



OPEN ACCESS

EDITED BY

Alexandra Gogou,
Hellenic Centre for Marine Research
(HCMR), Greece

REVIEWED BY

Xiangjin Shen,
Chinese Academy of Sciences (CAS), China
Sergio Balzano,
Anton Dohrn Zoological Station Naples, Italy

*CORRESPONDENCE

Julie Lattaud,
✉ julie.lattaud@unibas.ch

RECEIVED 29 March 2024

ACCEPTED 08 May 2024

PUBLISHED 03 June 2024

CITATION

Lattaud J, Martin C, Lloren R, Zborovsky B and Dubois N (2024), Temperature and nutrients control the presence and distribution of long-chain diols in Swiss lakes.
Front. Earth Sci. 12:1409137.
doi: 10.3389/feart.2024.1409137

COPYRIGHT

© 2024 Lattaud, Martin, Lloren, Zborovsky and Dubois. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Temperature and nutrients control the presence and distribution of long-chain diols in Swiss lakes

Julie Lattaud^{1,2*}, Céline Martin³, Ronald Lloren³,
Beata Zborovsky¹ and Nathalie Dubois³

¹Swiss Federal Institute of Technology Zurich, Geological Institute, Department of Earth Sciences, Zurich, Switzerland, ²Department of Environmental Sciences, University of Basel, Basel, Switzerland, ³Eawag, Department of Surface Waters Research and Management, Dübendorf, Switzerland

Long-chain diols are biomarkers commonly used in the marine realm to reconstruct several environmental parameters such as sea surface temperature and salinity. However, they are also produced in lacustrine and slow-flowing river environments, a characteristic that has proved to be useful to trace past riverine inputs in coastal sedimentary records. So far, their use in lacustrine settings is sparse as their controls are not well-known. Previous studies in two lakes have shown that long-chain diol distribution is linked to changes in temperature (in a small Spanish alpine lake), but also to water column stratification (in a large deep Swiss lake). To understand the controls on i) the presence of long-chain diols in lakes, and ii) the distribution of long-chain diol isomers, surface sediments from 52 Swiss lakes were studied. Long-chain diols are present in 57% of the lakes, and machine learning (i.e., random forest model) showed that their presence is mainly controlled by mean annual air temperature, sodium and potassium concentrations and area of the lakes. Long-chain diol isomer relative distribution seems to react to temperature, nutrient (here nitrate) and oxygen concentrations in the lakes. This new insight was tested on a short sedimentary core from Lake Zurich, and compared with other biomarker proxies (based on branched and isoprenoid glycerol dialkyl glycerol tetraethers), as well as with historical record of nutrient contents and temperature. Variations in the long-chain diol index (LDI) mirror measured temperature, but also reacted to changes in nutrients and oxygenation in the lake. This study highlights the potential of long-chain diols as a proxy to trace both nutrients and temperature in lakes, potentially on geological timescales.

KEYWORDS

long-chain diols, GDGT, random forest, lake Zurich, Swiss lakes

1 Introduction

Long-chain diols are biomolecules ubiquitous in marine sediments (e.g., [de Bar et al., 2020](#)), that are biosynthesized by unicellular algae (Eustigmatophytes, *Probooscia spp.*, *Apedinella radians*) (e.g., [Volkman et al., 1999](#); [Rampen et al., 2007](#); [Rampen et al., 2014](#); [Villanueva et al., 2014](#)). In freshwater environments, long-chain diols have been detected in lakes and rivers, but their production seems to be limited to standing water with low current ([Cranwell et al., 1987](#); [Robinson et al., 1989](#); [Xu et al., 2007](#); [Shimokawara et al., 2010](#);

