



## UvA-DARE (Digital Academic Repository)

### Evaluating the applicability of the VERHIB model to a 2600-year peat sequence from Central Germany

Thomas, C.L.; Jansen, B.; van Loon, E.E.; Wiesenberg, G.L.B.

**DOI**

[10.1016/j.apgeochem.2025.106338](https://doi.org/10.1016/j.apgeochem.2025.106338)

**Publication date**

2025

**Document Version**

Final published version

**Published in**

Applied Geochemistry

**License**

CC BY

[Link to publication](#)

**Citation for published version (APA):**

Thomas, C. L., Jansen, B., van Loon, E. E., & Wiesenberg, G. L. B. (2025). Evaluating the applicability of the VERHIB model to a 2600-year peat sequence from Central Germany. *Applied Geochemistry*, 184, Article 106338. <https://doi.org/10.1016/j.apgeochem.2025.106338>

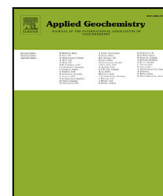
**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*



## Evaluating the applicability of the VERHIB model to a 2600-year peat sequence from Central Germany

Carrie L. Thomas <sup>a,b,c</sup>, Boris Jansen <sup>b</sup>, E. Emiel van Loon <sup>b</sup>, Guido L.B. Wiersberg <sup>a</sup>

<sup>a</sup> Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

<sup>b</sup> Institute of Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands

<sup>c</sup> Department of Ecology, Radboud University, Heyendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands

### ARTICLE INFO

Editor - Joyanto Routh

#### Keywords:

Peatland  
Lipid biomarkers  
Vegetation reconstruction  
Source apportionment  
VERHIB model

### ABSTRACT

Plant-derived biomarkers have been used extensively for source identification of organic matter in soil, peat, and sediments. These are chemical components, primarily of leaf waxes, that are naturally more resistant to degradation than most organic molecules and can be preserved in depositional archives for hundreds to thousands of years. However, using biomarkers for vegetation reconstruction can be complicated because individual compounds or compound classes do not necessarily indicate specific plant sources. The composition across multiple compounds must be holistically evaluated to develop accurate reconstructions. The VERHIB model was developed to enable a range of *n*-alkane and *n*-alkanol data to be evaluated simultaneously and estimate past vegetation development using the biomarker signature of leaves and roots from present-day specimens of the plant species of interest compared to the preserved, mixed signature of a soil or peat core. In this study, we have applied the VERHIB model to data gathered from the Beerberg peatland in Central Germany. A previous study characterized plant macrofossils, pollen, *n*-alkane, *n*-alkanol, and *n*-fatty acid composition in the Beerberg peat. We have collected modern plant samples from the peatland and measured their biomarker composition to use as input for the model. Despite many overlapping biomarker signatures across plant species and parts, our results show that the model could recreate a reasonable vegetation development pattern for most of the peat core if *n*-fatty acid data were included alongside *n*-alkane data. The model had difficulty recreating the transition from poor fen vegetation to *Sphagnum* bog, which was evident in the plant macrofossil records, so further calibration is needed. This was the first attempt at considering *n*-fatty acid data in a reconstruction using the VERHIB model; previous reconstructions only included *n*-alkanes or a combination of *n*-alkanes and *n*-alkanols. Our study shows that *n*-fatty acids are a valuable compound class to add to the VERHIB model and provides recommendations for future development.

### 1. Introduction

Soil organic matter (SOM) is a key sink in the global carbon cycle, storing about four times the amount of carbon (1700 PgC) as the entire stock of terrestrial vegetation (450 PgC) (Erb et al., 2018; Jackson et al., 2017). To combat climate change, it is imperative to protect and optimize this sink to reduce the amount of carbon in the atmosphere. Pursuant to that goal, it is essential to gain a deeper understanding of how SOM is formed and subsequently cycles (Schmidt et al., 2011). Better identification of the sources of SOM could aid in improved process understanding and potential regulation of these processes to enable increased carbon sequestration. As the remains of past vegetation are the primary contributor of organic matter to soil, reconstructing vegetation development through time is one way to identify sources of SOM. Furthermore, vegetation reconstruction provides insight into

past climate patterns and can be used to improve climate models and predict how ecosystems will respond to climate change (Lavorel et al., 2007). This is especially true in the case of vegetation reconstructions that do not include climate data as input or boundary conditions.

One tool used extensively for source identification of OM in lacustrine and marine sediments, as well as soil and peat archives, is “molecular fossils” or biomarkers (Peters et al., 2005). These are chemical components of previous life forms that are preferentially preserved following deposition due to their relatively stable molecular compositions (Eglinton and Logan, 1991). Many biomarkers have been connected to specific sources, particularly those derived from plant waxes including long-chain *n*-alkanes, *n*-alkanols, and *n*-fatty acids (Diefendorf and Freimuth, 2017). Of these compounds, *n*-alkanes are the most well-studied, though *n*-alkanols and *n*-fatty acids are

\* Corresponding author at: Department of Ecology, Radboud University, Heyendaalseweg 135, 6525 AJ, Nijmegen, The Netherlands.  
E-mail address: [carrie.thomas@ru.nl](mailto:carrie.thomas@ru.nl) (C.L. Thomas).

increasingly included in biomarker analyses (Diefendorf and Freimuth, 2017; Jansen and Wiesenberg, 2017). Generally, ranges of certain chain lengths can be attributed to different plant life forms: short-chain homologues (e.g., C<sub>17</sub>–C<sub>21</sub> *n*-alkanes) are characteristic of aquatic algae (Cranwell et al., 1987) while mid-chain homologues (e.g., C<sub>21</sub>–C<sub>25</sub> *n*-alkanes) derive from submerged and floating aquatic macrophytes (Barnes and Barnes, 1978; Cranwell, 1984; Ficken et al., 2000), and long-chain homologues (e.g., C<sub>27</sub>–C<sub>33</sub> *n*-alkanes) characterize terrestrial vegetation (Eglinton and Hamilton, 1967). These plant-derived biomarkers have been found to be well-preserved over time in a range of sediment types and even withstand transportation processes such as water or wind erosion (Castañeda and Schouten, 2011; Bush and McInerney, 2013; Diefendorf and Freimuth, 2017).

Numerous studies have identified specific biomarker distribution patterns that have allowed past vegetation shifts to be determined (e.g., Nott et al., 2000; Jansen et al., 2006; Schellekens and Buurman, 2011). However, single *n*-alkane, *n*-alkanol, or *n*-fatty acid compounds are not typically indicative of individual plant sources, and interpreting the data requires holistically considering the biomarkers across a range of chain lengths (Schwark et al., 2002; Inglis et al., 2022). Additionally, these biomarkers are present not only in leaves but also in other plant organs, including roots, which typically have a different chemical composition (e.g., Mueller et al., 2012; Mambelli et al., 2011). This can obscure the process of source apportionment due to the overlap in chemical signatures of aboveground and belowground biomass and the potential for belowground biomass to overprint the signature of previously incorporated aboveground biomass, specifically in deeper soil horizons (Goetze et al., 2010). Further, many potential climatic, environmental, and developmental variables can affect the biomarker composition of plant waxes. Therefore, the biomarker signature measured at one location of a species may be very different from that of the same species in another environment or even in a different growth stage (Jansen and Wiesenberg, 2017). Recent studies highlight the continuous incorporation of plant wax components into soils, as opposed to only from litterfall, which points to the need to further disentangle the incorporation pathways of organic matter into soil (Srivastava and Wiesenberg, 2018).

Despite these uncertainties regarding incorporation and source assessment of SOM, the degradation continuum of wax biomarkers from fresh plant biomass via organic horizons toward mineral soil still highlights the robustness of the approach that vegetation composition is reflected in mineral soils through the biomarker composition, and this consequently enables source apportionment of SOM (Hirave et al., 2020; Thomas et al., 2021). Additionally, due to the preferential degradation of plant remains, including pollen and macrofossils, in oxic environments, the more recalcitrant molecular fossils can provide more reasonable output in contexts where there is a severely degraded or sparse pollen or macrofossil record, such as loess-paleosol sequences (Zech et al., 2012) as well as sediments in desert environments (Knief et al., 2020; Jha et al., 2024b). Indeed, the use of biomarker compositions alongside proxies such as pollen or macrofossils enables more nuanced vegetation reconstructions than paleobotanical data on its own, as shown in our previous work (Thomas et al., 2023b).

The Vegetation Reconstruction with the Help of Inverse modeling and Biomarkers (VERHIB) model was developed to make deciphering biomarker data easier and enable more quantitative interpretations (Jansen et al., 2010). The model was developed using MATLAB and includes *n*-alkanes and *n*-alkanols. In brief, VERHIB is a linear regression model with input parameters that include the biomarker composition of plant leaves and roots as well as the biomarker composition of a sedimentary archive, such as a soil or peat core. An inverse model is used to solve for the proportion of biomarkers in the sedimentary archive that each plant species has contributed, serving as an estimate of the vegetation composition through time (Jansen et al., 2010). Compared

to other modeling approaches used with biomarkers, such as end-member modeling (e.g., Zech et al. (2012)), which rely on end-members having distinctly different average compositions, the VERHIB model allows for the consideration of more plant sources with potentially overlapping biomarker signals. The model has already been applied successfully in a handful of locales, including the Andes using both peat and soil cores (Jansen et al., 2010, 2013) and further investigated in other archives, such as polycyclic driftsand sequences (van Mourik and Jansen, 2013) and plaggic Anthrosols (van Mourik et al., 2016).

Previously, we analyzed a core from the Beerberg peatland in Thuringia, Germany, developing a high-resolution data set with biomarker, pollen, and macrofossil observations and a robust age-depth model (Thomas et al., 2023b). Because this data set contains both high-resolution pollen and macrofossil data as well as a wide range of (equally high-resolution) biomarkers (*n*-alkanes, *n*-alkanols, *n*-fatty acids), it will allow us to evaluate VERHIB. Earlier work with the VERHIB model had lower-resolution biomarker data and only pollen records available for comparison. Furthermore, the VERHIB model has yet to be applied using *n*-fatty acid data. Using the Beerberg data we can compare vegetation reconstructions by VERHIB with those derived from pollen and macrofossils at an equal temporal resolution and also assess the utility of including *n*-fatty acids for this purpose.

Therefore, the objectives of this study were to

- establish a reference library of plant biomarker signatures (*n*-alkanes, *n*-alkanols, *n*-fatty acids) at Beerberg peatland,
- apply the VERHIB model to the biomarker data from the Beerberg peatland, including the *n*-fatty acid composition, and compare it to the previously completed multi-proxy analysis of the Beerberg peatland development, and
- evaluate the sensitivity of the VERHIB result across a range of parameter choices.

## 2. Materials and methods

### 2.1. Study site and sampling

The Beerberg peatland (50°39'32" N, 10°44'36" E, 983 m) is an ombrotrophic peat bog atop the Grosser Beerberg and is part of the Vessertal-Thuringian Forest Biosphere Reserve in central Germany (Fig. 1). Between the period 2010 and 2019, the mean monthly minimum temperature was 3.6 °C, the mean monthly maximum temperature was 11.0 °C, and the mean monthly precipitation was 90.2 mm (extracted from WorldClim 2.1; Fick and Hijmans (2017)). Although there was a short period of drainage in the past, the Beerberg peatland has been under protection for decades and is, therefore, relatively undisturbed (see Thomas et al. (2023b) for more site details).

In October 2019, we took core samples of the peatland as well as vegetation samples. Core samples were taken using two Russian peat corers (5 cm diameter, 0.5 m depth, Eijkelkamp, Giesbeek, the Netherlands; 7 cm diameter, 1 m depth, self-made) and by alternately coring two small hummocks about 20 cm apart to a depth of 340 cm. The cores were aligned by sampling depth and stratigraphy; overlapping sections (90–100, 175–190, and 290–320 cm) were treated as replicates for the elemental and biomarker analyses and the average results were reported. Plant samples (16 species) were taken within a five-meter radius of the coring site, and multiple samples (3–4) of each species were mixed in their respective sample bags to avoid an outside effect of individual sample heterogeneity (Jansen et al., 2006).

Four *Sphagnum* moss species were identified (*S. angustifolium*, *S. magellanicum*, *S. capillifolium*, and *S. fuscum*), along with *Polytrichum strictum*. A few shrubs were found: *Calluna vulgaris*, *Empetrum nigrum*, *Oxycoccus palustris*, *Vaccinium myrtillus*, *Vaccinium uliginosum*, and *Vaccinium vitis-idaea*. Additionally, *Eriophorum vaginatum* was present. The peatland was surrounded by a spruce-dominated wooded area, but on the peat itself, a few smaller trees were growing, including *Pinus sylvestris*, *Picea abies*, *Betula pendula*, and *Betula pubescens*. The overall vegetation assemblage is typical for a temperate raised bog in Europe (Rydin et al., 2013).

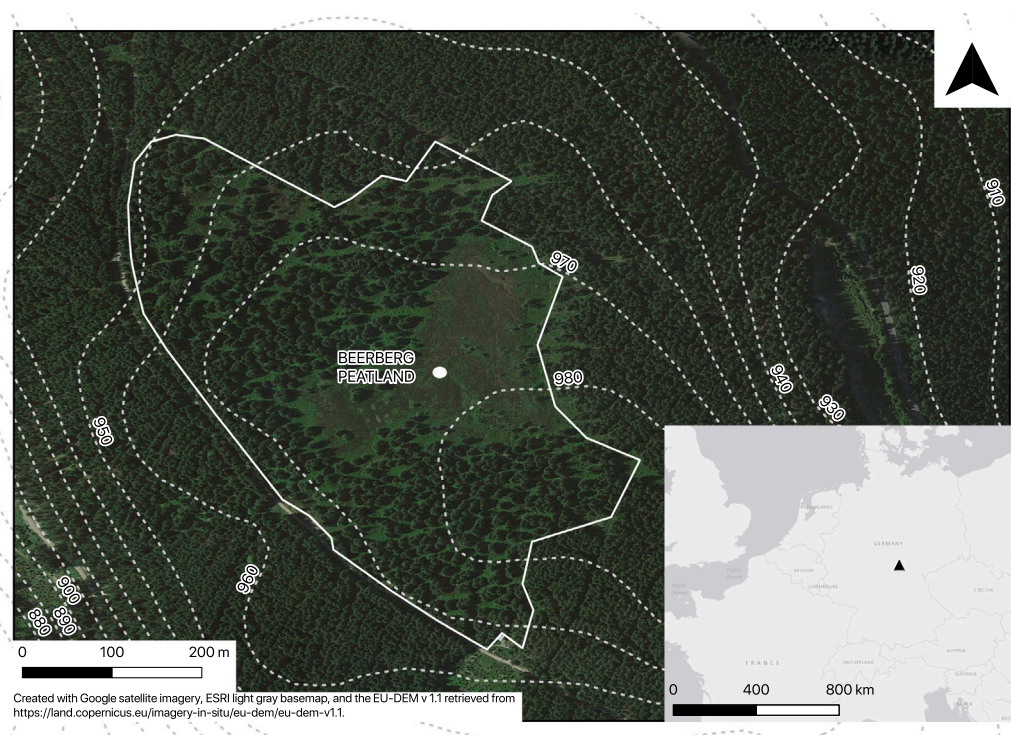


Fig. 1. Map showing sampling spot in Beerberg peatland with contour lines and inset showing location in continental Europe. Source: Reproduced from Thomas et al. (2023b) under the Creative Commons Attribution 4.0 License.

## 2.2. Sample preparation

The plant samples were kept cold with ice packs until they could be stored at  $-20\text{ }^{\circ}\text{C}$ . Prior to freeze-drying, the samples were separated into their constituent parts as well as possible. Separated into leaves, woody aboveground tissue (hereafter referred to as “stems”), and roots were *C. vulgaris*, *V. myrtillus*, *V. uliginosum*, *V. vitis-idaea*, *E. nigrum* (there was not enough root material to analyze for *E. nigrum*), *B. pendula*, and *B. pubescens*. Separated into needles, stems, and roots were *P. sylvestris* and *P. abies*. Separated into aboveground (non-woody) biomass and roots were *O. palustris* and *E. vaginatum*. The five moss species (*P. strictum*, *S. angustifolium*, *S. capillifolium*, *S. fuscum*, and *S. magellanicum*) were measured as whole plant samples. All samples were freeze-dried to a constant weight and then homogenized using a ball mill.

The subsampling procedure of the peat core is described in detail in Thomas et al. (2023b). Briefly, the core was split into 5 cm intervals, except for a few visibly distinct layers at 10–12 cm, 170–172 cm, 270–272 cm, 325–327 cm, 327–328 cm, and 337.5–340 cm. The samples were freeze-dried to a constant weight and homogenized using a ball mill.

## 2.3. Elemental analysis

Carbon and nitrogen concentrations (C, N) and stable C isotope ( $\delta^{13}\text{C}$ ) values of both the freeze-dried and milled bulk peat and plant samples were measured using an Elemental Analyzer-Isotope Ratio Mass Spectrometer (EA-IRMS; Flash 2000 HT-Plus, linked by ConFlo IV to DELTA V Plus IRMS; Thermo Fisher Scientific). Two replicates were analyzed for each sample. Measurement calibration was performed using caffeine (Merck, Germany) and a soil reference material from a Chernozem (Harsum, Germany; see Black Carbon Reference Materials, <https://www.geo.uzh.ch/en/units/2b/Services/BCmaterial/Environmental-matrices.html>).

## 2.4. Biomarker extraction and analysis

The procedure followed for the biomarker extraction and analysis is described in full in Thomas et al. (2023b). Soxhlet extraction was performed on the milled plant samples following Wiesenberg and Gocke (2015). Dichloromethane (DCM):methanol (MeOH) (93:7, v/v) was used as the solvent, and each extraction ran for approximately 30 h. The total lipid extracts were sequentially separated into three fractions using a glass column with silica 60 + 5% potassium hydroxide (KOH), 63–200  $\mu\text{m}$ : a low polarity fraction containing *n*-alkanes and *n*-alkanols eluted with DCM, a fraction containing *n*-fatty acids eluted with DCM:formic acid (99:1, v/v), and a final fraction containing high polarity and high molecular weight compounds eluted with DCM:MeOH (1:1, v/v). The low polarity fraction was further separated into aliphatic, aromatic, and heterocompound fractions using a pasteur pipette containing activated silica gel and the solvents *n*-hexane, *n*-hexane:DCM (1:1, v/v), and DCM:MeOH (93:7,v/v).

Internal standards of deuterated *n*-tetracosane ( $\text{D}_{50}\text{C}_{24}$ , Cambridge Isotope Laboratories, Inc.), deuterated *n*-octadecanol ( $\text{D}_{37}\text{C}_{18}$ , Cambridge Isotope Laboratories, Inc.), and deuterated *n*-eicosanoic acid ( $\text{D}_{39}\text{C}_{20}$ , Cambridge Isotope Laboratories, Inc.) were added to the *n*-alkane, *n*-alkanol, and *n*-fatty acid fractions, respectively. The concentrations of these standards varied but were typically around 0.1 mg/mL. 50  $\mu\text{L}$  (5  $\mu\text{g}$ ) were added to each sample. Before measurement, the *n*-fatty acid fraction was methylated using a boron trifluoride-methanol solution (CAS No. 373-57-9, Sigma-Aldrich), and the *n*-alkanol fraction was silylated using *N,O*-bis(trimethylsilyl)-acetamide (BSA) (CAS No. 10416-59-8, Sigma-Aldrich). The fractions were dried and the *n*-alkane fraction was dissolved in *n*-hexane and the *n*-fatty acid and *n*-alkanol fractions were dissolved in DCM:MeOH (93:7,v/v).

Quantification was performed using a gas chromatograph (GC) (Agilent 7890B) equipped with a multimode inlet and a flame ionization detector (FID). Identification was performed on a GC (Agilent 6890N) equipped with a split-splitless injector in splitless mode coupled to a MS (Agilent 5973). Both instruments had a DB-5MS column (50 m  $\times$  0.2 mm  $\times$  0.33  $\mu\text{m}$ ) and 1.5 m de-activated pre-column and

used helium as the carrier gas (1 ml min<sup>-1</sup>). The GC oven temperature for *n*-alkanes was held at 70 °C for 4 min and increased to 320 °C at a rate of 5 °C min<sup>-1</sup> held for 50 min. For *n*-fatty acids and *n*-alkanols, the temperature was held at 50 °C for 4 min, then increased to 150 °C at a rate of 4 °C min<sup>-1</sup>, and finally increased to 320 °C at a rate of 3 °C min<sup>-1</sup> held for 40 min. The GC-MS was operated in electron ionization mode at 70 eV and scanned from m/z 50–550. Individual compounds were identified by comparison of mass spectra with those of external standards (*n*-alkanes: ASTM D5442, Supelco 502251; ASTM D2887, Supelco 48882; *n*-alkanols: C<sub>20</sub>, CAS No. 629-96-9, Supelco; C<sub>24</sub>, CAS No. 506-51-4, Dr. Ehrenstorfer; C<sub>26</sub>, CAS No. 506-52-5, Tokyo Chemical Industry; *n*-fatty acids: C<sub>20</sub>, CAS No. 112-85-6, Supelco; C<sub>28</sub>, CAS No. 506-48-9, Sigma-Aldrich; C<sub>30</sub>, CAS No. 506-50-3, Sigma-Aldrich) and from the NIST and Wiley mass spectra library.

