Quaternary Science Reviews 351 (2025) 109204

Contents lists available at ScienceDirect

Quaternary Science Reviews

journal homepage: www.elsevier.com/locate/quascirev

# Comparing pollen and *n*-alkane carbon isotope records in a tropical lacustrine environment

Xennephone Hadeen [a,1], Cassandra Rowe [a], Michael Brand [a], Rainy Comley [a], Sourav Das [b], Christopher M. Wurster [a,2], Costijn Zwart [a], Michael I. Bird [a,*]

[a] College of Science and Engineering, ARC Centre of Excellence for Indigenous and Environmental Histories and Futures, James Cook University, Cairns, QLD, 4870, Australia

[b] School of Electrical Engineering, Computing and Mathematical Sciences, Science & Engineering, Curtin University Bentley, Building 314, Wark Avenue, WA, 6102, Australia

ARTICLE INFO

Handling editor: P Rioual

*Keywords:*
Compound specific isotope
Biomarker
Tropical
Photosynthetic pathway
Palaeoecology
Palaeolimnology

ABSTRACT

Both pollen and the carbon isotope ($\delta^{13}C$) compositions of long chain *n*-alkanes (*n*-$C_{29\text{-}31}$) are widely used to reconstruct changes in past vegetation in the tropics. Both approaches are able to infer the proportions of tree/shrub ($C_3$) and grass ($C_4$) biomass, with changes in these proportions generally interpreted to be driven by changes in hydroclimate. Both pollen and *n*-alkane $\delta^{13}C$ records are subject to biases in production and transport to a site of deposition that may influence the interpretation of the record. Here we compare detailed 150,000-year pollen and *n*-alkane carbon isotope ($\delta^{13}C$) records from a lake in monsoonal northern Australia. There is a broad agreement between the two records at some times, with both identifying major wet periods of high tree/shrub $C_3$ terrestrial biomass during marine isotope stages 5 and 1. There are significant differences between the two records at other times. At times during glacial marine isotope stages 2 and 6, $C_4$ grass pollen comprised >80% of terrestrial pollen, whereas the *n*-$C_{29\text{-}31}$ $\delta^{13}C$ values indicate dominantly $C_3$ terrestrial biomass.

These differences are the result of changes in the relative abundance and $\delta^{13}C$ values of alkane inputs from within the lagoon itself that impact the $\delta^{13}C$ value of the *n*-$C_{29\text{-}31}$ 'terrestrial' alkanes. The drivers of these changes include (i) changes in the biomass and $\delta^{13}C$ value of aquatic vegetation (floating and submerged) that result from dramatic changes in lake level (ii) the changing importance of groundwater-derived dissolved inorganic carbon as a substrate for photosynthesis in the lake, (iii) changes in the proportion of sedge biomass in and around the lake that also accompany changes in lake level, and (iv) changes in the mix of $C_3$ and $C_4$ species comprising the sedge biomass in and around the lake. The significance of contributions of *n*-alkanes from aquatic and sedge sources to chain lengths usually considered to derive from terrestrial plants, may be underestimated in small, high productivity lacustrine environments, particularly in carbonate terranes.

## 1. Introduction

Records of past vegetation change provide valuable insights into past climate and land use change in the tropics. Two proxies extensively used to infer change in vegetation over time are pollen and *n*-alkane biomarkers. Both are widely applied to lacustrine and marine sedimentary records and have provided valuable insights into the temporal course and dynamics of past vegetation, climate and land use change in the tropics (e.g. Rommerskirchen et al., 2006; Sarkar et al., 2015; Rowe et al., 2019; Deza-Araujo et al., 2020; Bird et al., 2024; Li et al., 2024).

Terrestrial and aquatic plants produce palynomorphs that are identifiable to the genera or family that produced them. Changes in the representation of different palynomorphs in a sedimentary record are therefore interpretable as changes in the plant communities delivering pollen to the site of accumulation (Prentice, 1988). Palaeoecological interpretation of the plant communities identified then enables inference as to the climate required to sustain those communities. In the tropics, hydroclimate (rainfall amount and seasonal distribution) exerts

* Corresponding author.
*E-mail address:* michael.bird@jcu.edu.au (M.I. Bird).
[1] Lead author
[2] Current address: Isotrace NZ Ltd, 167 High St, Dunedin, New Zealand.

https://doi.org/10.1016/j.quascirev.2025.109204
Received 6 November 2024; Received in revised form 13 January 2025; Accepted 16 January 2025
Available online 21 January 2025

a dominant control on vegetation (Rowe et al., 2019). In the humid tropics, natural vegetation is dominated by trees and shrubs (all using the $C_3$ photosynthetic pathway), but as rainfall decreases, or becomes more seasonal, ecosystems include a variable component of grass biomass (dominantly using the $C_4$ photosynthetic pathway) along a continuum of tree cover from closed forest to mixed tree-grass savanna to grassland. The pollen spectra in a sedimentary record can identify these communities, and changes in community structure over time can be used to infer changes in hydroclimate, in the tropics often linked to changes in monsoon strength (Ali et al., 2019).

*n*-alkanes are linear chain molecules composed of repeating subunits of single-bond carbon and hydrogen atoms. They are stable compounds due to the hydrogen and carbon atoms bonding covalently, making them non-exchangeable after death and deposition in sediments at temperatures below 150 °C (Schimmelmann et al., 1999; Sachse et al., 2006). Most *n*-alkane producing organisms produce a range of chain lengths, often centred on one or two high abundance *n*-alkanes, but with lesser amounts of both higher and lower chain-lengths also synthesised by the same organism (Naeher et al., 2022). For terrestrial vascular plants the highest abundance tends to occur at *n*-$C_{27\text{-}29}$ for trees and shrubs, and *n*-$C_{29\text{-}31}$, sometimes higher, for grasses and sedges (Eglinton and Hamilton, 1967; Massimo, 1996; Bush and McInerney, 2013; Naeher et al., 2022; Liu et al., 2022a,b).

While information on source and past environment can be obtained from chain-length distribution in sediments alone, significantly more information can be accessed through measurement of the compound-specific stable isotope composition of carbon in each of the suite of alkanes present in a sample ($\delta^{13}C_n$ values; where n is the number of carbons in an individual n-alkane). In the tropics and sub-tropics, and warm arid regions this is particularly useful because the two dominant forms of terrestrial vegetation present – trees and grass – use different photosynthetic pathways. Trees and shrubs exclusively use the $C_3$ pathway, resulting in mean *n*-alkane $\delta^{13}C$ values (for *n*-$C_{29}$) of $-37.6 \pm 2.8$ ‰ for tropical environments (Liu et al., 2022a,b). Grasses and sedges use both the $C_3$ pathway and the $C_4$ pathway, but grass species that use the $C_4$ pathway tend to dominate in warm, and wet to dry, environments, with mean *n*-alkane $\delta^{13}C$ values (for *n*-$C_{29}$) of $-21.4 \pm 1.7$ ‰ for 'tropical' environments (Liu et al., 2022a,b). Thus, in the same way that pollen spectra are used to reconstruct past vegetation, the $\delta^{13}C$ values of long-chain n-alkanes in sedimentary records have been interpreted as dominantly reflecting changes in terrestrial vegetation between 'tree' and 'grass' end-members, and thereby hydroclimate in the past (e.g. Rommerskirchen et al., 2006; Garcin et al., 2014; Sarkar et al., 2015).

Both pollen and *n*-alkane derived records of vegetation change are subject to potential biases. Both represent a mix of a local and a regional signal. It has been known for a long time that n-alkanes can be transported as aerosols over long distances before deposition (Gagosian et al., 1981; Conte and Weber, 2002). It has also been known for a long time that pollen can also be transported airborne over a long distance before deposition (e.g. Hooghiemstra et al., 2006). Thus, both proxies will incorporate both a local and a region signal of vegetation change.

In addition to dispersal potential, different species produce different amounts of both n-alkanes and pollen per individual plant. The production of *n*-alkanes varies between species, with grasses tending to produce lower alkane abundance per leaf than many trees at *n*-$C_{29}$ and *n*-$C_{31}$ (Diefendorf and Freimuth, 2017; Liu et al., 2013), but against this, around half of primary production in tropical savannas is attributable to grasses which are mostly leaf (Lloyd et al., 2008). In the case of pollen, low productivity and targeted dispersal are associated with specialised modes of pollination, notably insect-(entomophily), animal-(zoophily) and self-pollination (autogamy). With efficient modes of transport by wind, abundant pollen production is not required to ensure successful pollination (Bennett and Willis, 2001; Vanderhoorn et al., 2024). In northern Australia, for example, the most important canopy dominants (e.g. *Eucalyptus, Acacia, Melaleuca*) are pollinated by insects, marsupials, birds, insects and bats (Mariani et al., 2016), influencing their relative pollen representation in the sedimentary record compared to the high dispersal (wind pollinated, anemophily) pollen potential of grass (Luplau, 2015).

