Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

A.J. de Vreng

MSc. thesis
Earth Science

Supervisors:
Dr. J.M. van Mourik
Dr. B. Jansen
Prof. Dr. K. Kalbitz

March 2011
Abstract

Historical land use systems as shifting cultivation and plaggen agriculture on nutrient poor sandy soils caused serious degradation of soil and vegetation, resulting in Late Holocene sand drifting. Locally the coversand landscape transformed into a driftsand landscape. On sheltered deposition sites polycyclic driftsand profiles developed. They are valuable geo-ecological records of alternating instable (sand drifting) and stable (soil formation) phases in landscape development.

Previous research has been done on the reconstruction of the landscape genesis, based on pollen analysis and radiocarbon dating of polycyclic driftsand deposits. Recently, the application of Optically Stimulated Luminescence (OSL) dating improved the geochronology. The principles of pollen transport and deposition does not allow to select the plant species from pollen spectra, responsible for the on-site production of soil organic matter during soil formation in stable periods.

A possible proxy for this, that could add information, is the application of straight-chain lipids analysis. In this pilot study we introduced the biomarker method to separate local and regional species, found in pollen spectra. Different species often have their own specific patterns of alcohols and alkanes of different carbon chainlengths. Therefore, comparing the straight-chain lipid patterns of both (expected) vegetation species and selected soil samples is a means to assess the landscape history. In order to do this the recently developed VERHIB model was used in combination with a cluster analysis.

The results obtained from VERHIB and the cluster analysis suggest a dynamic sequence of dominant vegetation species, which in majority support the existing theories about the landscape history of polycyclic deposits in the Netherlands. Some results (e.g. the distribution of Pinus throughout the profile) are much harder to link to earlier research and are likely to be incorrect.

Although a couple of problems arose during the analytical process, the application of straight-chain lipids is found to be certainly a very promising technique. Nevertheless, a lot more research has to be carried out to draw more solid conclusions on the strength of the biomarker technique in sandy environments and its impact on the ideas about the cultural-ecological history of polycyclic soils in the Netherlands.
Contents

Abstract .......................................................................................................................... 3

1. Introduction .................................................................................................................. 7
2. Objectives ...................................................................................................................... 11
3. Research area .............................................................................................................. 13
   3.1. Summary of earlier work Defensiedijk ................................................................. 14
      3.1.1. Pollen analysis (profile 1984) ...................................................................... 14
      3.1.2. ¹³C dating (profiles 1984, 1986 and 2002) ................................................. 14
      3.1.3. Micromorphology (profiles 1986 and 2002) .................................................... 16
      3.1.4. OSL dating (profile 2002) .......................................................................... 17
4. Methodology ............................................................................................................... 19
   4.1. Field methodology ............................................................................................... 19
      4.1.1. Soil sampling ................................................................................................. 19
      4.1.2. Vegetation sampling ..................................................................................... 21
   4.2. Laboratory methodology ...................................................................................... 22
      4.2.1. Preparation .................................................................................................... 22
      4.2.2. ASE extraction ............................................................................................ 22
      4.2.3. Clean-up ........................................................................................................ 23
      4.2.4. Derivatisation .............................................................................................. 23
      4.2.5. GC/MS analysis ........................................................................................... 24
   4.3. Data analysis ...................................................................................................... 25
      4.3.1. Lipid quantification ...................................................................................... 25
      4.3.2. Cluster analysis ......................................................................................... 25
      4.3.3. VERHIB model ......................................................................................... 26
5. Results and discussion .............................................................................................. 29
   5.1. Lipid quantification .............................................................................................. 29
   5.2. Cluster analysis .................................................................................................... 34
   5.3. VERHIB model .................................................................................................. 37
      5.3.1. Output analysis ............................................................................................ 38
      5.3.2. Alternative runs .......................................................................................... 40
   5.4. Relation to data from previous studies ............................................................... 42
6. Conclusions .............................................................................................................. 43

References ..................................................................................................................... 45

APPENDICES .............................................................................................................. 49
Appendix A: Photos of Defensiedijk 2008 profile ....................................................... 49
Appendix B: Matlab code .............................................................................................. 53
Appendix C: GC/MS results ......................................................................................... 59
1. Introduction

Late glacial aeolian coversand dominates the surface geology of large parts in the east and the south of the Netherlands. During prehistoric and early historic time, forest grazing, wood cutting and shifting cultivation gradually transformed natural forest into heath land. Part of this heath land is later utilized in the plaggen agriculture.

In extensive parts of North-western Europe this plaggen agriculture was being carried out starting about 3000 years ago, increasing from the early middle Ages on. This went on until the introduction of chemical fertilizers in the beginning of the 20\textsuperscript{th} century made the use of plaggen management obsolete (e.g Spek, 2004).

Heather, forest and grass sods were used for animal bedding in stables and afterwards applied to the fields. This caused deep humus rich plaggen soils to be formed near the villages (e.g Blume & Leinweber, 2004), with significantly improved fertility. Large parts of The Netherlands are influenced by plaggen management. More than 221,000 ha of land has soils with a humic sand cover thicker than 50 cm, and more than 196,000 ha has soils with a cover from 30 to 50 cm thick (Pape, 1970). These soil are generally classified as Plaggic Podzols and Plaggic Anthrosols (IUSS Working Group WRB, 2007).

As an effect of the plaggen agriculture, serious degradation of soil and vegetation occurred in the areas where the removal of plaggen took place. This resulted in the comeback of sand drifting and locally the coversand landscape transformed into a driftsand landscape. Characteristic for this landscape are soils consisting of polycyclic driftsand deposits, which can be utilized as geo-ecological records of alternating instable and stable phases in landscape development. Every cycle of a polycyclic profile reflects a period of landscape instability (deposition) and landscape stability (soil development).

In the past quite some research has been done on the historic reconstruction of cultural sandy soils in the Netherlands using radiocarbon dating, Optically Stimulated Luminescence (OSL) dating and pollen records. Due to the complex composition of soil organic matter, the interpretation of conventional radiocarbon dating results has never been satisfactory (van
Mourik et al., 1995). This is because some of the components of soil organic material developed in situ, while other (often older) components were transported into the area with the sods. To overcome this problem OSL dating was introduced (e.g. Bokhorst et al., 2005). This technique dates the last time quartz grains were exposed to light.

Another source of uncertainty is in the origin of the pollen influx. Wind dispersal of the pollen prior to deposition might have a significant effect (e.g. Okubo and Levin, 1989). Therefore finding a proxy that is not influenced by this wind dispersal is essential to draw more solid conclusions concerning past vegetation distributions. Lipid biomarkers is such a proxy as it is based on root (no wind dispersal) and leaf (only very local wind dispersal) biomass of the vegetation.

Lipids are important components in soil organic matter (SOM) due to their hydrophobic nature (Quenea et al., 2006). As a result they are of major importance for various soil features such as aggregate stability, water retention and fertility. These compounds derive from the epicuticular waxes in leaves as well as waxes from roots and are usually present in distinctive odd-over-even carbon chain length predominance for the n-alkanes and an even-over-odd predominance for the n-alcohols (Jansen et al., 2006a). Solvent extractable lipids are usually regarded as specific indicators of terrestrial higher plant input (Rieley et al., 1991; van Bergen et al., 1997), as different plant species often have distinctive lipid compositions. If lipids are well preserved within the soil, peat or sediment and the stratification is still intact, the lipid composition within the soil profile can be used to determine past vegetation distributions and related dominating biogeochemical processes. Until recently the focus of research on lipids as biomarkers has been on lake sediments and peat bogs (e.g. Ficken et al., 1998; Ishiwatari et al., 2005). The application of lipid biomarkers in terrestrial soils is a comparatively unexplored terrain. In one of the few related studies Jansen et al. (2006a) utilize ratios of straight-chain lipids of different chain-lengths to differentiate between páramo and forest species in order to reconstruct past upper forest line positions in Ecuador as part of the RUFLE project. It shows that in this particular case, n-alkanes and n-alcohols compositions of the two different types of vegetation are the best suitable for differentiating between them. N-fatty acids, wax esters and n-aldehydes appear to be less suitable as
vegetation biomarkers due to their apparent lower specificity ($n$-fatty acids) and absence from a significant number of the species investigated (wax esters and $n$-aldehydes).

Bioturbation results in mixing of old and younger soil organic matter (SOM). When this process is dominant the applicability of lipid biomarkers is very limited. However, soil profiles of cultural sandy soils in the Netherlands show clear stratification, which together with the results of both radiocarbon and OSL dating suggests that bioturbation is not a dominant process in this system. As the $n$-alkanes and $n$-alcohols of interest ($C_{20}$-$C_{36}$, as these are typical for higher plant input) also have a hydrophobic nature (Quenea et al., 2004), and have shown to have a relatively high resistance to biodegradation, it is expected that they remain in situ after deposition for a reasonably long time. However it should be noted that roots grow downwards into the profile and their lipids might therefore be younger than the surrounding soil material.

