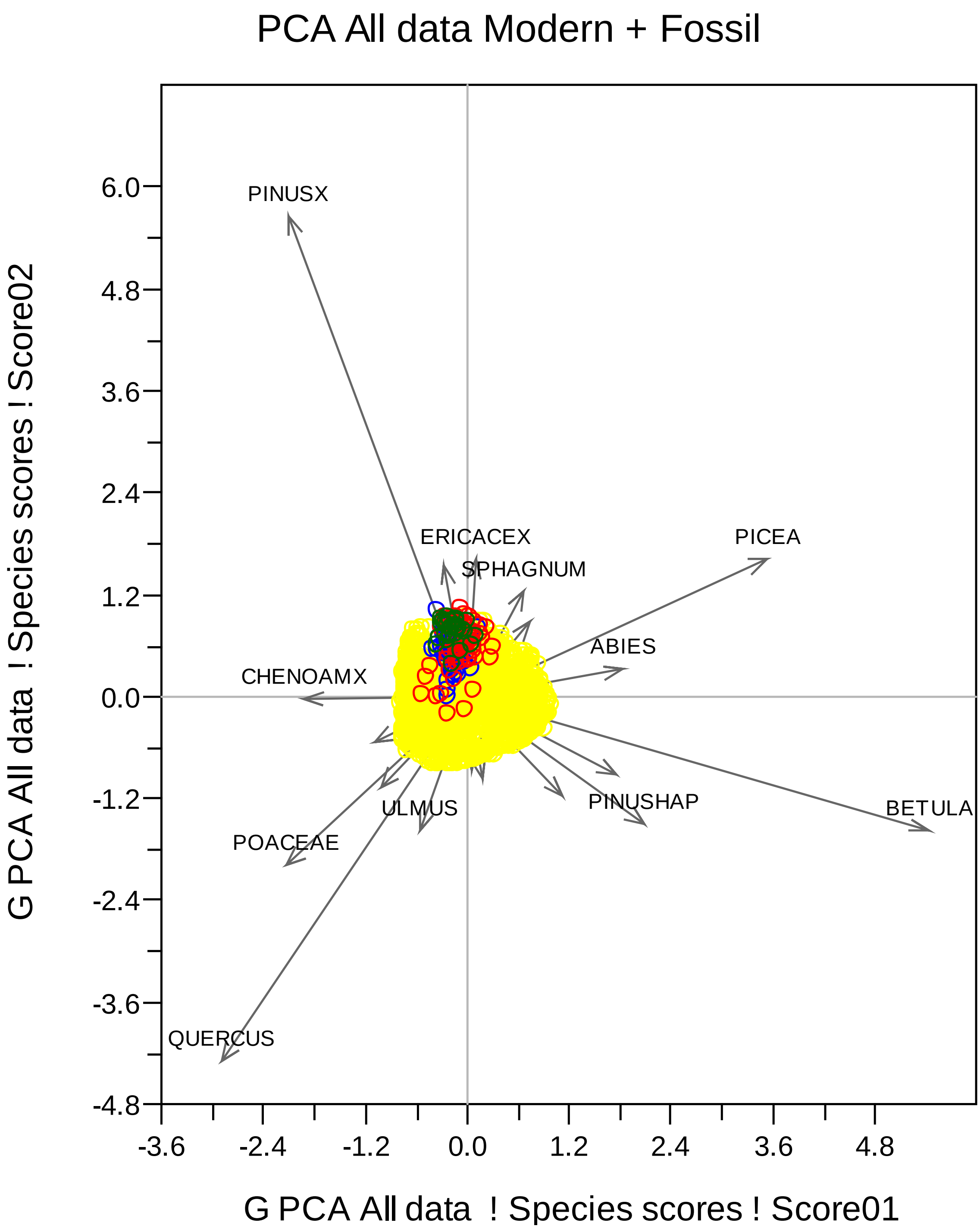
**Hank**



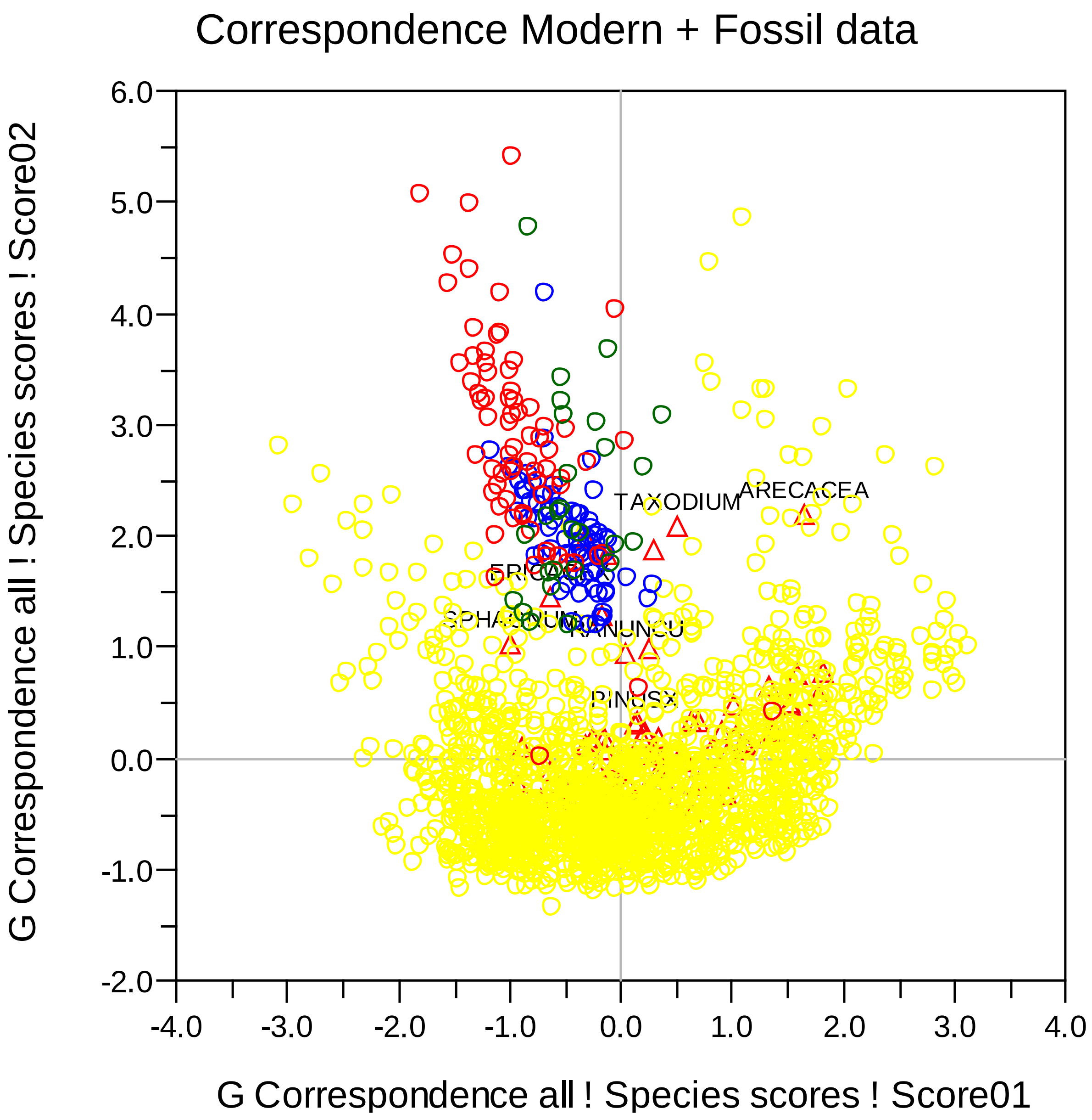
Taxodium
Polypodium
Picea
Abies
%
0 10 20 30 40
0 10 20 30 40 50
0 10 20
2.48143
2.68143
2.88143
3.08143
3.28143
3.48143
3.68143
3.88143
4.08143
4.28143
4.48143
4.68143
4.88143
5.08143
5.28143
Age (Ma)

## F. Principal Component Analyse Modern and Fossil data

Principal Component Analyse including all the data



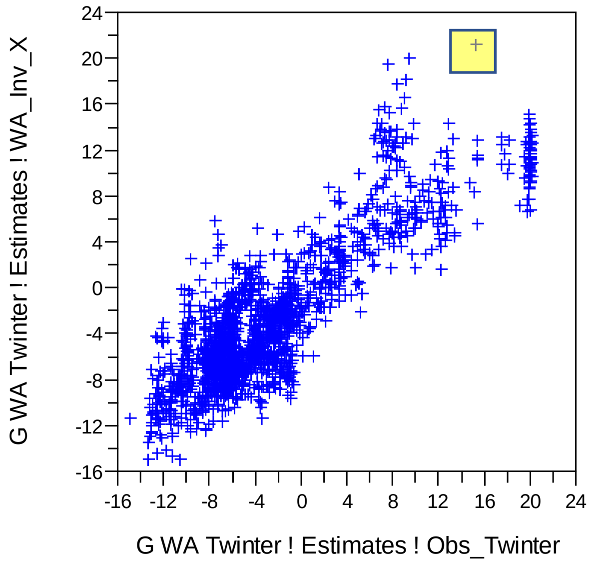
G. Correspondence graph with all the data