Zhang et al., 2011; Romero-Viana et al., 2012; Rampen et al., 2014; 2014; Atwood et al., 2014; Villanueva et al., 2014; Zhang et al., 2015; Lattaud et al., 2018; van Bree et al., 2018; Häggi et al., 2019; Zhang et al., 2019; García-Alix et al., 2020; Wang et al., 2022). Their production is enhanced during period of lake stratification (van Bree et al., 2018; Lattaud et al., 2021). However, very few freshwater systems have been investigated (Rampen et al., 2014; Lattaud et al., 2021), and their ubiquity in lakes and rivers remains uncertain. The main long-chain diols detected in lakes are the C₃₂ 1,15-diol as well as the C₃₀ 1,15-diol. The latter is used in the long-chain diol index (LDI) proxy, which is the ratio of C₃₀ 1,15-diol over the C₂₈ and C₃₀ 1,13-diols. LDI is correlated with sea surface temperature in the ocean (Rampen et al., 2012; de Bar et al., 2020). Long-chain diols have also been tested against lake temperature in lake surface sediments, and used to reconstruct Holocene lake temperature in a sedimentary record in a small lake from the southern Iberian Peninsula (García-Alix et al., 2020; Toney et al., 2020). However, other studies, in deeper larger lakes more globally distributed, did not find this relation with temperature (Rampen et al., 2014; Lattaud, et al., 2021). Instead of being a proxy for temperature, Lattaud et al. (2021) proposed that long-chain diol isomer distribution is controlled by thermal stratification in large lakes such as Lake Geneva (Switzerland). Long-chain diol production reveals two peaks during the stratification period of the lake water column: June and August/October. 18S rRNA analysis did not allow for the clear identification of the producer(s). Hence, it is possible that two communities of long-chain diols producers exist or that a unique producer biosynthesizes different long-chain diol isomers depending on environmental conditions. The lack of producer identification using 18S rRNA was also true for Balzano et al. (2018), although in the marine realm. In the culture of several eustigmatophytes, temperature of the media impacted long-chain diol distribution in a single species and across species (Rampen et al., 2014; Balzano et al., 2017). This indicates that long-chain diol distribution changes to both species temperature adaptation and community shift under temperature change are possible.

Nevertheless, long-chain diols have the potential to become a very useful proxy for continental paleoenvironmental studies. However, if long-chain diols are to become a useful proxy, we need to get study the ecological preferences and niches of their producers. Machine learning methods, such as random forest model, have the potential to unravel complex relationships between proxy and environmental parameters and give clues on which direction the community should take experimental approaches. Since long-chain diols are relatively novel proxies in lacustrine environments, this approach will help unravel complex relationships.

Unlike long-chain diols, that are fairly novel biomarker in continental settings, glycerol dialkyl glycerol tetraethers (i.e., GDGTs, either isoprenoids iso, or branched, br) are biomarkers commonly used with known constraints. BrGDGTs are biomarkers from the membrane of bacteria which are used to reconstruct mean annual air temperature, soil or water pH, but have also been linked to anoxia in lakes (e.g., Peterse et al., 2012; De Jonge et al., 2014; De Jonge et al., 2019; Halamka et al., 2021; Martin et al., 2019; Raberg et al., 2021; O'Beirne et al., 2023). IsoGDGTs are produced by Archaea, and can be used to reconstruct lake

temperature if the lake studied is deep (e.g., Baxter et al., 2021; Sinninghe Damsté et al., 2022).

Here we propose to 1) test the presence of long-chain diols against common environmental parameters, and 2) the potential controls on isomer distribution in 52 Swiss lakes and in global lakes (from Rampen et al., 2014). A Random forest model was trained on the dataset to determine which environmental parameters explain their presence in the lakes. Statistical analyses were applied to understand long-chain diol isomer distribution. In addition, a short sediment core from Lake Zurich covering the last 100 years was investigated using long-chain diols and several glycerol dialkyl glycerol tetraether (GDGT)-based proxies, which were compared with instrumental records to assess their limits and accuracy in such environments. Long-chain diols are biomolecules easily analysed in polar fractions, and have the potential to become reliable proxies for reconstructing continental environmental parameters.