For this pilot study we selected a previously investigated polycyclic driftsand sequence, profile Defensiedijk. The pollen diagram suggests the presence of plant species during stable periods with soil formation. The separation of in situ species and regional species can be based on the comparison of the pollen spectra with a carefully composed dataset of biomarkers.
2. Objectives

As few is known about the effectiveness of the lipid biomarker approach applied on cultural sandy landscapes, the primary goal of this research will be to determine its potential as an added resource of information (proxy) regarding landscape development. To achieve this goal, the following research questions have been addressed:

- Are straight-chain lipids well preserved within the organic soil layers of interest in a meaningful chronological order?
- Are the straight-chain lipid compositions of the plant species responsible for the dominant biomass input in the soil records in the area unique? Is it possible to distinguish between groups of heather, grasses and trees? Is further identification on species level achievable?
- Can the straight-chain lipid composition within the soil samples be linked to the different plant(group) specific compositions (both qualitative and quantitative)?
- What is the relation between the results from the biomarker approach and the existing palynological information derived from the analysis of fossil pollen?

The secondary objective of the research is to gain more insight into the landscape development in cultural sandy landscapes. The following question has been formulated to achieve this objective:

- What conclusions can be drawn by combining the results of the biomarker approach and the existing palynological information about the change in abiotic conditions that were present during the transition of the coversand landscape into a driftsand landscape?
3. Research area

Landforms and soils around the city of Weert are representative for the cultural landscapes that developed on chemically poor Late Weichselian aeolian coversands in NW Europe (van Mourik, 1988). Historically, there was a close relation between the development of Plaggic Anthrosols (IUSS Working Group WRB, 2007) and driftsand deposits in these cultural landscapes (Bokhorst et al., 2005).

Profile Defensiedijk is situated in the Weerterbergen, SE of the city of Weert. Figure 3.1 shows a fragment of a historical map (1900 AD) with the characteristic land use of cultural landscapes on chemically poor sandy soils. Around the city of Weert, arable fields (Plaggic Anthrosols) are visible. More to the west we can see the extensive heath (Podzols) with complexes of land dunes (Arenosols) and the first generation of pine plantations. Profile Defensiedijk is considered as an important paleo-ecological record of stable and instable periods in the development of the cultural landscape. In 1984, the profile was sampled for pollen analysis. Also samples were taken for radiocarbon dating, applied on bulk samples. In 1986 the same profile was resampled for fractionated radiocarbon dating and soil micromorphology. Finally, in 2002 samples were taken for OSL dating.

Figure 3.1: historical map of the research area (adapted from van Mourik et al. (2010))
3.1. Summary of earlier work Defensiedijk

The main reason why Defensiedijk was chosen as the area of interest of this research is that the abovementioned research was already carried out there, which gives the opportunity to get a good idea of the applicability of the lipid biomarker approach. The following section will therefore summarize the findings of the previous research done at the Defensiedijk profile (van Mourik et al., 2010).

3.1.1. Pollen analysis (profile 1984)

Pollen extractions of samples from all horizons using the pollen determination key of Moore et al. (1991) were carried out. This resulted in a pollen diagram (figure 3.2). This diagram gives a glance at the landscape evolution of the profile Defensiedijk and its surrounding. It shows the degradation from forest into heath as an effect of deforestation (zone 1), the result of and the plaggen removal, with a replacement of heather by grasses (zone 2) and the introduction of pine trees (plantation of Pinus started after 1550 AD) to stabilize the driftsand landscape (zone 3). Zone 4 represents the most recent period of sand drifting. Only since 1995 the area is stabilizing under a vegetation of grasses.

3.1.2. $^{14}$C dating (profiles 1984, 1986 and 2002)

Conventional radiocarbon ages of bulk samples of buried A horizons of the 1984 profile were performed in order to interpret the chronology of pollen zones and sand deposits by the CIO (Centrum voor Isotopen Onderzoek), Rijks Universiteit Groningen, The Netherlands, following the methods of Mook and Streurman (1983). The $^{14}$C dates of bulk samples (profile 1984) did not provide a clear geochronology (see table 3.1), as the top of the profile is clearly dated too old. Therefore the profile was resampled in 1986 for fractionated $^{14}$C dating (van Mourik et al., 1995) consisting of three fractions: fulvic acids (FUL), humic acids (HAC) and humin (HUM). The biological decomposition rate of FUL is relatively high; this fraction migrates easily downwards through the soil profile. This makes them unsuitable for dating and is therefore discarded. The biological decomposition rate of HUM is lower than that of FUL. Compared to FUL, HUM is immobile and reliable for dating purposes. The $^{14}$C age of HAC will be the closest to the moment of burying of the soil. HUM will accumulate during an active period of soil development; therefore, $^{14}$C ages of this fraction will overestimate the time of fossilization of the soil. It is assumed that the differences
The numbers under "Zone" refer to the zones described in 3.1.1. “S” stands for stable period; “D” for a period in which deposition was important.

Figure 3.2: pollen diagram of the Defensiedijk 1984 profile (from van Mourik et al. (2010))

<table>
<thead>
<tr>
<th>Zone</th>
<th>Depth in cm</th>
</tr>
</thead>
<tbody>
<tr>
<td>1S</td>
<td>190</td>
</tr>
<tr>
<td>2D</td>
<td>150</td>
</tr>
<tr>
<td>2S</td>
<td>125</td>
</tr>
<tr>
<td>3S</td>
<td>100</td>
</tr>
<tr>
<td>4S</td>
<td>75</td>
</tr>
<tr>
<td>4D</td>
<td>50</td>
</tr>
<tr>
<td>3D</td>
<td>25</td>
</tr>
</tbody>
</table>

LogD = logarithmic value of amounts of pollen grains per ml sediment

$^{14}C$ years BP
between the ages of HUM and HAC increase during an active period of soil formation. The big difference in \(^{14}\)C ages between HUM and HAC in the 2A, 3A and 4A horizons confirms this. Although the radiocarbon dating gives valuable insights in the development of the soil profile, the dates are not very reliable. This is best shown by comparing the HAC ages of the 1986 and 2002 profiles of horizons 4A and 4B. The 2002 horizons are dated 1000-2000 years younger, probably caused by influence of the above lying horizons (the distance between horizons 3A and 4A is considerably smaller in the 2002 profile in comparison with the 1986 profile).

### Table 3.1: \(^{14}\)C ages of different samples taken at profile Defensiedijk (from Van Mourik et al. (2010))

<table>
<thead>
<tr>
<th>Depth (cm)</th>
<th>Horizon</th>
<th>Bulk profile 1984 (a)</th>
<th>HUM profile 1986 (a)</th>
<th>HAC profile 1986 (a)</th>
<th>Depth (cm)</th>
<th>HAC profile 2002 (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25 - 27</td>
<td>2A</td>
<td>1130 ± 60</td>
<td>3230 ± 110</td>
<td>410 ± 45</td>
<td>69 - 71</td>
<td>410 ± 35</td>
</tr>
<tr>
<td>127 - 129</td>
<td>3A (top)</td>
<td>1075 ± 30</td>
<td>1350 ± 50</td>
<td>1365 ± 25</td>
<td>128 - 130</td>
<td>1230 ± 35</td>
</tr>
<tr>
<td>129 - 131</td>
<td>3A (bottom)</td>
<td>1900 ± 110</td>
<td>1675 ± 30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>173 - 175</td>
<td>4A (top)</td>
<td>3920 ± 40</td>
<td>4110 ± 90</td>
<td>3615 ± 35</td>
<td>154 - 156</td>
<td>2645 ± 40</td>
</tr>
<tr>
<td>175 - 176</td>
<td>4A (bottom)</td>
<td>4430 ± 165</td>
<td>3965 ± 40</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>185 - 186</td>
<td>4B (top)</td>
<td></td>
<td>3985 ± 35</td>
<td></td>
<td>170 - 172</td>
<td>2080 ± 35</td>
</tr>
<tr>
<td>187 - 188</td>
<td>4B (middle)</td>
<td>3535 ± 80</td>
<td>3730 ± 35</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>189 - 190</td>
<td>4B (bottom)</td>
<td></td>
<td>3700 ± 50</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.1.3. Micromorphology (profiles 1986 and 2002)

For micromorphological analysis, undisturbed samples were taken in Kubiena boxes for the production of thin sections (7 cm/4 cm/20 μm) (Brewer, 1976). The main goal of this analysis was to assess the reliability of the radiocarbon dating. Some of the main findings are the presence of channels throughout the profile and sin-sedimentary charcoal particles in the 2C horizon. These signs of bioturbation explain the contamination of driftsand with SOM from older soil horizons. Also pollen grains are found in sin-sedimentary transported aggregates, which present another complication in the interpretation of pollen diagrams in buried humic horizons.
3.1.4. OSL dating (profile 2002)

Unlike $^{14}$C dating, luminescence dating uses the constituent mineral grains of the sediment itself, and it makes direct determination of sedimentation ages possible (e.g., Aitken, 1998). The profile Defensiedijk had previously been sampled for thermoluminescence (TL) dating using feldspar, but the age results lacked precision (Dijkmans et al., 1992). More recently, optically stimulated luminescence (OSL) proved to be better suited to date sediments, as it was successfully applied to Holocene (Derese et al., 2010) and late Pleistocene sediments (Vandenberghe et al., 2004) and soils in cultural landscapes (Bokhorst et al., 2005). Therefore, in 2002, the profile Defensiedijk was resampled for OSL dating. OSL dating on quartz grains was performed in the luminescence dating laboratory at Ghent University. The methodology, luminescence characteristics of the samples and OSL dating results have previously been presented by Vandenberghe et al. (2005).