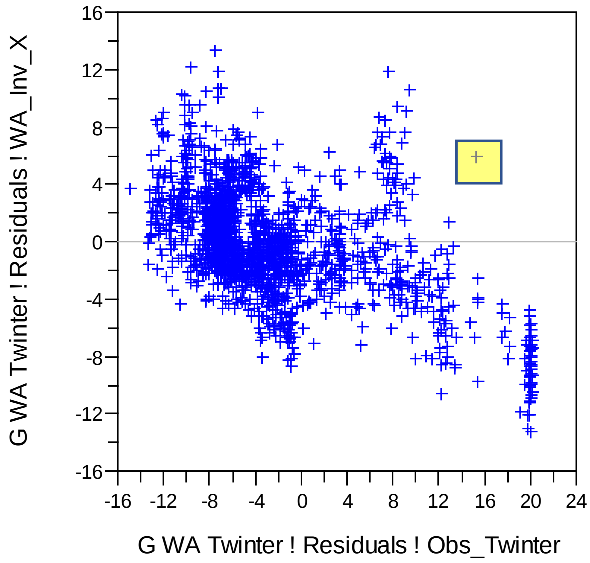


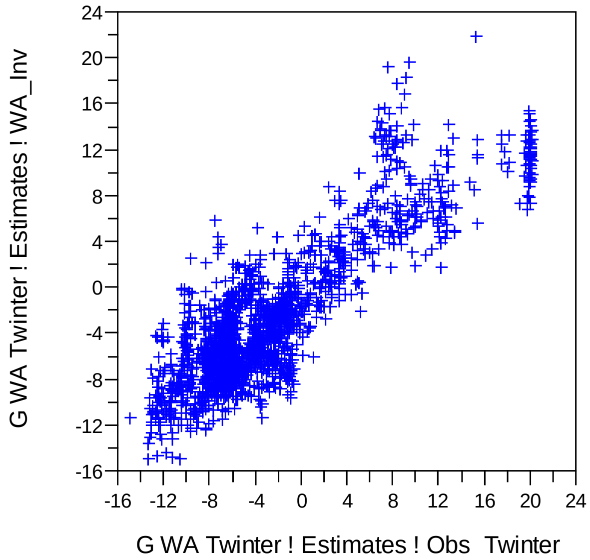
## H. Observed graphs WA and WA-PLS

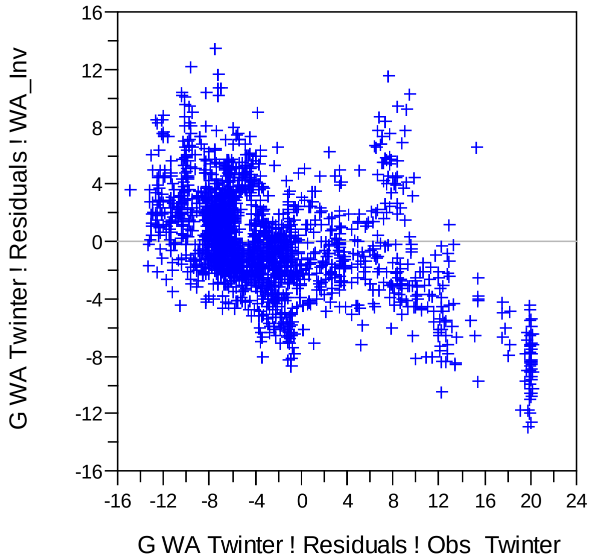
**WA Winter temperature Inverse deshrinking**