One complication in interpreting *n*-alkane proxy records that does not impact pollen records is the fact that most species produce *n*-alkanes with multiple chain lengths that overlap with other taxonomic groups. Therefore, while most trees (and grasses) produce *n*-$C_{29}$ to *n*-$C_{31}$ in greatest abundance (Diefendorf and Freimuth, 2017; Liu et al., 2018), they also produce n-alkanes from *n*-$C_{25}$ to *n*-$C_{35}$ in lower abundance. While it is generally thought that sedges are 'dominantly $C_4$', only 11 out of 107 genera that were assessed by Bruhl and Wilson (2007) include some species that utilise the $C_4$ pathway, with most using the $C_3$ pathway, and sedges can be either terrestrial or fully aquatic. Sedges produce n-alkanes in abundance down to at least *n*-$C_{25}$, while submerged aquatic plants can produce n-alkanes up to *n*-$C_{35}$ and lilies, for example, have *n*-alkane distributions that are similar the terrestrial plants. These complications are particularly problematic in terms of interpreting changes in $\delta^{13}C$ values as a simple measure of changes in the proportion of $C_3$ ('tree') and $C_4$ ('grass') biomass in the terrestrial landscape over time, and thereby as directly indicating changes in hydroclimate.

In general, both pollen and n-alkane records are subject to similar and different potential biases, that can influence the interpretation of each in terms of vegetation change, but the two 'types' of proxy record are rarely directly compared. Here we compare a pollen record with an *n*-alkane abundance and carbon isotope record from a sediment sequence spanning the last 150,000 years, collected from a relatively small lagoon in the mesic savanna of northern Australia. Our intent is to determine the degree to which the n-alkane and pollen records yield a similar and/or different interpretation of vegetation change in the area.

## 2. Study site and methods

### 2.1. Study site

Girraween lagoon (12° 31′ 3.6″ S, 131° 04′ 50.7″ E; 25 m above sea level; Fig. 1) is a perennial waterbody currently 4.5 m in depth and 45 ha in area. The lagoon was formed from a sinkhole collapse within a small catchment of 917 ha (Rowe et al., 2019). Drilling for water extraction has confirmed 'dolomitic clay', sandstone and dolomite with cavities underlie the area, beneath lateritic cover and beginning at 4–9 m depth (bores RN41159 and RN43239; https://nrmaps.nt.gov.au/knowyourbore_desktop.html).

The site is in the core monsoon region of northern Australia, toward the southern margin of influence of the Indo-Australian summer monsoon. Since its formation, the lagoon has contained permanent water, and has therefore archived a rare, high-quality, record of vegetation and climate change in the strongly seasonal Australian tropics. Most (90%) of the ~1700 mm annual rainfall is delivered in the wet season between November and April, currently supporting a mesic savanna that has developed over the course of the Holocene in the broader area with a number of concentric wetland communities bordering the lagoon (Fig. 1; Rowe et al., 2019). Previous studies of the terrestrial and lagoonal flora in the Darwin region have identified 111 terrestrial species, 40 of which have been recorded at Girraween Lagoon, and 49 fully aquatic or aquatic/terrestrial species, 26 of which have been recorded at Girraween Lagoon (Lamche and Schult, 2012). Whilst both $C_3$ and $C_4$ grass species are represented in the region, Hattersley (1983) found that >90% of the grass biomass is produced by species that utilise the $C_4$ photosynthetic pathway. During fieldwork from 2015 to 2018, multiple vegetation surveys were carried out around the lagoon with leaf and flower samples of the major extant taxa opportunistically collected for pollen and alkane analyses. Representative images of the main vegetation zones are provided in Fig. 2.

A 19.4 m core was collected from the centre of the lagoon using a floating platform with hydraulic corer (Rowe et al., 2019). The record is

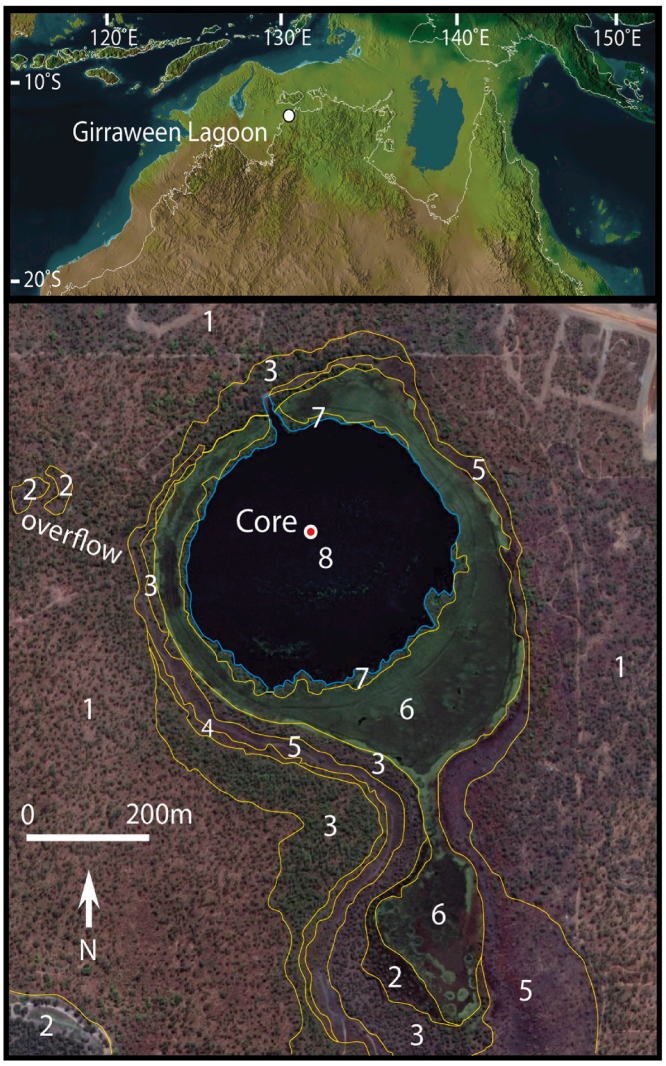


**Fig. 1.** Top - Location of Girraween lagoon with current coastline mapped onto the landmass of Sahul, exposed by sea level fall at the LGM. Bottom: Vegetation zones around the lagoon, 1 *Eucalyptus* woodland to open forest, 2 *Melaleuca* seasonal swamp, 3 Mixed association with monsoonal forest and/or riparian inclinations, 4 *Banksia* transition community 5, grassland, 6 sedgeland, 7 *Melaleuca cajuputi* dominated low open woodland, 8 open-water with *Nymphaea* species and numerous submerged taxa. See Rowe et al. (2019) for detailed community structure. Base image: Google © 2023 Maxar Technologies.

composed of alternating organic-poor green to grey clays deposited during dry periods of low water level when the water body retreated into the confines of the ~1–2 ha sinkhole itself, with no marginal area of shallow water, and therefore limited opportunity for in-lake photosynthetic carbon production. These sediments alternate with organic-rich black peaty sediments formed during wet periods of high-water level similar to today or higher, when the lagoon surface area expands by 1–2 orders of magnitude across a shallow marginal area that enables high rates of organic carbon fixation from rooted floating and submerged aquatic plants (Bird et al., 2024). An absolute chronology is provided by 12 radiocarbon ages (Rowe et al., 2019, 2021) and 24 optically stimulated luminescence ages (Bird et al., 2024). The entire record is continuous and spans the last 150 kyr.

## 2.2. Analytical

### 2.2.1. Initial handling

Core sections were split in half, described and samples of known volume taken at 5 or 10 cm intervals. Sediment samples were weighed, freeze-dried, and reweighed samples to calculate dry bulk density, and ground to a powder. Dried leaf samples were finely ground in a disc mill.

### 2.2.2. Pollen

Sediment samples of 2 $cm^3$ were taken and maintained wet prior to preparation for pollen analysis. The samples were not aliquots of the material used for the n-alkane studies, but sampling resolution is similar, equivalent or adjacent to the intervals used for geochemical analysis. Sample preparation followed standard techniques as outlined in Bennett and Willis (2001) and detailed in Brown (2008). We selected chemical preparations (including $Na_4P_2O_7$, KOH, HCL, acetolysis, and ethanol washes) to disperse sedimentary materials, and progressively remove humic acids, calcium carbonates, bulk organics and cellulose, silicates, as well as render pollen ornamentations more visible. A *Lycopodium* spike was added prior to laboratory preparations to determine relative concentrations of pollen and then samples were sieved at 125 μm and 7 μm to obtain a final preparation for counting.

Pollen identification was based on regionally appropriate image and slide libraries built from collections at the lagoon and online resources such as the *Australasian Pollen and Spore Atlas* (apsa.anu.edu.au). Pollen sums averaged 300 grains, with an average of 12% of pollen grains not identified either because of damage, or the taxa was not represented in the available libraries. We present the proportion of total grass pollen as a percentage of total dryland pollen taxa in each sample with taxa in the dryland non-grass group listed in Bird et al. (2024).

We also calculated the percentage of wetland to total pollen, and within the wetland taxa, *Cyperaceae* taxa (*Cyperus, Eleocharis* and *Fimbristylis* spp), and lily (*Nymphaea* and *Nymphoides* spp.) as a percentage of the total wetland pollen. The wetland taxa not included in these two groups are dominated by wetland-affiliated tree/shrub species such as *Melalauca* spp and other aquatic macrophyte taxa. For the full Girraween sediment sequence, a sum of 30 wetland and 97 dryland pollen-plants types were recorded. Sub-categories within this overall division incorporated: Myrtaceae wetland canopy (7), sedges (3), wet ground associated herbaceous (10), Pteridophytes (8), aquatics (2), and Myrtaceae dryland canopy (9), sclerophyll woody sub-canopy (22), monsoonal forest associated (25), liana and mistletoe (6), Poaceae (2), dry-herbaceous (sub-shrubs), 16), plus dry-herbaceous (forbs, 17).