The results of the OSL dating are shown in table 3.2. The ages for the samples from the 1C, 2Ah and 2C are in chronological order and therefore stratigraphical consistent with their position. Conversely, the samples of 3Ah and 3E show an age inversion. Field observations show a complicated sedimentary structure, which might be caused by short distance re-sedimentation, explaining the age inversion. The sample of 4Bh was taken from the coversand unit. Coversand deposition is generally believed to have stopped in the Late Pleistocene, which implicates that the age of the sample of 4Bh of 9200 years must be regarded as (at least 3000 years) too young. As discussed before, the micromorphological analysis shows some signs of bioturbation; contamination through this bioturbation of vertically transported younger mineral grains could have caused the OSL age of the 4Bh sample to be underestimated. It cannot be ruled out completely that other parts of the profile suffer from similar over- or underestimations of OSL ages.

Nevertheless, based on the OSL dating results, the chronology of the polycyclic sequence became more clear (table 3.3). Multiple phases of sand drifting in the last approximately 1300

<table>
<thead>
<tr>
<th>Depth (cm)</th>
<th>Horizon</th>
<th>OSL age (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>45</td>
<td>1C1</td>
<td>82 ± 8</td>
</tr>
<tr>
<td>55</td>
<td>1C2</td>
<td>100 ± 10</td>
</tr>
<tr>
<td>70</td>
<td>1C3</td>
<td>90 ± 10</td>
</tr>
<tr>
<td>105</td>
<td>2Ah</td>
<td>350 ± 30</td>
</tr>
<tr>
<td>105</td>
<td>2C1</td>
<td>590 ± 50</td>
</tr>
<tr>
<td>125</td>
<td>2C1</td>
<td>670 ± 60</td>
</tr>
<tr>
<td>137,5</td>
<td>2C2</td>
<td>1300 ±100</td>
</tr>
<tr>
<td>145</td>
<td>3Ah</td>
<td>5800 ±500</td>
</tr>
<tr>
<td>152,5</td>
<td>3E</td>
<td>4700 ±400</td>
</tr>
<tr>
<td>162,5</td>
<td>4Bh</td>
<td>9200 ±800</td>
</tr>
</tbody>
</table>
years can now be distinguished, and it shows signs of an extra phase of landscape instability roughly 5000 years ago.

Table 3.3: OSL based geochronology of profile Defensiedijk (from van Mourik et al. (2010))

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Sedimentation OSL ages (a)</th>
<th>Soil formation OSL ages (a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Driftsand</td>
<td>100 - recent</td>
<td>Initial</td>
</tr>
<tr>
<td>3. Driftsand/micro-podzols</td>
<td>1200 - 350</td>
<td>350 - 150</td>
</tr>
<tr>
<td>2. Driftsand/podzols</td>
<td>5000 - 4500</td>
<td>4500 - 1200</td>
</tr>
<tr>
<td>1. Coversand/podzols</td>
<td>&gt; 9200</td>
<td>9200 - 5000</td>
</tr>
</tbody>
</table>
4. Methodology

In order to achieve the desired objectives the following methodology, consisting of fieldwork, labwork and data analysis, was followed.

4.1. Field methodology

The field methodology consisted of collecting samples of both humic soil horizons of the Defensiedijk profile and plant tissue samples of the species that were expected to have played a key role in the landscape development. This fieldwork was carried out in November 2008.

4.1.1. Soil sampling

At the Defensiedijk site the points of interest were the different stages of vegetation development and the related conditions in between the various stadiums of sand drifting. Consequently, samples were collected only of the soil horizons rich in soil organic matter. When thickness of the organic horizons was sufficient, separate samples were taken from both upper and lower parts of the horizon. For this, a soil pit of approximately 1 m$^2$ surface and 1.5-2 meter deep was dug. The pit (figure 4.1; additional pictures can be found in appendix A) was dug near the pits for the radiocarbon and OSL age analysis and pollen analysis. It was considered that by doing this, linking data of old and new pits would be rather straightforward (by using depths and boundaries of soil horizons). The stratification of the soil profile was recorded (table 4.1 and figure 4.2), and 9 samples were taken. They were collected into airtight glass containers to prevent contamination prior to analyses.

Figure 4.1: Pit of Defensiedijk 2008 profile
When comparing stratifications of the pit dug for this research with the ones done for previous research it appeared that they were not as similar as anticipated beforehand. Apparently, the driftsand landscape was highly dynamic and consequently local differences in soil formation are big. Both depth and thickness of the various soil horizons in the 2008 profile are significantly different from the previously described profiles. Because of this, it will not be possible to link data with great certainty. Table 4.2 shows the most likely relationship between the sampled horizons of the 2008 Defensiedijk profile and the 1984 profile (see figure 3.2).
Table 4.2: Probable relationship between 2008 and 1984 Defensiedijk profiles

<table>
<thead>
<tr>
<th>2008 horizon</th>
<th>1984 horizon</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Ah)</td>
<td>1C</td>
</tr>
<tr>
<td>3(Ah)</td>
<td>2Ah</td>
</tr>
<tr>
<td>4(Ah)</td>
<td>2Cl</td>
</tr>
<tr>
<td>5Ah</td>
<td>2C2</td>
</tr>
<tr>
<td>7Ah</td>
<td>2C6</td>
</tr>
<tr>
<td>8Ah</td>
<td>3Ah</td>
</tr>
<tr>
<td>9Ah</td>
<td>4Ah</td>
</tr>
<tr>
<td>9Bh</td>
<td>4Bh</td>
</tr>
</tbody>
</table>

4.1.2. Vegetation sampling

The biomass input of the sampled soil profile is expected to be of a certain set of plant species. In order to be able to determine which of these species or groups of species were mainly responsible for the biomass input at the different parts of the profile, a reference set was assembled. Combining both information from the collected palynological information of the sample locations and literature (Spek, 2004) the species mentioned in table 4.3 were selected to be sampled and incorporated in this study.

Table 4.3: Sampled vegetation species

<table>
<thead>
<tr>
<th>Scientific name</th>
<th>English name</th>
<th>Dutch name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calluna vulgaris</td>
<td>Heather</td>
<td>Struikheide</td>
</tr>
<tr>
<td>Molinia caerulea</td>
<td>Moor Grass</td>
<td>Pijlenstrooitje</td>
</tr>
<tr>
<td>Corynephorus canescens</td>
<td>Grey hair-grass</td>
<td>Buntgras</td>
</tr>
<tr>
<td>Deschampsia flexuosa</td>
<td>Wavy hair-grass</td>
<td>Bochtige smeet</td>
</tr>
<tr>
<td>Betula pendula</td>
<td>Silver Birch</td>
<td>Ruwe berk</td>
</tr>
<tr>
<td>Pinus sylvestris</td>
<td>Scots Pine</td>
<td>Grove den</td>
</tr>
<tr>
<td>Quercus robur</td>
<td>Pedunculate Oak</td>
<td>Zomereik</td>
</tr>
<tr>
<td>Bryophyta (exact species unknown)</td>
<td>moss</td>
<td>mos</td>
</tr>
<tr>
<td>Austrova (exact species unknown)</td>
<td>lienen</td>
<td>korstmos</td>
</tr>
</tbody>
</table>

It has to be said that at the time of the fieldwork no complete knowledge was obtained about these key species, and therefore the selection of species later appeared not to be ideal. Sampled species Quercus and Betula were not found in the pollen record of the Defensiedijk profile. It would have been better if they were replaced by Corylus and Alnus. However, all four of these species belong to the group of deciduous trees and the lipid patterns can
therefore expected to be similar, as lipid patterns of related species are often to some extent similar.

4.2. Laboratory methodology

In order to prepare for the GC/MS analysis, a complex multi-step laboratory procedure was followed. It consisted of the following stages: sample preparation, the extraction of the straight-chain lipids from the sample, the purification of the extract from polar components and possible traces of moisture, and the derivatisation of thus obtained extract. Each of these steps is explained in detail below. The description of the GC/MS analysis follows.