*The yellow/blue boxes are the outliers*



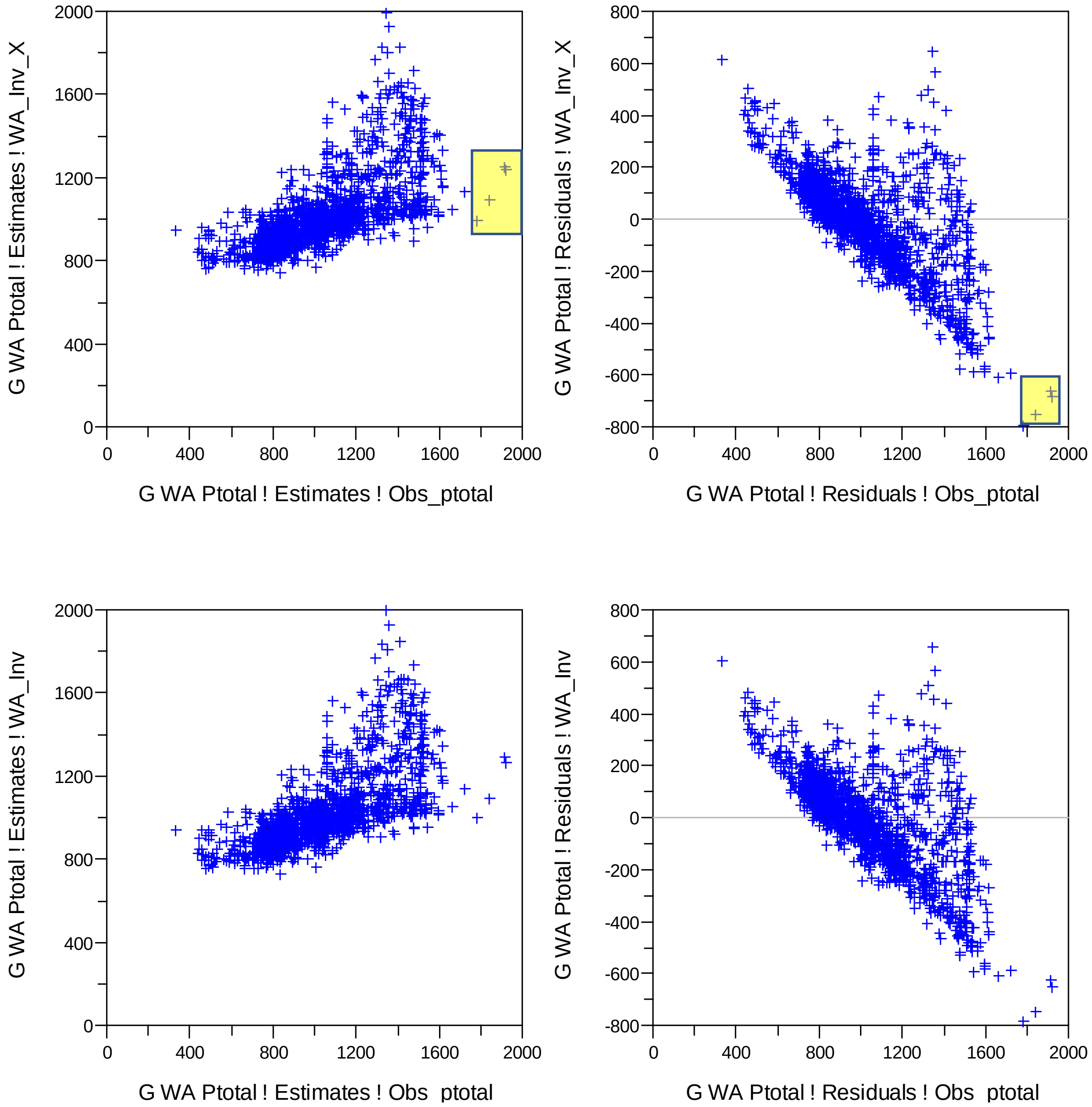



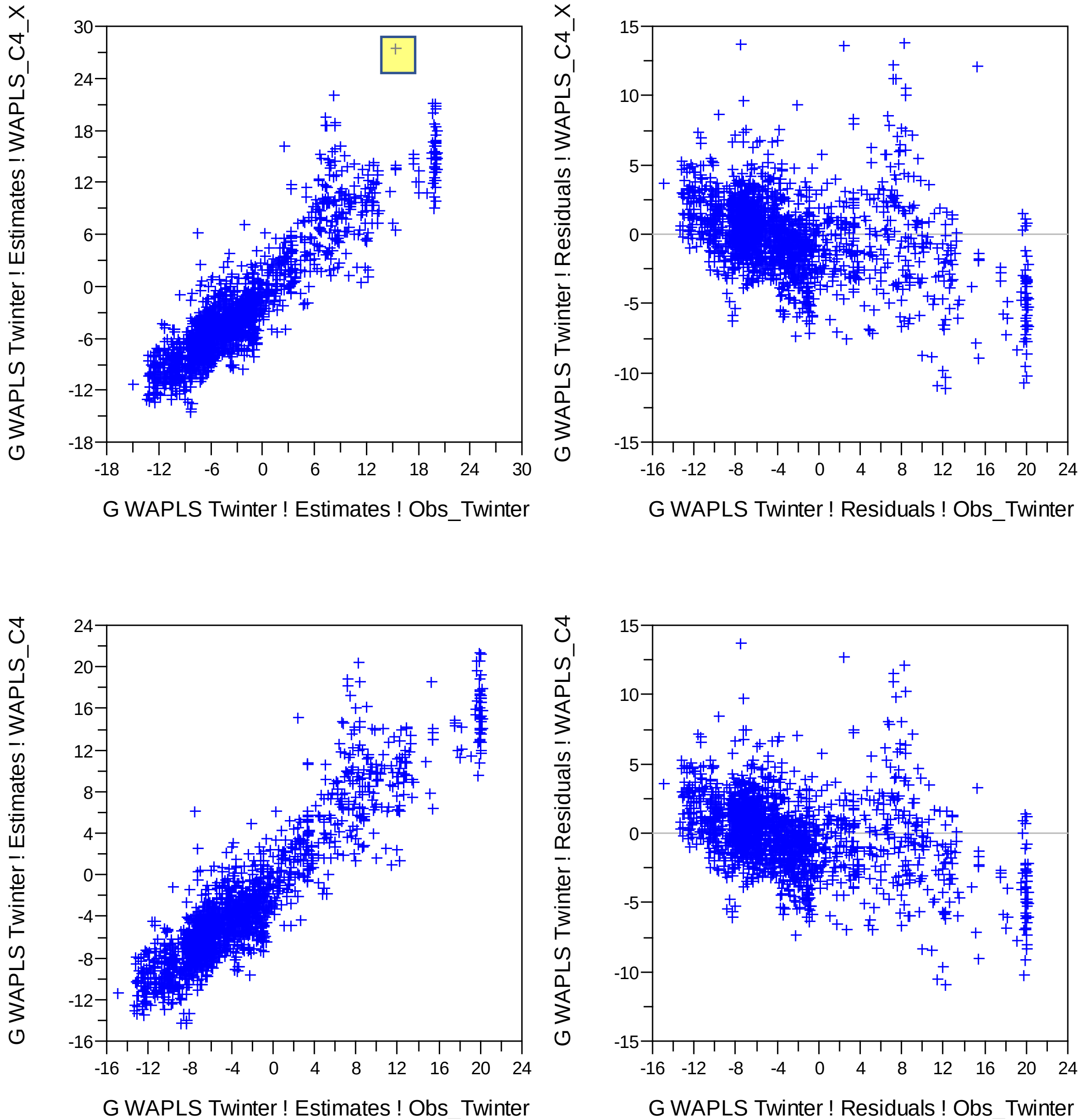
**WA Total precipitation inverse deshrinking**