### 2.2.3. n-alkane abundance carbon isotope composition

Aliquots of 10–50 g of freeze-dried homogenised sediment at 5–20 cm intervals (depending on total organic carbon content), or 1–2 g of homogenised leaf material were extracted in dichloromethane:methanol (9:1) at 100 °C for 15 min in a MARS 6 microwave digestion system. The resulting total lipid extract was transferred into a 50 ml glass tube and evaporated to dryness under nitrogen, followed by transfer in hexane to a 4 ml glass vial. Alkanes were separated on a Restek 12 position solid-phase extraction vacuum SPE manifold using 6 ml Chroma bond SPE glass columns (Macherey-Nagel) filled with 1.5 g of silica gel (0.040 mm–0.063 mm mesh size), via stainless steel connectors and valves (Macherey-Nagel) to avoid plastic contamination. Once the columns were cleaned with n-hexane, dichloromethane, and acetone, we transferred the total lipid extract onto the column in hexane, and eluted the alkanes with a further 10 mL n-hexane. Next, we added 10 μg of 5a-androstane (Sigma-Aldrich) as an internal standard.

Most samples had multiple additional compounds that interfered with the measurements of the *n*-alkanes. Sulphur was removed by loading the sample onto a standard long-stem glass pipette used as a column, half-filled with activated copper powder <425 μm (Sigma-Aldrich with a purity of 99.5%). The copper was activated by washing with 6 ml of HCl and then rinsing with 10 ml Milli-Q water, followed by 6 ml acetone and 6 ml dichloromethane. Each sample was then taken up in dichloromethane, transferred onto the copper pipette column, and rinsed five times with dichloromethane. Alkenes were removed from the samples with silver nitrate $AgNO_3$ on silica gel (Sigma-Aldrich),


**Fig. 2.** Representative images of vegetation zones delineated in Fig. 1 around Girraween Lagoon Photos by Xennephone Hadeen and Cassandra Rowe (2018.

activated in an oven at 105 °C for 1 h, packed into log stem glass pipettes ¾ full, and kept out of direct UV light. After cleaning the packed column with acetone, dichloromethane, and n-hexane, the sample was transferred in *n*-hexane onto the $AgNO_3$ column and eluted five times with *n*-hexane.

Branched alkanes were removed using urea adduction after drying each sample into an 8 ml screw-top vial. Three solutions were then added to the vial: (*i*) 300 μl 10% urea solution (Ultra-Pure Urea >98%; Thermo-Fischer Scientific), (*ii*) 300 μl acetone, and (*iii*) 300 μl pentane (99%; Scharlau). The vial was capped, homogenised with a vortex mixer, placed into a freezer for 30 min, then evaporated to dryness under $N_2$. The urea crystallised as a precipitate with *n*-alkanes taken into channels in the urea structure while excluding branched alkanes. The extract was rinsed with 1 ml of n-hexane, leaving behind the clean urea

crystals containing the *n*-alkanes. This procedure was repeated twice more. Then 500 μl of ultraclean HPLC Plus grade water (Sigma-Aldrich), 500 μl methanol, and 1.5 ml *n*-hexane was added to dissolve the urea, with the vial contents again homogenised by vortex mixer. The immiscible n-hexane (containing the *n*-alkanes) was allowed to separate above the water and methanol and then transferred into another 8 ml vial, repeated in triplicate. The entire process was repeated 4–5 times to eliminate all branched alkanes.

Alkanes were identified and quantified on 1 μl aliquots using a Shimadzu QP2010 GC-MS equipped with a mass selective detector and a flame ionisation detector coupled via an electronic split interface. Quantification was achieved by determining the peak area of each *n*-alkane in a sample with reference to the peak area of the internal standard 5a-androstane. The GC-MS also enabled confirmation of the elimination of impurities.

We determined the stable carbon isotope composition ($\delta^{13}$C value) of the purified *n*-alkanes in triplicate using a Trace 1310 gas chromatograph coupled to a Thermo-Scientific Delta V$^{Plus}$ isotope ratio mass spectrometer. The mass spectrometer measured the stable isotope composition of all *n*-alkanes present in sufficient quantity in each sample. We calibrated $\delta^{13}$C values through reference to standards (Eicosane; Sigma-Aldrich, Androstane; Sigma-Aldrich, Docosane; Sigma-Aldrich, and Squalane; Alfa Aesar) to the international VPDB reference standard for carbon isotopes. Triplicate analysis enabled calculation of the uncertainty of each measurement for each n-alkane in each sample and the mean error across the 2585 individual n-alkane measurements was ±0.14 ‰ (1σ).

#### 2.2.4. Data handling

We calculated average values for the two most abundant odd-numbered alkanes that are representative of each of the following broad groups of alkane-producing organisms (based on Naeher et al., 2022). Each of these two alkane groups is separated by a minimum of three carbon numbers from the next group of two alkanes to minimise, but not eliminate, tailing of n-alkanes produced by one group into the next: (i) algal/bacterial *n*-$C_{17\text{-}19}$ ($\delta^{13}C_{Al}$) (ii) aquatic plants *n*-$C_{23\text{-}25}$ ($\delta^{13}C_{Aq}$) (iii) terrestrial plants *n*-$C_{29\text{-}31}$ ($\delta^{13}C_{Te}$). We added a fourth group (see section 3.2) for sedges *n*-$C_{35\text{-}37}$ ($\delta^{13}C_{Se}$). For each group the error is the deviation from the mean of the two analyses, each analysis having been repeated in triplicate to provide a mean. For most samples, both of the alkanes in a group were present in sufficient quantity to enable a measurement. Not all samples produced a sufficient abundance of all alkanes for analysis.

Where a value for one alkane of a pair was missing, the single value available for that sample was adjusted up or down by the average deviation from the mean of all the other analyses in the group for which both alkanes were present, with the uncertainty for the result taken to be the average deviation from the mean for all analyses in the group for which values for both alkanes were obtained. The corrections amounted to an adjustment of the addition or subtraction of (i) 0.17 ‰ applied to 37% of the 237 samples from the *n*-$C_{17\text{-}19}$ algal group, (ii) 0.1 ‰ applied to 4% of the 273 samples from the *n*-$C_{23\text{-}25}$ aquatic plant group, (iii) no missing analyses of the 273 samples from the *n*-$C_{29\text{-}31}$ terrestrial plant group and (iv) 0.9 ‰ applied to 10% of the 267 samples from the *n*-$C_{35\text{-}37}$ sedge plant group.

We also calculated two indices to assess the potential magnitude of non-terrestrial *n*-alkane contributions to the *n*-$C_{29\text{-}31}$ terrestrial plant group (and vice versa). These are defined as:

$$Aq^*(\%) = (nC_{23-25} \ / \ (n-C_{23-25} + n-C_{29-31})) * 100$$

Where n-$C_{23\text{-}25}$ is the relative abundance of n-alkanes primarily attributed to the aquatic plant group and n-$C_{29\text{-}31}$ is the relative abundance of n-alkanes primarily attributed to the terrestrial plant group (proposed as a proportional measure P(aq) by Ficken et al., 2000), and,

$$Se^*(\%) = (n-C_{35-37} \ / \ (nC_{35-37} + n-C_{29-31})) * 100$$

Where *n*-$C_{35\text{-}37}$ is the relative abundance of n-alkanes primarily attributed to the sedge plant group and *n*-$C_{29\text{-}31}$ is the relative abundance of n-alkanes primarily attributed to the terrestrial plant group.

High values Aq* and/or Se* indicate periods of maximum potential to significantly impact the abundance and isotope composition of the adjacent terrestrial alkane group, low values indicate periods where the terrestrial group may significantly impact the abundance and isotope composition of the adjacent groups.

#### 2.2.5. Modelling

A non-linear model for Terrestrial $\delta^{13}C$ ($Y_t$) was tested against several potential explanatory features, Age (in ka; $T_t$) (for trend), aquatic $\delta^{13}C(X_{1t})$, sedge $\delta^{13}C(X_{2t})$, and their respective abundances ($X_{3t}$ *and* $X_{4t}$). Generalized additive models (GAM; Wood and Augustin, 2002) are the non-linear (smooth) extension of the generalized linear models (GLM). In standard implementation of a GAM this implies replacing linear coefficients $bX$ with penalized regression splines $f(X)$ (Wood and Augustin, 2002). For the present study and with an approximately Gaussian distributed response $Y_t$ the GAM is defined as:

$$Y_t = \alpha + \beta T_t + f_1(X_{1t}) + f_2(X_{2t}) + f_3(X_{3t}) + f_4(X_{4t}) + \epsilon_t, \epsilon_t \sim (0, \sigma^2)$$

For model fitting a penalized version of the maximum likelihood method, the restricted maximum likelihood (REML; Corbeil et al., 1976), is used as default. For alternatives the user is referred to Wood and Augustin (2002). This approach prevents overfitting of the data by controlling the smoothness of the functions $f_i(.), i = 1, 2, 3, 4,$ using the generalized cross-validation method of Wahba (1985). Note that it is likely that $Y_t$ is a mixture of multiple distributions and a key assumption is that the differences are only in the mean effect.