4.2.1. Preparation

The preparation of the soil and vegetation samples consisted in essence of the same stages. The major difference in the procedure was that the vegetation samples, before putting into the containers which would later be place into the freezer, had to be cut into small pieces and divided into roots and leaves. For obvious reasons this was impossible to perform on the moss and lichen samples. Both soil and vegetation samples were put into the freezer at a temperature of approximately -10°C and stayed there overnight. Afterwards, the sample was dehydrated using a freeze-dryer. The next step was grinding the samples with the use of a grinding machine relevant for the sample kind. The soil samples were ground into particles of a size smaller than 2μm using a planetary ball mill, while the vegetation samples were ground into particles of a size smaller than 0.2mm using a rotor mill. Finally, the ground material was homogenised and placed into clean 20ml plastic containers.

4.2.2. ASE extraction

In order to analyse the lipid composition of every sample, the first step is to extract the lipids from the soil. The most established technique to do this is Soxhlet extraction (eg. Naafs et al., 2004). However, Soxhlet extraction is very time consuming (16 hours per sample) and takes relatively large volumes of solvent (250 mL or more). A relatively new alternative is the technique of accelerated solvent extraction (ASE) (Richter et al., 1996). Basically it uses raised temperature and pressure to shorten extraction time to less than one hour. Jansen et al. (2006a) show that extraction performance of ASE is on the same level with Soxhlet
extraction. Therefore ASE extraction was chosen as the method to extract the lipids from the soil in this study.

Roughly 100mg of every sample (the exact weight was registered) was extracted by the method of Accelerated Solvent Extraction (ASE) employing a Dionex 200 ASE extractor with 11 ml extraction cells and dichloromethane/methanol (DCM/MeOH) (93:7 v/v) as the extractant. SiO$_2$ was used to fill the extraction cells. The extractions were performed at a temperature of 75°C and a pressure of 17x10$^6$ Pa with a heating phase of 5 minutes and a static extraction time of 20 minutes; these settings gave good results according to Jansen et al. (2006b). The product of the extraction was approximately 20ml of each sample extract collected in a glass vial.

4.2.3. Clean-up

The DCM/MeOH extractant was removed from the extract using a rotary evaporator. Subsequently, the residue was redissolved in approximately 2-5ml DCM/2-propanol (2:1 v/v). Afterwards, this solution was filtered through a Pasteur pipette filled with defatted cotton wool, 0.5cm Na$_2$SO$_4$(s) and 2cm SiO$_2$(s). This filtering is performed to remove potential remnants of moisture (with Na$_2$SO) and very polar components (with SiO$_2$). The filtered extract was then collected in 5ml glass vials which were weighed in advance. Then the DCM/2-propanol was removed by evaporation at 30°C, accelerated by applying a mild stream of N$_2$(g) over the opened vials. The next step of the clean-up procedure consisted in weighing the vials again in order to determine the amount of residue of the different samples. After this, the residues were redissolved in 1ml DCM/2-propanol.

4.2.4. Derivatisation

In order to get a distinct signal while preventing pollution of the GC/MS a final concentration of roughly 300mg residue per litre solvent was selected. After GC/MS analysis it showed that this concentration was too low to get distinct signals of the lipids. Therefore the concentrations were raised to approximately 3g residue per litre solvent. To get to these concentrations it was calculated for every sample how much extract was needed and in how much cyclo-hexane it had to be dissolved. The calculated amounts of cleaned-up extracts were then dispensed into a new 5ml vial. Then, known amounts of an internal standard,
which consisted of a mix of deuterated lipids (d42-(*-C20 alkane, d41-(*-C20 alcohol and d39-(*-C20 fatty acid) in a solution of 320mg/l were added to the vials, depending on the final volume (20µl internal standard per 1ml final volume). After this the DCM/2-propanol was again removed by evaporation at 30°C, supported by a stream of N₂(g). To the obtained residue 50µl BSTFA (N,O-bis(trimethylsilyl) trifluoroacetamide) containing 1% TMCS (trimethylchlorosilane) was added. The closed vials were then heated at 70°C for 1 hour to derivatise all free hydroxyl and carboxylic-acid groups to their corresponding trimethylsilyl (TMS) ethers and esters. After cooling down they were dried another time by applying the N₂(g) stream at 30°C to remove the surplus of BSFTA. Finally, the residues were redissolved in the calculated amounts of cyclo-hexane.

4.2.5. GC/MS analysis

Gas chromatography-mass spectrometry (GC/MS) analyses were performed on a ThermoQuest Trace GC 2000 gas chromatograph connected to a Finnigan Trace MS quadrupole mass spectrometer. Separation took place by on-column injection of 1.0µl of the derivatised extracts on a 30-m Restek Rtx®-5Sil MS column with an internal diameter of 0.25mm and film thickness of 0.1µm, using He as a carrier gas. Temperature programming consisted of an initial temperature of 50°C for 2 minutes, heating at 40°C/min to 80°C, holding at 80°C for 2 minutes, heating at 20°C/min to 130°C, immediately followed by heating at 4°C/min to 350°C and finally holding at 350°C for 10 minutes. The subsequent MS detection in full scan mode used a mass-to-charge ratio (m/z) of 50-650 with a cycle time of 0.65s and followed electron impact ionisation with an ionisation energy of 70eV (Jansen et al., 2006b)
4.3. Data analysis

After chromatograms were obtained from the GC/MS analysis, a quantification of the different lipid compounds was made. To assess how the lipid patterns differentiate between the vegetation samples, a cluster analysis was performed. Finally a model was used to reconstruct the vegetation history based on the lipid patterns.

4.3.1. Lipid quantification

Resulting GC/MS peaks in the chromatograms for \( n \)-alkanes and \( n \)-alcohols were interpreted and peak areas calculated. The selected compounds were identified from the chromatograms based on their mass spectra and their retention times, using the peak area from selected dominant fragment ions for each class of components. The fragment ions were represented by: \( m/z \) 57 for the \( n \)-alkanes and \( m/z \) 75 for the \( n \)-alcohols. Absolute concentrations of the different compounds were calculated in \( \mu g/g \) soil sample by assuming that the response of the GC/MS is the same for the added deuterated standards and the natural compounds in the samples. Haussmann (2006) showed that this assumption is valid for \( n \)-alkanes and \( n \)-alcohols with carbon chain-lengths in the range of interest for this study (C16 to C36). The selected fragment ions for the internal standards were represented by: \( m/z \) 66 for \( d_4 \)-\( n \)-alkane and \( m/z \) 76 for \( d_4 \)-\( n \)-C20 alcohol.

4.3.2. Cluster analysis

In order to assess the distinctiveness of the patterns of the lipids found in the vegetation samples a cluster analysis was carried out. This also gives an impression whether the lipid data of the different species can be upscaled to the vegetation types mosses, grasses, heather and (deciduous and coniferous) trees. Ideally, a cluster analysis would result in clusters containing the different vegetation species belonging to a single vegetation type.

Clustering was done by doing a hierarchical cluster analysis in PASW Statistics 18 (SPSS Inc., 2009) using Ward’s method (Ward, 1963). The data to be clustered was the relevant lipid concentrations (alcohols with even-numbered carbon chain lengths and alkanes with uneven-numbered carbon chain lengths) of both leaf and root samples. The lipid concentrations
were standardized to ranges from 0 to 1 prior to clustering. The squared Euclidean distance was used as a measure of similarity as this measure is the recommended distance measure for Ward’s method (Everitt et al., 2001).

In some occasions the outcome of clustering is depended on the order of the clustered data. To check for this so called robustness, the cluster analysis was done twice with differently ordered data.

4.3.3. VERHIB model

Vegetation types are reconstructed by linking the signals of both the vegetation species and the soil samples. Linking the signals is not very straightforward and difficult to do by hand. Therefore, Jansen et al. (2010) developed the statistical model VERHIB (VEgetation Reconstruction with the Help of Inverse modelling and Biomarkers) in MATLAB (Mathworks, 2007). It was designed and successfully applied to determine whether a shift in the upper forest line in Ecuador took place. However, its universal design makes application on different datasets as the one obtained in this research possible.

The VERHIB model consists of a linear regression model to describe the way in which a certain vegetation development over time at a certain location results in accumulation of biomarkers (n-alkanes and n-alcohols) in a suitable archive (the soil samples at various depths). An inversion of the forward model is used to reconstruct paleo-vegetation on the basis of the observed accumulated biomarker signal (Jansen et al., 2010).

Furthermore, there are a number of model parameters that have to be selected in order to run the model. They have been selected in the following ways:

- **Incorporation of leaf and root data.** For every single species it can be selected whether to use the leaf and/or root data. It was chosen to incorporate all the data into the model. For both moss and lichen there were no separate samples taken for leaf and root material. Here, the obtained biomarker data was introduced in VERHIB as being both leaf and root material.