*The yellow/blue boxes are the outliers*



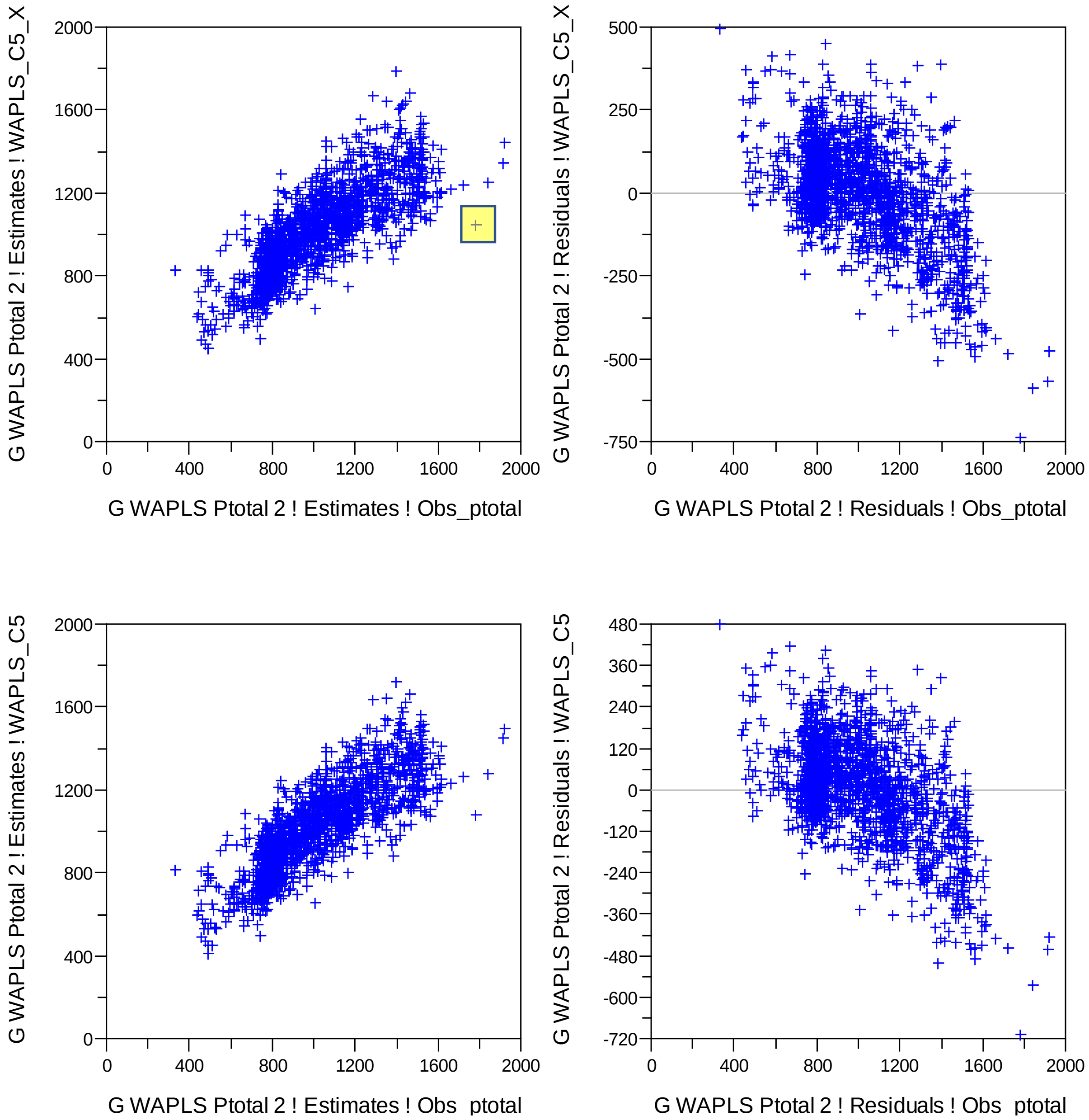
**WAPLS Winter temperature C4**

*The yellow/blue boxes are the outliers*



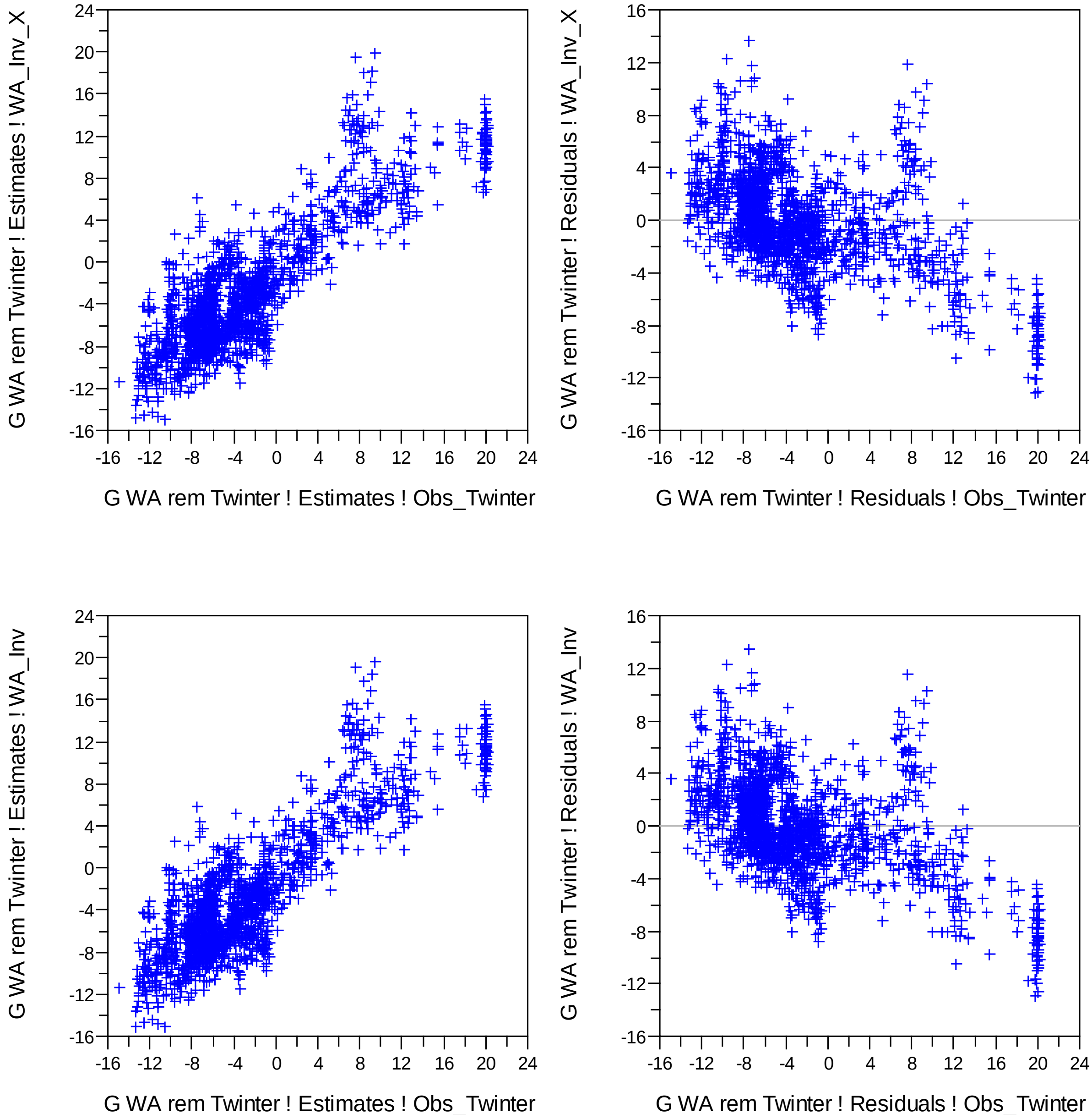
**WAPLS Total precipitation C5**

*The yellow/blue boxes are outliers*

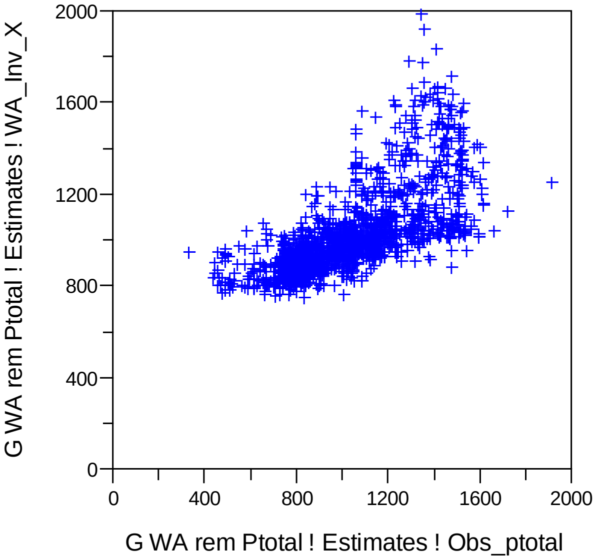


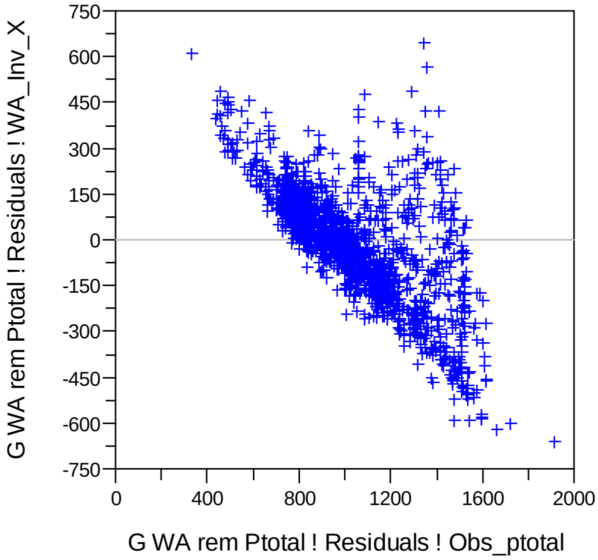
I. Observed graphs WA and WAPLS with removed samples

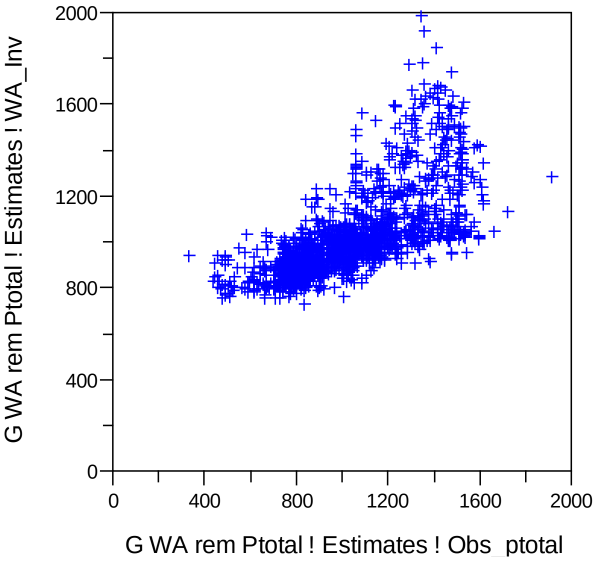
**WA Winter temperature inverse deshrinking**