## 3. Results

All data are provided in Data Set S1 and the Supplementary Information (sections 1-3).

### 3.1. Pollen

The pollen results are discussed in their palaeoecological context elsewhere (Rowe et al., 2019, 2021, 2024; Bird et al., 2024), and here serve to provide the 'benchmark' against which to assess vegetation change as inferred from the *n*-alkane analyses. Across the entire record ($C_4$-dominated) grass pollen representation varies widely. Extended periods occur when the percentage of grass pollen in the total terrestrial pollen sum is above 80% and up to 99%, punctuated by extended periods below 80% including shorter periods where values are below 50% and as low as 3%.

Pollen from aquatic-affiliated taxa also ranges widely from <5% to >80% of total pollen, with periods of high aquatic pollen representation usually corresponding to periods of high total organic carbon, interpreted as periods of high water level and large lagoon surface area. Within the wetland group, the sedges tend to be most abundant, with discrete intervals where lilies are most abundant. All other wetland taxa combined tend to be present in lower abundance than the sedge taxa, with some exceptions during periods of inferred high water level in Marine Isotope Stage (MIS) 5 and earlier. Fig. 3 displays the three pollen indices derived from the wetland, as well as the *Poaceae* (grass) and *Cyperaceae* (sedge) subdivisions of pollen taxa that are present in the core, aggregated into groups as described in section 2.2.4. All indices show substantial variation through the record.

### 3.2. Plant n-alkanes

The *n*-alkane distribution for all modern plant samples is provided in

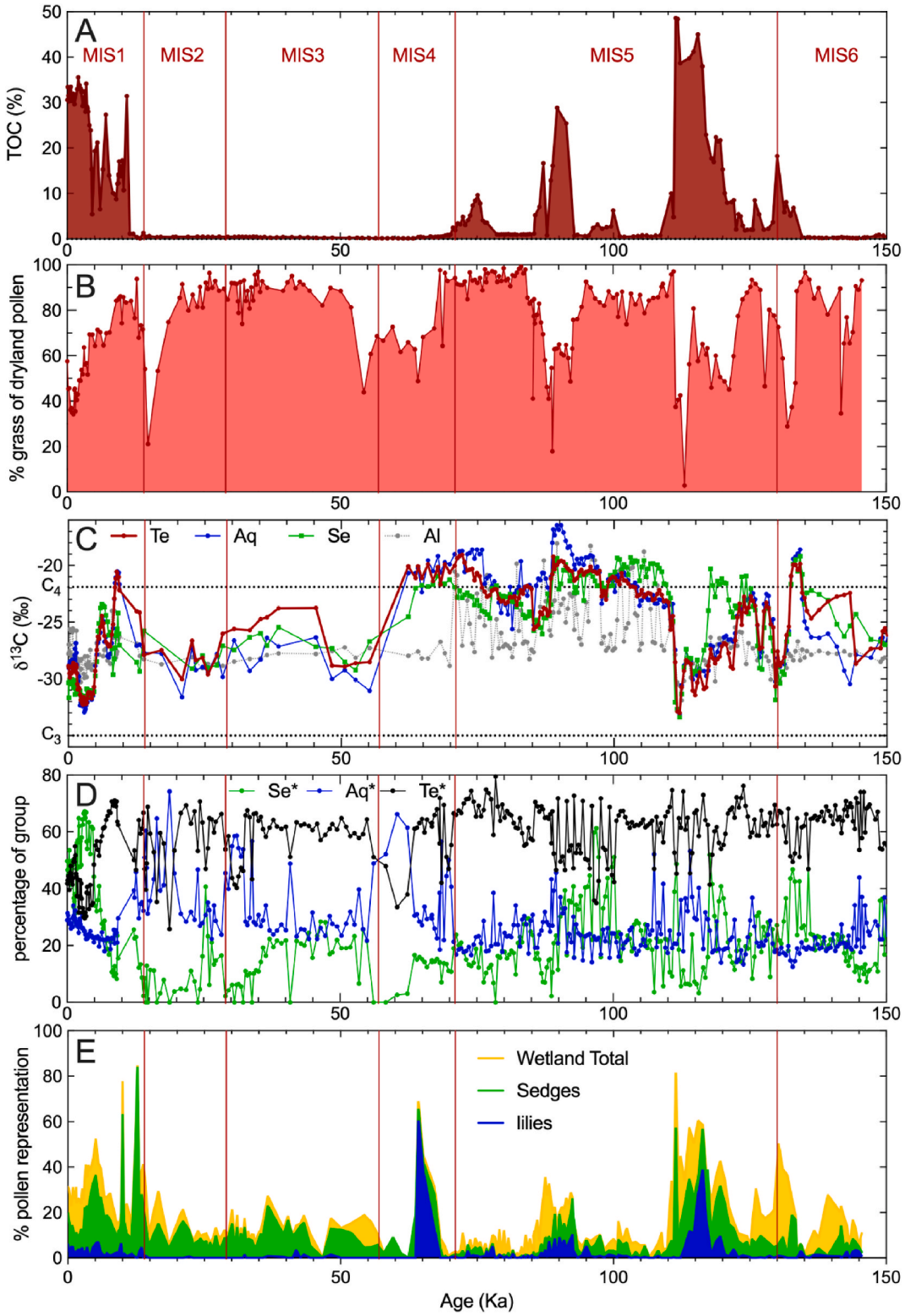


**Fig. 3.** All proxies records used in this study from the Girraween sediment core. (A) Total organic carbon, (B) grass pollen as a percentage of total dryland pollen, (C) $\delta^{13}C$ values for the Te = terrestrial, Aq = aquatic, Se = sedge and Al = Algal/bacterial groups as defined in section 2.2.4, with the global averages for $C_3$ and $C_4$ vegetation from Liu et al. (2022a,b) also shown as horizontal dotted lines (D) Aq* and Se* indices as defined in section 2.2.4 and (E) The proportion of lily, sedge and 'other' taxa as a percentage of total pollen.

Supplementary Table S1 and representative distributions for aquatic species (sedges, lilies and submerged aquatics) are shown in Fig. 4. For the fully terrestrial flora, Figs. S1 and S2 show the results for four $C_4$ grasses and two $C_3$ grasses (Table S3 lists all major species identified, and their photosynthetic pathway), while Figs. S3–S5 show the results for four shrub and seven tree species, all $C_3$. The n-alkane distributions for these taxa are typical of terrestrial species globally (Liu et al., 2022a, b), with greatest abundance generally at *n*-$C_{29}$ or *n*-$C_{31}$, with some shrub