2 Material and methods

2.1 Sites and sampling

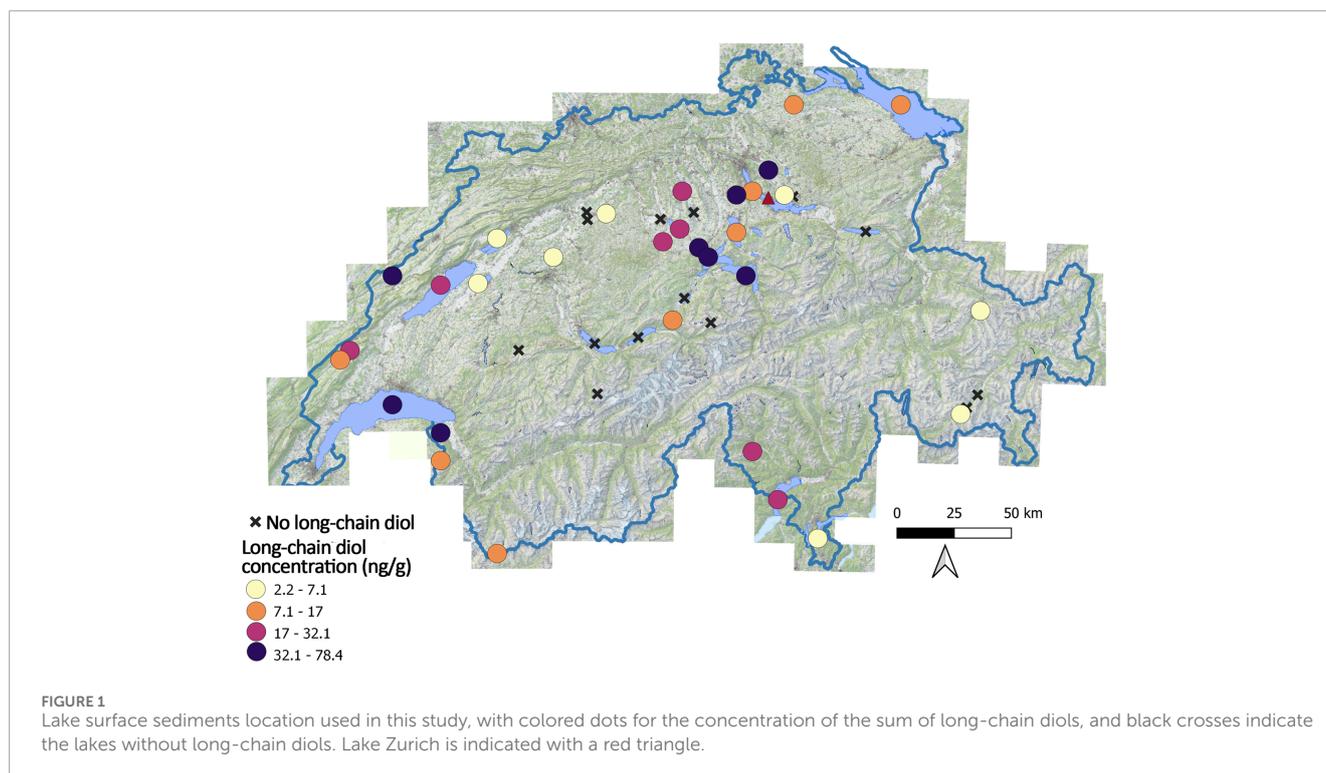
Surface sediments (0–2 cm) from 52 Swiss lakes were subsampled from short gravity cores retrieved between 2011 and 2019 (Figure 1). The lakes have been chosen so they cover as much physico-chemical parameters as possible and represent the entirety of the Swiss lake diversity (See list in Table 1). All sediment samples have been freeze-dried before extraction (Martin et al., 2024 submitted to the same issue).

Lake Zurich is a Swiss pre-alpine, mesotrophic lake with a surface area of 67.3 km², and a maximum depth of 136 m (Jankowski et al., 2006). The lake is divided in two by a moraine sill and the upper basin acts as a sediment trap for sediment brought by the Linth River. The lake is monomictic or dimictic depending on the prevailing winter conditions, but it recently tends to become oligomictic due to climate change (Livingstone, 2003; Friedrich et al., 2014; North et al., 2014). A weather station around Zurich Lake (SMA Zurich Fluntern 47°22.7'N; 8°33.9'E) was monitored since 1864 (Federal Office of Meteorology and Climatology MeteoSwiss, 2023) for air temperature and precipitation. Monitoring data and previous studies indicate that the lake was strongly eutrophic until the 1970s (Naeher et al., 2016), when phosphate concentration finally decreased (Bossard et al., 2001; Jankowski et al., 2006). The short sediment core studied (52 cm long) originates from the lower lake and was obtained in 2020 (ZH-20-11, N 47°16.995E 8° 35.624) with a gravity corer at the depth of 137 m. The upper 28 cm were laminated (Figure 2), while underneath turbidites intersect with laminated intervals. The upper section was sliced every centimeter and freeze-dried.