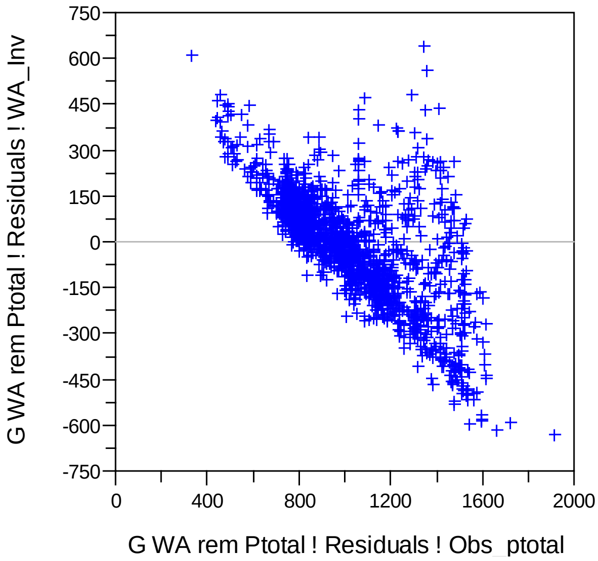


## WA Total precipitation inverse deshrinking

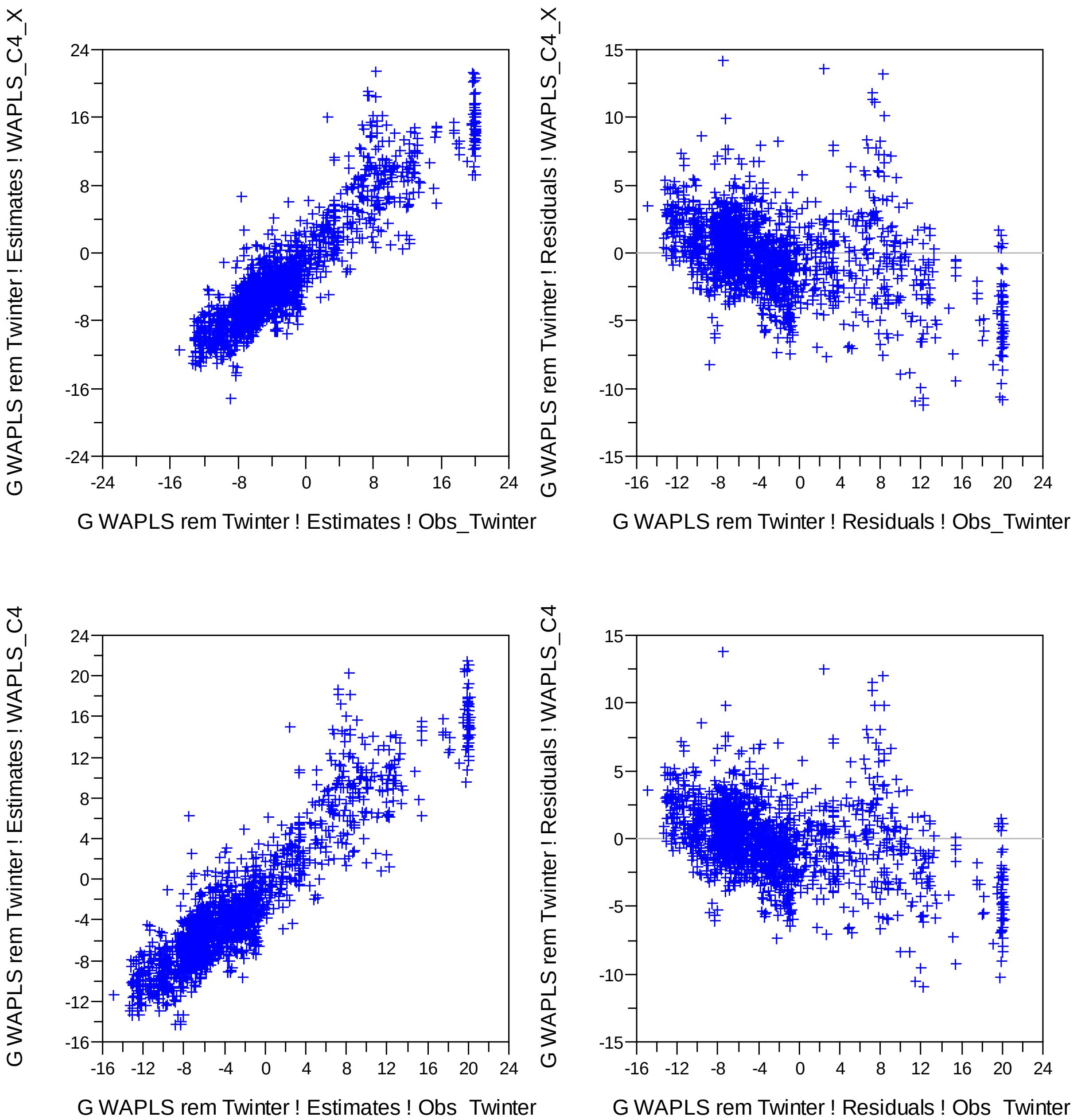


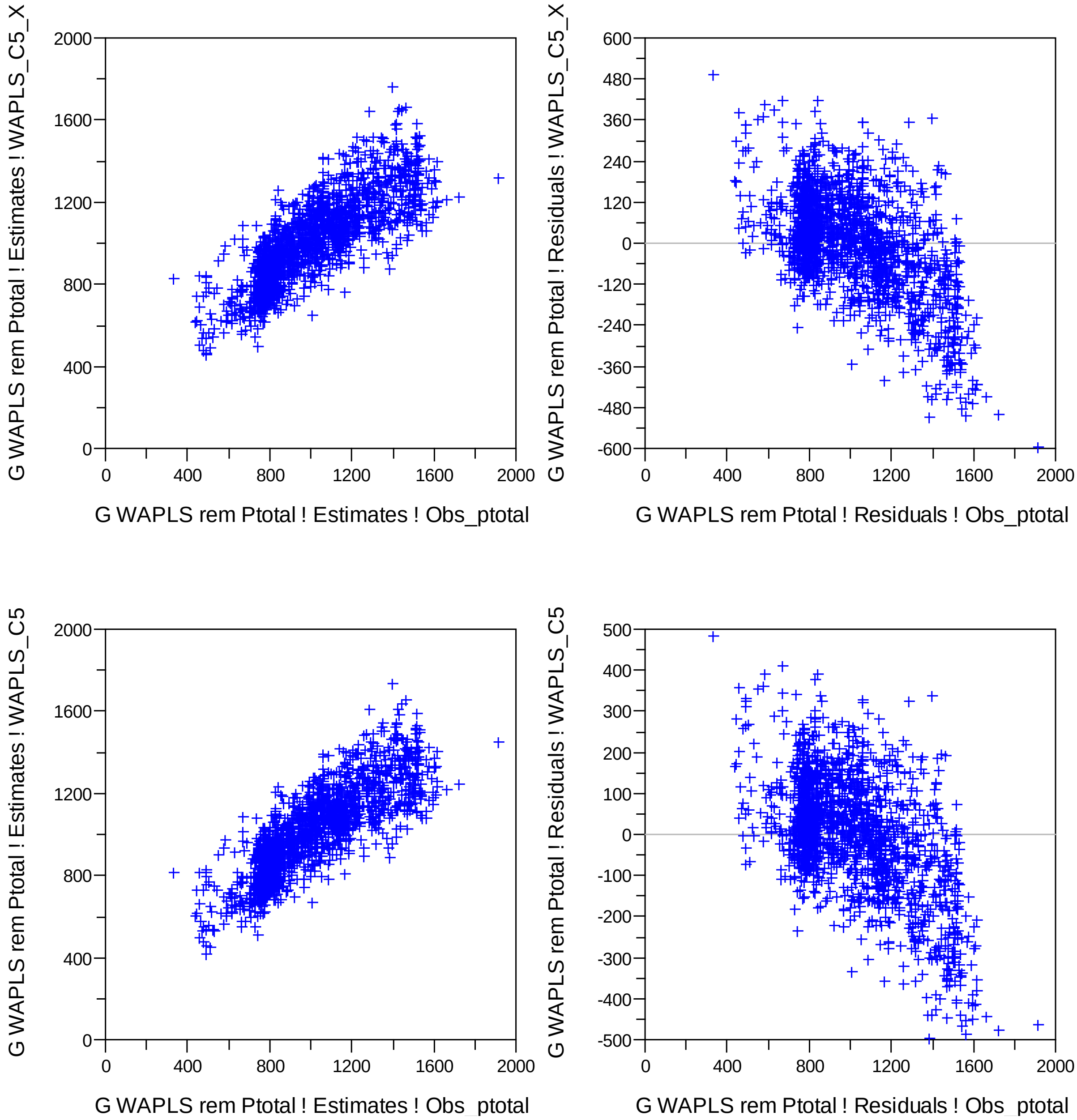



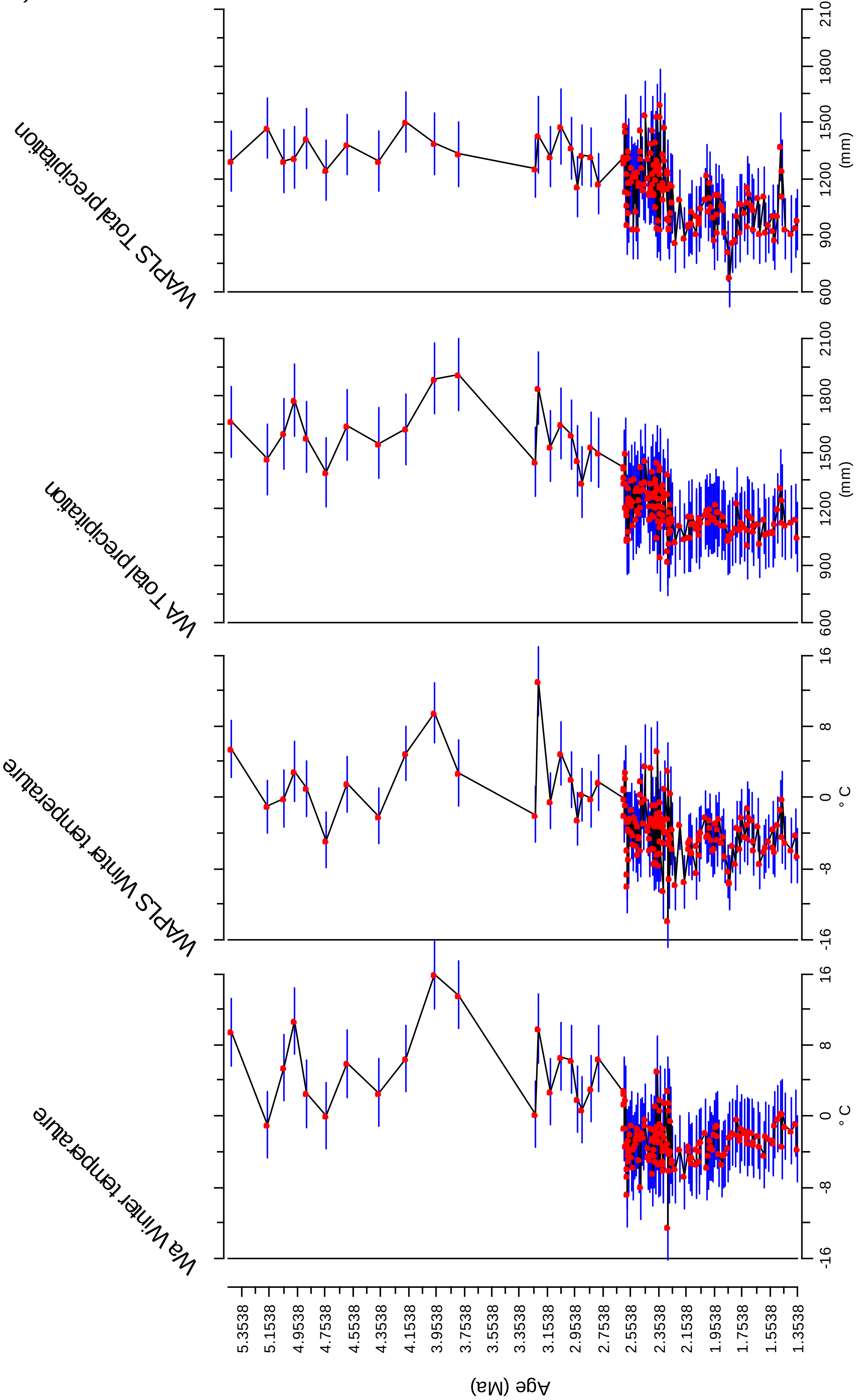
**WAPLS Winter temperature C4**



## WAPLS Total precipitation C5



## J. First climate reconstruction



K. Calibration Sample number – age

| Sample nr. | Age (Ma) | | Age (Ma) | | Age (Ma) | | |
|---|---|---|---|---|---|---|---|
| 1 | 2.116098 | 44 | 2.424891 | 87 | 1.57887 | 130 | 2.13393 |
| 2 | 2.119061 | 45 | 2.431959 | 88 | 1.59743 | 131 | 2.16844 |
| 3 | 2.130577 | 46 | 2.443502 | 89 | 1.62904 | 132 | 2.19599 |
| 4 | 2.255519 | 47 | 2.450815 | 90 | 1.6331 | 133 | 2.2276 |
| 5 | 2.258005 | 48 | 2.456692 | 91 | 1.67109 | 134 | 2.26124 |
| 6 | 2.260425 | 49 | 2.474284 | 92 | 1.67138 | 135 | 2.28676 |
| 7 | 2.272114 | 50 | 2.482949 | 93 | 1.68414 | 136 | 2.30387 |
| 8 | 2.273652 | 51 | 2.485803 | 94 | 1.70067 | 137 | 2.3146 |
| 9 | 2.275655 | 52 | 2.488636 | 95 | 1.71326 | 138 | 2.32127 |
| 10 | 2.277604 | 53 | 2.49423 | 96 | 1.714505 | 139 | 2.34128 |
| 11 | 2.279502 | 54 | 2.500189 | 97 | 1.71575 | 140 | 2.35317 |
| 12 | 2.284025 | 55 | 2.505537 | 98 | 1.7382 | 141 | 2.36187 |
| 13 | 2.285751 | 56 | 2.51074 | 99 | 1.75635 | 142 | 2.37335 |
| 14 | 2.28743 | 57 | 2.513913 | 100 | 1.76418 | 143 | 2.37666 |
| 15 | 2.312246 | 58 | 2.519455 | 101 | 1.76447 | 144 | 2.38072 |
| 16 | 2.31425 | 59 | 2.530914 | 102 | 1.78622 | 145 | 2.40798 |
| 17 | 2.316169 | 60 | 2.541043 | 103 | 1.79753 | 146 | 2.48143 |
| 18 | 2.321493 | 61 | 2.549707 | 104 | 1.82351 | 147 | 2.53071 |
| 19 | 2.322326 | 62 | 2.556837 | 105 | 1.84074 | 148 | 2.58182 |
| 20 | 2.323146 | 63 | 2.562435 | 106 | 1.84219 | 149 | 2.58788 |
| 21 | 2.334567 | 64 | 2.56659 | 107 | 1.85031 | 150 | 2.59091 |
| 22 | 2.335763 | 65 | 2.569487 | 108 | 1.86887 | 151 | 2.59576 |
| 23 | 2.337516 | 66 | 2.571419 | 109 | 1.88192 | 152 | 2.60242 |
| 24 | 2.338917 | 67 | 2.572799 | 110 | 1.89729 | 153 | 2.78867 |
| 25 | 2.342445 | 68 | 2.574169 | 111 | 1.91846 | 154 | 2.84333 |
| 26 | 2.343159 | 69 | 2.576215 | 112 | 1.92455 | 155 | 2.90471 |
| 27 | 2.346052 | 70 | 2.579773 | 113 | 1.9318 | 156 | 2.93157 |
| 28 | 2.349784 | 71 | 2.585846 | 114 | 1.93615 | 157 | 2.975 |
| 29 | 2.352887 | 72 | 2.595611 | 115 | 1.95152 | 158 | 3.05091 |
| 30 | 2.360342 | 73 | 2.605367 | 116 | 1.95674 | 159 | 3.13 |
| 31 | 2.363895 | 74 | 2.615122 | 117 | 1.97501 | 160 | 3.21691 |
| 32 | 2.367609 | 75 | 1.35383 | 118 | 1.97704 | 161 | 3.23709 |
| 33 | 2.372486 | 76 | 1.36978 | 119 | 1.98139 | 162 | 3.79 |
| 34 | 2.377632 | 77 | 1.39356 | 120 | 1.99038 | 163 | 3.97 |
| 35 | 2.380848 | 78 | 1.43851 | 121 | 2.00865 | 164 | 4.17 |
| 36 | 2.384162 | 79 | 1.4591 | 122 | 2.01329 | 165 | 4.37 |