## 2.5. Data analysis

### 2.5.1. Biomarker measurements

All biomarker measurements were calculated in units of microgram per gram dry sample weight. Several equations have been developed during past research to better evaluate biomarker concentrations and their indications regarding vegetation as well as climate and environmental conditions and potential degradation following deposition. One such index measurement is the Carbon Preference Index (CPI); for *n*-alkanes, this measures the proportion of odd chain lengths to even chain lengths, and for *n*-alkanols and *n*-fatty acids, this measures the proportion of even chain lengths to odd chain lengths. A CPI greater than 1 indicates that the biomarkers are likely derived from a higher terrestrial plant; CPIs closer to or less than 1 are likely to be degraded organic matter or have a microbial source (Cranwell, 1981). The CPI formula we used for the *n*-alkanes is from Marzi et al. (1993):

$$CPI_{ALK} = \frac{(\sum_{i=n}^m C_{2i+1} + \sum_{i=n+1}^{m+1} C_{2i+1})}{2(\sum_{i=n+1}^{m+1} C_{2i})} \quad (1)$$

Where *n* = the starting chain length divided by 2 and *m* = the ending chain length divided by 2. We used *n* = 10 and *m* = 16 for the *n*-alkanes, meaning that chain lengths between 21 and 33 are included in the calculation. For the *n*-alkanols and *n*-fatty acids, the formula was slightly adjusted to measure the even to odd chain length predominance:

$$CPI_{ALC/FA} = \frac{(\sum_{i=n}^m C_{2i} + \sum_{i=n+1}^{m+1} C_{2i})}{2(\sum_{i=n+1}^{m+1} C_{2i+1})} \quad (2)$$

We used *n* = 10 and *m* = 14 for the *n*-alkanols, so that chain lengths from 20 to 28 are included, and *n* = 10 and *m* = 16 for the *n*-fatty acid, thereby including chain lengths 20 to 32.

The Average Chain Length (ACL) is another index used to help interpret biomarker data. It is the average of the individual chain lengths (Poynter et al., 1989) and has previously been used to make a distinction between vegetation types (e.g., Zhou et al., 2005) and as an indicator for environmental conditions such as drought (e.g., Crausbay et al., 2014; Wüthrich et al., 2017). The formula for the *n*-alkanols and *n*-fatty acids was again slightly adjusted, and the same *n* and *m* values were used as for the CPI equations.

$$ACL_{ALK} = \frac{\sum_{i=n}^m (2i + 1) * C_{2i+1}}{\sum_{i=n}^m C_{2i+1}} \quad (3)$$

$$ACL_{ALC/FA} = \frac{\sum_{i=n}^m (2i) * C_{2i}}{\sum_{i=n}^m C_{2i}} \quad (4)$$

$P_{aq}$  and  $P_{wax}$  are proxies developed to determine different sources of *n*-alkanes.  $P_{aq}$  was initially used to determine the proportion of aquatic macrophytes to terrestrial plant inputs in sediments (Ficken et al., 2000) and subsequently used in peatlands to infer past water levels, with higher values of  $P_{aq}$  implying a higher water table (e.g.,

Andersson et al., 2011; Zheng et al., 2007; Zhou et al., 2005).  $P_{wax}$  is used to determine the input of terrestrial plants (Zheng et al., 2007).

$$P_{aq} = \frac{C_{23} + C_{25}}{C_{23} + C_{25} + C_{29} + C_{31}} \quad (5)$$

$$P_{wax} = \frac{C_{27} + C_{29} + C_{31}}{C_{23} + C_{25} + C_{27} + C_{29} + C_{31}} \quad (6)$$

Other significant indicative ratios between biomarker chain lengths have been identified in previous studies. For *n*-alkanes, these include the use of  $C_{23}/(C_{23}+C_{29})$  (Ronkainen et al., 2013) and  $C_{25}/(C_{25}+C_{29})$  (Vonk and Gustafsson, 2009) for distinguishing contributions of *Sphagnum* from other plants and  $C_{23}/C_{25}$  (Bingham et al., 2010; McClymont et al., 2008) and  $C_{23}/(C_{27}+C_{31})$  (Andersson et al., 2011) as an indicator of a shift in dominant *Sphagnum* species related to changes in water table depth and surface moisture. For *n*-alkanols, shifts in the ratio  $C_{22}/C_{24}$  have been correlated with shifts from forested to non-forested vegetation (Jansen et al., 2008; Zheng et al., 2007) and  $(C_{22}+C_{24})/(C_{26}+C_{28})$  (Zheng et al., 2011) has been correlated with changes in temperature. For *n*-fatty acids, the ratio  $C_{15}/(C_{15}+C_{16})$  has also been related to shifts in temperature (Zheng et al., 2007) and we found in our previous study that  $C_{24}/C_{28}$  could be related to shifts in *Sphagnum* abundance (Thomas et al., 2023b). All ratios were calculated for both the plant and peat core samples.

### 2.5.2. Statistical analysis

Principal component analysis (PCA) was performed using the biomarker concentrations, scaled to unit variance, and the calculated ratios (Eqs. (1) to (6)) to determine which linear combination of these could best separate the plant and peat core samples and which contributed the most variance. This was done for all three compound classes together as well as for each of them separately. The PCAs were performed using the R packages “FactoMineR” (Husson et al., 2023) and “factoextra” (Kassambara and Mundt, 2020). The PCAs were used to describe the structure of the biomarker data from the Beerberg samples in addition to the univariate methods.

## 2.6. Modeling with VERHIB

The VERHIB model is fully described in Jansen et al. (2010). Briefly, the forward model is based on the equation:

$$\begin{bmatrix} b_1(d,t) \\ \vdots \\ b_i(d,t) \end{bmatrix} = \begin{bmatrix} l f_1(d) l c_{1,1} & \cdots & l f_j(d) l c_{1,j} & r f_1(d) r c_{1,1} & \cdots & r f_j(d) r c_{1,j} \\ \vdots & & \ddots & \vdots & & \ddots \\ l f_1(d) l c_{i,1} & \cdots & l f_j(d) l c_{i,j} & r f_1(d) r c_{i,1} & \cdots & r f_j(d) r c_{i,j} \end{bmatrix} \begin{bmatrix} l m_1(t) \\ \vdots \\ l m_j(t) \\ r m_1(t) \\ \vdots \\ r m_j(t) \end{bmatrix} \quad (7)$$

in which  $b_i(d,t)$  is the mass of biomarker *i* accumulated at depth *d* of the core during time *t*,  $l c_{i,j}$  is the concentration of biomarker *i* in the leaf of plant *j*,  $r c_{i,j}$  is the concentration of biomarker *i* in the root of plant *j*,  $l f_j(d)$  is the fraction of leaf mass of plant *j* accumulating at depth *d*, and  $r f_j(d)$  is the same for root mass,  $l m_j(t)$  is the leaf mass of plant *j* during time *t* and  $r m_j(t)$  is the root mass of plant *j* during time *t*. The leaf to root mass ratio is a parameter that can be specified for each plant species. Additionally, each plant's leaf and root biomass must be greater than or equal to zero as a constraint.

The inverse model is based on the matrix equation:

$$b = A p \quad (8)$$

Where *b* is a vector of the biomarker composition in each core depth layer, *A* is a matrix containing known constants including the values from Eq. (7), and *p* is a matrix of the parameters to be solved for. In order to reach a unique solution, Tikhonov regularization (Tikhonov

and Arsenin, 1977) is applied and vector  $b$  includes two regularization coefficients that can be used to indicate the importance of smoothness constraints. The equation is solved using a least-squares approach. Jansen et al. (2010) quantitatively tested the robustness of the model using an artificial dataset in which data was omitted and/or Gaussian noise was added and the model was able to reconstruct the original vegetation pattern from these datasets with limited uncertainty.

To evaluate the application of the model to the Beerberg dataset, we calculated the Pearson correlation coefficient between the predicted core biomarker composition ( $y$ ) and the actual values ( $x$ ) using the equation:

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (9)$$

### 3. Results

#### 3.1. Biomarker and chemical composition of modern plants

We report the results of the biomarker measurements from the modern plant samples for  $n$ -alkanes  $C_{21}$ – $C_{33}$ ,  $n$ -alkanols  $C_{20}$ – $C_{28}$ , and  $n$ -fatty acids  $C_{20}$ – $C_{32}$ . The total abundance of the three compound classes, the chain length of their most abundant homologues, and their respective CPI and ACL measurements, as well as the  $P_{aq}$  and  $P_{wax}$  values for the  $n$ -alkanes and the C%, N%, and  $\delta^{13}C$  values, are shown in Table 1 for each plant species and part. The complete dataset of measurements is available at <https://doi.pangaea.de/10.1594/PANGAEA.961142> (Thomas et al., 2023a). Selected parameters are shown in the boxplots in Fig. 2, grouped by plant functional type (Moss, Sedge, Shrubs, and Trees) and plant part (Moss, Leaves/non-woody aboveground biomass, Needles, Stems, Roots).

The average C and N concentration and stable  $^{13}C$  isotope values for the moss samples were 45.4%, 0.7%, and  $-28.9\text{‰}$ , respectively (Table 1). For the shrub leaf and aboveground samples, the averages were 50.5%, 0.9%, and  $-29.7\text{‰}$ . The averages for the shrub stems were 49.2%, 0.5%, and  $-29.0\text{‰}$ . The averages for the shrub roots were 48.0%, 0.5%,  $-28.2\text{‰}$ . There was only one sedge species, *E. vaginatum*, and the values for its aboveground biomass were 46.6%, 1.1%, and  $-26.6\text{‰}$  (Table 1). The values for the roots were 45.6%, 1.0%, and  $-25.8\text{‰}$ . The averages for the tree needles were 49.4%, 1.1%, and  $-28.2\text{‰}$ , and for the tree leaves were 48.8%, 1.3%, and  $-28.2\text{‰}$ . The tree stem averages were 50.3%, 0.8%, and  $-28.6\text{‰}$ ; the tree root averages were 48.2%, 0.9%, and  $-27.3\text{‰}$  (Table 1).

Across all species and plant parts, some general trends could be identified. Largely, the  $n$ -fatty acids were the compound class with the highest absolute concentration (115  $\mu\text{g/g}$  – 3363  $\mu\text{g/g}$ ), usually greater than the  $n$ -alkane (5  $\mu\text{g/g}$  – 2408  $\mu\text{g/g}$ ) and  $n$ -alkanol (5  $\mu\text{g/g}$  – 535  $\mu\text{g/g}$ ) absolute concentrations by an order of magnitude (Table 1, Fig. 2). Between the  $n$ -alkanes and  $n$ -alkanols,  $n$ -alkanes were typically more abundant in leaves/non-woody aboveground biomass (46  $\mu\text{g/g}$  – 2408  $\mu\text{g/g}$ ). In contrast,  $n$ -alkanols were primarily the more abundant compound class in stem and root samples (5  $\mu\text{g/g}$  – 535  $\mu\text{g/g}$ ). Of the  $n$ -alkanes,  $C_{31}$  was most often the most abundant homologue ( $C_{max}$ ) in 21 out of 35 samples, followed by  $C_{29}$  in 8 samples (Table 1). For  $n$ -alkanols,  $C_{28}$  was the most frequent  $C_{max}$  in 11 of 35 samples, closely followed by  $C_{24}$  in 9 samples. For  $n$ -fatty acids,  $C_{24}$  was the most frequent  $C_{max}$  in 14 of 35 samples, followed by  $C_{20}$  and  $C_{28}$ , both occurring as  $C_{max}$  in 6 samples (Table 1).

The  $CPI_{ALK}$  tended to be higher in aboveground biomass samples (3–42.6) (excluding the coniferous needles) than in the roots (5.6–33.5) (Fig. 2). The  $ACL_{ALK}$  also tended to be higher in aboveground biomass samples (26.2–31.3), though tree roots (29.9) had a higher average than both deciduous tree leaves (28.15) and coniferous needles (28.2) (Fig. 2a). The highest average values for both  $C_{23}/C_{25}$  and  $C_{23}/(C_{27}+C_{31})$  were for the coniferous tree needles and the tree

stems, followed by the moss samples. This also held for average  $P_{aq}$  values, while the highest average  $P_{wax}$  value was for shrub leaves, followed by tree roots and shrub stems and roots (Fig. 2a). The highest average values for both  $C_{23}/(C_{23}+C_{29})$  and  $C_{25}/(C_{25}+C_{29})$  were for the tree stems and the moss samples, followed by the coniferous needles. The  $CPI_{ALC}$  average value was highest in tree stems (43.05), deciduous leaves (28.95), and coniferous needles (24.8), while the  $ACL_{ALC}$  average value was highest in deciduous tree leaves (26) and shrub leaves (25.6) (Fig. 2b). The highest average values for the  $n$ -alkanol  $C_{22}/C_{24}$  ratio were in tree roots and stems while the lowest were in the moss samples and *E. vaginatum*. The highest values for the  $n$ -alkanol  $(C_{22}+C_{24})/(C_{26}+C_{28})$  ratio were for tree stems and coniferous needles. The highest average values for  $C_{24}/C_{28}$  were in the moss samples, followed by *E. vaginatum*, then coniferous needles (Fig. 2b). The  $CPI_{FA}$  average value was highest in tree roots (8.95), and shrub stems had the highest  $ACL_{FA}$  average value (26.16) (Fig. 2c). The highest average values for  $C_{15}/(C_{15}+C_{16})$  were in the moss samples. The highest average values for  $C_{24}/C_{28}$  are in the tree stems and roots, followed by the coniferous needles and moss samples (Fig. 2c).

#### 3.2. PCAs of modern plant data

PCAs were performed using the biomarker data: the absolute concentrations of the plant-derived chain lengths and the values calculated for the ratios described in the Methods section. First, all identified biomarkers and ratios were included, with PC1 explaining 25.1% of the variance and PC2 21.6% (Fig. 3a). The top six variables contributing to the first and second PCs were  $P_{aq}$ ,  $P_{wax}$ ,  $C_{25}$   $n$ -alkane,  $C_{26}$  and  $C_{28}$   $n$ -alkanols, and the  $n$ -alkane ratio  $C_{23}/(C_{23}+C_{29})$ . There are distinct clusters of leaves, moss, and roots, while the stem samples are spread out, likely due to the high  $P_{aq}$  values of the tree stem samples. Separate PCAs were also performed for each biomarker group and their respective ratios. The  $n$ -alkane PCA identified a PC1 accounting for 41.4% of the variance and PC2 for 23.1%, with  $C_{25}$   $n$ -alkane,  $P_{wax}$ ,  $P_{aq}$ ,  $ACL_{ALK}$ ,  $C_{23}/(C_{23}+C_{29})$ , and  $C_{25}/(C_{25}+C_{29})$  as the top six contributors (Fig. 3b). While the leaves and moss were separated well by the first two axes, the roots cluster completely overlapped with that of the moss and the stem cluster completely overlapped that of the moss and half of the leaf cluster. In the PCA of  $n$ -alkanol concentrations and ratios, PC1 accounted for 32.7% and PC2 accounted for 30% of the variance with the top six contributors being  $C_{20}$ ,  $C_{22}$ ,  $C_{26}$ , and  $C_{28}$   $n$ -alkanol,  $ACL_{ALK}$ , and the  $(C_{22}+C_{24})/(C_{26}+C_{28})$  ratio (Fig. 3c). The leaves and root clusters were separated, but the roots cluster overlapped with that of the moss and the stems cluster overlapped the other three. The PCA with  $n$ -fatty acids identified PC1 accounting for 38.4% of the variance and PC2 accounting for 27.7%, with the top six contributors being  $C_{20}$ ,  $C_{22}$ ,  $C_{24}$ , and  $C_{30}$   $n$ -fatty acids,  $CPI_{FA}$ , and  $ACL_{FA}$  (Fig. 3d). The leaves and roots clusters were separated, but the moss cluster overlapped a small part of each with the stems cluster overlapping part of the leaf and moss clusters.

#### 3.3. PCAs of Beerberg core data

The  $n$ -alkane,  $n$ -alkanol, and  $n$ -fatty acid concentrations measured in a core from Beerberg peatland have been previously reported in Thomas et al. (2023b,a). The same sequence of PCAs was performed to examine the data from the peat core. The PCA including all ratios and biomarker concentrations identified PC1 as explaining 51.6% of the variance and PC2 as explaining 12.8% (Fig. 4a). The top six variables contributing to the first and second PCs were  $C_{29}$  and  $C_{31}$   $n$ -alkanes,  $C_{22}$  and  $C_{28}$   $n$ -alkanols,  $P_{wax}$ , and the  $C_{25}/(C_{25}+C_{29})$  ratio. The  $n$ -alkane PCA identified a PC1 accounting for 59.8% of the variance and PC2 for 23%, with the  $C_{25}/(C_{25}+C_{29})$  ratio,  $P_{wax}$ ,  $P_{aq}$ ,  $ACL_{ALK}$ , and  $C_{29}$  and  $C_{31}$   $n$ -alkanes as the top six contributors (Fig. 4b). In the PCA of  $n$ -alkanol concentrations and ratios, PC1 accounted for 55.1% and PC2 accounted for 28.3% of the variance with the top six contributors

**Table 1**  
Elemental and biomarker composition of modern plants at Beerberg peatland.

Species	PFT	Plant part	C	N	$\delta^{13}C$	$C_{max}$	Total( $\mu g/g$ )	CPI	ACL	$P_{eq}$	$P_{max}$	$C_{max}$	Total( $\mu g/g$ )	CPI	ACL	$C_{max}$	Total( $\mu g/g$ )	CPI	ACL
<i>Calluna vulgaris</i>	Evergreen shrub	Leaves	51.9 ± 0.1	0.75 ± 0	-30.4 ± 0.2	31	2056	14.9	31.3	0.02	0.98	28	163	10.4	25.3	32	2543	7.8	27.1
		Stem	47.9 ± 0.2	0.33 ± 0	-30.4 ± 0	31	370	11.6	30.6	0.10	0.91	28	228	15.2	25.7	28	1453	8.0	26.5
		Roots	47.5 ± 0.4	1.05 ± 0	-28.7 ± 0.1	33	813	12.7	31.7	0.03	0.97	28	243	10.0	25.2	24	699	1.5	24.2
<i>Empetrum nigrum</i>	Evergreen shrub	Leaves	53.0 ± 0.6	0.66 ± 0	-29.1 ± 0	31	2408	34.1	30.8	0.00	1.00	26	181	7.2	25.8	28	377	8.0	27.4
		Stem	51.2 ± 0.7	0.31 ± 0	-28.6 ± 0.1	29	1725	35.3	30.1	0.00	1.00	28	169	28.4	26.7	30	677	17.3	27.5
<i>Oxycoccus palustris</i>	Evergreen shrub	Aboveground	50.4 ± 1.2	0.60 ± 0	-31.1 ± 0.1	29	275	13.0	29.3	0.02	0.98	26	120	11.9	25.3	24	715	6.0	25.9
		Roots	50.1 ± 0.3	0.55 ± 0.1	-29.9 ± 0.1	29	25	8.9	29.0	0.11	0.91	24	201	10.3	25.0	24	682	5.9	24.5
<i>Vaccinium myrtillus</i>	Deciduous shrub	Leaves	48.8 ± 0.7	1.41 ± 0	-29.9 ± 0	31	461	10.7	30.6	0.01	0.99	24	284	14.5	24.8	20	502	4.3	25.5
		Stem	48.9 ± 0	0.72 ± 0	-30.3 ± 0	31	131	3.0	29.7	0.06	0.94	24	143	13.7	24.2	30	465	5.3	26.8
		Roots	47.1 ± 0.7	0.22 ± 0	-29.2 ± 0.2	31	5	NA	29.7	0.15	0.88	28	64	35.8	27.2	26	587	3.5	25.2
<i>Vaccinium uliginosum</i>	Deciduous shrub	Leaves	48.4 ± 0.7	1.14 ± 0	-29.7 ± 0.1	27	131	5.6	26.3	0.87	0.66	28	309	6.2	26.4	28	3363	8.6	25.5
		Stem	48.8 ± 0	0.79 ± 0	-28.1 ± 0.1	27	21	7.8	27.5	0.54	0.71	28	75	11.7	25.1	20	436	6.1	23.9
		Roots	47.7 ± 0	0.40 ± 0	-26.9 ± 0.2	31	7	16.5	28.9	0.35	0.74	26	20	6.4	25.3	20	222	3.3	22.7
<i>Vaccinium vitis-idaea</i>	Evergreen shrub	Leaves	50.7 ± 1.0	1.05 ± 0	-28.2 ± 0.1	29	316	8.0	28.1	0.23	0.82	26	58	3.9	26.0	20	403	4.7	25.6
		Stem	49.0 ± 0.4	0.36 ± 0	-27.5 ± 0	29	644	4.6	28.7	0.14	0.89	24	113	5.0	23.5	30	309	4.2	26.1
		Roots	47.4 ± 0.6	0.25 ± 0	-26.5 ± 0	31	36	5.6	29.2	0.17	0.85	26	46	12.2	26.0	26	170	5.1	26.2
<i>Eriophorum vaginatum</i>	Sedge	Aboveground	46.6 ± 0.4	1.08 ± 0	-26.6 ± 0.1	31	1540	33.1	31.3	0.01	0.99	26	108	4.4	25.5	24	296	3.2	24.5
		Roots	45.6 ± 0.2	0.99 ± 0	-25.8 ± 0.1	31	69	14.1	29.7	0.13	0.88	24	94	9.0	25.3	24	307	5.0	24.2
<i>Pinus sylvestris</i>	Evergreen coniferous tree	Needles	49.9 ± 0.7	2.30 ± 0	-27.5 ± 0.1	29	24	6.4	28.6	0.26	0.78	24	194	8.4	24.9	22	1024	9.7	23.2
		Stem	48.3 ± 0.1	0.64 ± 0	-27.6 ± 0	31	5	4.5	29.3	0.17	0.85	22	195	89.9	22.5	22	2359	23.6	22.2
		Roots	48.3 ± 0.4	0.87 ± 0	-26.6 ± 0.1	31	5	6.9	30.1	0.15	0.87	22	13	7.3	22.2	24	1037	17.2	22.8
<i>Picea abies</i>	Evergreen coniferous tree	Needles	49.6 ± 1.2	0.50 ± 0	-30.5 ± 0	29	13	4.7	27.8	0.40	0.66	22	28	41.2	22.8	22	559	8.7	24.0
		Stem	51.4 ± 0	0.45 ± 0	-30.5 ± 0.1	23	19	5.5	26.8	0.54	0.52	22	142	3.6	22.6	24	2966	23.9	22.8
		Roots	48.2 ± 0.3	0.56 ± 0	-26.5 ± 0.1	29	111	33.5	30.1	0.01	1.00	22	23	0.9	23.2	20	382	13.2	22.2
<i>Betula pendula</i>	Deciduous tree	Leaves	50.9 ± 0.3	1.08 ± 0	-28.7 ± 0	31	955	22.8	27.0	0.20	0.85	28	312	26.3	25.0	28	1172	5.1	26.3
		Stem	51.6 ± 0.1	1.10 ± 0	-26.8 ± 0	27	494	42.6	26.2	0.85	0.58	22	480	43.4	24.8	28	800	11.1	25.6
		Roots	48.3 ± 0.4	0.91 ± 0	-26.9 ± 0	31	16	11.3	30.4	0.03	0.97	26	5	NA	24.1	20	115	1.2	22.3
<i>Betula pubescens</i>	Deciduous tree	Leaves	47.9 ± 0.6	2.01 ± 0.1	-30.4 ± 0.1	31	1255	12.5	29.3	0.16	0.88	28	218	31.6	27.0	28	1356	4.4	26.1
		Stem	50.0 ± 1.2	1.03 ± 0	-29.5 ± 0.1	25	499	32.6	26.5	0.76	0.48	28	535	35.3	25.6	26	1566	7.4	24.5
		Roots	48.2 ± 0.9	1.20 ± 0	-29.1 ± 0.1	31	35	14.4	29.0	0.32	0.75	22	62	15.6	23.7	24	464	4.2	23.6
<i>Polytrichum strictum</i>	Moss		45.3 ± 0.4	0.74 ± 0	-28.7 ± 0	31	46	6.4	30.0	0.13	0.89	28	23	5.5	25.9	24	433	5.2	24.6
<i>Sphagnum angustifolium</i>	Moss		46.4 ± 0.2	0.78 ± 0	-30.2 ± 0.4	31	955	13.1	31.1	0.04	0.96	24	190	11.7	25.3	24	1201	5.4	26.4
<i>Sphagnum capillifolium</i>	Moss		48.3 ± 4.4	0.81 ± 0.1	-29.1 ± 0	31	95	5.6	27.7	0.44	0.62	24	124	9.5	25.1	24	380	4.2	24.9
<i>Sphagnum fuscum</i>	Moss		43.4 ± 0	0.38 ± 0	-27.9 ± 0.1	31	200	15.5	29.7	0.36	0.7	26	137	12.6	25.2	24	236	6.3	25.5

being  $C_{22}$ ,  $C_{24}$ , and  $C_{28}$   $n$ -alkanol,  $ACL_{ALC}$ , the  $C_{22}/C_{24}$  ratio, and the  $C_{22}+C_{24}/C_{26}+C_{28}$  ratio (Fig. 4c). The PCA with  $n$ -fatty acids identified PC1 accounting for 60.3% of the variance and PC2 accounting for 17.7%, with the top six contributors being  $C_{22}$ ,  $C_{24}$ ,  $C_{26}$ ,  $C_{28}$ ,  $C_{30}$ , and  $C_{32}$   $n$ -fatty acids (Fig. 4d).