- **Leaf-root ratio.** An expected leaf-root ratio can be chosen for every single species. As this ratio is not determined in this study, it is set as the standard value of 3 for all
species. This implicates that the contribution of leaf-lipids to the soil is considered to be three times as large as the contribution of root-lipids.

- **Depth duration.** For every sampled horizon it can be set how long it took to develop its thickness. As the sampled horizons in this study are not continuous, setting this parameter did not seem very useful. On top of this, testing VERHIB with different values for this did not affect the modelled output at all. Therefore, for every soil sample the depth duration was set to 200 years, to have some kind of realistic value.

- **Leaf depth.** For every species it can be set how much the leaves mix throughout the soil horizon. As the horizons are mostly more than 10 cm apart from each other, it was assumed that no mixing takes place.

- **Root depth.** A similar parameter is incorporated to take into account the mixing of roots with older layers. Grasses, mosses and heather have only very shallow roots and mixing is therefore set as zero. Pinus is known to root deeply, and therefore the parameter was set so that 20% of a horizon’s signal is modelled one soil horizon below it and 10% was modelled two soil horizon below it. Deciduous trees root less deeply than pine trees, hence these values were chosen as 10% and 5%, respectively.

- **Include chem.** It can be chosen which specific compounds to base the model on. As the typical vegetation lipid compounds are considered to be alcohols with even numbered carbon chain-lengths and alkanes with uneven carbon chain-lengths, alcohols C16,18,20,22,24,26,28,30,32,34,36 and alkanes C17,19,21,23,25,27,29,30,31,33,35 were used as model input. Additionally, the model was also run only based on alcohols and alkanes separately, to assess the importance of combining both with vegetation reconstruction in mind.

The VERHIB model appeared to be very sensitive of differences in concentrations between samples (both vegetation and soil samples). As it is desired to base the modelled results on the difference in patterns, rather than on the differences in concentrations, first a normalization of the data was carried out. Per sample the total concentration of biomarkers was determined. Every compound concentration was then divided by this total concentration.
The Matlab code of VERHIB used for this analysis is appended as appendix B. This was thought to be useful as the model might be altered and extended in the future, and possible confusion about the used version of VERHIB is hereby avoided.
5. Results and discussion

5.1. Lipid quantification

Figures 5.1a and 5.1b show the results for the vegetation samples of the GC/MS analysis. These only show the results of the third analysis of the vegetation samples. The first two
analyses did not lead to satisfactory results, but have been included as appendix C. The first analysis resulted in unexpected high concentrations of alkanes with even-numbered carbon chainlengths in most of the vegetation samples. Therefore it was decided to carry out a second identical analysis. This second analysis resulted in significant lower concentrations of alkanes with even-numbered carbon chainlengths. This shows that somewhere during the

Figure 5.1b: Results of the third GC/MS analysis of the vegetation samples
first analysis, the samples must have been contaminated with alkanes not part of the vegetation samples. Unfortunately, it was not possible to determine when exactly the contamination took place. Although the results of the second analysis seemed to be better than those of the first analysis, they were still unsatisfactory. Many of the peaks within the chromatograms were difficult or even impossible to quantify as they were very low compared to the noise within the chromatograms. Therefore a third analysis was carried out. This time the concentration of the vegetation samples was raised from roughly 300mg to 3g vegetation residue per litre solvent. This resulted in more distinguishable peaks, which seem to be relatively accurate. However, due to having to repeat the procedure two times, no time was left to assess the repeatability of the analysis with the increased concentrations. But as the results of the third analysis appeared to be significantly better than those of the first two analyses, modelling and statistical analysis was only based on the results of the third analysis.

Similarly, two GC/MS analyses were carried out of the soil samples. As the concentrations of lipids in soil samples were generally lower than those in vegetation samples, results of the first (300 mg soil residue per litre solvent) analysis showed few peaks, which were often inaccurate because of the high noise to signal ratio. The second analysis (with roughly 3g soil residue per litre solvent), showed much better results. Figure 5.2 shows these results. Appendix C shows the results of the first analysis of the soil samples, additionally.

The concentrations of the lipids were based on the assumption that the response of the GC/MS is the same for the added deuterated standards and the natural compounds in the samples as is explained in the methodology description (4.3.1). However, from the results of the GC/MS analysis it seems that this assumption might not be correct. Figure 5.3 shows the ratio between the measured concentrations of the added deuterated standards d-20 alcohol and d-20 alkane in the different. In theory (with a perfectly clean GC/MS), this ratio should be the same for all samples, as the standards were added from a single solution with deuterated standards. But in practice the ratio had a mean of 1.9, a standard deviation of 1.2 and minimum and maximum values of 0.8 and 5.6, respectively. An explanation for a difference in these ratios could be the gradual pollution of the GC/MS throughout a measuring sequence. However, when looking at the continuous sequence of the soil samples,
Figure 5.2: Results of the second GC/MS analysis of the soil samples
no trend of a gradual change of ratio could be observed. Therefore, basing the concentrations of the lipids on the measured deuterated concentrations, seemed not to be very accurate. However, there was no alternative. For the separate patterns of both alcohols and alkanes of different carbon chainlengths, the absolute concentrations are not important though.

Figure 5.3: Ratio between both added deuterised compounds
5.2. Cluster analysis

Hierarchical clustering of both leaf and root lipid (alcohols and alkane) concentrations resulted in the clusters shown in the dendrogram shown below (figure 5.4). The proximity matrix, resulting from calculating the squared Euclidean distances, which was used to carry out the clustering is shown as table 5.1. Sample order did not influence the cluster analysis, thus the clustering can be considered as robust.

Although the obtained dendrogram gives a quick insight on the similarities of the lipid patterns of the different species, looking in detail at the squared Euclidean distances reveals some additional information.

Of all the sampled species, Betula and Quercus were by far the most similar species. Based on the lipids patterns they could not be distinguished at all. This must be taken into account when analysing the VERHIB modelling results. Assuming that other deciduous trees also have very similar patterns, it seems not to be such a serious omission that no samples of Alnus and Corylus were analyzed. But there is no absolute certainty that this assumption can be made.

Calluna’s lipid patterns were most closely related to Betula and Quercus. Squared Euclidian distances to the other species are relatively big.

Lichen’s patterns were most closely related to those of Quercus, closely followed by Moss and Betula. Moss was most closely related to Deschampsia, Molinia and Lichen. So, although Lichen did not form a cluster with Moss, they are definitely related.

According to squared Euclidian distance, Molina was most closely related to Deschampsia, Betula, Quercus and Corynephorus. This suggests that lipid patterns of grasses are quite similar, but also that deciduous trees have such similar patterns, that makes a clear distinction between grasses and deciduous trees difficult.

This can also be seen when focusing on Corynephorus and Molinia. Corynephorus had
relatively small squared Euclidian distances with Molina, Betula and Quercus. Deschampsia had the smallest squared Euclidian distances with Molinia, Quercus, Moss and Betula.

Table 5.1: Proximity matrix of the vegetation samples, based on leaves and roots signals

<table>
<thead>
<tr>
<th>Case</th>
<th>Lichen</th>
<th>Moss</th>
<th>Calluna</th>
<th>Molinia</th>
<th>Corynephorus</th>
<th>Deschampsia</th>
<th>Pinus</th>
<th>Betula</th>
<th>Quercus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lichen</td>
<td>0.00</td>
<td>2.467</td>
<td>3.515</td>
<td>4.163</td>
<td>4.016</td>
<td>3.702</td>
<td>5.049</td>
<td>2.705</td>
<td>2.229</td>
</tr>
<tr>
<td>Moss</td>
<td>2.467</td>
<td>0.00</td>
<td>3.681</td>
<td>2.014</td>
<td>2.659</td>
<td>1.960</td>
<td>3.901</td>
<td>2.522</td>
<td>2.528</td>
</tr>
<tr>
<td>Calluna</td>
<td>3.515</td>
<td>3.681</td>
<td>0.00</td>
<td>2.279</td>
<td>2.663</td>
<td>3.126</td>
<td>3.472</td>
<td>1.946</td>
<td>0.981</td>
</tr>
<tr>
<td>Molinia</td>
<td>4.163</td>
<td>2.014</td>
<td>2.279</td>
<td>0.00</td>
<td>1.836</td>
<td>1.351</td>
<td>2.459</td>
<td>1.450</td>
<td>1.489</td>
</tr>
<tr>
<td>Corynephorus</td>
<td>4.016</td>
<td>2.659</td>
<td>2.663</td>
<td>1.836</td>
<td>0.00</td>
<td>2.467</td>
<td>2.716</td>
<td>1.939</td>
<td>1.995</td>
</tr>
<tr>
<td>Deschampsia</td>
<td>3.702</td>
<td>1.960</td>
<td>3.126</td>
<td>1.351</td>
<td>2.467</td>
<td>0.00</td>
<td>3.336</td>
<td>2.096</td>
<td>1.959</td>
</tr>
<tr>
<td>Pinus</td>
<td>5.049</td>
<td>3.901</td>
<td>3.472</td>
<td>2.459</td>
<td>2.716</td>
<td>3.336</td>
<td>0.00</td>
<td>2.492</td>
<td>2.563</td>
</tr>
<tr>
<td>Betula</td>
<td>2.705</td>
<td>2.522</td>
<td>0.946</td>
<td>1.450</td>
<td>1.939</td>
<td>2.096</td>
<td>2.492</td>
<td>0.00</td>
<td>0.084</td>
</tr>
<tr>
<td>Quercus</td>
<td>2.229</td>
<td>2.528</td>
<td>0.981</td>
<td>1.489</td>
<td>1.995</td>
<td>2.563</td>
<td>0.084</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Figure 5.4: Dendrogram, resulting from the cluster analysis of all the data
Pinus had the biggest average squared Euclidian distance to the other species. It had the closest relation with Molinia, Betula and Quercus.