| 37 | 2.388728 | 80 | 1.46606 | 123 | 2.04345 | 166 | 4.59 |
| 38 | 2.389895 | 81 | 1.47331 | 124 | 2.05331 | 167 | 4.75 |
| 39 | 2.397113 | 82 | 1.4939 | 125 | 2.05534 | 168 | 4.89 |
| 40 | 2.400854 | 83 | 1.51159 | 126 | 2.06056 | **169** | 4.97 |
| 41 | 2.40597 | 84 | 1.5229 | 127 | 2.08115 | 170 | 5.05 |
| 42 | 2.412559 | 85 | 1.52464 | 128 | 2.08521 | 171 | 5.17 |
| 43 | 2.419342 | 86 | 1.56205 | 129 | 2.11392 | 172 | 5.43 |

L. Fossil pollen species combined with modern pollen species

| **Hank** | **Noordwijk** | **Petten** | **Whitemore** |
|---|---|---|---|
| *Abies* | *Abies* | *Abies* | *Abies* |
| - | *Acer* | *Acer* | *Acer +* |
| *Alisma* | - | - | Not used |
| *Alnus* | *Alnus* | *Alnus* | *Alnus +* |
| *Armeria* | - | - | *Armeria* |
| *Artemisia* | *Artemisia* | *Artemisia* | *Artemisia* |
| - | - | *Asteraceae* | *Aster* |
| - | *Azolla* | - | Not used |
| *Betula* | *Betula* | *Betula* | *Betula* |
| - | *Brassicaceae* | - | *Brassicaceae* |
| *Calluna* | - | - | *Ericaceae* |
| *Carpinus* | *Carpinus* | *Carpinus* | *Ostrycar* |
| *Carya* | *Carya* | *Carya* | *Carya* |
| *Caryophyllaceae* | *Caryophyllaceae* | *Caryophyllaceae* | *Caryophyllaceae* |
| - | *Castanea* | - | *Castanea* |
| *Cedrus* | *Cedrus* | - | *Larix* |
| *Celtis* | - | - | *Celtis* |
| | *Cerebropollenites* | - | Not used |
| *Chenopodiaceae* | *Chenopodiaceae* | *Chenopodiaceae* | *Chenopodiaceae* |
| - | *Cicatricosisporites* | - | Not used |
| - | *Classopollis spp* | - | Not used |
| *C. Tubulifloraea* | *C. Tubulifloraea* | - | Not used |
| *C. Liguliforaea* | *C. Liguliflorea* | - | Not used |
| *Corylus* | *Corylus* | *Corylus* | *Corylus* |
| - | *Cupressaceae* | - | *Cupressaceae* |
| *Cyperaceae* | *Cyperaceae* | *Cyperaceae* | *Cyperaceae* |
| - | *Dipsacaceae* | - | Not used |
| *Dryopteris* | *Dryopteris* | *Dryopteris* | Not used |
| - | *Elaeagnus* | - | *Elaegnaceae* |
| - | *Engelhardia* | - | Not used |
| - | *Ephedra* | - | *Ephedra* |
| *Equisetum* | - | - | *Equisetum* |
| *Ericales* | *Ericales* | *Ericales* | *Ericales* |
| *Eucommia* | *Eucommia* | - | Not used |
| *Euphorbia* | - | - | *Euphorbiaceae* |
| *Fagus* | *Fagus* | *Fagus* | *Fagus* |
| *Filipendula* | - | - | Not used |
| *Fraxinus* | *Fraxinus* | *Fraxinus* | *Fraxinus +* |
| - | *Gramineae* | - | Not used |
| *Hedera* | *Hedera* | - | Not used |
| - | *Illex* | - | *Aquifolicaeae* |
| - | *Isoëtes* | - | Not used |
| *Juglans* | - | *Juglans* | *Juglans +* |
| - | - | *Juniperus* | Not used |
| - | - | *Liliaceae* | *Liliaceae* |
| *Liquidambar* | *Liquidambar* | - | *Liquidambar* |
| *Lonicera* | - | - | Not used |
| *Lycopodium (exoot)* | *Lycopodium (exoot)* | *Lycopodium (exoot)* | *Lycopodium +* |