### 3.4. Modeling

We used the VERHIB model to run various scenarios and answer the following:

(a) Does adding  $n$ -fatty acid data help to create a more accurate and unique solution?

(b) Does adding root data help to create a more accurate and unique solution?

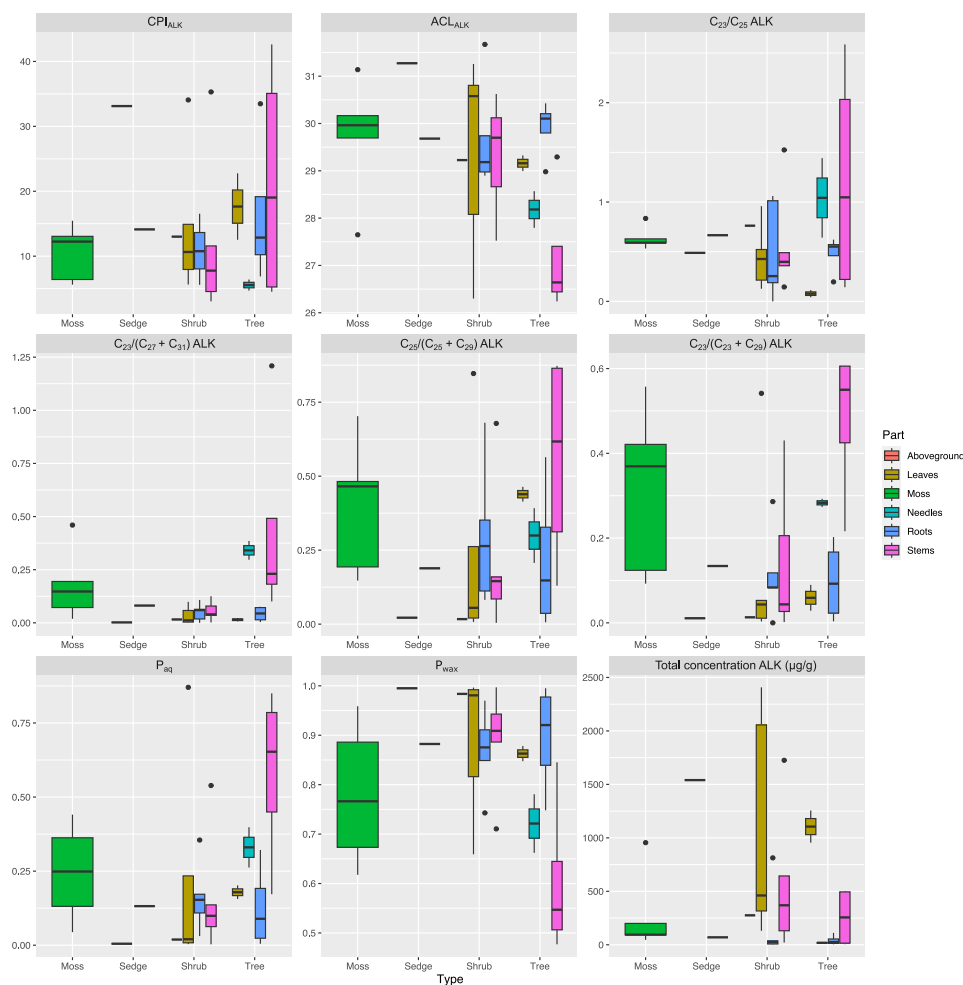
We compared the results to the macrofossil analysis included in Thomas et al. (2023b), as macrofossils are often more representative of in situ vegetation than pollen analysis (e.g., Parducci et al., 2015).

For the first group of scenarios (see Table 2 and Figs. 5a–5d and 6), we included all of the chain lengths, not only plant-derived, and we did not include the root biomarker data. The two regularization parameters were kept at 0.1, and the maximum iterations were set at 2000. The plants were grouped as followed: Group 1: *S. angustifolium*, *S. magellanicum*, *S. capillifolium*, *S. fuscum*, Group 2: *C. vulgaris*, *V. uliginosum*, *V. myrtillus*, *V. vitis-idaea*, *O. palustris*, *E. nigrum*, Group 3: *E. vaginatum*, Group 4: *P. strictum*, Group 5: *P. sylvestris*, Group 6: *P. abies*, and Group 7: *B. pendula* and *B. pubescens*. Run 1 included all three compound classes, Run 2 included  $n$ -alkanes and  $n$ -alcohols, Run 3 included only  $n$ -alkanes, and Run 4 included  $n$ -alkanes and  $n$ -fatty acids. In this group of scenarios, Run 1 had a Pearson correlation coefficient between the predicted and actual core measurement values of 0.7476, Run 2 of 0.9963, Run 3 of 0.9995, and Run 4 of 0.7367. When comparing the VERHIB reconstructions to the plant macrofossil

analysis, if  $n$ -fatty acid data was excluded (Figs. 5b and 5c), the contribution of tree species was overestimated, particularly *P. abies*, as was the contribution of *S. capillifolium* and *S. magellanicum*. However, all of the runs underestimated the presence of *E. vaginatum*, especially in the lowest part of the core. Fig. 6 shows that the highest residuals typically occurred in the lowest depth layers.

In the second group of scenarios (Table 2 and Figs. 7a–7d and 8), we again included all of the chain lengths and the root biomarker data, with a leaf:root mass ratio of 5:1 (based off of the ratios seen in the measured data). The two regularization parameters were kept at 0.1, and the maximum iterations were set at 2000. The plant groupings remained the same. Run 5 included all three compound classes, Run 6 included  $n$ -alkanes and  $n$ -alcohols, Run 7 included only  $n$ -alkanes, and Run 8 included  $n$ -alkanes and  $n$ -fatty acids. Run 5 had a Pearson correlation coefficient between the predicted and actual core measurement values of 0.7436, Run 6 of 0.9962, Run 7 of 0.9952, and Run 8 of 0.9255. The results for this group of scenarios were similar to the first in that *E. vaginatum* was underestimated in all of the scenarios, compared to what was evident in the macrofossil analysis. In Fig. 7a, the contribution of *P. sylvestris* is overestimated in the lowest depth layers. In Fig. 7b, the estimated contribution of the plant species did not vary throughout the core as expected, and the contribution of *Sphagnum* mosses was underestimated. Fig. 8 shows again that the highest residuals typically occurred in the lowest depth layers.

For the third group of scenarios (see Table 2 and Figs. 9a–9d and 10), we included only plant-derived chain lengths and did not include the root biomarker data. The parameters, maximum iterations, and plant groupings were the same as the previous runs. Run 9 included all three compound classes, Run 10 included  $n$ -alkanes and  $n$ -alcohols, Run 11 included only  $n$ -alkanes, and Run 12 included  $n$ -alkanes and  $n$ -fatty acids. In this group of scenarios, Run 9 had a Pearson correlation coefficient between the predicted and actual core measurement values



**Fig. 2a.** Boxplots of *n*-alkane variables in modern plant parts (CPI, ACL,  $C_{23}/C_{25}$ ,  $C_{23}/(C_{27}+C_{31})$ ,  $C_{25}/(C_{25}+C_{29})$ ,  $C_{23}/(C_{23}+C_{29})$ ,  $P_{aa}$ ,  $P_{wax}$ , and total concentration). Plant/part: moss-whole plant (n=5), sedge-aboveground biomass (n=1), sedge-roots (n=1), shrub-aboveground biomass (n=1), shrub-leaves (n=5), shrub-roots (n=5), shrub-stems (n=5), tree-leaves (n=2), tree-needles (n=2), tree-roots (n=4), tree-stems (n=4).

**Table 2**

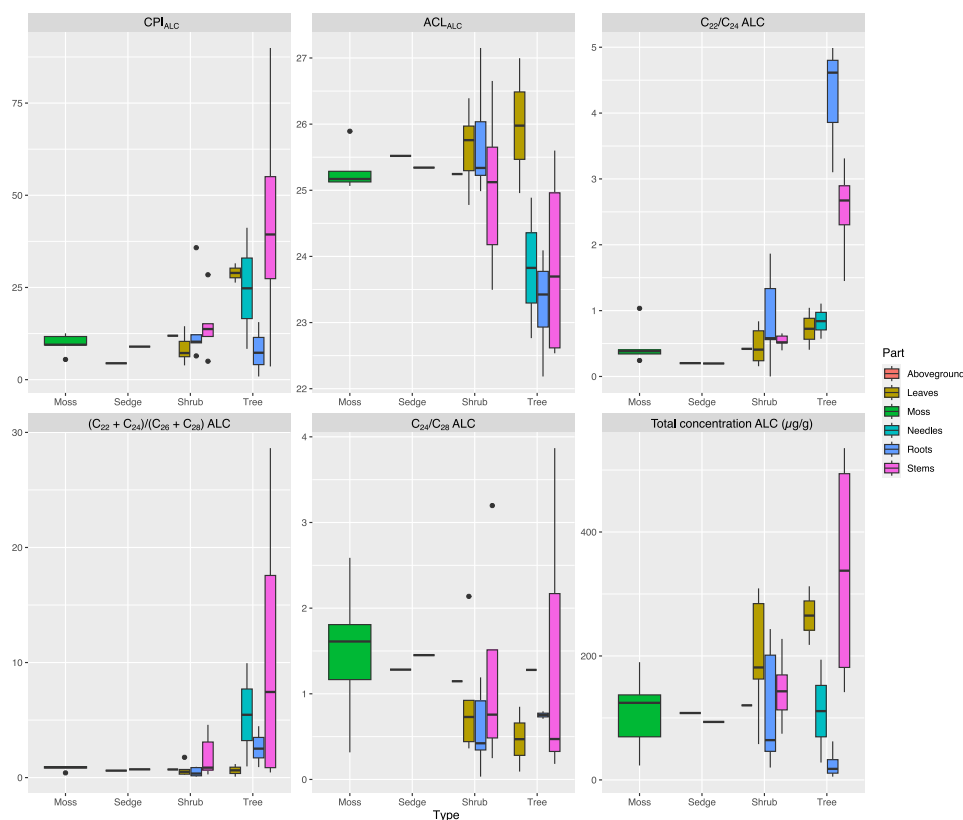
Details of VERHIB runs and Pearson correlation coefficients. Additional runs can be seen in [Table A.3](#).

Run #	Compounds included	Leaf:root ratio	Reg (depth)	Reg (group)	Max iterations	Pearson coefficient (soil chem values)
1	ALK: C <sub>19</sub> –C <sub>33</sub> , ALC: C <sub>15</sub> –C <sub>28</sub> , FA: C <sub>14</sub> –C <sub>32</sub>	1:0	0.1	0.1	2000	0.7476
2	ALK: C <sub>19</sub> –C <sub>33</sub> , ALC: C <sub>15</sub> –C <sub>28</sub>	1:0	0.1	0.1	2000	0.9963
3	ALK: C <sub>19</sub> –C <sub>33</sub>	1:0	0.1	0.1	2000	0.9995
4	ALK: C <sub>19</sub> –C <sub>33</sub> , FA: C <sub>14</sub> –C <sub>32</sub>	1:0	0.1	0.1	2000	0.7367
5	ALK: C <sub>19</sub> –C <sub>33</sub> , ALC: C <sub>15</sub> –C <sub>28</sub> , FA: C <sub>14</sub> –C <sub>32</sub>	5:1	0.1	0.1	2000	0.7436
6	ALK: C <sub>19</sub> –C <sub>33</sub> , ALC: C <sub>15</sub> –C <sub>28</sub>	5:1	0.1	0.1	2000	0.9962
7	ALK: C <sub>19</sub> –C <sub>33</sub>	5:1	0.1	0.1	2000	0.9952
8	ALK: C <sub>19</sub> –C <sub>33</sub> , FA: C <sub>14</sub> –C <sub>32</sub>	5:1	0.1	0.1	2000	0.9255
9	Plant-derived: ALK, ALC, FA	1:0	0.1	0.1	2000	0.9783
10	Plant-derived; ALK, ALC	1:0	0.1	0.1	2000	0.9972
11	Plant-derived; ALK	1:0	0.1	0.1	2000	0.9995
12	Plant-derived; ALK, FA	1:0	0.1	0.1	2000	0.9759
13	Plant-derived: ALK, ALC, FA	5:1	0.1	0.1	2000	0.9681
14	Plant-derived; ALK, ALC	5:1	0.1	0.1	2000	0.9971
15	Plant-derived; ALK	5:1	0.1	0.1	2000	0.9937
16	Plant-derived; ALK, FA	5:1	0.1	0.1	2000	0.9755

of 0.9783, Run 10 of 0.9972, Run 11 of 0.9995, and Run 12 of 0.9759. Again, all of the runs underestimated the proportion of *E. vaginatum*. In [Figs. 9a](#) and [9d](#), the contribution of *S. magellanicum* is overestimated. In [Figs. 9b](#) and [9c](#), the contribution of *P. sylvestris* and *P. abies* is overestimated. [Fig. 10](#) shows that the highest residuals typically

occurred in the lowest depth layers, though there are also high residuals in the upper depth layers for Run 11 and throughout the core for Run 10.

For the last group of scenarios (see [Table 2](#) and [Figs. 11a–11d](#) and [12](#)), we included only the plant-derived chain lengths and the root



**Fig. 2b.** Boxplots of *n*-alkanol variables in modern plant parts (CPI, ACL,  $C_{22}/C_{24}$ ,  $(C_{22}+C_{24})/(C_{26}+C_{28})$ ,  $C_{24}/C_{28}$ , and total concentration). Plant/part: moss-whole plant ( $n=5$ ), sedge-aboveground biomass ( $n=1$ ), sedge-roots ( $n=1$ ), shrub-aboveground biomass ( $n=1$ ), shrub-leaves ( $n=5$ ), shrub-roots ( $n=5$ ), shrub-stems ( $n=5$ ), tree-leaves ( $n=2$ ), tree-needles ( $n=2$ ), tree-roots ( $n=4$ ), tree-stems ( $n=4$ ).

biomarker data, with a leaf:root mass ratio of 5:1. The parameters, maximum iterations, and plant groupings were kept the same as the previous runs. Run 13 included all three compound classes, Run 14 included *n*-alkanes and *n*-alcohols, Run 15 included only *n*-alkanes, and Run 16 included *n*-alkanes and *n*-fatty acids. In this group of scenarios, Run 13 had a Pearson correlation coefficient between the predicted and actual core measurement values of 0.7476, Run 14 of 0.9963, Run 15 of 0.9995, and Run 16 of 0.7367. The primary difference between the reconstructions of this group (Figs. 11a–11d) with the last group (plant-derived, no root data) (Figs. 9a–9d) is that the estimated contribution of *P. strictum* increased in all the scenarios. The residual patterns also stayed very similar (Fig. 12).

## 4. Discussion

### 4.1. Biomarker composition of modern plants

The current species present at the Beerberg peatland generally represent an ombrotrophic bog plant community in central Europe. Although previously published biomarker data is not available for all of the investigated plant species and parts, we could compare our data to those of a few studies.

#### 4.1.1. *Sphagnum* mosses

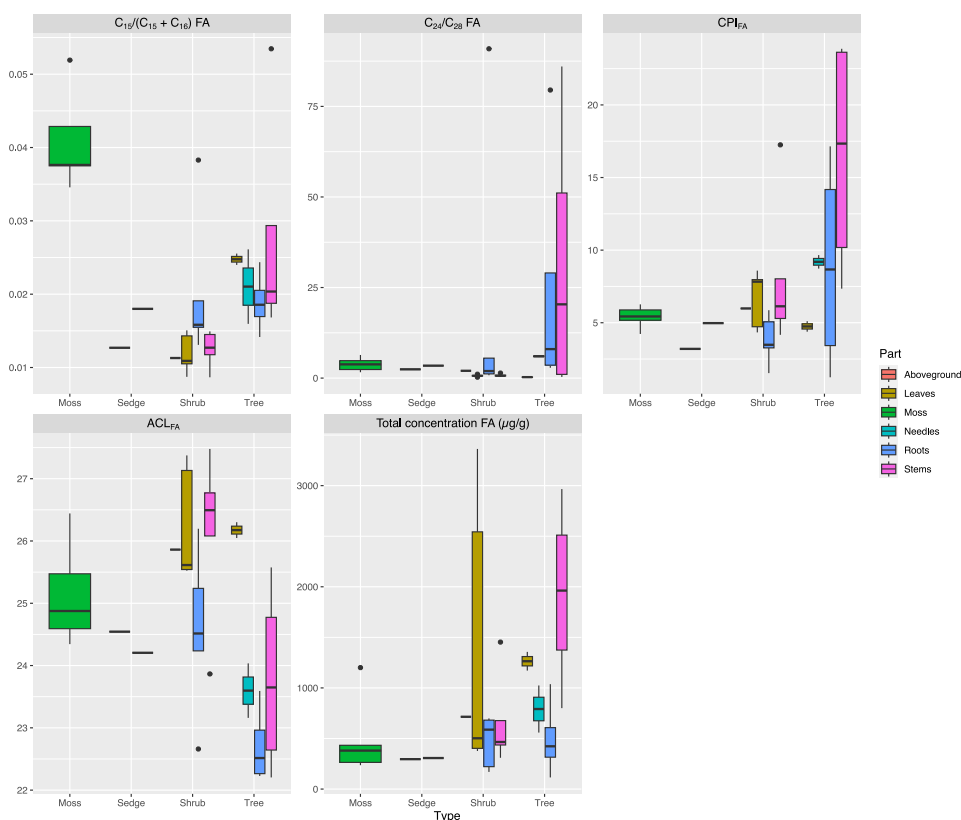
Previous studies have established variations in *n*-alkane biomarker compositions of *Sphagnum* moss species related to the moisture conditions of their typical microhabitat Nott et al. (e.g., 2000), Nichols et al. (e.g., 2006), Bingham et al. (e.g., 2010). Certain *Sphagnum* species such grow more typically in hummocks, raised areas further from the water table, while other species are found in flatter, wetter areas known as lawns or hollows (Rydin et al., 2013). Hummock species have been shown to contain higher amounts of longer chain lengths ( $C_{25}$  and  $C_{31}$ ),

while hollow species are usually dominated by  $C_{23}$  (Bingham et al., 2010). As one major function of plant waxes is to reduce evapotranspiration (Eglinton and Hamilton, 1967), moisture availability to the plant affects the chain lengths of the waxes.