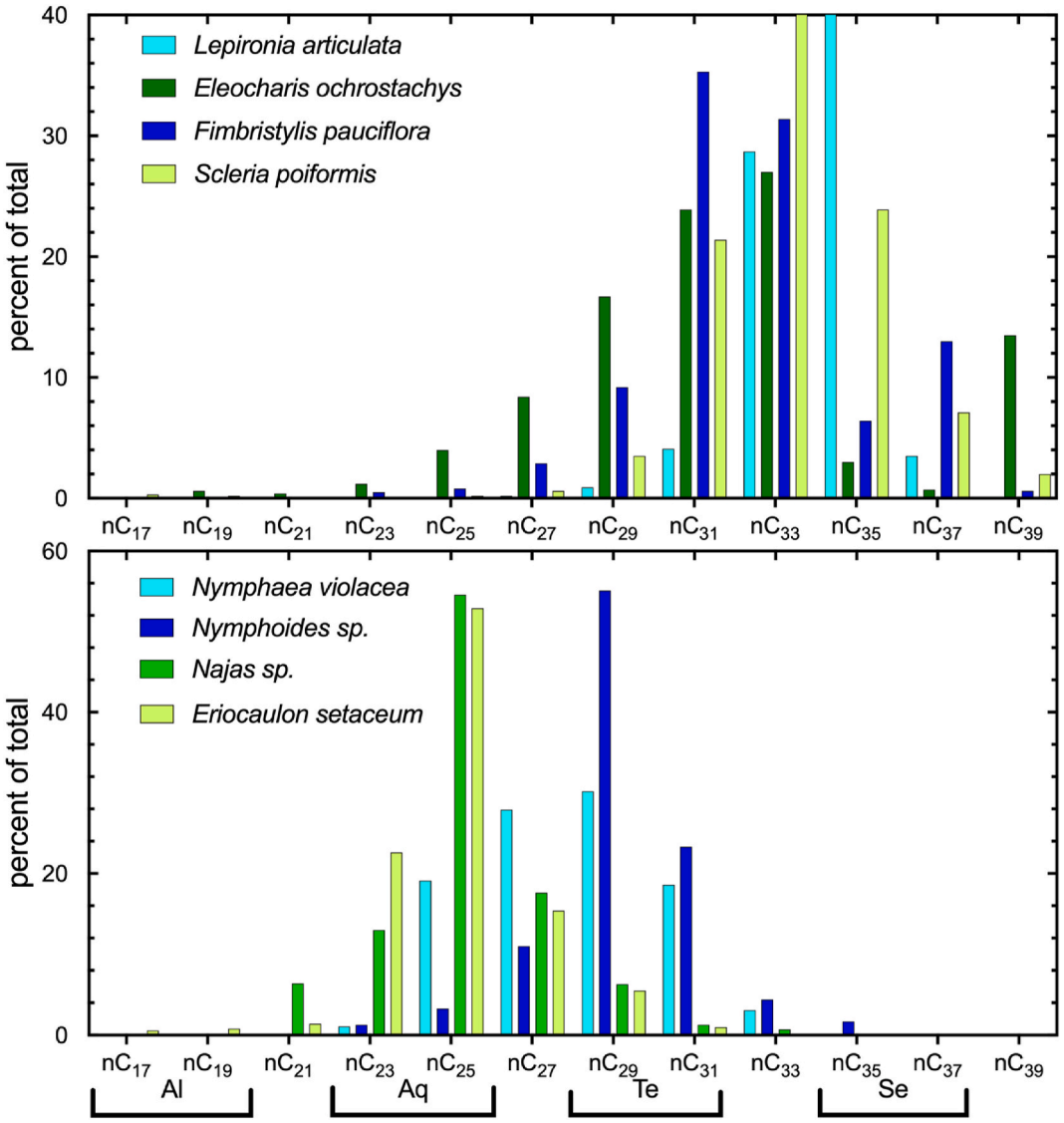


**Fig. 4.** Odd-numbered *n*-alkane distributions for representative aquatic plants in Girraween Lagoon from *n*-$C_{17}$ to *n*-$C_{39}$. 'Percent of total' refers to the percent of each alkane as a function of the sum of all *n*-alkanes from *n*-$C_{17}$ to *n*-$C_{39}$, calculated from the peak area of each alkane measured sequentially during a GC analysis of a single sample (see methods). Top Panel: *Cyperaceae* samples. Bottom panel: two emergent aquatic plants in blue and two submerged aquatic plants in green. Carbon numbers used to calculate $\delta^{13}C$ values for the Te = terrestrial, Aq = aquatic, Se = sedge and Al = Algal/bacterial groups as defined in section 2.2.4 also shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

and tree species also producing significant *n*-$C_{33}$ and most grass species also producing significant quantities of *n*-$C_{33}$ and *n*-$C_{35}$. All species also produce shorter n-alkanes generally down to n-$C_{23\text{-}25}$ in successively lower amounts.

The *n*-alkane distributions of five abundant (non-*Cyperaceae*) aquatic plants are shown in Figs. S6 and S7. The water lilies (*Nymphaea* and *Nymphoides;* Fig. 4 lower panel) and *Azolla pinnata* (all $C_3$) show a distribution largely indistinguishable from tree/shrubs, with peak abundance at *n*-$C_{29}$, decreasing in abundance up to *n*-$C_{33\text{-}35}$ and down to *n*-$C_{23}$. In contrast, the fully submerged *Najas* spp and *Eriocaulon setaceum* (Fig. 4 lower panel) exhibit a 'typical' aquatic plant distribution, with a peak in abundance at *n*-$C_{25}$ decreasing in abundance at short and longer chain lengths, but still producing minor amounts of *n*-$C_{31\text{-}33}$. In some cases, the shorter chain-length alkanes (<~*n*$C_{21}$) may represent a contribution from algae and bacteria growing on the leaves that were not completely removed before analysis.

The results from four species from the sedge (*Cyperaceae*) family are shown in Fig. 4 (upper panel), with a further seven examples provided in Figs. S8 and S9, with photosynthetic pathway, and habitat preference listed in Table S2. One species, *Eleocharis sundaica* ($C_3$), has an *n*-alkane distribution similar to terrestrial plants with peak abundance at *n*-$C_{29}$, and five exhibit a peak in abundance at *n*-$C_{31}$ similar to terrestrial grasses. In contrast to the terrestrial and other aquatic flora, seven of eleven species (five $C_3$, two $C_4$) show significant production of ~5 to >10%) of *n*-$C_{37}$ and four (two $C_3$ and two $C_4$) also produce 1–3% *n*-$C_{39}$. In some cases, the shorter chain-length alkanes present (<~*n*-$C_{21}$) may represent a contribution from algae and bacteria growing on the leaves that were not completely removed before analysis.

### 3.3. Sediment n-alkane distributions

Four representative examples of sediment alkane distributions in the Girraween record are shown in Fig. 5 and these demonstrate a wide range of characteristics. Times of low lake level and hence a very restricted wetland/dampland margin due to retreat of the water body into a steep-walled sinkhole (Bird et al., 2024) are characterised by a dominance of shorter *n*-alkanes, generally most abundant at *n*-$C_{19}$. Times of high lake level with an extensive wetland/dampland marginal

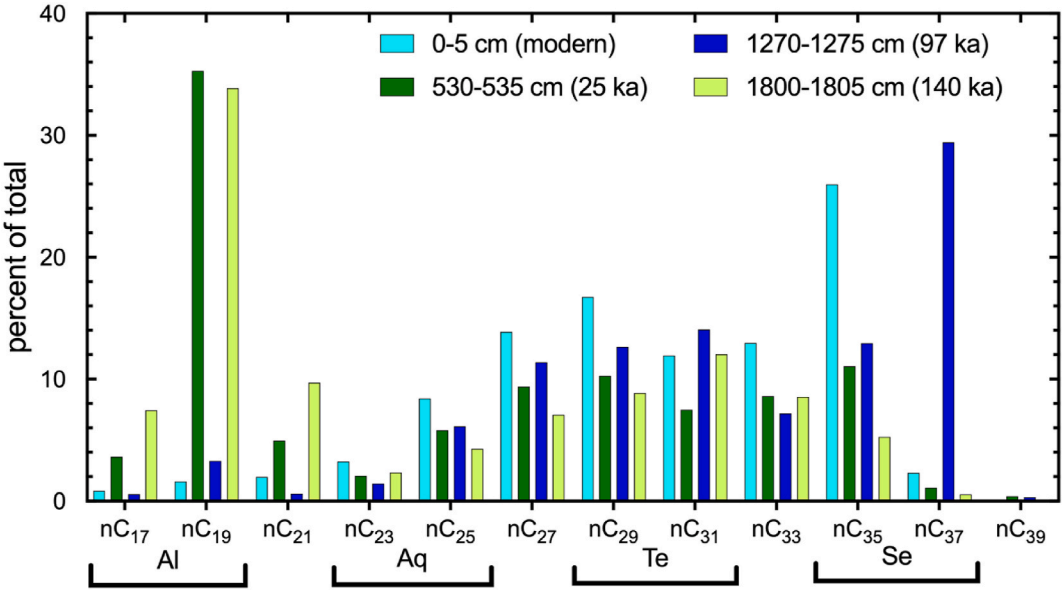


**Fig. 5.** Four representative examples of Girraween sediment showing odd numbered *n*-alkane abundance from *n*-$C_{17}$ to *n*-$C_{39}$. 'Percent of total' refers to the percent of each alkane as a function of the sum of all *n*-alkanes from *n*-$C_{17}$ to *n*-$C_{39}$, calculated from the peak area of each alkane measured sequentially during a GC analysis of a single sample (see methods). Examples with dominant short chain lengths (in green) are from low TOC intervals of the core, examples with dominant long chain lengths (in blue) are from high TOC intervals of the core. Carbon numbers used to calculate $\delta^{13}C$ values for the Te = terrestrial, Aq = aquatic, Se = sedge and Al = Algal/bacterial groups as defined in section 2.2.4 also shown. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

zone, are characterised by the highest abundance of chain lengths in a broad range from *n*-$C_{31}$ up to *n*-$C_{39}$ in some cases.

Aq* and Se* indicate the degree to which it might be expected that alkanes produced by the aquatic and/or sedge groups can influence the terrestrial vegetation group and vice versa. Aq* ranges from ~20 to >50%, while Se* is generally >10% but includes several excursions to lower values and also to values > 50%, with a tendency to exhibit anti-correlation with Aq* in some intervals.

### 3.4. Sediment n-alkane carbon isotope composition

The carbon isotope values for the full record, for the four groups of chain lengths identified in section 2.2.4 are shown in Fig. 3C. Across all groups individual $\delta^{13}C$ values range from −33.8 ‰ (*n*$C_{35\text{-}37}$ sedge group), approaching the global mean value for $C_3$-derived alkanes (Liu and An, 2020; Liu et al., 2022a,b) to −16.5 ‰ (*n*$C_{23\text{-}25}$ aquatic group), higher than the mean global $\delta^{13}C$ value for $C_4$-derived alkanes but within the range for $C_4$ reported globally (Liu and An, 2020; Liu et al., 2022a,b).

The two alkanes that make up each group have $\delta^{13}C$ values that are strongly and linearly correlated for the terrestrial group ($\delta^{13}C_{Te}$ $r^2$ = 0.92) and the ($\delta^{13}C_{Aq}$ $r^2$ = 0.99), less strongly correlated for the algal group ($\delta^{13}C_{Al}$ $r^2$ = 0.54) and sedge group ($\delta^{13}C_{Se}$ $r^2$ = 0.77). Fig. 3C shows that temporal trends in the $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ are broadly similar to the $\delta^{13}C_{Te}$ record, but values for $\delta^{13}C_{Aq}$ and/or $\delta^{13}C_{Se}$ can be higher or lower than $\delta^{13}C_{Te}$ values by several per mil, in one or the other group, or both. When both $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ values diverge from the $\delta^{13}C_{Te}$ record, the divergence of each can be in either a similar or opposite direction.