In general, the groups of vegetation that the different species belong to, were reflected in the cluster analysis. However the relationships between the species belonging to the different vegetation groups were not very strong. The uniqueness of the lipid patterns of the researched species seems not to be ideal. Distinctions will be hard to make with great certainty.
5.3. VERHIB model

The output of the VERHIB model, based on alcohols with even-numbered carbon chainlengths and alkanes with uneven-numbered carbon chainlengths of both leaves and root data is shown in figure 5.5. It should be noted that VERHIB will always produce the best fit of the lipid patterns of the included vegetation species to the lipid patterns of the soil samples, even if the input data is flawed or important vegetation species are missing.

![VERHIB output](image)

**Figure 5.5: visualization of VERHIB model output, based on all data**

In the following section an attempt has been made to explain the modelled output by the lipid patterns of both vegetation and soil samples. Values of less than 10% of modelled vegetation coverage are not discussed here, as they are difficult to interpret and statistically weak.
5.3.1. Output analysis

Soil sample 1 (0-2 cm) was modelled to consist almost completely (95%) of lipids produced by Lichen. When the patterns of soil sample 1 are compared with the patterns of the Lichen sample it can be seen why this is. These patterns look very similar, both for alkanes and alcohols. Seemingly, lichens consist of quite some extractable lipids, which made up the majority of the lipids that were found in the upmost soil horizon. The photos of the top of the Defensiedijk profile (appendix A, e.g. photo 1) show that the soil profile is mostly covered by lichens (or mosses, they are hard to distinguish on photos), and thus confirms the outcome of the model.

The second soil sample (34-36 cm) consisted according to the model output of 49% Molinia, 21% Pinus and 20% Quercus. The high peak of alcohol C28 is the main reason why this sample was modelled to consist of high amounts of Molinia. Quercus can be explained by additional lower peaks of alcohols C22, C24 and C26. The 20% modelled Pinus is hard to interpret; most characteristic peaks of Pinus were also present in the soil sample, but its very characteristic high peaks of alcohols C22 and C24 were only present in low amounts in the soil sample.

Soil sample 3 (63-64 cm) was build up by 31% Corynephorus, 25% Quercus, 22% Lichen and 19% Pinus according to VERHIB. The high peak of C26 was responsible for the 31% of modelled Corynephorus. All major lipid compounds of Quercus, Lichen and Pinus were also present in the sample, which explains their presence in the output model. Although, the low alcohol peaks of C22 and C24 makes the presence of Pinus in this sample also questionable.

VERHIB’s output modelled soil sample 4 (77-78 cm) to exist mainly (78%) of Calluna and some Pinus and Lichen (both 11%). Calluna is characterised by its large peaks of alkanes C31 and C33. These peaks were also present in this sample, which explains its high modelled presence. The main lipid compounds of Lichen and Pinus were also present in the soil sample, but it is impossible to further interpret these low amounts of modelled coverage.
Soil sample 5 (115-115.5 cm) consisted according to the model output of 44% Pinus and 32% Quercus. Pinus can be explained by the relative high peaks of alcohols C22 and C24 that are not present in the other species. The peaks of alkanes C25, C27, C29 and C31 were responsible for the modelled Quercus.

The modelled output suggests that soil sample 6 (130-132 cm) consisted of Calluna (56%) and Pinus (44%). This can be explained by the presence of their characteristic peaks (alkanes C31 and C33 for Calluna, and alcohols C22 and C24 for Pinus).

Soil sample 7 (132-134 cm), which was sampled directly below soil sample 6 also had a similar composition according to the modelled output; 42% of Pinus and 35% of Calluna.

Soil sample 8 (151-154 cm) had a high peak of alcohol C26. As this is characteristic for Corynephorus, its modelled coverage of 67% can be explained by this. The 28% of Pinus must have been modelled because the alcohols C22 en C24 were present. However, these peaks were again low and might not be able to explain for these amounts of Pinus.

Finally, soil sample 9 (168-172 cm) consisted mainly of Molinia (70%) and some Pinus (21%). The single high peak of alcohol C28 explains Molinia. The Pinus signal could also be seen here, but similar problems as with soil samples 2, 3 and 8 make this uncertain.

The VERHIB model suggests that Pinus played a big role in the landscape development. However, pollen diagrams show that Pinus was only present in recent history. The first explanation that came to mind is that through the deep rooting of Pinus, more recent biomass of Pinus roots contaminates the biomarker signal of the older soil horizons. But when looking at the biomarker patterns of root and leaf biomass, it can be seen that both are very different and that the root signal is very weak. It also showed that its leaf pattern is very broad with presence of alcohols with chain lengths of 22, 24, 26, 28, 30 and 32. Most soil horizons also contained these alcohols, which caused VERHIB to model Pinus as the main vegetation in these horizons. Most likely there was another source besides Pinus that was not taken into account in this study, which was responsible for the presence of alcohols with chain lengths of 22 to 32.
5.3.2. Alternative runs

Of main interest for the research was to run the model with all available data. However, running the model with only part of the data gave a lot of insight in the distinguishness of the lipid patterns and the related strength of the previously discussed model output. Figure 5.6 shows the output of VERHIB when only alcohols (left) or only alkanes (right) were used as model input. The results were very different from the combined model output (figure 5.4). For example Lichen was completely absent in the surface layer in the model output based solely on the alcohol patterns, while this was the main constituent in the combined model. The most likely explanation for the difference in model output is that the patterns of both alcohols and alkanes of the tested vegetation species on their own are not distinct enough to reconstruct the paleo-vegetation based on extractable lipids within the sampled soil profile, although the simple fact that only half of the data was used when running the model for only one component class, is also likely to impair the quality of the output. It also cannot be ruled out completely that the lipid signals of one component class (either alcohols or alkanes) explained the species distribution very well, while the other component class performed very poorly. In this way the combined model would give less accurate output as the one for the well performing component class.

Another noticeable thing is that a lot more Pinus was modelled when only taking into account alcohols or alkanes. It seems that Pinus has the most general, broad pattern for both alcohols and alkanes, which caused VERHIB to model the remaining part of the pattern (the
part that cannot be explained by the lipid patterns of other species) as Pinus. This explains the presence of Pinus in the combined model output, when its pattern cannot easily be distinguished. It would be very interesting to see if adding an extra component class (e.g. fatty acids) results in smaller amounts of Pinus in the output model, as it seems to be important to use as much (relevant) data as possible as model input in order to obtain the most accurate model output.

Figure 5.7 shows VERHIB’s output when only using either leaf lipids (left) or root lipids (right) are used as model input. Again the results here are quite different from the results obtained with the combined model. Especially when only using only root data as input, the results are almost incomparable with the combined model. This is partly due to the higher importance that is given to the leaf patterns compared to the root patterns (3:1), but also because the leaf lipid patterns are more distinctive than the root lipid patterns.

![Figure 5.7: VERHIB model output, based on leaf data (left) and root data (right)](image-url)
5.4. Relation to data from previous studies

When the results from the VERHIB model are compared to the results from the previously carried out pollen analysis (see figure 3.2), the following observations can be made:

The relative abundance of deciduous trees (Alnus and Corylus) pollen found in the lower part of the profile was not seen in the output of the VERHIB model. Instead, the VERHIB model output suggests that the in situ vegetation consisted mainly of grasses. One reason for this could be the absence of both Alnus and Corylus as vegetation samples. However, it is expected that Betula and Quercus will have more similar lipid patterns to Alnus and Corylus in comparison to grasses (Molinia, Corylus and Deschampsia). But the cluster analysis also showed that grasses and deciduous trees have relatively similar lipid patterns and are therefore hard to distinguish. This shows that that the use of lipid patterns (of alkanes and alcohols) as a single proxy does not always explain the vegetation history correctly.