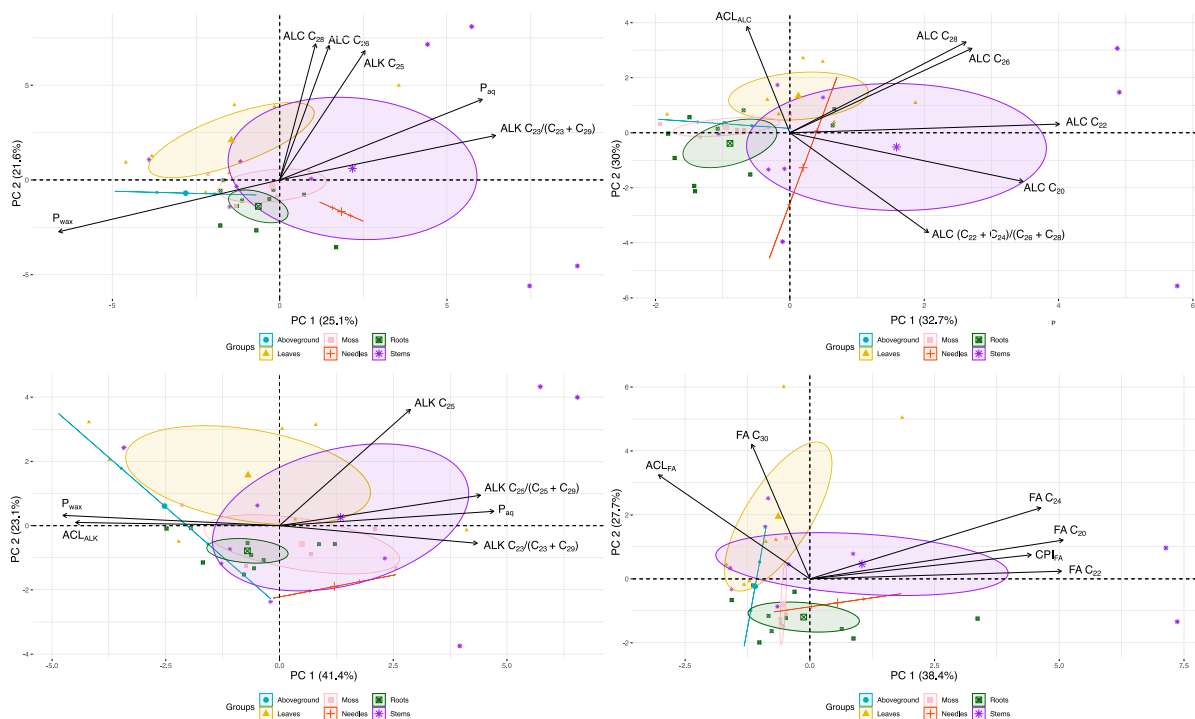
Of the *Sphagnum* species, we collected at the Beerberg peatland, *S. fuscum* and *S. capillifolium* are typical hummock species while *S. angustifolium* is more likely to be found in lawns or hollows and *S. magellanicum* is more adaptable and can be found in a wider range of conditions (Rydin et al., 2013). For all of these species, we measured a  $C_{max}$  of  $C_{31}$  for the *n*-alkanes. For *S. fuscum*, *S. magellanicum*, and *S. angustifolium*, the second most abundant chain length was  $C_{33}$  while for *S. capillifolium*, it was  $C_{29}$ . Some of these results contrast previously published data with *S. fuscum* being reported with  $C_{25}$  as a dominant chain length in multiple studies in the British Isles (Corrigan et al., 1973; Ficken et al., 1998) and Finland and Sweden (Vonk and Gustafsson, 2009; Bingham et al., 2010) and *S. angustifolium* measured with dominant chain lengths of  $C_{23}$  (Bingham et al., 2010) and  $C_{27}/C_{29}$  (Vonk and Gustafsson, 2009) in Finland and Sweden, respectively.

*S. magellanicum* has more mixed results, with  $C_{25}$  reported as the dominant chain in the Netherlands (Baas et al., 2000), the British Isles (Corrigan et al., 1973; Bingham et al., 2010), Finland (Bingham et al., 2010), and Sweden (Vonk and Gustafsson, 2009) and  $C_{31}$  the dominant chain length in Denmark (Bingham et al., 2010) and at another site in Great Britain (Nott et al., 2000; Xie et al., 2004). *S. capillifolium* has also been reported in different studies as having either  $C_{25}$  in Sweden (Vonk and Gustafsson, 2009) or  $C_{31}$  in the British Isles (Ficken et al., 1998; Nott et al., 2000; Xie et al., 2004) as a dominant chain length.

A likely cause for the differences in biomarker composition and the higher proportion of longer-chain lengths in our samples is a difference in climate. Since 2015, central Europe has been experiencing historic droughts (Büntgen et al., 2021), with the period of



**Fig. 2c.** Boxplots of *n*-fatty acid variables in modern plant parts ( $C_{15}/(C_{15}+C_{16})$ ,  $C_{24}/C_{28}$ , CPI, ACL, and total concentration). Plant/part: moss-whole plant (n=5), sedge-aboveground biomass (n=1), sedge-roots (n=1), shrub-aboveground biomass (n=1), shrub-leaves (n=5), shrub-roots (n=5), shrub-stems (n=5), tree-leaves (n=2), tree-needles (n=2), tree-roots (n=4), tree-stems (n=4).



**Fig. 3.** (a) PCA of plant samples using all plant-derived biomarker absolute concentrations, odd  $C_{21}$ – $C_{33}$  *n*-alkanes, even  $C_{20}$ – $C_{28}$  *n*-alkanols, and even  $C_{20}$ – $C_{32}$  *n*-fatty acids, and the ACL and CPI for all of the compound groups, and the other ratios listed in the Methods section. The ellipses and lines represent a 95% confidence interval for each plant part group. (b) PCA of plant samples using the odd  $C_{21}$ – $C_{33}$  *n*-alkanes, the ACL and CPI, and the other *n*-alkane ratios listed in the Methods section. The ellipses and lines represent a 95% confidence interval for each plant part group. (c) PCA of plant samples using the even  $C_{20}$ – $C_{28}$  *n*-alkanols, the ACL and CPI, and the other ratios listed in the Methods section. The ellipses and lines represent a 95% confidence interval for each plant part group. (d) PCA of plant samples using the even  $C_{20}$ – $C_{32}$  *n*-fatty acids, and the ACL and CPI. The ellipses and lines represent a 95% confidence interval for each plant part group.

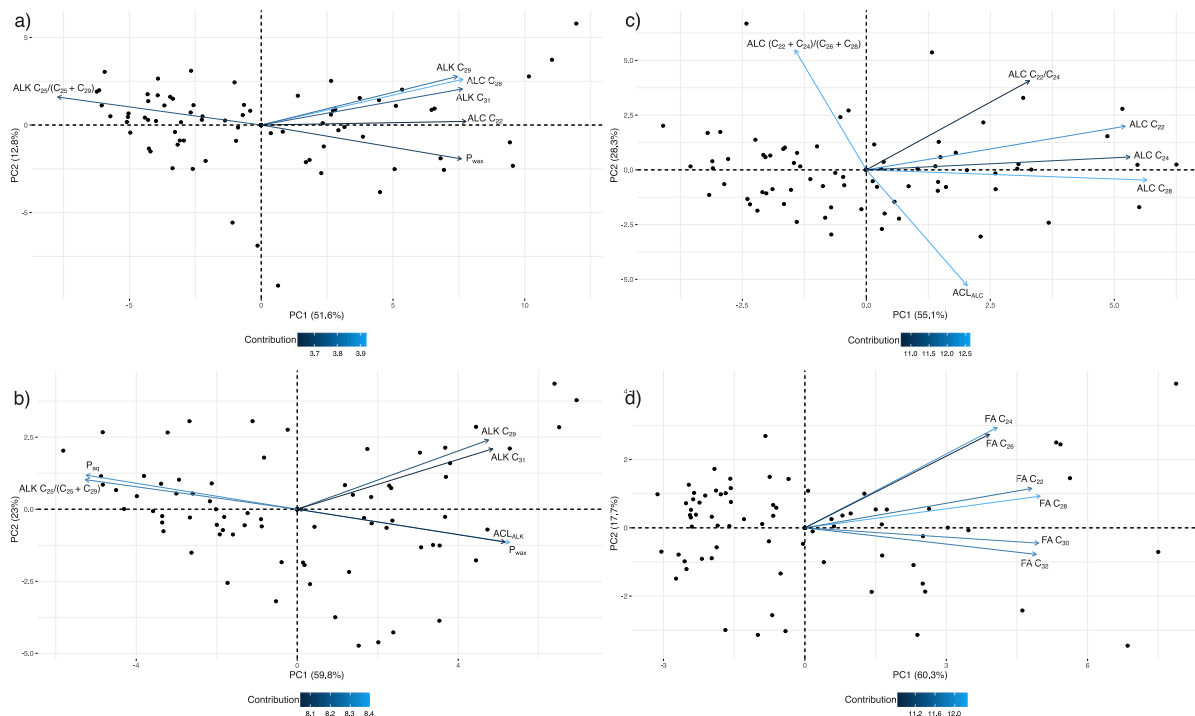


Fig. 4. (a) PCA of core samples using all plant-derived biomarker absolute concentrations, odd  $C_{21}$ – $C_{33}$   $n$ -alkanes, even  $C_{20}$ – $C_{28}$   $n$ -alkanols, and even  $C_{20}$ – $C_{32}$   $n$ -fatty acids, and the ACL and CPI for all of the compound groups, and the other ratios listed in the Methods section. (b) PCA of core samples using the odd  $C_{21}$ – $C_{33}$   $n$ -alkanes, the ACL and CPI, and the other  $n$ -alkane ratios listed in the Methods section. (c) PCA of core samples using the even  $C_{20}$ – $C_{28}$   $n$ -alkanols, the ACL and CPI, and the other ratios listed in the Methods section. (d) PCA of core samples using the even  $C_{20}$ – $C_{32}$   $n$ -fatty acids, and the ACL and CPI.

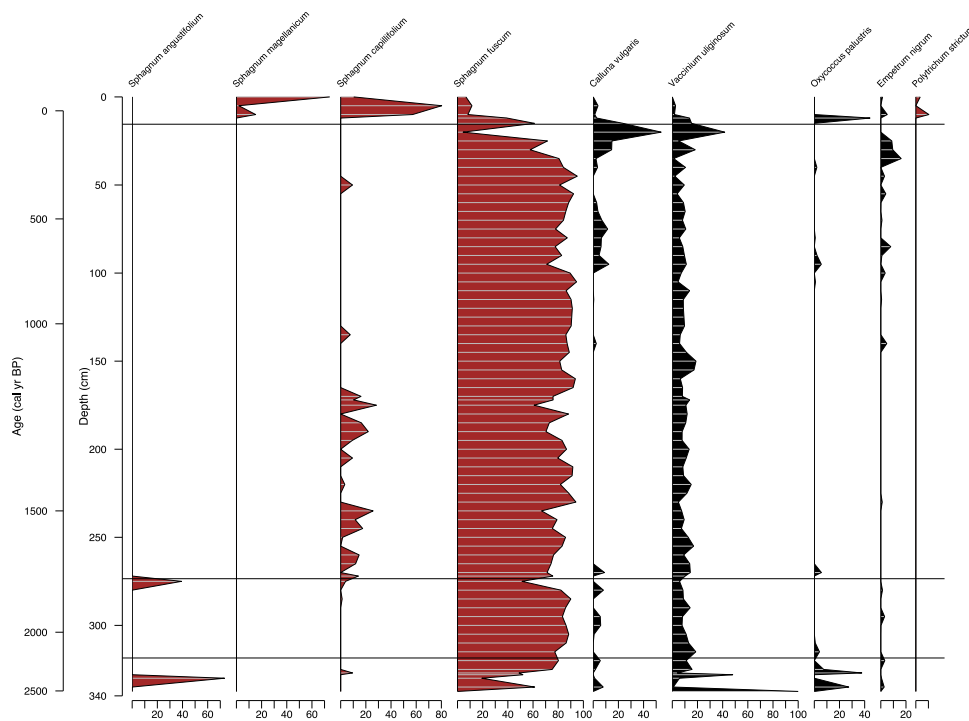
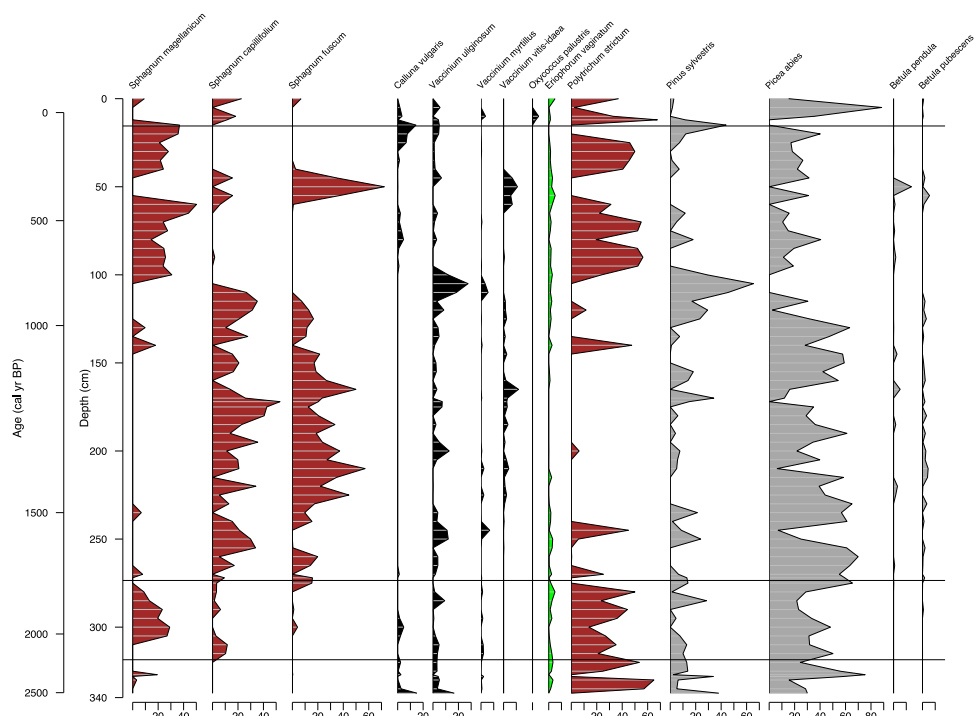
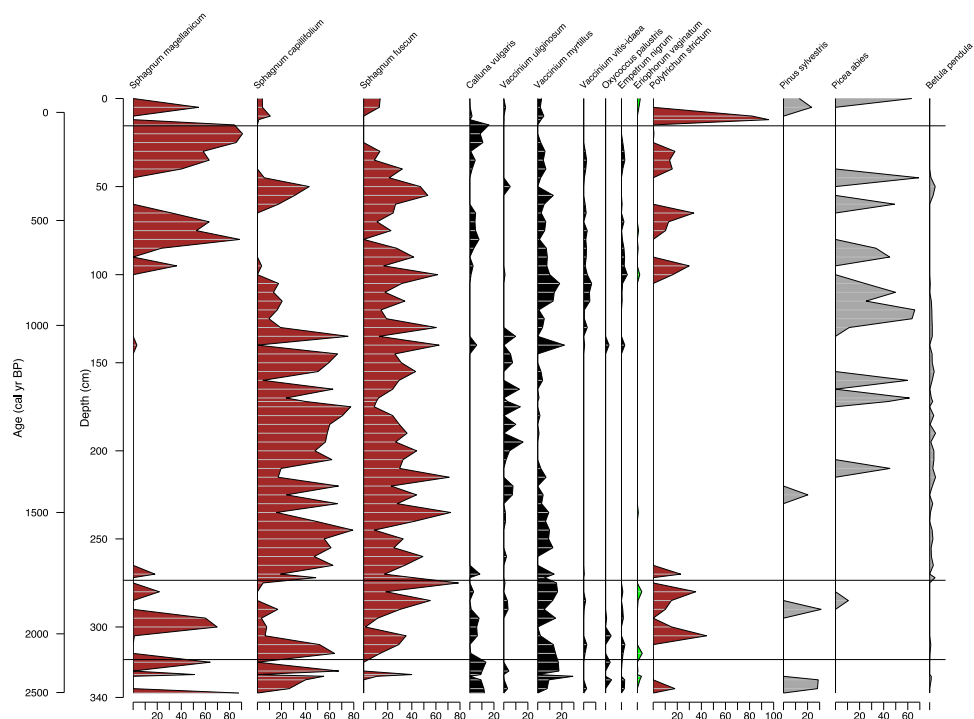


Fig. 5a. VERHIB reconstruction for Run 1, showing calculated percentages of vegetation biomass input. Run 1 includes only leaf measurements and all chain length measurements for  $n$ -alkanes,  $n$ -alkanols, and  $n$ -fatty acids. The brown graphs are moss species, and the black graphs are shrub species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 5b.** VERHIB reconstruction for Run 2, showing calculated percentages of vegetation biomass input. Run 2 includes only leaf measurements and all chain length measurements for *n*-alkanes and *n*-alkanols. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 5c.** VERHIB reconstruction for Run 3, showing calculated percentages of vegetation biomass input. Run 3 includes only leaf measurements and all chain length measurements for *n*-alkanes. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.

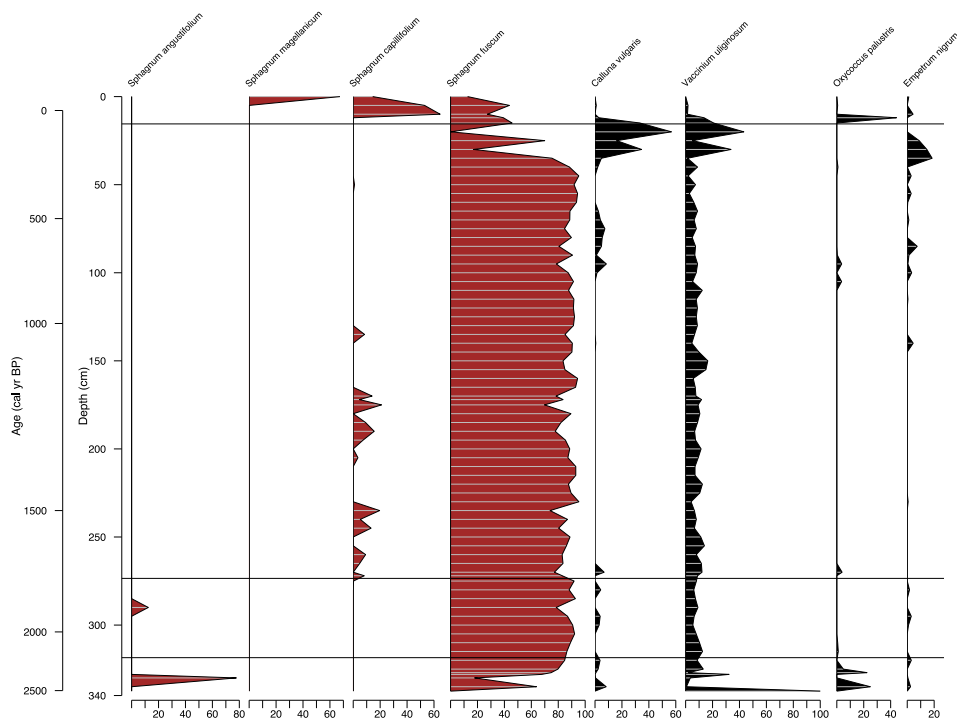


Fig. 5d. VERHIB reconstruction for Run 4, showing calculated percentages of vegetation biomass input. Run 4 includes only leaf measurements and all chain length measurements for *n*-alkanes and *n*-fatty acids. Brown graphs are moss species, and black graphs are shrub species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.