Across the entire record $\delta^{13}C_{Te}$ values range from −33 to −19 ‰, $\delta^{13}C_{Aq}$ values from −33 to −16.5 ‰ and $\delta^{13}C_{Se}$ values from −33.4 to −19 ‰. In contrast, $\delta^{13}C_{Al}$ values do not track the generally coherent trends between the other groups through the record. $\delta^{13}C_{Aq}$ values exhibit long periods when $\delta^{13}C_{Aq}$ values are generally lower than −26 ‰ and as low as −32 ‰, with a period during Marine Isotope Stage 5, between 70 and 110 ka, when $\delta^{13}C_{Aq}$ values abruptly changed over relatively short, multiple periods from the low values characteristic of the rest of the record, to values as high as −18 ‰.

### 3.5. GAM modelling

Fig. 6 shows the results of the non-linear calibration model for Terrestrial $\delta^{13}C$ ($Y_t$) as response fitted accounting for features, Age (ka) ($T_t$) (for trend), aquatic $\delta^{13}C(X_{1t})$, sedge $\delta^{13}C(X_{2t})$, and respective abundances ($X_{3t}$ *and* $X_{4t}$). The estimates and corresponding uncertainty (p-values and standard errors) are given in Table S4. We observe that Age (ka) ($p < 0.001$), aquatic $\delta^{13}C$ ($p < 0.001$), and sedge $\delta^{13}C$ ($p < 0.001$), are statistically significant. There is a strong match ($r^2 = 0.98$) between the mean effects of the two series indicating that we are able to reconstruct the terrestrial $\delta^{13}C$ using a time dependent regression model involving only the aquatic and sedge records for abundance and $\delta^{13}C$ values. Further detail is provided in Figs. S10–S12.

## 4. Discussion

### 4.1. Comparison of the alkane and pollen records

Given the complications in interpretation of both the pollen and *n*-alkane records identified in the introduction, with the additional caveat that the two sample sets were not always from identical core intervals, concordance in gross trends in the % grass and $\delta^{13}C_{Te}$ records is gratifyingly reasonable, with the major changes in MIS-5/6 (before 65 ka) and the Holocene (after ~10 ka) represented in both records. The agreement between pollen and $\delta^{13}C_{Te}$ records is not good in some places and where there is disagreement, there tends to be a divergence between either or both of the $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ values from the $\delta^{13}C_{Te}$ values in the same sample, particularly at times of high Aq* and/or Se*.

We explore some of these periods below.

(i) 3–5 ka – a period of monotonic decrease in grass pollen, that coincides with a broad minimum, then rise, in $\delta^{13}C_{Te}$ values at a time of particularly high Aq*. This suggests a significant contribution of macrophyte, and to a lesser extent sedge-derived

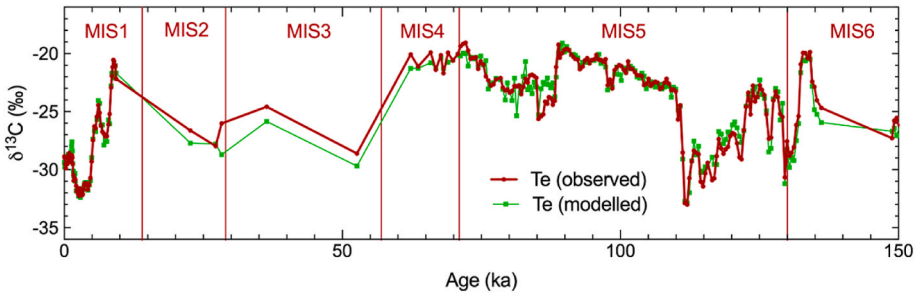


**Fig. 6.** Comparison between the observed $\delta^{13}C_{Te}$ record (in red) and GAM-modelled $\delta^{13}C_{Te}$ results (green). Note model values could only be provided where data was available for all of Aq*, Se*, $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$. Further information provided in the Supplementary information. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

carbon, both with a low $\delta^{13}C$ value to the carbon numbers from which $\delta^{13}C_{Te}$ is calculated.

(ii) 25–45 ka: a period when uniformly high grass pollen abundance is not reflected in high $\delta^{13}C_{Te}$ values, at a time of broadly average Aq* and Se*, and high lily pollen abundance. Both $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ values through this period are 3–5‰ lower than the corresponding $\delta^{13}C_{Te}$ values. Contributions from the aquatics and the sedges have therefore likely lowered the $\delta^{13}C_{Te}$ values. An alternative explanation could be that, during what was the Last Glacial Maximum through MIS-2, cooler temperatures (by 3–4°C; Li et al., 2023), favoured an increase in $C_3$ grass biomass, as the proportion of $C_3$ to $C_4$ species is correlated to temperature (Hattersley, 1983). This might be part of the explanation for the low $\delta^{13}C_{Te}$ values, but decreased rainfall (Rowe et al., 2024) and lower atmospheric $CO_2$ concentrations during this time interval should have provided a competitive advantage to $C_4$ species (Sage, 2004), that would likely offset most if not all of any temperature effect.

(iii) 72–93 ka: This interval represents a period where $\delta^{13}C_{Te}$ values suggest both more and less $C_4$ carbon compared with the pollen record. The period includes times when $\delta^{13}C_{Se}$ values tend to be similar to or up to 3‰ lower than $\delta^{13}C_{Te}$ values, while $\delta^{13}C_{Aq}$ values are similar to or higher than $\delta^{13}C_{Te}$ values by up to 5‰. Aquatic pollen representation is high, Aq* is average to high, and Se* ranges erratically from high to low. This was a time when the sinkhole was not filled by sediment and hence was connected laterally to the carbonate aquifer. This appears to have resulted in the high $\delta^{13}C_{Aq}$ values through access to carbonate-derived DIC, but the low $\delta^{13}C_{Te}$ values appear to be the result of the low $\delta^{13}C_{Se}$ values suggesting an increase in the proportion of $C_3$ species in the sedge biomass.

(iv) 112–125 ka: The $\delta^{13}C_{Te}$ values are uniformly low during this interval, and do not capture the rise in grass pollen centred on 115 ka, likely because of the high proportion of aquatic pollen at this time, including from lilies, which have low $\delta^{13}C_{Aq}$ values. This is despite a brief period where $\delta^{13}C_{Se}$ values are up to 7 ‰ higher than both $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Te}$ values, but the sedges are not well represented in the pollen spectra at this time and Aq* is slightly below average. This is also a time when the lagoon waters were connected to the carbonate aquifer via the sinkhole walls. However, this is also a time of increased water level and hence expanded lake surface area to beyond the sinkhole itself, leading to limited impact of carbonate-derived DIC from the aquifer on $\delta^{13}C_{Aq}$ values.

(v) 130–150 ka: As with MIS 2 $\delta^{13}C_{Te}$ values imply less $C_4$ carbon than is evident in the abundant grass pollen through most of this interval, likely driven in particular by $\delta^{13}C_{Aq}$ values that are up to 8‰ lower than $\delta^{13}C_{Te}$ values.

In all the time periods discussed above, it appears that terrestrial vegetation does not exert the primary control on $\delta^{13}C_{Te}$ values. The conclusion that there is significant control from sedge and aquatic *n*-alkanes on $\delta^{13}C_{Te}$ values is underscored by the fact that for much of the record, $\delta^{13}C_{Te}$ values can be closely modelled (r = 0.98) from Aq*, Se*, $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ values alone (Fig. 6). This is partly due to tailing of the terrestrially-derived *n*-alkanes into the adjacent aquatic and sedge *n*-alkane distributions. However, given the divergence of $\delta^{13}C_{Aq}$ and $\delta^{13}C_{Se}$ values from the $\delta^{13}C_{Te}$ values in some parts of the record, it is likely that the over-riding control on $\delta^{13}C_{Te}$ values across several intervals of the last 150 kyr is from in-lagoon carbon production. At some times, $\delta^{13}C_{Te}$ values are not well predicted by the results from the adjacent groups, suggesting periods where terrestrial carbon input exerts a more significant and direct control on $\delta^{13}C_{Te}$ values. For example, between 25 and 45 ka, the GAM model predicts lower $\delta^{13}C_{Te}$ values than observed (Fig. 6), implying a high $\delta^{13}C$ alkane component derived from terrestrial vegetation is making a significant and identifiable contribution the $\delta^{13}C_{Te}$ values. This consistent with the high abundance of grass pollen in this period (Fig. 3B).

### 4.2. The significance of sedges in the Girraween record

Sedges (*Cyperaceae*), which can be both aquatic and terrestrial in habitat, are a significant contributor to biomass in many terrestrial, lacustrine, swamp and seasonal 'dampland' areas (Starr and Ford, 2009). At Girraween, the sedges are currently represented by both $C_3$ and $C_4$ species, some fully aquatic, some fully terrestrial (see Table S3). From a global compilation, *Cyperaceae* showed the greatest range in $\delta^{13}C_{n\text{-}C29}$ values of any family from −37 to −22 ‰ (Liu et al., 2022a,b). Indeed, at least one species (*Eleocharis vivipara*) can change between the $C_3$ and the $C_4$ pathway depending on habitat (Ueno, 2001).