The two periods of high abundance of Ericaceae (Calluna) pollen in the (lower) middle of the profile were very well reflected in the output of VERHIB; it models mainly Calluna in the soil samples at depths 77-78 cm, 130-132 cm and 132-34 cm, which correspond with the peaks in heather in the pollen diagrams.

The replacement of pollen of heather by pollen of grasses after the removal of heather in the upper (middle) part of the Defensiedijk profile is also visible in the output of VERHIB (34-36 cm, 63-64 cm).

The big peak of Pinus pollen in the top of the pollen diagram could not be seen in the model output. Possibly this is caused by the very strong signal of Lichen that might blend the signal of Pinus.
6. Conclusions

Application of lipid biomarkers as an added proxy in the reconstruction of landscape development in cultural sandy landscapes appeared to be a promising methodology. This research showed that the two most important conditions on which the lipid biomarker approach depends are met. Firstly, both the signal of the n-alkanes and the signal of the n-alcohols were found throughout the whole soil profile in somewhat similar concentrations. This implicates that the decomposition of these (lipid) plant compounds is slow enough to draw conclusions on the plant composition of organic rich soil layers up to at least a few thousand years. Secondly, combining the ratios between the carbon chain length of both the n-alkanes and the n-alcohols of the different investigated species, it is found to be possible to distinguish the group of vegetation (grasses, heather and trees). Differentiation on a species level appeared much more difficult and is only feasible for a few of the studied vegetation species.

Generally, the results of the biomarker analysis support the alleged history of landscape development, based on previous (palynological) studies, of the surroundings of the studied polycyclic soil profile. In the cases where the results of the biomarker analysis point in different directions than based on the results of previous research it is most likely not because the landscape history is different than thought before, but because of the lack of quantity and quality of obtained data used in the biomarker analysis.

To make sure the results of this study are valid and reliable, additional research has to be carried out. The trustworthiness of the found lipid patterns of both soil and vegetation samples must be improved to draw more solid conclusions.
References


Jansen, B., Nierop, K.G.J., Kotte, M.C., de Voogt, P., Verstraten, J.M. (2006a) The applicability of accelerated solvent extraction (ASE) to extract lipid biomarkers from soils. Applied Geochemistry, 21, pp. 1006-1015


APPENDICES

Appendix A: Photos of Defensiedijk 2008 profile

Figure A.1: Photo of the top (0-39 cm) of the studied Defensiedijk 2008 profile.
Figure A.2: Photo of the part of the Defensiesijk 2008 profile between a depth of 29 and 82 cm. The irregularity of the soil horizons is clearly visible.
Figure A.3: Between a depth of 67 and 99 cm

Figure A.4: Between a depth of 124 and 151 cm
Appendix B: Matlab code

B.1. verhib.m

```Matlab
function [plantmass,soilchem,pred,soilchem,rootdepth,soilgroup] = verhib(varargin)

% [plantmass,soilchem,pred,soilchem,rootdepth,soilgroup] = ...
% verhib(inputxls,regdepth,reggroup)
% or
% [plantmass,soilchem,pred,soilchem,rootdepth,soilgroup] = ...
% verhib(soilchem,leafchem,rootchem,leafdepth,...
% plantgroups,leafmass,rootmass,leafgroup,rootgroup);
% if logit is not found, add path to optimization toolbox

% [par,pred,soilchem,rootchem,leafdepth,rootdepth,predid,ylpg,ylg,soilchem,predid]
% = verhib('verhib_g15',0.1,0.1)
if isstr(varargin{1})
    inputxls = varargin{1};
    regdepth = varargin{2};
    reggroup = varargin{3};
    [soilchem,leafchem,rootchem,leafdepth,rootdepth,plantgroups,soilchem,leafchem,rootchem,leafdepth,rootdepth,soilchem,predid,ylpg,ylg,soilchem,predid] = ...
        verhib(inputxls,regdepth,reggroup);
else
    soilchem = varargin{1};
    leafchem = varargin{2};
    rootchem = varargin{3};
    leafdepth = varargin{4};
    rootdepth = varargin{5};
    plantgroups = varargin{6};
    leafchem = varargin{7};
    rootchem = varargin{8};
    rootchem = varargin{9};
    leafchem = varargin{10};
    rootchem = varargin{11};
end

% build matrices for regression function
[E,f,G,h,C,d,minmass,maxmass] = mk_verhib_matrices(soilchem,leafchem,rootchem,leafdepth,rootdepth,plantgroups,soilchem,leafchem,rootchem,leafdepth,rootdepth,soilchem,predid,ylpg,ylg,soilchem,predid);
%
% the actual regression
% set tolerance and maximum number of iterations to reasonable levels for this problem
% options = optimset('TolFun',1e-4,'MaxIter',1000);
% [par,predid] = lsqlin(E,f,G,h,C,d,minmass,maxmass,maxmass/length(plantmass),options);
%%
% ref output
[plantmass,soilchem,pred,soilchem,rootdepth,soilgroup] = mk_verhib_output(length(plantmass),species,chemid,soilchem,predid);
%
% write output
if xlsInOut
    write_verhib_xls(inputxls,regdepth,reggroup,depthlbl_num,species,chemid,soilchem,predid);
else
    disp('Output not written to xls file');
end
```

B.2. read_verhib_xls.m

```Matlab
function [soilchem,leafchem,rootchem,soildepth,rootdepth,soilchem,leafchem,rootchem,leafdepth,rootdepth,soilchem,predid,ylpg,ylg,soilchem,predid] = read_verhib_xls(inputxls)

%% load data from 'inputxls'
[dat,lbl] = xlsread(inputxls,'inclu dec hem');
endcol = size(dat,2);
inchem = logical([0 dat]);
chemid = lbl(1,inchem);
[dat,lbl] = xlsread(inputxls,'plant groups');
useleaf = logical(dat(:,6));
useroot = logical(dat(:,7));
```

Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

B.3. mk_verhib_matrices.m

function [f,G,h,C,d] = mk_verhib_matrices(soilchem,leafchem,rootchem,leafdepth,rootdepth,plantgroups,usleaf,usroot,depthlimit,var,regdepth,reggroups,deptlimit)

% where
% soilchem  is the matrix with soil chemical data (nrcomp,nrlayer)
% leafchem  is the matrix with leaf chemical data (nrcomp,nrplant)
% rootchem  is the matrix with root chemical data (nrcomp,nrplant)
% leafdepth  is the matrix with leaf depth relation (nrlayers,nrplant)
% rootdepth  is the matrix with root depth relation (nrlayers,nrplant)
% plantgroups is the matrix with plant groups (nrgroups,nrplants)
% reggroups is the regularization parameter for plant groups
% regdepth  is the regularization parameter for root groups
% deptlimit specifies the number of layers to consider for spatial partitioning
% var is the number of variables
% and 0,1,0 specifies low,medium,high regularization
% rootgroups is the regularization parameter for root groups
% leafgroups is the regularization parameter for plant groups
% leafdepth specifies the number of layers to consider for spatial partitioning
% soilgroups is the regularization parameter for soil groups
% soilcomp specifies the number of components
% nrlayer specifies the number of layers
% nrplant specifies the number of plants
% plantmass is the plant mass
%
% output matrices and vectors are for the least squares problem:
% \[ (E \cdot f = C \cdot d) \] subject to \[ G \cdot x \geq h \] and \[ C \cdot x = d \]
% where
% E is the constraint matrix
% f is the vector of unknowns
% C is the constraint matrix
% d is the vector of constraints
% G is the inequality constraint matrix
% h is the vector of inequality constraints
% x is the vector of unknowns
%
% input matrices and vectors are for the least squares problem:
% \[ (E \cdot f = C \cdot d) \] subject to \[ G \cdot x \geq h \] and \[ C \cdot x = d \]
% where
% E is the constraint matrix
% f is the vector of unknowns
% C is the constraint matrix
% d is the vector of constraints
% G is the inequality constraint matrix
% h is the vector of inequality constraints
% x is the vector of unknowns
%
% output matrices and vectors are for the least squares problem:
% \[ (E \cdot f = C \cdot d) \] subject to \[ G \cdot x \geq h \] and \[ C \cdot x = d \]
% where
% E is the constraint matrix
% f is the vector of unknowns
% C is the constraint matrix
% d is the vector of constraints
% G is the inequality constraint matrix
% h is the vector of inequality constraints
% x is the vector of unknowns
%
% input matrices and vectors are for the least squares problem:
% \[ (E \cdot f = C \cdot d) \] subject to \[ G \cdot x \geq h \] and \[ C \cdot x = d \]
% where
% E is the constraint matrix
% f is the vector of unknowns
% C is the constraint matrix
% d is the vector of constraints
% G is the inequality constraint matrix
% h is the vector of inequality constraints
% x is the vector of unknowns
%
% determine the dimension of the problem
[nrcomp,nrlayers] = size(soilchem);