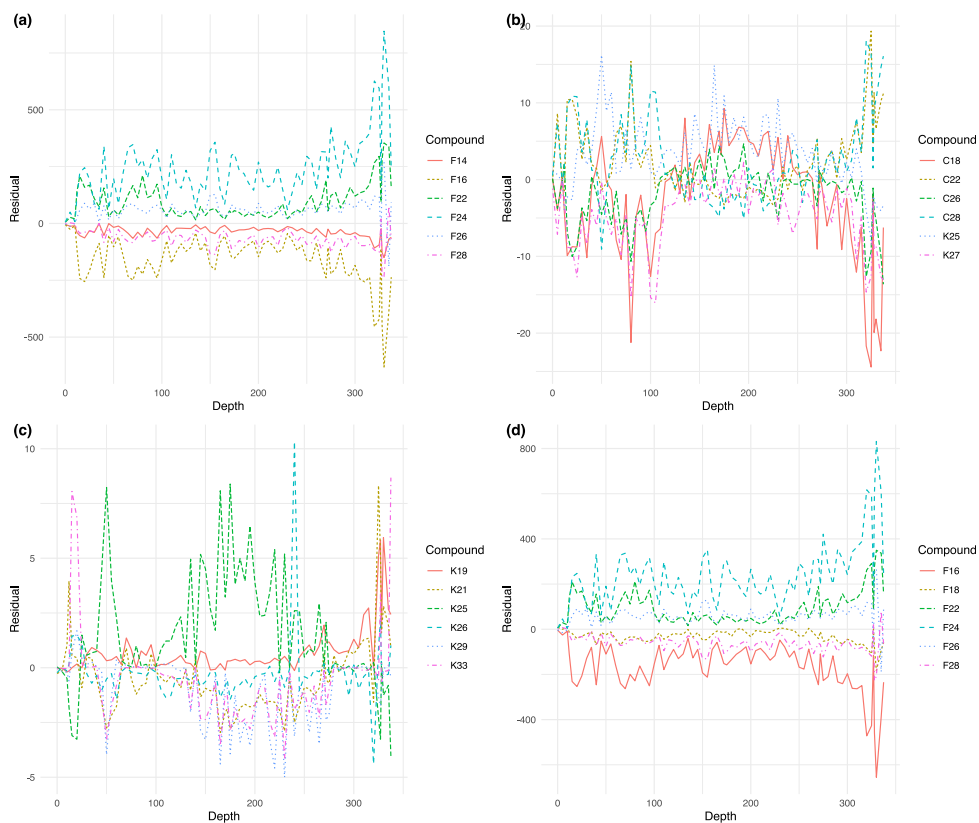
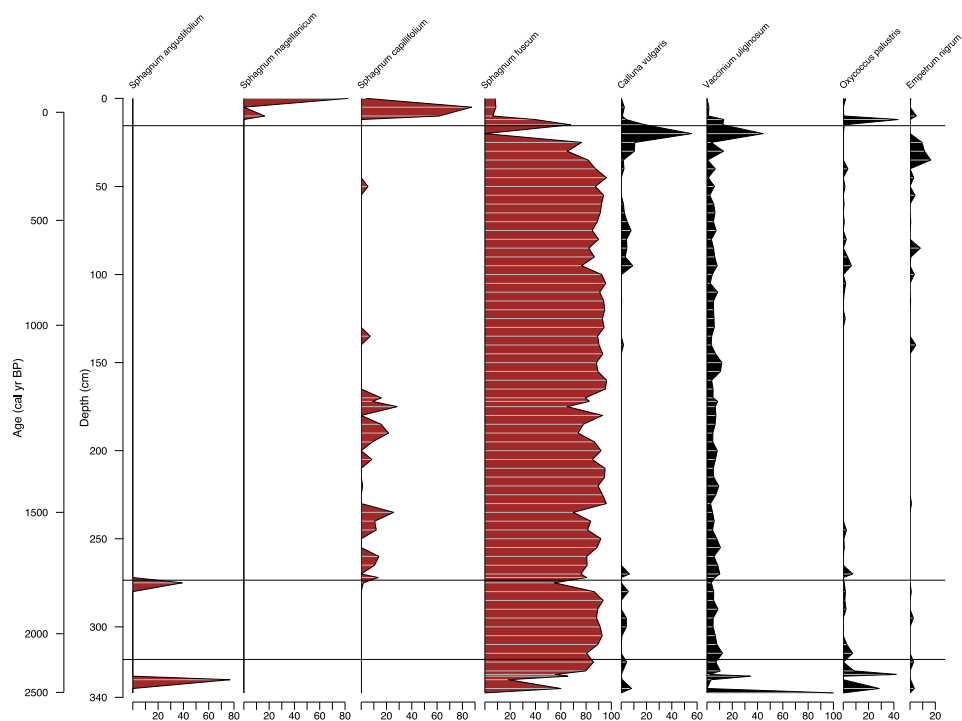
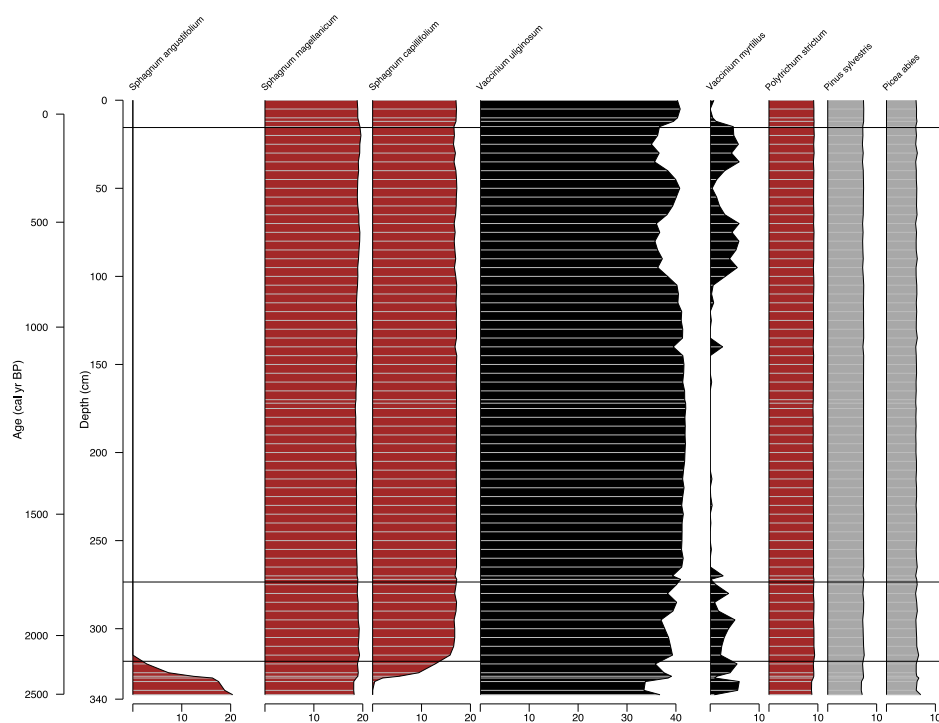


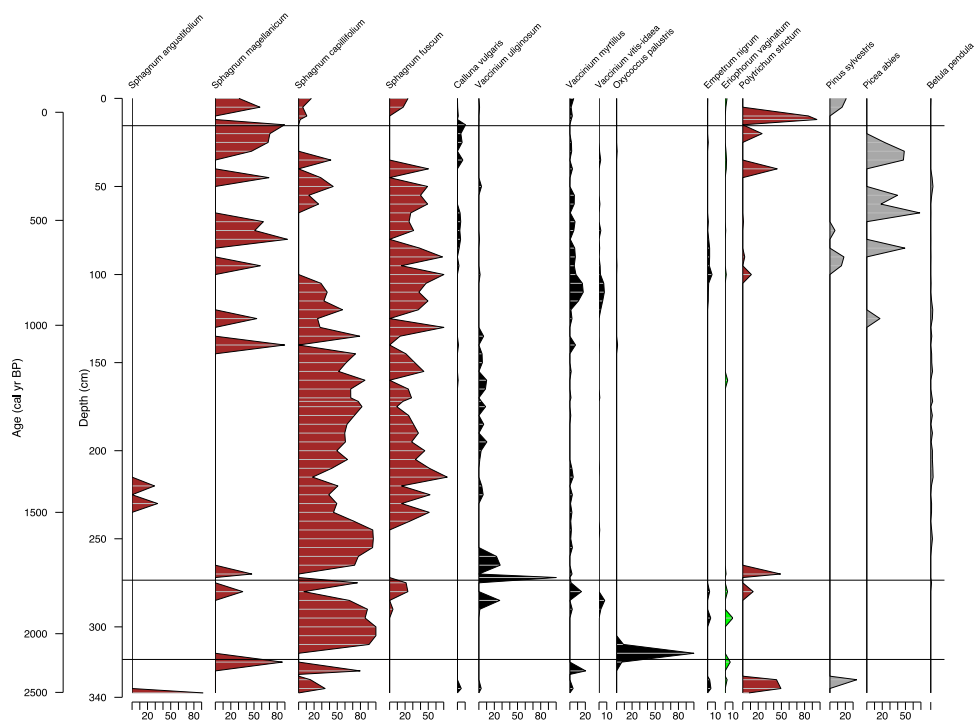
Fig. 6. Core measurement residuals versus depth for Run 1(a), Run 2(b), Run 3(c), and Run 4(d), showing compounds with the highest residual values.



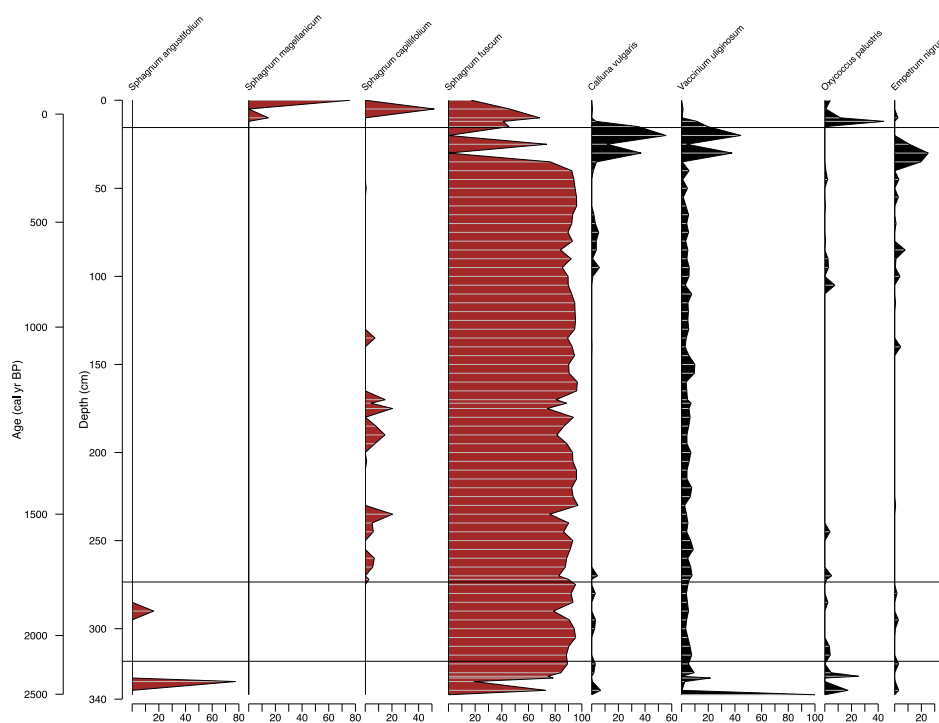
**Fig. 7a.** VERHIB reconstruction for Run 5, showing calculated percentages of vegetation biomass input. Run 5 includes leaf and root measurements (ratio leaf:root is 5:1) and all chain length measurements for *n*-alkanes, *n*-alkanols, and *n*-fatty acids. Brown graphs are moss species, and black graphs are shrub species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 7b.** VERHIB reconstruction for Run 6, showing calculated percentages of vegetation biomass input. Run 6 includes leaf and root measurements (ratio leaf:root is 5:1) and all chain length measurements for *n*-alkanes and *n*-alkanols. Brown graphs are moss species, black graphs are shrub species, and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 7c.** VERHIB reconstruction for Run 7, showing calculated percentages of vegetation biomass input. Run 7 includes leaf and root measurements (ratio leaf:root is 5:1) and all chain length measurements for *n*-alkanes. The brown graphs are moss species, the black graphs are shrub species, the green graph is *E. vaginatum*, and the gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 7d.** VERHIB reconstruction for Run 8, showing calculated percentages of vegetation biomass input. Run 8 includes leaf and root measurements (ratio leaf:root is 5:1) and all chain length measurements for *n*-alkanes and *n*-fatty acids. The brown graphs are moss species, and the black graphs are shrub species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.

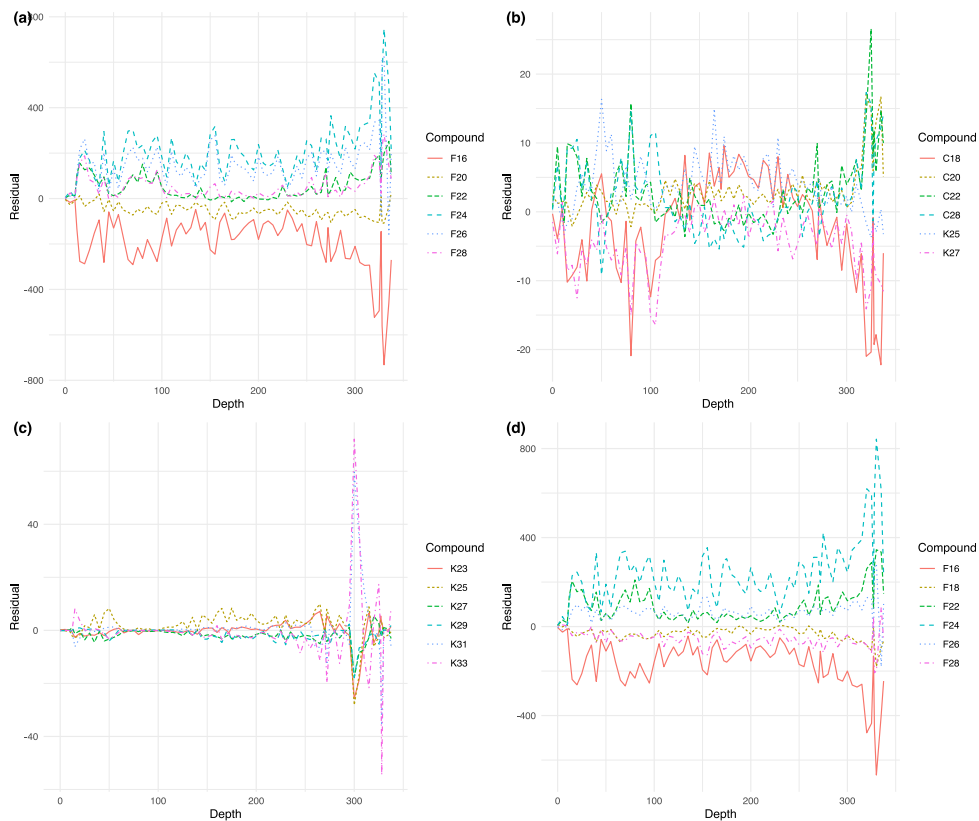


Fig. 8. Core measurement residuals versus depth for Run 5(a), Run 6(b), Run 7(c), and Run 8(d), showing compounds with the highest residual values.

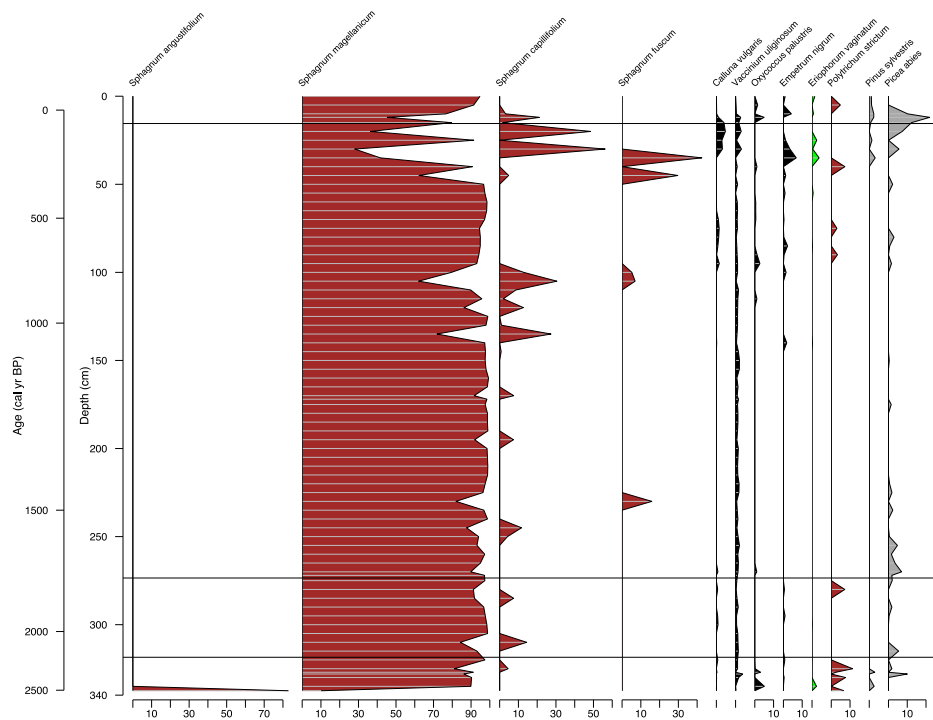
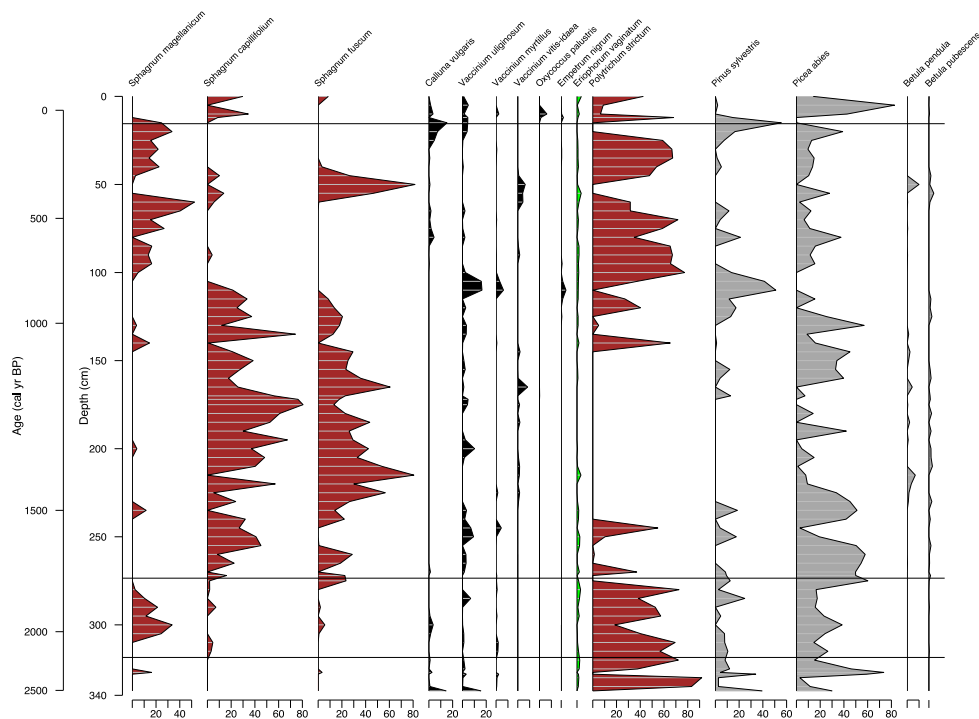
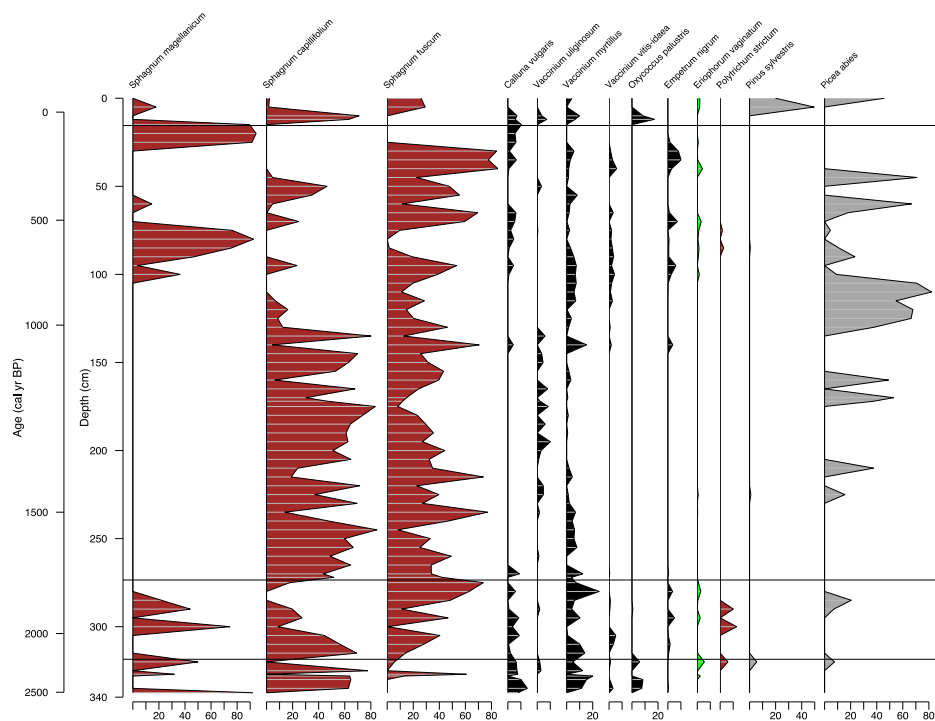


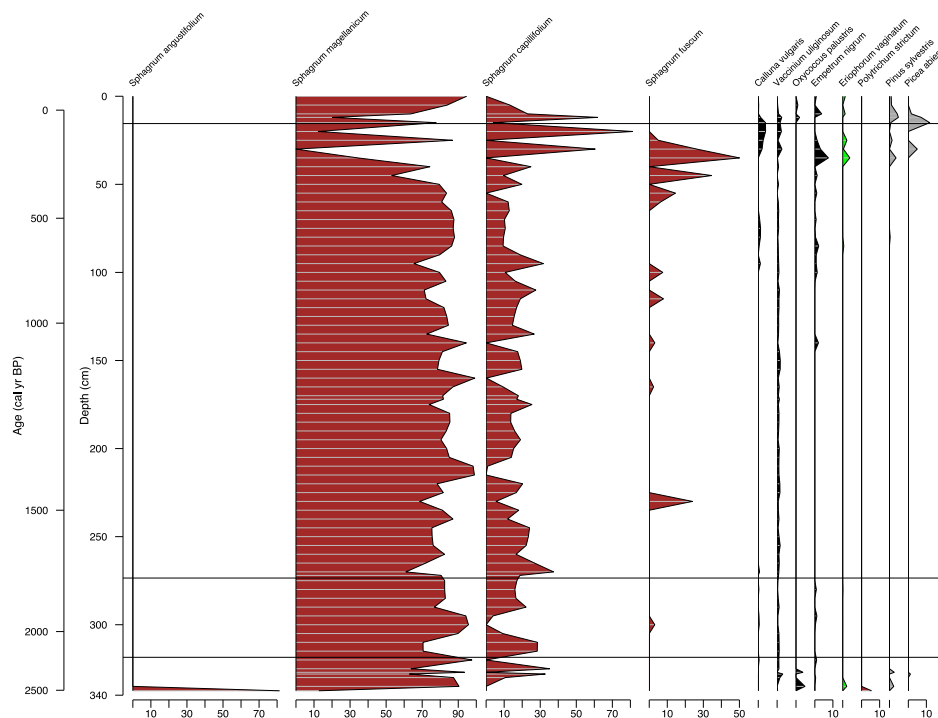
Fig. 9a. VERHIB reconstruction for Run 9, showing calculated percentages of vegetation biomass input. Run 9 includes only leaf measurements and only plant-derived chain length measurements for *n*-alkanes, *n*-alkanols, and *n*-fatty acids. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