| | | | |
|---|---|---|---|
| - | *Lythrum* | - | Not used |
| - | - | *Malvaceae* | *Malvaceae* |
| *Myrica* | *Myricaceae* | - | *Myrica* |
| - | *Myriophyllum* | - | Not used |
| - | *Nymphaea* | *Nympheae* | Not used |
| *Nyssa* | *Nyssa* | *Nyssa* | *Nyssa* |
| *Osmunda* | *Osmunda* | *Osmunda* | *Osmundaceae* |
| *Ostrya* | - | *Ostrya* | *Ostryacar* |
| *Parthenociccus* | - | - | Not used |
| - | - | *Pilularia* | Not used |
| *Picea* | *Picea* | *Picea* | *Picea +* |
| *Pinus* | *Pinus* | *Pinus* | *Pinus* |
| *Pinus haplo* | *Pinus haplo* | *Pinus haplo* | *Pinus haplo* |
| - | *Plantago* | - | *Plantaginaceae* |
| - | *Plantago spp* | - | Not used |
| - | *Platycarya* | - | Not used |
| *Poaceae* | - | *Poaceae* | *Poaceae* |
| *Polypodium* | - | - | *Polypodiaceae* |
| - | *Potamogeton* | - | Not used |
| *Pteridium* | *Pteridium* | *Pteridium* | *Pteridium* |
| *Pterocarya* | *Pterocarya* | *Pterocarya* | Not used |
| *Quercus* | *Quercus* | *Quercus* | *Quercus* |
| *Ranunculus* | *Ranunculaceae* | - | *Ranunculaceae* |
| - | *Reveesia* | - | Not used |
| - | *Rhus* | - | Not used |
| - | *Rosaceae* | - | *Rosaceae* |
| *Rumex* | *Rumex* | - | *Rumex* |
| - | *Salix* | *Salix* | *Salix* |
| *Sanguisorba* | - | - | *Sanquisorba* |
| - | *Sapotaceae* | - | Not used |
| *Sciadopitys* | *Sciadopitys* | - | Not used |
| *Sequoia* | *Sequoia* | - | Not used |
| - | *Spergula* | - | *Caryophyllaceae* |
| *Sphagnum* | *Sphagnum* | *Sphagnum* | *Sphagnum* |
| - | - | *Succisa* | *Caprifoliaceae* |
| - | *Symplocos* | - | Not used |
| *Taxodium* | *Taxodium* | *Taxodium* | *Taxodium* |
| *Taxus* | - | - | *Taxus* |
| - | - | *Thalictrum* | *Thalictrum* |
| *Thelypteris* | - | - | Not used |
| *Tilia* | *Tilia* | *Tilia* | *Tilia* |
| | *Trudopollis* | | Not used |
| *Tsuga* | *Tsuga* | *Tsuga* | *Tsuga +* |
| - | *Typhaceae* | - | Not used |
| *Ulmus* | *Ulmus* | *Ulmus* | *Ulmus* |
| - | *Umbelliferaea* | - | Not used |
| *Urtica* | - | *Urtica* | *Urticaceae* |
| - | *Viscum* | - | *Arceuthobium* |
| - | *Vitis* | - | Not used |

## M. Percentage Trees & Bisaccates

| Trees | English name | Bisaccates | English name |
|---|---|---|---|
| *Acer* | Japanese maple | *Abies* | Firs |
| *Alnus* | Alder | *Cupressa* | Conifer |
| *Anacardia* | Cashew family | *Picea* | Spruce |
| *Arecaceae* | Palm | *Pinus* | Pine |
| *Betula* | Birch | *Pinus Haploxyon* | Pine |
| *Carpinus* | Hormbeam | *Tsuga* | Hemlock tree |
| *Carya* | Hickory | | |
| *Celtis* | Nettle tree | | |
| *Corylus* | Hazel | | |
| *Fagus* | Beech | | |
| *Juglans* | Walnut tree | | |
| *Fraxinus* | Ash | | |
| *Liquidambar* | American Storax | | |
| *Nyssa* | Tupelo | | |
| *Quercus* | Oak | | |
| *Salix* | Willow | | |
| *Taxodium* | Taxus | | |
| *Tilia* | Lime tree | | |
| *Ulmus* | Elm | | |

| Sample nr | Sum trees | Sum Bisaccates | Pollensum | %Trees | %Bisaccates |
|---|---|---|---|---|---|
| **1** | 10 | 12 | 22 | 45,45455 | 54,54545 |
| **2** | 14 | 2 | 16 | 87,5 | 12,5 |
| **3** | 52 | 23 | 75 | 69,33333 | 30,66667 |
| **4** | 5 | 23 | 28 | 17,85714 | 82,14286 |
| **5** | 39 | 103 | 142 | 27,46479 | 72,53521 |
| **6** | 43 | 113 | 156 | 27,5641 | 72,4359 |
| **7** | 107 | 55 | 162 | 66,04938 | 33,95062 |
| **8** | 116 | 39 | 155 | 74,83871 | 25,16129 |
| **9** | 50 | 87 | 137 | 36,49635 | 63,50365 |
| **10** | 32 | 113 | 145 | 22,06897 | 77,93103 |
| **11** | 70 | 72 | 142 | 49,29577 | 50,70423 |
| **12** | 71 | 76 | 147 | 48,29932 | 51,70068 |
| **13** | 76 | 85 | 161 | 47,20497 | 52,79503 |
| **14** | 74 | 82 | 156 | 47,4359 | 52,5641 |
| **15** | 38 | 138 | 176 | 21,59091 | 78,40909 |
| **16** | 28 | 111 | 139 | 20,14388 | 79,85612 |
| **17** | 17 | 119 | 136 | 12,5 | 87,5 |
| **18** | 59 | 89 | 148 | 39,86486 | 60,13514 |
| **19** | 51 | 82 | 133 | 38,34586 | 61,65414 |
| **20** | 62 | 52 | 114 | 54,38596 | 45,61404 |
| **21** | 40 | 75 | 115 | 34,78261 | 65,21739 |
| **22** | 60 | 44 | 104 | 57,69231 | 42,30769 |

| | | | | | |
|---|---|---|---|---|---|
| **23** | 148 | 22 | 170 | 87,05882 | 12,94118 |
| **24** | 56 | 65 | 121 | 46,28099 | 53,71901 |
| **25** | 85 | 50 | 135 | 62,96296 | 37,03704 |
| **26** | 111 | 40 | 151 | 73,50993 | 26,49007 |
| **27** | 88 | 44 | 132 | 66,66667 | 33,33333 |
| **28** | 68 | 67 | 135 | 50,37037 | 49,62963 |
| **29** | 61 | 91 | 152 | 40,13158 | 59,86842 |
| **30** | 38 | 75 | 113 | 33,62832 | 66,37168 |
| **31** | 59 | 49 | 108 | 54,62963 | 45,37037 |
| **32** | 25 | 76 | 101 | 24,75248 | 75,24752 |
| **33** | 36 | 61 | 97 | 37,1134 | 62,8866 |
| **34** | 43 | 61 | 104 | 41,34615 | 58,65385 |
| **35** | 49 | 60 | 109 | 44,95413 | 55,04587 |
| **36** | 37 | 68 | 105 | 35,2381 | 64,7619 |
| **37** | 32 | 64 | 96 | 33,33333 | 66,66667 |
| **38** | 31 | 66 | 97 | 31,95876 | 68,04124 |
| **39** | 38 | 99 | 137 | 27,73723 | 72,26277 |
| **40** | 29 | 74 | 103 | 28,15534 | 71,84466 |
| **41** | 41 | 80 | 121 | 33,8843 | 66,1157 |
| **42** | 36 | 91 | 127 | 28,34646 | 71,65354 |
| **43** | 16 | 68 | 84 | 19,04762 | 80,95238 |
| **44** | 19 | 41 | 60 | 31,66667 | 68,33333 |
| **45** | 15 | 93 | 108 | 13,88889 | 86,11111 |
| **46** | 16 | 77 | 93 | 17,2043 | 82,7957 |
| **47** | 18 | 77 | 95 | 18,94737 | 81,05263 |
| **48** | 25 | 69 | 94 | 26,59574 | 73,40426 |
| **49** | 25 | 129 | 154 | 16,23377 | 83,76623 |
| **50** | 27 | 77 | 104 | 25,96154 | 74,03846 |
| **51** | 42 | 69 | 111 | 37,83784 | 62,16216 |
| **52** | 38 | 72 | 110 | 34,54545 | 65,45455 |
| **53** | 40 | 67 | 107 | 37,38318 | 62,61682 |
| **54** | 37 | 51 | 88 | 42,04545 | 57,95455 |
| **55** | 28 | 67 | 95 | 29,47368 | 70,52632 |
| **56** | 29 | 36 | 65 | 44,61538 | 55,38462 |
| **57** | 18 | 41 | 59 | 30,50847 | 69,49153 |
| **58** | 34 | 92 | 126 | 26,98413 | 73,01587 |
| **59** | 63 | 72 | 135 | 46,66667 | 53,33333 |
| **60** | 32 | 50 | 82 | 39,02439 | 60,97561 |
| **61** | 26 | 88 | 114 | 22,80702 | 77,19298 |
| **62** | 30 | 110 | 140 | 21,42857 | 78,57143 |
| **63** | 27 | 117 | 144 | 18,75 | 81,25 |
| **64** | 19 | 91 | 110 | 17,27273 | 82,72727 |
| **65** | 19 | 98 | 117 | 16,23932 | 83,76068 |
| **66** | 27 | 114 | 141 | 19,14894 | 80,85106 |
| **67** | 39 | 156 | 195 | 20 | 80 |