The proportion of *Cyperaceae* biomass attributable to each photosynthetic pathway at Girraween Lagoon varies over time in response to changes in lake level (Fig. 3E). This is because lake level interacts with local, shallow basin topography to change the proportion of areas of shallow but fully submerged habitat preferred by $C_3$ species, and seasonally 'damp' regions, preferred by $C_4$ species. In addition, the localisation of sedge biomass around and in the lake means that sedge-derived *n*-alkanes are likely to be over-represented relative to the surrounding fully terrestrial vegetation, and will therefore exert a significant impact on the $\delta^{13}C$ values of chain lengths usually ascribed to terrestrial vegetation. The total range of $\delta^{13}C_{Se}$ values in the Girraween record varies considerably over time (Fig. 3) and these variations at times of high Se* will impact $\delta^{13}C_{Te}$ values in a way that is not linked to the broader terrestrial vegetation change.

### 4.3. Interpretation of the $\delta^{13}C_{Al}$ record

While not central to the interpretation of the record in terms of terrestrial vegetation change, the $\delta^{13}C_{Al}$ record is also presented in Fig. 3C. For most of the last 150 ka $\delta^{13}C_{Al}$ values are stable and low,

generally below −27 ‰. In contrast, the period between 72 and 112 ka is marked by erratic, short-lived increases in $\delta^{13}C_{Al}$ values by up to 7 ‰. This period coincides with several periods when $\delta^{13}C_{Aq}$ values were also elevated. We cannot offer a definitive explanation for this behaviour. We do note that this was an 'intermediate' period in the filling of the sinkhole by sediment, prior to the time when sediments completely infilled the sinkhole, severing the direct connection between the aquifer and lagoon waters. It may be that the contribution of DIC supplied from the carbonate aquifer became particularly sensitive to changes in water level through changing lagoon surface area during this time.

### 4.4. Implications for interpreting n-alkane carbon isotope records in the tropics

Girraween lagoon, in some respects, is a special case in that it is (i) small in area, (ii) has had a variable input of carbonate-derived DIC in the past and (iii) can be highly productive, diluting the 'terrestrial' alkane signal. In addition, at times of high-water level, very small changes in water level can dramatically impact the areas available for colonisation by lilies and other emergent macrophytes, fully submerged macrophytes and sedges. Therefore, the extreme changes in carbon isotope values we report here, across most of the range reported globally, due in part to local changes in hydrology, may not be as much of an issue in other places - for example in marine records, or records from large lakes with a small and/or invariant littoral zone.

We do note several instances where it may be that the impact of sedges has been more significant than recognized. The first is Lake Towuti in equatorial Indonesia (Konecky et al., 2016) where an increase in *n*-alkanoic acid carbon isotope composition was taken to indicate a major increase in $C_4$-derived biomass, attributed to glacial aridity and an increase in terrestrial grass biomass. However, the pollen record of Stevenson (2018) from Lake Towuti does not indicate an expansion of grass at any time, and instead finds an increase in sedge pollen during the last glacial maximum, possibly as a result of an expanded area of shoreline around the lake. This indicates that even in a large lake, the impact of local factors may be significant.

Rommerskirchen et al. (2003) compared the pollen and *n*-alkane carbon isotope records from marine cores on a latitudinal transect off the southern Atlantic coast of Africa and found the two were in generally good agreement. However, even in these marine records, sedge pollen was generally over 5% and up to 28% of total pollen. There are large swamp areas in the alkane and pollen source areas for the cores (the Congo basin and Okavango delta for example). Therefore, it is possible that the alkanes derived from a (changing) mix of $C_3$ and $C_4$ sedge species, could influence the interpretation of the isotope record (derived in that study from n-$C_{27\text{-}33}$) as one attributable solely to changes in the proportions of 'land plant vegetation of different biosynthetic types'.

## 5. Conclusions

This study compares the pollen record with *n*-alkane chain length and carbon isotope records in a continuous sedimentary record spanning 150 ka. The records have been generated from a small lagoon that is subject to multiple influences that impact the productivity and carbon isotope composition of aquatic biomass, as well as influences on terrestrial vegetation that relate to regional scale hydroclimate changes over the last glacial-interglacial cycle.

While there is broad agreement between the representation of grass ($C_4$) pollen and the carbon isotope composition of the long chain *n*-alkanes usually used to infer changes in the abundance of $C_4$ biomass in terrestrial vegetation, there are significant differences in detail. These differences relate to large changes in the $\delta^{13}C$ values of *n*-alkane chain lengths that contribute to the 'terrestrial' isotope composition but derive from biomass in the lagoon itself, associated (i) changes in the proportion and isotope composition of macrophyte vegetation, (ii) access of aquatic taxa to groundwater-derived DIC as a substrate for photosynthesis, (iii) changes in the abundance of sedge biomass and (iv) changes in the mix of $C_3$ and $C_4$ species making up the sedge biomass.

Alkanes produced by both the aquatic and sedge groups of plants make a significant contribution to the chain lengths ascribed to terrestrial vegetation (represented here as $\delta^{13}C_{Te}$ values), and at times can significantly modify the carbon isotope composition of those chain lengths to both higher and lower carbon isotope values. It therefore appears important in this case to 'consider the lilies' (Bradbury, 1953; Chapman et al., 2001), the submerged macrophytes and the sedges in terms of their impact on the interpretation of $\delta^{13}C_{Te}$ values as a measure of broad scale terrestrial vegetation (and thereby hydroclimate) change.

We propose a new index - Se* - that, in conjuction with Aq* (or $P_{aq}$), and in the absence of pollen information, shows promise in being able to assess the likelihood that variation in $\delta^{13}C_{Aq}$ or $\delta^{13}C_{Se}$ values could impact $\delta^{13}C_{Te}$ values (and vice versa). Correcting for the additional influences on $\delta^{13}C_{Te}$ values we identify is likely difficult because (i) emergent macrophytes, as in this study, may have a chain length distribution that is very similar to terrestrial plants (ii) correcting $\delta^{13}C_{Te}$ values for any change in $\delta^{13}C_{Aq}$ or $\delta^{13}C_{Se}$ values would require a knowledge of the full chain length distributions of the organisms contributing to those groups and this may change over time. In some circumstances an approximation may be possible, but if the intent is to infer terrestrial vegetation change care should be exercised in choosing a site where a minimal *n*-alkane contribution from in-lake productivity can be reliably assumed.

## Contributions

MIB, CMW, and CR designed the study. MIB, CR, CMW, MB, RC, XH, and CZ did field sampling and initial preparation, XH did the compound specific isotope analyses, CR the palynology. All authors contributed to the interpretation and the final form of the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

Funding: This work was funded by the Australian Research Council (CE170100015 and FL140100044). We thank Larrakia and Wulna Traditional Owners for permission to do this research on their traditional lands.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.quascirev.2025.109204.

## Data availability

All data and/or code is contained within the submission.

## References

Ali, S.N., Dubey, J., Shekhar, M., Morthekai, P., 2019. Holocene Indian Summer Monsoon variability from the core monsoon zone of India, a pollen-based review. Grana 58 (5), 311–327.

Bennett, K.D., Willis, K.J., 2001. Pollen. In: Smol, J.P., Birks, J.B., Last, W.M. (Eds.), Tracking Environmental Change Using Lake Sediments: Terrestrial, Algal, and Siliceous Indicators. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 5–32.

Bird, M.I., Brand, M., Comley, R., Fu, X., Hadeen, X., Jacobs, Z., Rowe, C., Wurster, C.M., Zwart, C., Bradshaw, C.J., 2024. Late Pleistocene emergence of an anthropogenic fire regime in Australia's tropical savannahs. Nat. Geosci. 17 (3), 233–240.

Bradbury, R., 1953. Fahrenheit 451. Ballantine Books, p. 153. New York, USA.