% number of plant species
[pngroups] = numgroups(plantgroups);

% output = landscape & biomarker
[x, y] = output(nocomp, nrlayers);

% set values leafroot, nrods to zero if the respective plant
% component is not included

% leafroot = leafroot.* sprouts(allincomp,plant); nrods = nrods.* sprouts(allincomp,plant);
if nrods(nrods>0), nrods = [0 0 0 0];
end

% if not empty, leafroot = leafroot.* sprouts(allincomp,plant);
% nrods = nrods.* sprouts(allincomp,plant);

% build the vector
ytime = zeros(2*nrlayers*nrods,1);

if isempty(pngroups)

groups = [];
else

groups = pngroups.* sprouts(allincomp,plant); nrods = nrods.* sprouts(allincomp,plant);
end

f = [solchem, ytime, groups];

% build the matrix

% the blocks for the biomarker mass in the plant parts

leafblocks = sprouts(leafchem, nrlayers, nrlayers);* spacetimeknitting(leafdepth./100, nrcomp, depthlimit);

rootblocks = sprouts(rootchem, nrlayers, nrlayers);* spacetimeknitting(rootdepth./100, nrcomp, depthlimit);

totalmass = [leafblocks rootblocks];

% the blocks to control transference distribution

if isempty(pngroups)

groupsblocks = []; end

E = [totalmass regdepth/*totalmass; groupsblocks];

regf = regdepth; regf is imposed on the system.

% the block for the biomarker mass in the plant groups

leafblocks = sprouts(leafchem, nrlayers, nrlayers);* spacetimeknitting(leafdepth./100, nrcomp, depthlimit);

rootblocks = sprouts(rootchem, nrlayers, nrlayers);* spacetimeknitting(rootdepth./100, nrcomp, depthlimit);

regdepth = [leafblocks rootblocks];

% the blocks to control transference distribution

if isempty(pngroups)

groupsblocks = []; end

E = [regf is imposed on the system; regdepth/*regf; groupsblocks];

regf is imposed on the system.
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

B.4. mk_verhib_output.m

function [partitionmatrix] = spacetime_knitting(partition_in,nrcomps,nrlag)
    \[ \text{partitionmatrix} = \text{blockdiag_lim(input,nrblk)} \]
end

function [output] = blockdiag_lim(input,nrblk)
    \[ \text{output} = \text{bspace}(	ext{input},
\]
end

% B.4. mk_verhib_output.m

function [planmass,soilchem_pred,re sidchem,re siddept h,re sidgr oup] = mk_verhib_output(nrcomps,nrlayer s,nrpla nts,nrgr oups,X ,y ,par,leafrootratio)
% totale lengte van y is:
%   nrcomps* nrlayers + nrplants*(nrlayers
1) + nrgroups* nrlayers
% residual = y
- X*par;
if length(par)==nrplants*nrlayers
% only leaves were considered - no roots
    totmass = par + par ./ repmat(leafrootratio,nrlayers,1);
    plantmass = reshape(totmass, nrplants, nrlayers, nrcomps);
    soilchem_pred = reshape(y(1:nrcomps* nrlayers) ,nrcomps,nrlayers);
    residchem = reshape(y(nrcomps* nrlayers+1:nrcomps* nrlayers+nrplants* nrlayers), nrcomps,nrlayers);
else % both leaves and roots were considered.
    % calculate the parameter output and calculate the relative contribution
    plantmass = reshape(y(1 : nrplants*nrlayers ) ,nrplants, nrlayers, nrcomps);
    soilchem pred = reshape(y(nrcomps* nrlayers+1:nrcomps* nrlayers+nrplants* nrlayers), nrcomps,nrlayers);
    residchem = reshape(y(nrcomps* nrlayers+1:nrcomps* nrlayers+nrplants* nrlayers), nrcomps,nrlayers);
end

end
MSc. thesis Earth Science A.J. de Vreng
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

```matlab
mldpsh = reshape( (brk + 1 : brk + nr_plants*nr_layers) , nr_plants, nr_layers);
rl = reshape( (brk + 1 : brk + nr_plants*nr_layers) , nr_plants, nr_layers);
mlgrop = reshape( (brk + 1 : brk + nr_groups*nr_layers) , nr_groups, nr_layers);
```

**B.5. write_verhib_xls**

```matlab
function write_verhib_xls( inputxls, regdepth, reggroup, depthlbl_num, species, chemlbl, plantmass, residchem)

warning("off");
% write the output (vegetation composition) to the source spreadsheet
% the regularisation parameters are used in the sheet name
outputsheet = [vegetation_depth_regchem_reggroup; num2str(regdepth), num2str(reggroup)];
xlswrite(inputxls, ["Dept h ( cm):"], outputsheet, B1);
xlswrite(inputxls, ["Spe cies name :"], outputsheet, A2);
xlswrite(inputxls, ["Spe cies"], outputsheet, A3);
xlswrite(inputxls, ["Plantmass"], outputsheet, B3);

% write the residuals to the source spreadsheet
% the regularisation parameters are used in the sheet name
outputsheet = [residual_depth_regchem_reggroup; num2str(regdepth), num2str(reggroup)];
xlswrite(inputxls, ["Dept h ( cm):"], outputsheet, A1);
xlswrite(inputxls, ["Dept hlbl num:"], outputsheet, A2);
xlswrite(inputxls, ["Residchem"], outputsheet, B2);
warning("on");
```

Appendix C: GC/MS results

Soil sample 1 (0-2 cm)

First analysis

Second analysis

Soil sample 1 (0-2 cm)

alcohols
alkanes

lipid concentration (μg/g soil)

Carbon chainlength
Soil sample 2 (34-36 cm)

First analysis

Second analysis

lipid concentration (μg/g soil)

Carbon chainlength
Soil sample 3 (63-64 cm)

First analysis

Second analysis

Carbon chainlength

lipid concentration (μg/g soil)

Soil sample 3 (63-64 cm)

alcohols

alkanes
Soil sample 4 (77-78 cm)

First analysis

Second analysis

[Graph showing lipid concentration (μg/g soil) vs. carbon chain length for alcohols and alkanes for two different analyses.]
Soil sample 5 (115-115.5 cm)

First analysis

Second analysis

Carbon chainlength

lipid concentration (μg/g soil)

alcohols

alkanes
Soil sample 6 (130-132 cm)

First analysis

Second analysis

Carbon chainlength
Soil sample 7 (132-134 cm)

First analysis

Second analysis

Carbon chainlength

lipid concentration (μg/g soil)

alcohols

alkanes
Soil sample 8 (151-154 cm)

First analysis

Second analysis

Carbon chainlength
Soil sample 9 (168-172 cm)

First analysis

- Alcohol concentration
- Alkanes concentration

Second analysis

- Alcohol concentration
- Alkanes concentration
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

---

**Lichen**

First analysis

[Graph showing lipid concentration (μg/g plant) vs. carbon chainlength (16-36)]

Second analysis

[Graph showing lipid concentration (μg/g plant) vs. carbon chainlength (16-36)]

Third analysis

[Graph showing lipid concentration (μg/g plant) vs. carbon chainlength (16-36)]
Moss
First analysis

Second analysis

Third analysis

Carbon chainlength
Calluna: leaves
First analysis

Second analysis

Third analysis

Carbon chainlength (μg/g plant)
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

**Calluna: roots**

**First analysis**

- **Carbon chainlength**
- **Lipid concentration (μg/g plant)**

**Second analysis**

**Third analysis**

**Carbon chainlength**
Molinia: leaves

First analysis

Second analysis

Third analysis

Carbon chainlength

lipid concentration (μg/g plant)
Molinia: roots
First analysis

Second analysis

Third analysis
Corynephorus: leaves

First analysis

Second analysis

Third analysis

Carbon chainlength

lipid concentration (μg/g plant)

alcohols

alkanes
Corynephorus: roots

First analysis

- alcohols
- alkanes

Carbon chainlength

Second analysis

Lipid concentration (μg/g plant)

Third analysis

Carbon chainlength
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands
**Deschampsia: roots**

*First analysis*

*Second analysis*

*Third analysis*

**Carbon chainlength**
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands.
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

**Pinus: roots**

First analysis

- Alcohol concentration (μg/g plant)
- Alkane concentration (μg/g plant)

Second analysis

- Alcohol concentration (μg/g plant)
- Alkane concentration (μg/g plant)

Third analysis

- Alcohol concentration (μg/g plant)
- Alkane concentration (μg/g plant)
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

Betula: leaves
First analysis

Second analysis

Third analysis
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands

Betula: roots
First analysis

Second analysis

Third analysis

Carbon chainlength

l lipid concentration (μg/g plant)
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands
Application of straight-chain lipids as biomarkers in the reconstruction of landscape development in a driftsand landscape in the Netherlands.