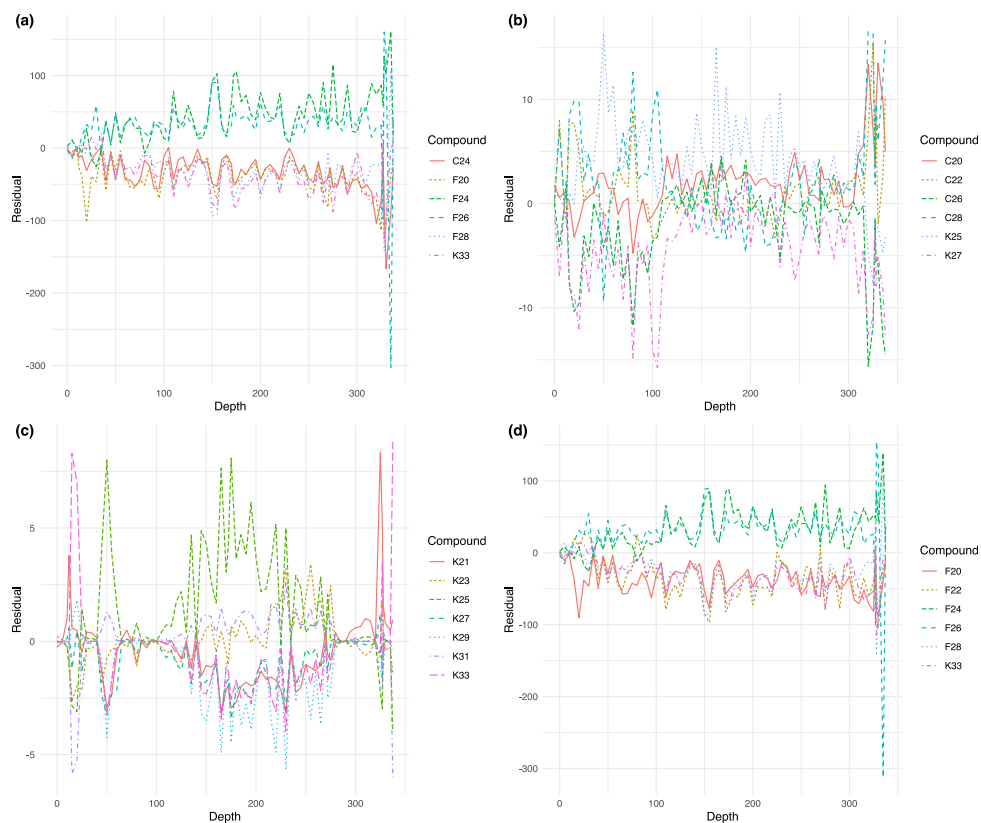
**Fig. 9b.** VERHIB reconstruction for Run 10, showing calculated percentages of vegetation biomass input. Run 10 includes only leaf measurements and only plant-derived chain length measurements for *n*-alkanes and *n*-alkanols. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



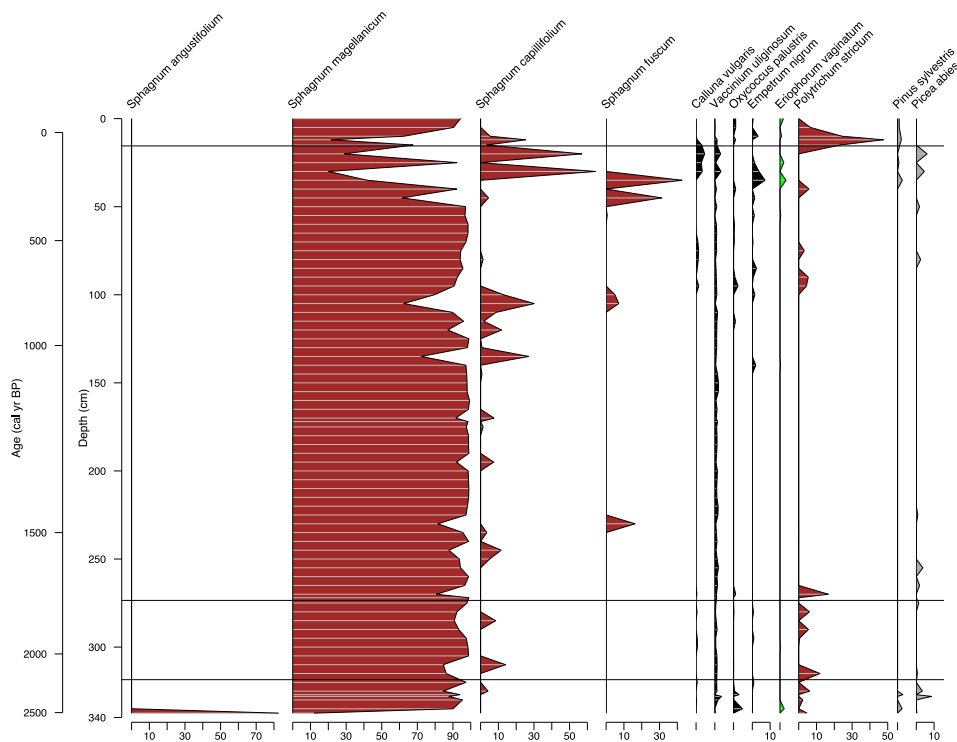
**Fig. 9c.** VERHIB reconstruction for Run 11, showing calculated percentages of vegetation biomass input. Run 11 includes only leaf measurements) and only plant-derived chain length measurements for *n*-alkanes. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



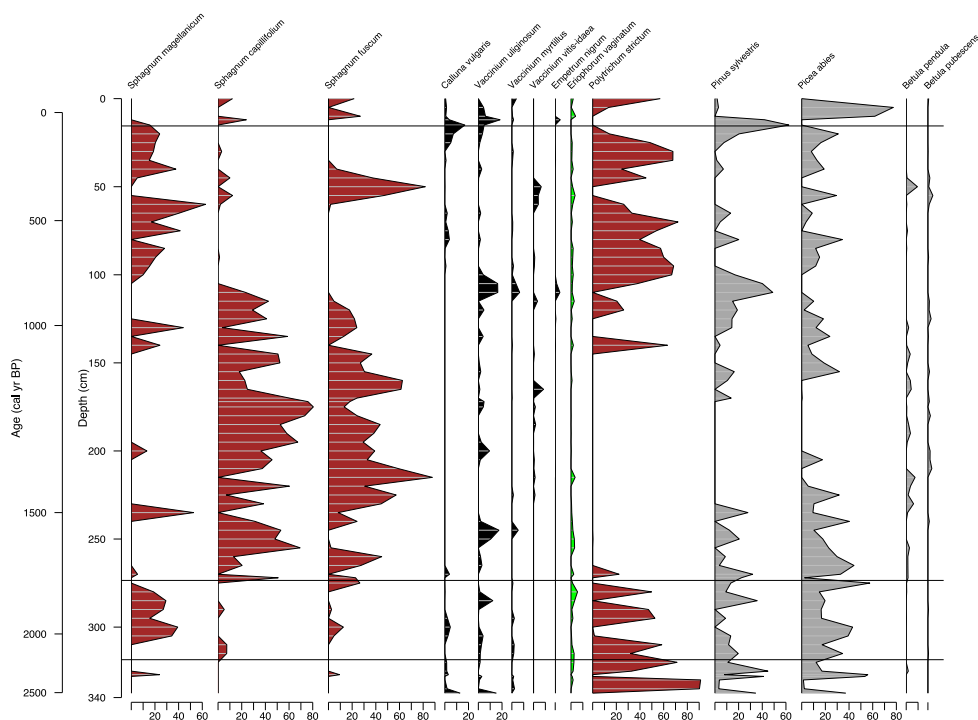
**Fig. 9d.** VERHIB reconstruction for Run 12, showing calculated percentages of vegetation biomass input. Run 12 includes only leaf measurements) and only plant-derived chain length measurements for *n*-alkanes and *n*-fatty acids. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



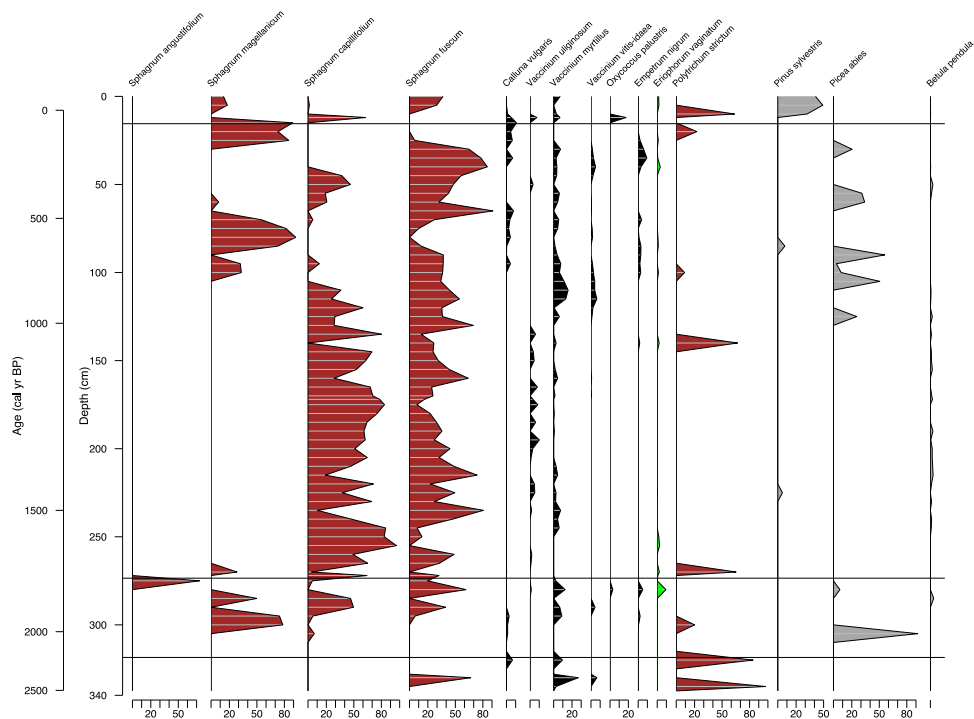
**Fig. 10.** Core measurement residuals versus depth for Run 9(a), Run 10(b), Run 11(c), and Run 12(d), showing compounds with the highest residual values.



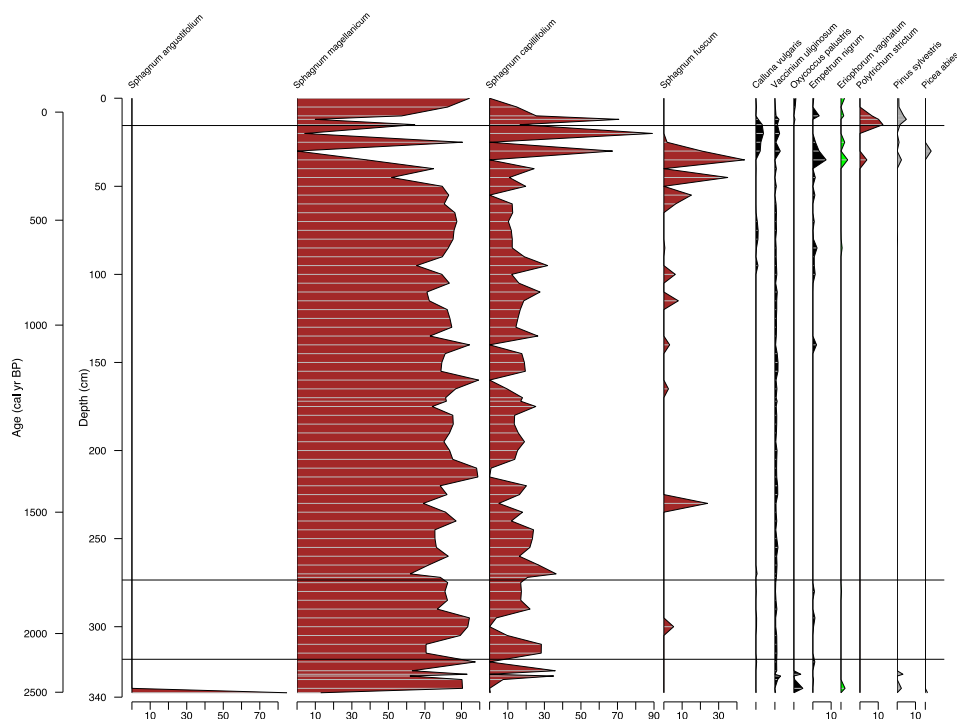
**Fig. 11a.** VERHIB reconstruction for Run 13, showing calculated percentages of vegetation biomass input. Run 13 includes leaf and root measurements (ratio leaf:root is 5:1) and only plant-derived chain length measurements for *n*-alkanes, *n*-alkanols, and *n*-fatty acids. Brown graphs are moss species; black graphs are shrub species; green graph is *E. vaginatum*; and gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 11b.** VERHIB reconstruction for Run 14, showing calculated percentages of vegetation biomass input. Run 14 includes leaf and root measurements (ratio leaf:root is 5:1) and only plant-derived chain length measurements for *n*-alkanes and *n*-alkanols. The brown graphs are moss species; the black graphs are shrub species; the green graph is *E. vaginatum*; and the gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 11c.** VERHIB reconstruction for Run 15, showing calculated percentages of vegetation biomass input. Run 15 includes leaf and root measurements (ratio leaf:root is 5:1) and only plant-derived chain length measurements for *n*-alkanes. The brown graphs are moss species, the black graphs are shrub species, the green graph is *E. vaginatum*; and the gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.



**Fig. 11d.** VERHIB reconstruction for Run 16, showing calculated percentages of vegetation biomass input. Run 16 includes leaf and root measurements (ratio leaf:root is 5:1) and only plant-derived chain length measurements for *n*-alkanes and *n*-fatty acids. The brown graphs are moss species, the black graphs are shrub species, the green graph is *E. vaginatum*; and the gray graphs are tree species. The phases identified in the Beerberg plant macrofossil analysis (Thomas et al., 2023b) are marked.

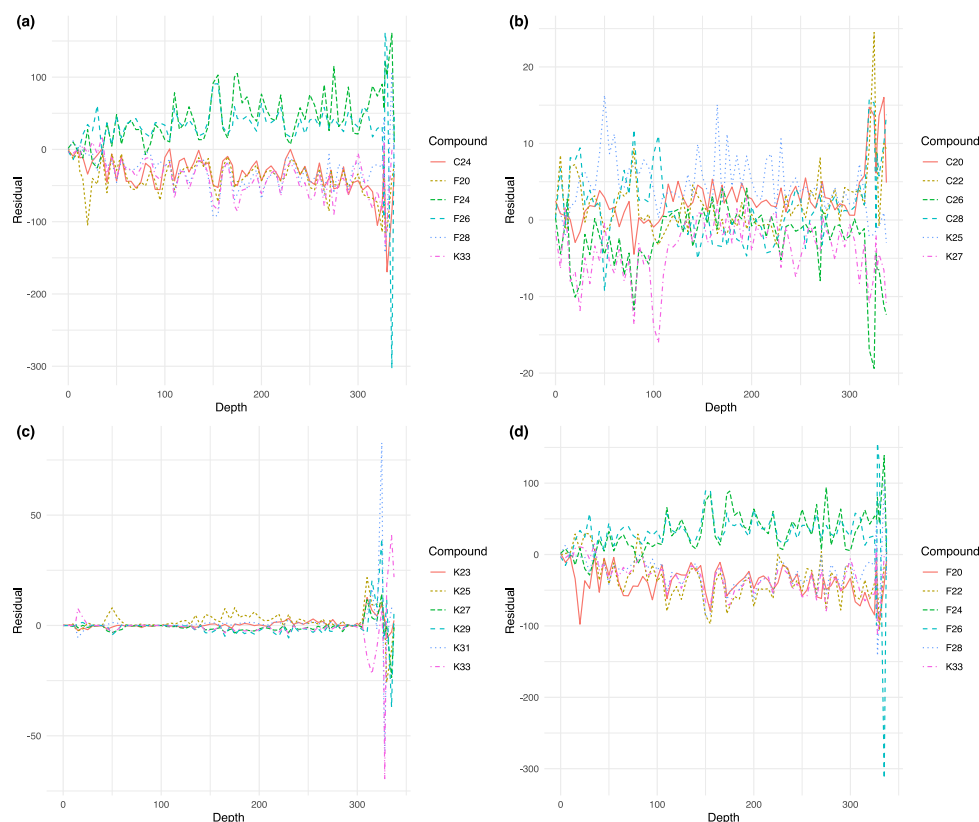


Fig. 12. Core measurement residuals versus depth for Run 13(a), Run 14(b), Run 15(c), and Run 16(d), showing compounds with the highest residual values.

2018–2019 being exceptionally warm with reduced precipitation (Hari et al., 2020). Therefore, it is likely that even the *Sphagnum* species situated in the wetter, hollow areas synthesized longer *n*-alkanes in response to reduced water availability.

As mentioned, there is far less available data for *n*-alkanols and *n*-fatty acids. However, we found measurements published in literature for *S. fuscum*, *S. capillifolium*, and *S. magellanicum*. We found a dominant chain length of  $C_{26}$  *n*-alkanol for *S. fuscum* and  $C_{24}$  *n*-alkanol for *S. capillifolium* (Table 1), which agrees with the findings of Ficken et al. (1998) (British Isles). Baas et al. (2000) (Netherlands) found that *S. magellanicum* had a  $C_{26}$  *n*-alkanol dominant chain length, contrasting our findings of  $C_{24}$  (Table 1). For the *n*-fatty acids, we found that each *Sphagnum* species had a dominant chain length of  $C_{24}$  (Table 1), which agrees with findings for *S. fuscum* (Corrigan et al., 1976; Ficken et al., 1998) and *S. capillifolium* (Ficken et al., 1998) in the British Isles. While Corrigan et al. (1976) also found a  $C_{24}$  dominant chain length for *S. magellanicum* in the British Isles, Baas et al. (2000) found  $C_{26}$  at a site in the Netherlands. As our *n*-alkanol and *n*-fatty acid findings are not as disparate from published data as those of the *n*-alkanes, potentially *n*-alkanol and *n*-fatty acid compositions are less sensitive to climate perturbations.

#### 4.1.2. Shrub species

While we could find at least one published study with *n*-alkanes results for each of the shrub leaves, there is still a relative paucity of data for the compositions of stems and roots as well as for *n*-alkanols and *n*-fatty acids. *Calluna vulgaris* is a common plant in peatland settings and has been measured in multiple studies, with results that generally agree with ours. Salasoo (1987b, 1987a, 1989) measured *C. vulgaris* in Germany (G), Norway (N), and Austria (A) at multiple altitudes (G: Wisselhorst- 65 m, Besenfeld- 815 m, Buchen- 440 m; N: Samsjöen- 530 m, Kjerringvatn- 260 m; A: Hebalpe- 1300 m, Axamer Lizum- 1900 m), finding dominant chain lengths to be either  $C_{31}$  or  $C_{33}$ . Pancost et al. (2002) found that *C. vulgaris* at Bargerveen in the Netherlands

had a  $C_{max}$  in the leaves of  $C_{33}$ , in a stem and leaves combination of  $C_{33}$ , and  $C_{31}$  in the roots. Ali et al. (2005) measured a stem and leaf combination of *C. vulgaris* from Great Britain, determining a dominant chain length of  $C_{31}$ . We also found  $C_{31}$  to be dominant in the leaves and stems but  $C_{33}$  dominant in the roots (Table 1). This causes a relatively high  $P_{wax}$  value for the roots, which could result in complications in interpreting peat cores.

Both Pancost et al. (2002) and Ali et al. (2005) also measured *n*-alkanols and *n*-fatty acids in *C. vulgaris*. Pancost et al. (2002) found  $C_{max}$  of the *n*-alkanols to be  $C_{28}$  in the leaves, a bimodal  $C_{24}$  and  $C_{28}$  in the stems and leaves combination, and  $C_{24}$  in the roots and  $C_{max}$  of the *n*-fatty acids to be  $C_{30}$  in the leaves,  $C_{24}$  in the stems and leaves, and  $C_{28}$  in the roots. Ali et al. (2005) measured a stem and leaves combination of *C. vulgaris* with a  $C_{max}$  of  $C_{22}$  for both the *n*-alkanols and *n*-fatty acids. Our results were more in line with those of Pancost et al. (2002) with a  $C_{max}$  of  $C_{28}$  for *n*-alkanols in the leaves, stems, and roots and  $C_{max}$  of  $C_{32}$  for *n*-fatty acids in the leaves,  $C_{28}$  in the stems, and  $C_{24}$  in the roots (Table 1).

Salasoo (1987b, 1989) also measured *n*-alkanes in *V. myrtillus*, *V. uliginosum*, and *V. vitis-idaea* in both Germany and Austria, as well as *O. palustris* in Austria. Based on  $C_{max}$  and  $ACL_{ALK}$ , the results are similar to our findings of *V. myrtillus* leaves having a dominant chain length of  $C_{31}$ , *V. uliginosum* having a  $C_{max}$  of 27 (though in the Austrian samples, the  $C_{max}$  was 25 (Salasoo, 1989)), and both *V. vitis-idaea* and *O. palustris* having a dominant chain length of  $C_{29}$  (Table 1). Ficken et al. (1998) measured *n*-alkanes, *n*-alcohols, and *n*-fatty acids in *V. vitis-idaea* and *E. nigrum*. The *n*-alkane results reflected ours, with *E. nigrum* having a relatively high *n*-alkane concentration as well as  $C_{31}$  as  $C_{max}$  and *V. vitis-idaea* having  $C_{29}$  as  $C_{max}$ . The *n*-alkanol results differed with Ficken et al. (1998) finding  $C_{30}$  as the  $C_{max}$  for both species compared to our  $C_{26}$  (Table 1). The *n*-fatty acid results for *E. nigrum* agree at  $C_{28}$  as  $C_{max}$ , while those for *V. vitis-idaea* diverge, with Ficken et al. (1998) measuring  $C_{30}$  as the dominant chain length and our study finding  $C_{20}$  (Table 1).

Overall, the shrub results generally agreed with the previously published data we reviewed, and the *n*-alkanes showed variation across species that were consistent across multiple locations. The consistency of plant wax composition in the shrubs with earlier studies in spite of the more recent drought conditions echoes results from a drought experiment measuring *n*-alkanes in *C. vulgaris* that found no changes in ACL and CPI over the course of the experiment (Srivastava and Wiesenberg, 2018).

#### 4.1.3. Tree species

On the Beerberg peatland, we found two angiosperm tree species growing (*B. pendula* and *B. pubescens*) and two gymnosperms (*P. sylvestris* and *P. abies*). As reported in other studies (Diefendorf et al., 2011; Bush and McInerney, 2013; Dasgupta et al., 2022; Jha et al., 2024a), we measured a higher abundance of lipids in the leaves of the angiosperms than the needles of the gymnosperms (Table 1, Fig. 2). The exact driver of this divergence remains poorly understood (Diefendorf et al., 2011; Jha et al., 2024a).