| | | | | | |
|---|---|---|---|---|---|
| **68** | 26 | 116 | 142 | 18,30986 | 81,69014 |
| **69** | 28 | 84 | 112 | 25 | 75 |
| **70** | 7 | 177 | 184 | 3,804348 | 96,19565 |
| **71** | 78 | 96 | 174 | 44,82759 | 55,17241 |
| **72** | 74 | 110 | 184 | 40,21739 | 59,78261 |
| **73** | 25 | 175 | 200 | 12,5 | 87,5 |
| **74** | 1 | 1 | 2 | 50 | 50 |
| **75** | 1 | 6 | 7 | 14,28571 | 85,71429 |
| **76** | 5 | 2 | 7 | 71,42857 | 28,57143 |
| **77** | 1 | 3 | 4 | 25 | 75 |
| **78** | 1 | 0 | 1 | 100 | 0 |
| **79** | 3 | 16 | 19 | 15,78947 | 84,21053 |
| **80** | 8 | 83 | 91 | 8,791209 | 91,20879 |
| **81** | 4 | 120 | 124 | 3,225806 | 96,77419 |
| **82** | 3 | 64 | 67 | 4,477612 | 95,52239 |
| **83** | 3 | 0 | 3 | 100 | 0 |
| **84** | 13 | 39 | 52 | 25 | 75 |
| **85** | 0 | 1 | 1 | 0 | 100 |
| **86** | 2 | 0 | 2 | 100 | 0 |
| **87** | 1 | 4 | 5 | 20 | 80 |
| **88** | 9 | 1 | 10 | 90 | 10 |
| **89** | 0 | 3 | 3 | 0 | 100 |
| **90** | 6 | 106 | 112 | 5,357143 | 94,64286 |
| **91** | 37 | 26 | 63 | 58,73016 | 41,26984 |
| **92** | 20 | 31 | 51 | 39,21569 | 60,78431 |
| **93** | 7 | 6 | 13 | 56 | 44 |
| **94** | 17 | 65 | 82 | 20,73171 | 79,26829 |
| **95** | 22 | 2 | 24 | 93,61702 | 6,382979 |
| **96** | 15 | 19 | 34 | 44,11765 | 55,88235 |
| **97** | 2 | 1 | 3 | 66,66667 | 33,33333 |
| **98** | 7 | 3 | 10 | 70 | 30 |
| **99** | 12 | 79 | 91 | 13,18681 | 86,81319 |
| **100** | 19 | 57 | 76 | 25 | 75 |
| **101** | 12 | 76 | 88 | 13,63636 | 86,36364 |
| **102** | 13 | 62 | 75 | 17,33333 | 82,66667 |
| **103** | 9 | 14 | 23 | 39,13043 | 60,86957 |
| **104** | 3 | 5 | 8 | 37,5 | 62,5 |
| **105** | 2 | 8 | 10 | 20 | 80 |
| **106** | 6 | 7 | 13 | 46,15385 | 53,84615 |
| **107** | 3 | 5 | 8 | 37,5 | 62,5 |
| **108** | 13 | 62 | 75 | 17,33333 | 82,66667 |
| **109** | 8 | 41 | 49 | 16,32653 | 83,67347 |
| **110** | 13 | 26 | 39 | 33,33333 | 66,66667 |
| **111** | 8 | 82 | 90 | 8,888889 | 91,11111 |
| **112** | 9 | 49 | 58 | 15,51724 | 84,48276 |

| | | | | | |
|---|---|---|---|---|---|
| **113** | 8 | 13 | 21 | 38,09524 | 61,90476 |
| **114** | 6 | 37 | 43 | 13,95349 | 86,04651 |
| **115** | 23 | 64 | 87 | 26,5896 | 73,4104 |
| **116** | 6 | 52 | 58 | 10,34483 | 89,65517 |
| **117** | 9 | 30 | 39 | 23,07692 | 76,92308 |
| **118** | 4 | 13 | 17 | 23,52941 | 76,47059 |
| **119** | 11 | 23 | 34 | 32,35294 | 67,64706 |
| **120** | 14 | 44 | 58 | 24,13793 | 75,86207 |
| **121** | 8 | 4 | 12 | 66,66667 | 33,33333 |
| **122** | 4 | 19 | 23 | 17,77778 | 82,22222 |
| **123** | 3 | 16 | 19 | 15,78947 | 84,21053 |
| **124** | 3 | 39 | 42 | 7,142857 | 92,85714 |
| **125** | 14 | 60 | 74 | 19,04762 | 80,95238 |
| **126** | 2 | 38 | 40 | 5 | 95 |
| **127** | 9 | 59 | 68 | 13,33333 | 86,66667 |
| **128** | 16 | 75 | 91 | 17,58242 | 82,41758 |
| **129** | 8 | 75 | 83 | 9,638554 | 90,36145 |
| **130** | 15 | 43 | 58 | 25,86207 | 74,13793 |
| **131** | 5 | 94 | 99 | 5,050505 | 94,94949 |
| **132** | 10 | 15 | 25 | 40 | 60 |
| **133** | 5 | 71 | 76 | 6,578947 | 93,42105 |
| **134** | 8 | 90 | 98 | 8,163265 | 91,83673 |
| **135** | 2 | 68 | 70 | 2,857143 | 97,14286 |
| **136** | 16 | 59 | 75 | 21,33333 | 78,66667 |
| **137** | 5 | 7 | 12 | 41,66667 | 58,33333 |
| **138** | 5 | 69 | 74 | 6,756757 | 93,24324 |
| **139** | 9 | 47 | 56 | 16,07143 | 83,92857 |
| **140** | 2 | 21 | 23 | 8,695652 | 91,30435 |
| **141** | 41 | 37 | 78 | 52,90323 | 47,09677 |
| **142** | 17 | 76 | 93 | 18,37838 | 81,62162 |
| **143** | 11 | 77 | 88 | 12,57143 | 87,42857 |
| **144** | 16 | 53 | 69 | 23,35766 | 76,64234 |
| **145** | 7 | 76 | 83 | 8,433735 | 91,56627 |
| **146** | 15 | 125 | 140 | 10,71429 | 89,28571 |
| **147** | 8 | 164 | 172 | 4,651163 | 95,34884 |
| **148** | 16 | 86 | 102 | 15,68627 | 84,31373 |
| **149** | 9 | 22 | 31 | 29,03226 | 70,96774 |
| **150** | 8 | 36 | 44 | 18,18182 | 81,81818 |
| **151** | 14 | 183 | 197 | 7,106599 | 92,8934 |
| **152** | 27 | 128 | 155 | 17,41935 | 82,58065 |
| **153** | 34 | 65 | 99 | 34,34343 | 65,65657 |
| **154** | 8 | 39 | 47 | 17,02128 | 82,97872 |
| **155** | 18 | 38 | 56 | 32,14286 | 67,85714 |
| **156** | 19 | 135 | 154 | 12,33766 | 87,66234 |
| **157** | 23 | 99 | 122 | 18,85246 | 81,14754 |

| | | | | | |
|---|---|---|---|---|---|
| **158** | 60 | 293 | 353 | 16,99717 | 83,00283 |
| **159** | 74 | 341 | 415 | 17,83133 | 82,16867 |
| **160** | 15 | 124 | 139 | 10,79137 | 89,20863 |
| **161** | 20 | 191 | 211 | 9,478673 | 90,52133 |
| **162** | 140 | 375 | 515 | 27,18447 | 72,81553 |
| **163** | 35 | 43 | 78 | 44,87179 | 55,12821 |
| **164** | 57 | 166 | 223 | 25,56054 | 74,43946 |
| **165** | 13 | 118 | 131 | 9,923664 | 90,07634 |
| **166** | 22 | 232 | 254 | 8,661417 | 91,33858 |
| **167** | 19 | 99 | 118 | 16,10169 | 83,89831 |
| **168** | 30 | 495 | 525 | 5,714286 | 94,28571 |
| **169** | 9 | 20 | 29 | 31,03448 | 68,96552 |
| **170** | 20 | 77 | 97 | 20,61856 | 79,38144 |
| **171** | 11 | 175 | 186 | 5,913978 | 94,08602 |
| **172** | 28 | 102 | 130 | 21,53846 | 78,46154 |