Brown, C.A., 2008. In: Riding, J.B., Warny, S. (Eds.), Palynological Techniques, 2nd Edition. American Association of Stratigraphic Palynologists, Dallas, Texas, p. 137.
Bruhl, J.J., Wilson, K.L., 2007. Towards a comprehensive survey of $C_3$ and $C_4$ photosynthetic pathways in *Cyperaceae*. Aliso: J. Sys. Flor. Botany 23 (1), 99–148.
Bush, R.T., McInerney, F.A., 2013. Leaf wax *n*-alkane distributions in and across modern plants: implications for paleoecology and chemotaxonomy. Geochem. Cosmochim. Acta 117, 161–179.
Chapman, C., Palin, M., Cleese, J., Jones, T., Gilliam, T., 2001. The Monty Python's Life of Brian (Of Nazareth): Screenplay. Methuen Publishing Ltd, London, England, p. 112.
Conte, M.H., Weber, J.C., 2002. Plant biomarkers in aerosols record isotopic discrimination of terrestrial photosynthesis. Nature 417 (6889), 639–641.
Corbeil, R.R., Shayle, R.S., 1976. Restricted maximum likelihood (REML) estimation of variance components in the mixed model. Technometrics 18 (1), 31–38.
Deza-Araujo, M., Morales-Molino, C., Tinner, W., Henne, P.D., Heitz, C., Pezzatti, G.B., Hafner, A., Conedera, M., 2020. A critical assessment of human-impact indices based on anthropogenic pollen indicators. Quat. Sci. Rev. 236, 106291.
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.
Eglinton, G., Hamilton, R.J., 1967. Leaf epicuticular waxes. Science 156, 1322–1327.
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, 745–749.
Gagosian, R.B., Peltzer, E.T., Zafiriou, O.C., 1981. Atmospheric transport of continentally derived lipids to the tropical North Pacific. Nature 291 (5813), 312–314.
Garcin, Y., Schefuß, E., Schwab, V.F., Garreta, V., Gleixner, G., Vincens, A., Todou, G., Séné, O., Onana, J., Achoundong, G., Sachse, D., 2014. Reconstructing $C_3$ and $C_4$ vegetation cover using *n*-alkane carbon isotope ratios in recent lake sediments from Cameroon, Western Central Africa. Geochem. Cosmochim. Acta 142, 482–500.
Hattersley, P.W., 1983. The distribution of $C_3$ and $C_4$ grasses in Australia in relation to climate. Oecologia 57, 113–128.
Hooghiemstra, H., Lézine, A.M., Leroy, S.A., Dupont, L., Marret, F., 2006. Late Quaternary palynology in marine sediments: a synthesis of the understanding of pollen distribution patterns in the NW African setting. Quat. Int. 148 (1), 29–44.
Konecky, B., Russell, J., Bijaksana, S., 2016. Glacial aridity in central Indonesia coeval with intensified monsoon circulation. Earth Planet Sci. Lett. 437, 15–24.
Lamche, G., Schult, J., 2012. Macrophyte Vegetation of Six Lagoons in the Darwin Region, NT. Aquatic Health Unit, Department of Land Resource Management, Darwin. Report No 11/2012D.
Li, T., Comley, R., Zhang, E., Zhou, Y., Zhou, X., Munksgaard, N.C., Zhu, Z., Haig, J., Zheng, F., Bird, M.I., 2023. Paleo-temperature inferred from brGDGTs over the past 18 cal ka BP from Lake Barrine, tropical NE Australia. Quat. Sci. Rev. 310, 108125.
Li, T., Zhou, Y., Wurster, C.M., Zhou, X., Zhao, Y., Comley, R., Munksgaard, N.C., Cernusak, L.A., Haig, J., Bird, M.I., 2024. Vegetation and precipitation change inferred from the $\delta^{13}C$ and $\delta^{2}H$ values of *n*-alkanes from lake sediment from 18 cal ka BP, tropical NE Australia. Quat. Sci. Rev. 337, 108807.
Liu, J., An, Z., 2020. Leaf wax *n*-alkane carbon isotope values vary among major terrestrial plant groups: different responses to precipitation amount and temperature, and implications for paleoenvironmental reconstruction. Earth Sci. Rev. 202, 103081.
Liu, W., Li, X., An, Z., Xu, L., Zhang, Q., 2013. Total organic carbon isotopes: a novel proxy of lake level from Lake Qinghai in the Qinghai–Tibet Plateau, China. Chem. Geol. 347, 153–160.
Liu, J., An, Z., Liu, H., 2018. Leaf wax *n*-alkane distributions across plant types in the central Chinese Loess Plateau. Org. Geochem. 125, 260–269.
Liu, H., Liu, J., Hu, J., Cao, Y., Xiao, S., Liu, W., 2022a. Systematical $\delta^{13}C$ investigations of TOC in aquatic plants, DIC and dissolved $CO_2$ in lake water from three Tibetan Plateau lakes. Ecol. Indicat. 140, 109060.
Liu, J., Zhao, J., He, D., Huang, X., Jiang, C., Yan, H., Lin, G., An, Z., 2022b. Effects of plant types on terrestrial leaf wax long-chain *n*-alkane biomarkers: implications and paleoapplications. Earth Sci. Rev. 235, 104248.
Lloyd, J., Bird, M.I., Vellen, L., Miranda, A.C., Veenendaal, E.M., Djagbletey, G., Miranda, H.S., Cook, G., Farquhar, G.D., 2008. Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: a quasi-global estimate. Tree Physiol. 28 (3), 451–468.
Luplau, R.R., 2015. Pollen Dispersal from *Banksia, Acacia*, and *Grevillea* Species at Neerabup National Park, Western Australia. Masters Thesis, Centre for Foresic Science. University of Western Australia. https://research-repository.uwa.edu.au/files/4711440/Luplau_Rebecca_2015.pdf.
Mariani, M., Connor, S.E., Theuerkauf, M., Kuneš, P., Fletcher, M.S., 2016. Testing quantitative pollen dispersal models in animal-pollinated vegetation mosaics: an example from temperate Tasmania, Australia. Quat. Sci. Rev. 154, 214–225.
Massimo, M., 1996. Chemotaxonomic significance of leaf wax alkanes in the *Gramineae*. Biochem. Systemat. Ecol. 24, 53–64.
Naeher, S., Cui, X., Summons, R.E., 2022. Biomarkers: molecular tools to study life, environment, and climate. Elements Intern. Magaz. Miner. Geochem. Petrol. 18 b (2), 79–85.
Prentice, C., 1988. Records of vegetation in time and space: the principles of pollen analysis. In: Vegetation History. Springer, The Netherlands, pp. 17–42. Dordrecht.
Rommerskirchen, F., Eglinton, G., Dupont, L., Güntner, U., Wenzel, C., Rullkötter, J., 2003. A north to south transect of Holocene southeast Atlantic continental margin sediments: Relationship between aerosol transport and compound-specific $\delta^{13}C$ land plant biomarker and pollen records. G-cubed 4 (12), 1101.
Rommerskirchen, F., Eglinton, G., Dupont, L., Rullkötter, J., 2006. Glacial/interglacial changes in southern Africa: compound-specific $\delta^{13}C$ land plant biomarker and pollen records from southeast Atlantic continental margin sediments. G-cubed 7 (8), Q08010.
Rowe, C., Brand, M., Hutley, L.B., Wurster, C., Zwart, C., Levchenko, V., Bird, M., 2019. Holocene savanna dynamics in the seasonal tropics of northern Australia. Rev. Palaeobot. Palynol. 267, 17–31.
Rowe, C., Wurster, C.M., Zwart, C., Brand, M., Hutley, L.B., Levchenko, V., Bird, M.I., 2021. Vegetation over the last glacial maximum at Girraween Lagoon, monsoonal northern Australia. Quat. Res. 102, 39–52.
Rowe, C., Brand, M., Wurster, C.M., Bird, M.I., 2024. Vegetation changes through stadial and interstadial stages of MIS 4 and MIS 3 based on a palynological analysis of the Girraween Lagoon sediments of Darwin, Australia. Palaeogeogr. Palaeoclimatol. Palaeoecol. 642, 112150.
Sachse, D., Radke, J., Gleixner, G., 2006. δD values of individual *n*-alkanes from terrestrial plants along a climatic gradient - implications for the sedimentary biomarker record. Org. Geochem. 37, 469–483.
Sage, R.F., 2004. The evolution of $C_4$ photosynthesis. New Phytol. 161 (2), 341–370.
Sarkar, S., Prasad, S., Wilkes, H., Riedel, N., Stebich, M., Basavaiah, N., Sachse, D., 2015. Monsoon source shifts during the drying mid-Holocene: biomarker isotope-based evidence from the core 'monsoon zone'(CMZ) of India. Quat. Sci. Rev. 123, 144–157.
Schimmelmann, A., Lewan, M.D., Wintsch, R.P., 1999. D/H isotope ratios of kerogen, bitumen, oil, and water in hydrous pyrolysis of source rocks containing kerogen types I, II, IIS, and III. Geochem. Cosmochim. Acta 63, 3751–3766.
Starr, J.R., Ford, B.A., 2009. Phylogeny and evolution in *Cariceae* (*Cyperaceae*): current knowledge and future directions. Bot. Rev. 75, 110–137.
Stevenson, J., 2018. Vegetation and climate of the last glacial maximum in sulawesi. In: O'Connor, S., Bulbeck, D., Meyer, J. (Eds.), *The Archaeology of Sulawesi: Current* Research on the Pleistocene to the Historic Period, vol. 48, pp. 17–29.
Ueno, O., 2001. Environmental regulation of $C_3$ and $C_4$ differentiation in the amphibious sedge *Eleocharis vivipara*. Plant Physiol. 127 (4), 1524–1532.
Vanderhoorn, J.M., Wilmshurst, J.M., Richardson, S.J., Etherington, T.R., Perry, G.L., 2024. Revealing the palaeoecology of silent taxa: selecting proxy species from associations in modern vegetation data. J. Biogeogr. 51, 1299–1314.
Wahba, G., 1985. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem. Ann. Stat. 13 (4), 1378–1402.
Wood, S.N., Augustin, N.H., 2002. GAMs with integrated model selection using penalized regression splines and applications to environmental modelling. Ecol. Model. 157 (2–3), 157–177.