Mayes et al. (1994) measured *n*-alkanes in *B. pubescens* leaves in Norway with a dominant chain length of  $C_{25}$ , followed by  $C_{23}$ . This is in direct contrast with our result of  $C_{31}$  as  $C_{max}$  (Table 1). However, Weber and Schwark (2020) analyzed *B. pubescens* and *B. pendula* leaves from Germany, finding for both species an *n*-alkane dominant chain length of  $C_{31}$ , *n*-alkanol  $C_{max}$  of  $C_{22}$ , and *n*-fatty acid  $C_{max}$  of  $C_{28}$ . This is similar to our results for both species.

Other studies have measured *n*-alkanes, *n*-alkanols, and *n*-fatty acids in coniferous needles. Maffei et al. (2004) found relatively low concentrations of *n*-alkanes in *P. sylvestris* and *P. abies* needles sampled in Italy, as well as  $C_{max}$  of  $C_{31}$  in both species. Ali et al. (2005) measured a needle and stem combination of *P. sylvestris*, finding a  $C_{max}$  of  $C_{27}$  for the *n*-alkanes and  $C_{22}$  for the *n*-alkanols and *n*-fatty acids. In North America, Bush and McInerney (2013) measured *n*-alkanes in *P. abies* and *P. sylvestris*, finding a  $C_{max}$  of  $C_{25}$  and  $C_{29}$ , respectively. In our study, we found  $C_{29}$  as the  $C_{max}$  for both in the needles (Table 1). Our results for *n*-alkanols and *n*-fatty acids were comparable to those of Ali et al. (2005).

As we found in the tree species at Beerberg peatland, higher *n*-alkane chain-lengths ( $C_{27}$ – $C_{35}$ ) are typically the dominant chain length in both angiosperm leaves and gymnosperm needles across a range of species and locations (Diefendorf et al., 2011; Dasgupta et al., 2022; Jha et al., 2024a). Therefore, it is difficult to elucidate from our limited tree species whether there was a potential drought response similar to that in some of the *Sphagnum* species.

#### 4.1.4. Overlap between plant species and parts

Our results from the biomarker analysis of the modern plant samples show that signatures across plant groups and parts exhibit some overlap, as found in previous studies (e.g., Mueller et al., 2012). This increases the difficulty of organic matter source apportionment based on biomarker distributions. In our assemblage of plants, the most potentially problematic overlaps occurred in the *n*-alkane measurements and ratios. The  $C_{23}/C_{25}$  ratio has been previously suggested to be connected to bog wetness because, as previously mentioned, it has been observed that *Sphagnum* species occupying drier hummock habitats (e.g., *S. fuscum* and *S. capillifolium*) are dominated by higher chain lengths such as  $C_{25}$  or  $C_{31}$ , while species in wetter hollow habitats are dominated by  $C_{23}$  (Bingham et al., 2010). Our study found that all the *Sphagnum* mosses were dominated by  $C_{31}$ , meaning that the  $C_{23}/C_{25}$  ratio is likely less effective at the Beerberg peatland.  $C_{31}$  is also used in the  $P_{aq}$  and  $P_{wax}$  ratios, and the dominance of  $C_{31}$  in *Sphagnum* mosses and some of the root samples likely contributed to a few of the unexpected results for the two ratios that would make it difficult to determine biomarker inputs accurately based only on the ratios. However, the  $P_{aq}$  and  $P_{wax}$  ratios were important contributors to variance in the PCAs of both the plant samples and the core samples, along with the  $C_{25}/(C_{25}+C_{29})$  ratios. Our results confirm the importance of having a localized library

of plant biomarker data to compare to sedimentary archives, as well as including a variety of index measurements or ratios in case the local biomarker signatures of the plants render the ratios less effective. Nonetheless, it is important to note that a local database of plant-derived biomarker signatures is not entirely sufficient on its own to characterize preserved organic matter as diagenetic processes, such as microbial reworking, can alter the biomarker signal and plants are not the only organisms contributing biomarkers to sedimentary archives.

#### 4.2. Evaluating the VERHIB model reconstruction

The VERHIB model results provide insight into the modeling process and the biomarker data. All the runs struggled to capture an accurate representation of the transitions seen not only in the plant macrofossil records but also with temporal offsets in the pollen record and the core elemental and biomarker data (Thomas et al., 2023b). Based on the plant macrofossils, four phases were identified: 2528–2251 cal yr BP (340–318.5 cm), 2251–1671 cal yr BP (318.5–273.5 cm), 1671–64 cal yr BP (273.5–15.5 cm), and 64 cal yr BP - Present (15.5–0 cm) (see Supplementary Material for method details of radiocarbon dating and macrofossil phase identification). In Phase I, the vegetation was predominately poor fen before going through paludification and transitioning to a *Sphagnum* bog in Phase II. The peatland was relatively stable through Phase III until Phase IV, when there was an increase in some fen species like *E. vaginatum* and a change in dominant *Sphagnum* species from *S. fuscum* to *S. medium/divinum* (Thomas et al., 2023b).

The reconstructions closest to the plant macrofossil record included all chain lengths and *n*-fatty acids (Run 1, Run 4, Run 5, Run 8) (Figs. 4a, 4d, 6a, 6d). These reconstructions estimated that *S. fuscum* was the dominant moss through most of the core, transitioning only in the last phase to a different species. Additionally, these reconstructions showed a higher proportion of Ericaceae shrubs at the bottom of the core and closer to the top. However, the proportion of *E. vaginatum* was underestimated, particularly at the base of the core, potentially the cause of the higher residuals at lower depths (Figs. 5a, 5d, 7a, 7d). When only plant-derived chain lengths were included, *S. fuscum* was not recognized as the dominant moss, and when *n*-fatty acid data was excluded, the proportion of tree species was overestimated. The VERHIB model was previously only used with *n*-alkane and *n*-alkanol data, so this is the first time that *n*-fatty acid data has been shown to be a valuable inclusion. In these reconstructions, including root data did not alter the result significantly, but this is logical as the peatland vegetation was dominated by *Sphagnum* moss, and there was likely not a lot of root input from other plants.

We additionally considered the Pearson correlation coefficients between the predicted peat core biomarker values and the actual values. What we found echoed the importance of including *n*-fatty acid data and all chain lengths, not only plant-derived. The Pearson correlation coefficients were over 0.9 for all model runs except 1, 4, and 5. This result shows that these runs (without the *n*-fatty acid data) were overparameterized and that it was only the regularization coefficients preventing the model from having a non-unique solution.

Although the VERHIB reconstructions deviated from the pollen and macrofossil analyses, particularly in missing the poor fen vegetation likely present in Phase I, the results are still promising and provide some insight into how the model might be adapted in future research.

#### 4.3. Future recommendations for the VERHIB model

There are a few opportunities for improving the VERHIB model that could increase its accuracy and applicability to more study sites, as well as its usability for a wider group of researchers. As MATLAB requires a license, developing a version of the VERHIB model in R would allow for more widespread use as R is open-source and more prevalent today in ecological and environmental research. An R version could be more easily expandable to include further compound classes,

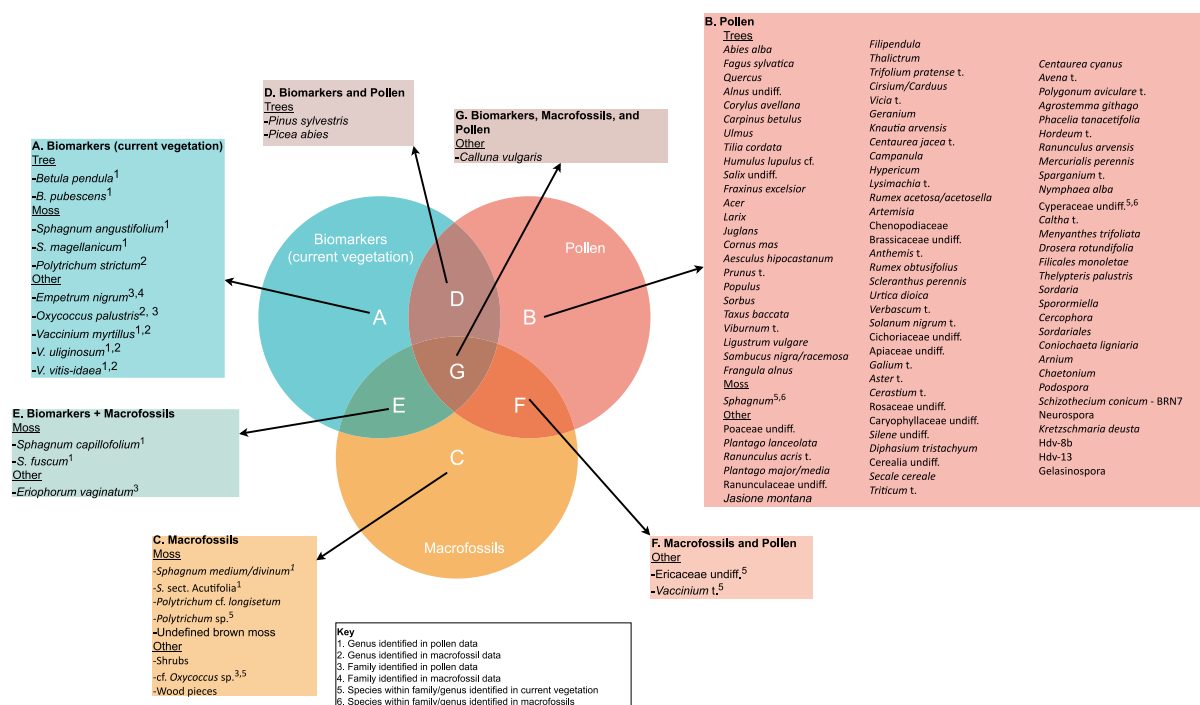


Fig. 13. Comparison of species present in current vegetation at Beerberg peatland with those found in the plant macrofossil and pollen records.

in addition to the *n*-fatty acids we added and tested here. Terpinoids have particularly high potential to improve the performance of the model as many studies have found that they have a high level of chemotaxonomic specificity (e.g., Otto and Simoneit, 2001; Diefendorf et al., 2014; Lemma et al., 2024). The addition of lignin monomers could also improve species specificity in the model as they also enable distinction between woody and non-woody plants as well as between angiosperms and gymnosperms (Hedges and Mann, 1979; Thevenot et al., 2010). Furthermore, considering the compound specific isotopes of plant waxes would facilitate more precise interpretations of both vegetation community and climate shifts (Diefendorf and Freimuth, 2017).

A further significant improvement for the model would be the addition of a module allowing plant macrofossil or pollen data to be considered. As seen in Fig. 13, there are a lot of overlaps between current vegetation species found at the peatland and with the plant macrofossil record. The pollen record contains many other species, but this is not unexpected as pollen is typically airborne, and the record is therefore a regional rather than local signal. Other models using biomarkers have been more recently developed with Bayesian approaches (Yang and Bowen, 2022) that allow for consideration of prior knowledge to inform the final estimate. Our study shows that many species and plant parts have overlapping biomarker signatures that complicate the source apportionment process. If the model had access to other proxy data in the form of plant macrofossils or pollen, the biomarker-based reconstruction could be more informed by which species are more likely to have been present at different points in time. There has also been promising research into using machine learning techniques with biomarker data from lake sediments (People et al., 2021), and this could be an additional avenue for improvement of the VERHIB model following the development of a more widely accessible version.

## 5. Conclusion

Our study echoes what has been found before, that while straight-chain lipid biomarkers are capable of serving as a proxy for paleovegetation and paleoclimate conditions, on their own, they are often not precise enough to provide a reconstruction that is specific at species or family level and work best when augmented by other proxies, such as pollen or macrofossils. The lack of precision results from overlapping biomarker signatures across plant species and plant organs, as well as the inherent uncertainty of using modern plant analogues to evaluate source apportionment for past vegetation growing under likely different climate conditions. Underscoring this, we found many examples of divergences from previously reported biomarker signals in our analysis of modern plants from the Beerberg peatland, including the key *Sphagnum* species. One avenue for mitigating the overlap in signatures would be to expand the biomarker compounds that could be included in the model, as some classes such as terpinoids or lignin monomers have been found to have improved chemotaxonomic specificity over the straight-chain lipids. The inclusion of compound specific isotope signatures could also aid in more specific source identification. A further potential next step for increasing the rigor and quantitiveness of multi-proxy studies using biomarkers would be to create a model capable of integrating plant macrofossil or pollen data into a biomarker-based reconstruction to allow for more probable and robust outcomes.

Additionally, it is essential to develop a library of plant biomarker signatures that is as localized as possible to the site under investigation. Furthermore, an easily accessible database of published plant biomarker signatures will enable researchers to better choose if they need to analyze plants at their site or if sufficient data is available from plant species in the region. This could also allow for a better understanding of how climate and environmental variables affect biomarker signatures, enabling correction for such effects when using modern plant analogues for a paleoreconstruction.



Finally, as the current study is limited to a *Sphagnum*-dominated peatland that remained relatively stable in terms of vegetation communities through its development, further testing of the VERHIB model in other environments and sedimentary archives is ongoing and shows promise.

### CRedit authorship contribution statement

**Carrie L. Thomas:** Writing – review & editing, Writing – original draft, Visualization, Investigation, Formal analysis, Data curation, Conceptualization. **Boris Jansen:** Writing – review & editing, Supervision, Conceptualization. **E. Emiel van Loon:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Guido L.B. Wiesenberg:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

### Funding support

This research has been supported by the Swiss National Science Foundation (grant no. 188684).

### Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Guido L. B. Wiesenberg reports financial support was provided by Swiss National Science Foundation. Carrie L. Thomas reports financial support was provided by swissuniversities. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Appendix. Supplementary material

#### A.1. Supplementary methods

##### A.1.1. Radiocarbon dating

Plant remains were handpicked from 12 depths (7.5, 16.5, 34.5, 54.5, 69.5, 124.5, 174.5, 258.5, 278.5, 293.5, 314.5, and 334–336 cm). These samples underwent an acid–alkali–acid cleaning process and were combusted at 900 °C to generate CO<sub>2</sub>, which was subsequently reduced to graphite. The carbon isotope composition was analyzed using accelerator mass spectrometry (AMS) at the Institute of Ion Beam Physics, part of the Swiss Federal Institute of Technology in Zurich, Switzerland, at the 0.2 MV MICADAS facility.

##### A.1.2. Age-depth model and phase identification

An age–depth model of the radiocarbon dates was constructed using the Bacon model (Blaauw and Christen, 2011), implemented via the rbacon package (Blaauw et al., 2021). For a detailed description of the model, see Thomas et al. (2023b). All dates mentioned in subsequent sections represent the mean values calculated by the age–depth model at 1 cm resolution.

Using macrofossil abundance data from the peat core (see Thomas et al. (2023b)), phases of differing vegetation composition were identified through constrained hierarchical clustering (CONISS; Grimm (1987)). We used the vegan (Oksanen et al., 2020) and rioja (Juggins, 2020) R packages. Using the *chclust* function, a CONISS analysis was performed on the Bray–Curtis dissimilarity of the macrofossil abundances; the clusters were constrained by depth. The number of phases was determined by the broken-stick model (MacArthur, 1957; Bennett, 1996) using the *bstick* function.

#### A.2. Supplementary results

See Table A.3.

#### Data availability

The data used has already been published at the Pangaea data repository and the DOI is provided in the manuscript. The code will be made available upon request.

#### References

- Ali, H.A.M., Mayes, R.W., Hector, B.L., Verma, A.K., Ørskov, E.R., 2005. The possible use of *n*-alkanes, long-chain fatty alcohols and long-chain fatty acids as markers in studies of the botanical composition of the diet of free-ranging herbivores. *J. Agric. Sci.* 143 (1), 85–95. <http://dx.doi.org/10.1017/S0021859605004958>.
- Andersson, R.A., Kuhry, P., Meyers, P., Zebühr, Y., Crill, P., Mörth, M., 2011. Impacts of paleohydrological changes on *n*-alkane biomarker compositions of a Holocene peat sequence in the eastern European Russian Arctic. *Org. Geochem.* 42 (9), 1065–1075. <http://dx.doi.org/10.1016/j.orggeochem.2011.06.020>.
- Baas, M., Pancost, R., Van Geel, B., Sissinghe Damsté, J.S., 2000. A comparative study of lipids in *Sphagnum* species. *Org. Geochem.* 31 (6), 535–541. [http://dx.doi.org/10.1016/S0146-6380\(00\)00037-1](http://dx.doi.org/10.1016/S0146-6380(00)00037-1).
- Barnes, M., Barnes, W., 1978. Organic compounds in lake sediments. In: *Lakes: Chemistry, Geology, Physics*. Springer, pp. 127–152.
- Bennett, K.D., 1996. Determination of the number of zones in a biostratigraphical sequence. *New Phytol.* 132 (1), 155–170. <http://dx.doi.org/10.1111/j.1469-8137.1996.tb04521.x>.
- Bingham, E.M., McClymont, E.L., Väiliranta, M., Mauquoy, D., Roberts, Z., Chambers, F.M., Pancost, R.D., Evershed, R.P., 2010. Conservative composition of *n*-alkane biomarkers in *Sphagnum* species: Implications for palaeoclimate reconstruction in ombrotrophic peat bogs. *Org. Geochem.* 41 (2), 214–220. <http://dx.doi.org/10.1016/j.orggeochem.2009.06.010>.
- Blaauw, M., Christen, J.A., 2011. Flexible paleoclimate age-depth models using an autoregressive gamma process. *Bayesian Anal.* 6 (3), <http://dx.doi.org/10.1214/11-BA618>.
- Blaauw, M., Christen, J.A., Aquino López, M.A., 2021. Rbacon: Age-depth modelling using Bayesian statistics. URL <https://CRAN.R-project.org/package=rbacon>. R package version 2.5.2.
- Büntgen, U., Urban, O., Krusic, P.J., Rybníček, M., Kolář, T., Kyncl, T., Ač, A., Koňasová, E., Čáslavský, J., Esper, J., Wagner, S., Saurer, M., Tegel, W., Dobrovolný, P., Cherubini, P., Reinig, F., Trnka, M., 2021. Recent European drought extremes beyond common era background variability. *Nat. Geosci.* 14 (4), 190–196. <http://dx.doi.org/10.1038/s41561-021-00698-0>.
- Bush, R.T., McInerney, F.A., 2013. Leaf wax *n*-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Acta* 117, 161–179. <http://dx.doi.org/10.1016/j.gca.2013.04.016>.
- Castañeda, I.S., Schouten, S., 2011. A review of molecular organic proxies for examining modern and ancient lacustrine environments. *Quat. Sci. Rev.* 30 (21–22), 2851–2891. <http://dx.doi.org/10.1016/j.quascirev.2011.07.009>.
- Corrigan, D., Kloos, C., O'Connor, C., Timoney, R., 1973. Alkanes from four species of *Sphagnum* moss. *Phytochemistry* 12 (1), 213–214. [http://dx.doi.org/10.1016/S0031-9422\(00\)84653-1](http://dx.doi.org/10.1016/S0031-9422(00)84653-1).
- Corrigan, D., Kloos, C., O'Connor, C., Timoney, R., 1976. Lipid components of *Sphagnum* mosses. *Planta Med.* 29 (03), 261–267. <http://dx.doi.org/10.1055/s-0028-1097660>.
- Cranwell, P.A., 1981. Diagenesis of free and bound lipids in terrestrial detritus deposited in a lacustrine sediment. *Org. Geochem.* 3 (3), 79–89. [http://dx.doi.org/10.1016/0146-6380\(81\)90002-4](http://dx.doi.org/10.1016/0146-6380(81)90002-4).
- Cranwell, P.A., 1984. Lipid geochemistry of sediments from Upton Broad, a small productive lake. *Org. Geochem.* 7 (1), 25–37. [http://dx.doi.org/10.1016/0146-6380\(84\)90134-7](http://dx.doi.org/10.1016/0146-6380(84)90134-7).
- Cranwell, P.A., Eglinton, G., Robinson, N., 1987. Lipids of aquatic organisms as potential contributors to lacustrine sediments-II. *Org. Geochem.* 11 (6), 513–527. [http://dx.doi.org/10.1016/0146-6380\(87\)90007-6](http://dx.doi.org/10.1016/0146-6380(87)90007-6).
- Crausbay, S., Genderjahn, S., Hotchkiss, S., Sachse, D., Kahmen, A., Arndt, S.K., 2014. Vegetation dynamics at the upper reaches of a tropical montane forest are driven by disturbance over the past 7300 years. *Arct. Antarct. Alp. Res.* 46 (4), 787–799. <http://dx.doi.org/10.1657/1938-4246-46.4.787>.
- Dasgupta, B., Ajay, A., Prakash, P., Sanyal, P., 2022. Understanding the disparity in *n*-alkane production among angiosperms and gymnosperms from the higher Himalayas: Inferences drawn from a machine learning approach. *Org. Geochem.* 171, 104463. <http://dx.doi.org/10.1016/j.orggeochem.2022.104463>.
- Diefendorf, A.F., Freeman, K.H., Wing, S.L., 2014. A comparison of terpenoid and leaf fossil vegetation proxies in Paleocene and Eocene Bighorn Basin sediments. *Org. Geochem.* 71, 30–42. <http://dx.doi.org/10.1016/j.orggeochem.2014.04.004>.

- Diefendorf, A.F., Freeman, K.H., Wing, S.L., Graham, H.V., 2011. Production of *n*-alkyl lipids in living plants and implications for the geologic past. *Geochim. Cosmochim. Acta* 75 (23), 7472–7485. <http://dx.doi.org/10.1016/j.gca.2011.09.028>.
- Diefendorf, A.F., Freimuth, E.J., 2017. Extracting the most from terrestrial plant-derived *n*-alkyl lipids and their carbon isotopes from the sedimentary record: A review. *Org. Geochem.* 103, 1–21. <http://dx.doi.org/10.1016/j.orggeochem.2016.10.016>.
- Eglinton, G., Hamilton, R.J., 1967. Leaf epicuticular waxes. *Science* 156, <http://dx.doi.org/10.1126/science.156.3780.1322>.
- Eglinton, G., Logan, G.A., 1991. Molecular preservation. *Philos. Trans. R. Soc. London [Biol]* 333 (1268), 315–328. <http://dx.doi.org/10.1098/rstb.1991.0081>.
- Erb, K.H., Kastner, T., Plutzer, C., Bais, A.L.S., Carvalhais, N., Fetzel, T., Gingrich, S., Haberl, H., Lauk, C., Niederscheider, M., Pongratz, J., Thurner, M., Luysaert, S., 2018. Unexpectedly large impact of forest management and grazing on global vegetation biomass. *Nature* 553 (7686), 73–76. <http://dx.doi.org/10.1038/nature25138>.
- Fick, S.E., Hijmans, R.J., 2017. Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* 37 (12), 4302–4315. <http://dx.doi.org/10.1002/joc.5086>.
- Ficken, K.J., Barber, K.E., Eglinton, G., 1998. Lipid biomarker,  $\delta^{13}\text{C}$  and plant macrofossil stratigraphy of a scottish montane peat bog over the last two millennia. *Org. Geochem.* 28 (3), 217–237. [http://dx.doi.org/10.1016/S0146-6380\(97\)00126-5](http://dx.doi.org/10.1016/S0146-6380(97)00126-5).
- Ficken, K.J., Li, B., Swain, D.L., Eglinton, G., 2000. An *n*-alkane proxy for the sedimentary input of submerged/floating freshwater aquatic macrophytes. *Org. Geochem.* 31 (7), 745–749. [http://dx.doi.org/10.1016/S0146-6380\(00\)00081-4](http://dx.doi.org/10.1016/S0146-6380(00)00081-4).
- Gocke, M.I., Kuzuyakov, Y., Wiesenberg, G.L.B., 2010. Rhizoliths in loess – evidence for post-sedimentary incorporation of root-derived organic matter in terrestrial sediments as assessed from molecular proxies. *Org. Geochem.* 41 (11), 1198–1206. <http://dx.doi.org/10.1016/j.orggeochem.2010.08.001>.
- Grimm, E.C., 1987. CONISS: a FORTRAN 77 program for stratigraphically constrained cluster analysis by the method of incremental sum of squares. *Comput. Geosci.* 13 (1), 13–35. [http://dx.doi.org/10.1016/0098-3004\(87\)90022-7](http://dx.doi.org/10.1016/0098-3004(87)90022-7).
- Hari, V., Rakovec, O., Markonis, Y., Hanel, M., Kumar, R., 2020. Increased future occurrences of the exceptional 2018–2019 Central European drought under global warming. *Sci. Rep.* 10 (1), 12207. <http://dx.doi.org/10.1038/s41598-020-68872-9>.
- Hedges, J.I., Mann, D.C., 1979. The characterization of plant tissues by their lignin oxidation products. *Geochim. Cosmochim. Acta* 43 (11), 1803–1807. [http://dx.doi.org/10.1016/0016-7037\(79\)90028-0](http://dx.doi.org/10.1016/0016-7037(79)90028-0).
- Hirave, P., Wiesenberg, G.L.B., Birkholz, A., Alewell, C., 2020. Understanding the effects of early degradation on isotopic tracers: implications for sediment source attribution using compound-specific isotope analysis (CSIA). *Biogeosciences* 17 (8), 2169–2180. <http://dx.doi.org/10.5194/bg-17-2169-2020>.
- Husson, F., Josse, J., Le, S., Mazet, J., 2023. FactoMineR: Multivariate exploratory data analysis and data mining. URL <https://cran.r-project.org/web/packages/FactoMineR/index.html>.
- Inglis, G.N., Bhattacharya, T., Hemingway, J.D., Hollingsworth, E.H., Feakins, S.J., Tierney, J.E., 2022. Biomarker approaches for reconstructing terrestrial environmental change. *Annu. Rev. Earth Planet. Sci.* 50 (1), 369–394. <http://dx.doi.org/10.1146/annurev-earth-032320-095943>.
- Jackson, R.B., Lajtha, K., Crow, S.E., Hugelius, G., Kramer, M.G., Piñeiro, G., 2017. The ecology of soil carbon: Pools, vulnerabilities, and Biotic and abiotic controls. *Annu. Rev. Ecol. Evol. Syst.* 48 (1), 419–445. <http://dx.doi.org/10.1146/annurev-ecolsys-112414-054234>.
- Jansen, B., De Boer, E.J., Cleef, A.M., Hooghiemstra, H., Moscol-Olivera, M., Tonneijck, F.H., Verstraten, J.M., 2013. Reconstruction of late Holocene forest dynamics in northern Ecuador from biomarkers and pollen in soil cores. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 386, 607–619. <http://dx.doi.org/10.1016/j.palaeo.2013.06.027>.
- Jansen, B., Haussmann, N.S., Tonneijck, F.H., Verstraten, J.M., De Voogt, P., 2008. Characteristic straight-chain lipid ratios as a quick method to assess past forest-páramo transitions in the Ecuadorian Andes. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 262 (3–4), 129–139. <http://dx.doi.org/10.1016/j.palaeo.2008.02.007>.
- Jansen, B., Nierop, K.G.J., Hageman, J.A., Cleef, A.M., Verstraten, J.M., 2006. The straight-chain lipid biomarker composition of plant species responsible for the dominant biomass production along two altitudinal transects in the Ecuadorian Andes. *Org. Geochem.* 37 (11), 1514–1536. <http://dx.doi.org/10.1016/j.orggeochem.2006.06.018>.
- Jansen, B., van Loon, E.E., Hooghiemstra, H., Verstraten, J.M., 2010. Improved reconstruction of palaeo-environments through unravelling of preserved vegetation biomarker patterns. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 285 (1–2), 119–130. <http://dx.doi.org/10.1016/j.palaeo.2009.10.029>.
- Jansen, B., Wiesenberg, G.L.B., 2017. Opportunities and limitations related to the application of plant-derived lipid molecular proxies in soil science. *Soil* 3 (4), 211–234. <http://dx.doi.org/10.5194/soil-3-211-2017>.
- Jha, D.K., Hirave, P., Ghosh, S., Dasgupta, B., Sanyal, P., 2024a. Does leaf wax isotopic characterisation of gymnosperms and angiosperms capture environmental gradients in Himalayas? *Org. Geochem.* 187, 104720. <http://dx.doi.org/10.1016/j.orggeochem.2023.104720>.
- Jha, D.K., Patalano, R., Ilgner, J., Achyuthan, H., Alsharekh, A.M., Armitage, S., Blinkhorn, J., Boivin, N., Breeze, P.S., Devra, R., Drake, N., Groucutt, H.S., Guagnin, M., Roberts, P., Petraglia, M., 2024b. Preservation of plant-wax biomarkers in deserts: implications for quaternary environment and human evolutionary studies. *J. Quat. Sci.* 39 (3), 349–358. <http://dx.doi.org/10.1002/jqs.3597>.
- Juggins, S., 2020. rioja: Analysis of Quaternary science data. URL <https://cran.r-project.org/package=rioja>. R package version 0.9-26.
- Kassambara, A., Mundt, F., 2020. factoextra: Extract and visualize the results of multivariate data analyses. URL <https://cran.r-project.org/web/packages/factoextra/index.html>.
- Knief, C., Bol, R., Amelung, W., Kusch, S., Frindte, K., Eckmeier, E., Jaeschke, A., Dunai, T., Fuentes, B., Mörchen, R., Schütte, T., Lücke, A., Klumpp, E., Kaiser, K., Rethemeyer, J., 2020. Tracing elevational changes in microbial life and organic carbon sources in soils of the Atacama Desert. *Glob. Planet. Change* 184, 103078. <http://dx.doi.org/10.1016/j.gloplacha.2019.103078>.
- Lavorel, S., Díaz, S., Cornelissen, J.H.C., Garnier, E., Harrison, S.P., McIntyre, S., Pausas, J.G., Pérez-Harguindeguy, N., Roumet, C., Urcelay, C., 2007. Plant functional types: Are we getting any closer to the holy grail? In: Canadell, J.G., Pataki, D.E., Pitelka, L.F. (Eds.), *Terrestrial Ecosystems in a Changing World*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 149–164. <http://dx.doi.org/10.1007/978-3-540-32730-1-13>.
- Lemma, B., Bromm, T., Zech, W., Zech, M., Nemomissa, S., Glaser, B., 2024. Terpenoid profiling of keystone plant species of the Bale Mountains, Ethiopia: Implications for chemotaxonomy and paleovegetation studies. *Biochem. Sys. Ecol.* 116, 104865. <http://dx.doi.org/10.1016/j.bse.2024.104865>.
- MacArthur, R.H., 1957. On the relative abundance of bird species. *Proc. Natl. Acad. Sci.* 43 (3), 293–295. <http://dx.doi.org/10.1073/pnas.43.3.293>.
- Maffei, M., Badino, S., Bossi, S., 2004. Chemotaxonomic significance of leaf wax *n*-alkanes in the Pinales (Coniferales). *J. Biol. Res.* 1, 3–19.
- Mambelli, S., Bird, J.A., Gleixner, G., Dawson, T.E., Torn, M.S., 2011. Relative contribution of foliar and fine root pine litter to the molecular composition of soil organic matter after in situ degradation. *Org. Geochem.* 42, <http://dx.doi.org/10.1016/j.orggeochem.2011.06.008>.
- Marzi, R., Torkelson, B., Olson, R., 1993. A revised carbon preference index. *Org. Geochem.* 20 (8), 1303–1306. [http://dx.doi.org/10.1016/0146-6380\(93\)90016-5](http://dx.doi.org/10.1016/0146-6380(93)90016-5).
- Mayer, R.W., Beresford, N.A., Lamb, C.S., Barnett, C.L., Howard, B.J., Jones, B.E.V., Eriksson, O., Hove, K., Pedersen, Ø., Staines, B.W., 1994. Novel approaches to the estimation of intake and bioavailability of radiocaesium in ruminants grazing forested areas. *Sci. Total Environ.* 157, 289–300. [http://dx.doi.org/10.1016/0048-9697\(94\)90592-4](http://dx.doi.org/10.1016/0048-9697(94)90592-4).
- McClymont, E.L., Mauquoy, D., Yeloff, D., Broekens, P., Van Geel, B., Charman, D.J., Pancost, R.D., Chambers, F.M., Evershed, R.P., 2008. The disappearance of *Sphagnum imbricatum* from Butterburn Flow, UK. *Holocene* 18 (6), 991–1002. <http://dx.doi.org/10.1177/0959683608093537>.
- Mueller, K.E., Polissar, P.J., Oleksyn, J., Freeman, K.H., 2012. Differentiating temperate tree species and their organs using lipid biomarkers in leaves, roots and soil. *Org. Geochem.* 52, 130–141. <http://dx.doi.org/10.1016/j.orggeochem.2012.08.014>.
- Nichols, J.E., Booth, R.K., Jackson, S.T., Pendall, E.G., Huang, Y., 2006. Paleohydrologic reconstruction based on *n*-alkane distributions in ombrotrophic peat. *Org. Geochem.* 37 (11), 1505–1513. <http://dx.doi.org/10.1016/j.orggeochem.2006.06.020>.
- Nott, C.J., Xie, S., Avsejs, L.A., Maddy, D., Chambers, F.M., Evershed, R.P., 2000. *n*-Alkane distributions in ombrotrophic mires as indicators of vegetation change related to climatic variation. *Org. Geochem.* 31 (2), 231–235. [http://dx.doi.org/10.1016/S0146-6380\(99\)00153-9](http://dx.doi.org/10.1016/S0146-6380(99)00153-9).
- Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., 2020. vegan: Community ecology package. URL <https://CRAN.R-project.org/package=vegan>. R package version 2.5-7.
- Otto, A., Simoneit, B.R., 2001. Chemosystematics and diagenesis of terpenoids in fossil conifer species and sediment from the Eocene Zeitz formation, Saxony, Germany. *Geochim. Cosmochim. Acta* 65 (20), 3505–3527. [http://dx.doi.org/10.1016/S0016-7037\(01\)00693-7](http://dx.doi.org/10.1016/S0016-7037(01)00693-7).
- Pancost, R.D., Baas, M., Van Geel, B., Sinnighe Damsté, J.S., 2002. Biomarkers as proxies for plant inputs to peats: an example from a sub-boreal ombrotrophic bog. *Org. Geochem.* 33 (7), 675–690. [http://dx.doi.org/10.1016/S0146-6380\(02\)00048-7](http://dx.doi.org/10.1016/S0146-6380(02)00048-7).
- Parducci, L., Välaranta, M., Salonen, J.S., Ronkainen, T., Matetovici, I., Fontana, S.L., Eskola, T., Sarala, P., Suyama, Y., 2015. Proxy comparison in ancient peat sediments: pollen, macrofossil and plant DNA. *Phil. Trans. R. Soc. B* 370 (1660), 20130382. <http://dx.doi.org/10.1098/rstb.2013.0382>.
- People, M.D., Tierney, J.E., McGee, D., Lowenstein, T.K., Bhattacharya, T., Feakins, S.J., 2021. Identifying plant wax inputs in lake sediments using machine learning. *Org. Geochem.* 156, 104222. <http://dx.doi.org/10.1016/j.orggeochem.2021.104222>.
- Peters, K.E., Walters, C.C., Moldovan, J.M., 2005. *The Biomarker Guide*. Cambridge University Press.
- Poynter, J.G., Farrimond, P., Robinson, N., Eglinton, G., 1989. Aeolian-derived higher plant lipids in the marine sedimentary record: Links with palaeoclimate. In: Leinen, M., Sarnthein, M. (Eds.), *Paleoclimatology and Paleometeorology: Modern and Past Patterns of Global Atmospheric Transport*. Springer Netherlands, Dordrecht, pp. 435–462. <http://dx.doi.org/10.1007/978-94-009-0995-3-18>.

- Ronkainen, T., McClymont, E.L., Väiliranta, M., Tuittila, E.S., 2013. The *n*-alkane and sterol composition of living fen plants as a potential tool for palaeoecological studies. *Org. Geochem.* 59, 1–9. <http://dx.doi.org/10.1016/j.orggeochem.2013.03.005>.
- Rydin, H., Jeglum, J.K., Bennett, K.D., 2013. *The Biology of Peatlands*, 2nd Edition OUP Oxford.
- Salasoo, I., 1987a. Alkane distribution in epicuticular wax of some heath plants in Norway. *Biochem. Sys. Ecol.* 15 (6), 663–665. [http://dx.doi.org/10.1016/0305-1978\(87\)90041-X](http://dx.doi.org/10.1016/0305-1978(87)90041-X).
- Salasoo, I., 1987b. Epicuticular wax hydrocarbons of Ericaceae in Germany. *Z. Naturforschung C* 42 (5), 499–501. <http://dx.doi.org/10.1515/znc-1987-0501>.
- Salasoo, I., 1989. Epicuticular wax alkanes of Ericaceae and Empetrum from alpine and sub-alpine heaths in Austria. *Plant Syst. Evol.* 163 (1–2), 71–79. <http://dx.doi.org/10.1007/BF00936154>.
- Schellekens, J., Buurman, P., 2011. *n*-Alkane distributions as palaeoclimatic proxies in ombrotrophic peat: The role of decomposition and dominant vegetation. *Geoderma* 164 (3–4), 112–121. <http://dx.doi.org/10.1016/j.geoderma.2011.05.012>, Publisher: Elsevier BV.
- Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. *Nature* 478 (7367), 49–56. <http://dx.doi.org/10.1038/nature10386>.
- Schwark, L., Zink, K., Lechterbeck, J., 2002. Reconstruction of postglacial to early Holocene vegetation history in terrestrial Central Europe via cuticular lipid biomarkers and pollen records from lake sediments. *Geology* 30 (5), 463. [http://dx.doi.org/10.1130/0091-7613\(2002\)030<0463:ROPTEH>2.0.CO;2](http://dx.doi.org/10.1130/0091-7613(2002)030<0463:ROPTEH>2.0.CO;2).
- Srivastava, K., Wiesenberg, G.L.B., 2018. Severe drought-influenced composition and  $\delta^{13}\text{C}$  of plant and soil *n*-alkanes in model temperate grassland and heathland ecosystems. *Org. Geochem.* 116, 77–89. <http://dx.doi.org/10.1016/j.orggeochem.2017.11.002>.
- Thevenot, M., Dignac, M.F., Rumpel, C., 2010. Fate of lignins in soils: A review. *Soil Biol. Biochem.* 42 (8), 1200–1211. <http://dx.doi.org/10.1016/j.soilbio.2010.03.017>.
- Thomas, C.L., Jansen, B., Czerwiński, S., Gałka, M., Knorr, K.H., van Loon, E.E., Egli, M., Wiesenberg, G.L.B., 2023a. Paleobotanical and biomarker records over 2600 years from Beerberg peatland (Central Germany). <http://dx.doi.org/10.1594/PANGAEA.961142>, Type: data set.
- Thomas, C.L., Jansen, B., Czerwiński, S., Gałka, M., Knorr, K.H., van Loon, E.E., Egli, M., Wiesenberg, G.L.B., 2023b. Comparison of paleobotanical and biomarker records of mountain peatland and forest ecosystem dynamics over the last 2600 years in central Germany. *Biogeosciences* 20 (23), 4893–4914. <http://dx.doi.org/10.5194/bg-20-4893-2023>.
- Thomas, C.L., Jansen, B., van Loon, E.E., Wiesenberg, G.L.B., 2021. Transformation of *n*-alkanes from plant to soil: a review. *Soil* 7 (2), 785–809. <http://dx.doi.org/10.5194/soil-7-785-2021>.
- Tikhonov, A.N., Arsenin, V.Y., 1977. *Solutions of Ill-Posed Problems*. Winston & Sons, Washington, DC.
- van Mourik, J.M., Jansen, B., 2013. The added value of biomarker analysis in palaeoecology; reconstruction of the vegetation during stable periods in a polycyclic driftsand sequence in SE-Netherlands. *Quat. Int.* 306, 14–23. <http://dx.doi.org/10.1016/j.quaint.2013.05.034>.
- van Mourik, J.M., Wagner, T.V., De Boer, J.G., Jansen, B., 2016. The added value of biomarker analysis to the genesis of plaggic Anthrosols; the identification of stable fillers used for the production of plaggic manure. *Soil* 2 (3), 299–310. <http://dx.doi.org/10.5194/soil-2-299-2016>.
- Vonk, J.E., Gustafsson, Ö., 2009. Calibrating *n*-alkane *Sphagnum* proxies in sub-Arctic Scandinavia. *Org. Geochem.* 40 (10), 1085–1090. <http://dx.doi.org/10.1016/j.orggeochem.2009.07.002>.
- Weber, J., Schwark, L., 2020. Epicuticular wax lipid composition of endemic European *Betula* species in a simulated ontogenetic/diagenetic continuum and its application to chemotaxonomy and paleobotany. *Sci. Total Environ.* 730, 138324. <http://dx.doi.org/10.1016/j.scitotenv.2020.138324>.
- Wiesenberg, G.L.B., Gocke, M.I., 2015. Analysis of lipids and polycyclic aromatic hydrocarbons as indicators of past and present (micro)biological activity. In: McGenity, T.J., Timmis, K.N., Nogales, B. (Eds.), *Hydrocarbon and Lipid Microbiology Protocols*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 61–91. <http://dx.doi.org/10.1007/978-3-642-28157-7>.
- Wüthrich, L., Bliedner, M., Schäfer, I., Zech, J., Shajari, F., Gaar, D., Preusser, F., Salazar, G., Szidat, S., Zech, R., 2017. Late quaternary climate and environmental reconstruction based on leaf wax analyses in the loess sequence of Möhlin, Switzerland. *E G Quat. Sci. J.* 66, <http://dx.doi.org/10.5194/egqsj-66-91-2017>.
- Xie, S., Nott, C.J., Avsejs, L.A., Maddy, D., Chambers, F.M., Evershed, R.P., 2004. Molecular and isotopic stratigraphy in an ombrotrophic mire for paleoclimate reconstruction. *Geochim. Cosmochim. Acta* 68 (13), 2849–2862. <http://dx.doi.org/10.1016/j.gca.2003.08.025>.
- Yang, D., Bowen, G.J., 2022. Integrating plant wax abundance and isotopes for paleovegetation and paleoclimate reconstructions: a multi-source mixing model using a Bayesian framework. *Clim. Past* 18 (10), 2181–2210. <http://dx.doi.org/10.5194/cp-18-2181-2022>.
- Zech, M., Rass, S., Buggle, B., Löscher, M., Zöller, L., 2012. Reconstruction of the late Quaternary paleoenvironments of the Nussloch loess paleosol sequence, Germany, using *n*-alkane biomarkers. *Quat. Res.* 78 (2), 226–235. <http://dx.doi.org/10.1016/j.yqres.2012.05.006>.
- Zheng, Y., Zhou, W., Liu, Z., Liu, X., 2011. The *n*-alkanol paleoclimate records in two peat deposits: a comparative study of the northeastern margin of the Tibetan Plateau and Northeast China. *Environ. Earth Sci.* 63 (1), 135–143. <http://dx.doi.org/10.1007/s12665-010-0676-2>.
- Zheng, Y., Zhou, W., Meyers, P.A., Xie, S., 2007. Lipid biomarkers in the Zoigé-Hongyuan peat deposit: Indicators of Holocene climate changes in West China. *Org. Geochem.* 38 (11), 1927–1940. <http://dx.doi.org/10.1016/j.orggeochem.2007.06.012>.
- Zhou, W., Xie, S., Meyers, P.A., Zheng, Y., 2005. Reconstruction of late glacial and Holocene climate evolution in southern China from geolipids and pollen in the Dingnan peat sequence. *Org. Geochem.* 36 (9), 1272–1284. <http://dx.doi.org/10.1016/j.orggeochem.2005.04.005>.