2.2 Methods

2.2.1 Lipid extraction

An Accelerator Solvent Extractor (ASE) was used to extract lipids from the surface sediment (0.4–2.4 g dry weight) with



dichloromethane (DCM): methanol (MeOH) (9: 1, v/v), as indicated in [Martin et al. 2024](#) (submitted to the same Research Topic). Whereas 1–7 g of freeze-dried sediments from Lake Zurich have been extracted following [Lattaud et al. \(2021b\)](#) with an EDGE (CEM) using DCM: MeOH (9: 1, v/v). All total lipid extracts were saponified with 10 mL of potassium hydroxide (KOH) (1.5 M) and 10 mL of MilliQ for 2 h at 70°C ([Lattaud et al., 2021](#)). 10 mL of MilliQ was added and the neutral fraction was liquid-liquid extracted with three times 10 mL of hexane, dried under N₂. All extracts were separated on a silica gel column (activated for 3 h at 450°C and deactivated with 1% MilliQ) into three fractions of increasing polarity. The solvent were different for the surface sediments (hexane, DCM and MeOH) and Lake Zurich sediments (hexane, hexane: DCM (1: 1, v/v) and DCM: MeOH (1: 1, v/v)). The polar fraction was filtered using a polytetrafluoroethylene (PTFE) 0.45 μm pore size filter prior to analysis. Two internal standards were added to all polar fractions before analysis: C₂₂ 7,16-diol (Interbioscreen) to quantify the long-chain diols and the C46 GDGT ([Huguet et al., 2006](#)) to quantify the GDGTs.

2.2.2 Lipid analysis

Long-chain diols were quantified on a gas chromatograph (GC, Agilent 7890B) coupled to a mass spectrometer (MS, Agilent 5977B) equipped with a DB-5MS column (30 m × 0.25 mm, 0.25 μm film thickness). GC-MS analyses were done at Eawag Dübendorf, in the Department of Surface Waters Research and Management. The temperature regime for the GC-oven was as follows ([Rampen et al., 2012](#)): held at 70°C for 1 min, increased to 130°C at 20°C min⁻¹, increased to 320°C at 4°C min⁻¹, held at 320°C during 25 min. The flow of hydrogen was held constant at 2 mL min⁻¹. The MS source temperature was held at 250°C and the MS quadrupole at 150°C.

The electron impact ionization energy of the source was 70 eV. The long-chain diols were quantified on selected ion monitoring (SIM) mode using their specific ions $m/z = 299.3$ (C₂₆ 1.12, C_{28:1} 1.14, C₂₈ 1.14), 313.3 (C₂₈ 1.13, C₃₀ 1.15), 327.3 (C₂₈ 1.12, C_{30:1} 1.14, C₃₀ 1.14) and 341.3 (C₃₀ 1.13, C₃₂ 1.15) ([Versteegh et al., 1997](#); [Rampen et al., 2012](#)). The internal standard was integrated on $m/z = 187.2$.

Long-chain diol index (LDI) [Eq. 1] is defined as follows ([Rampen et al., 2012](#)):

$$LDI = \frac{C_{30}1,15}{C_{30}1,15 + C_{30}1,13 + C_{28}1,13} \quad (1)$$

The temperature calibration tested is the one of [Rampen et al. \(2012\)](#), [Eq. 2], although it is only applied in marine settings (LT denotes lake temperature):

$$LT = LDI \times 30.769 - 3.329 \quad (2)$$

The GDGTs were analysed with high performance liquid chromatography (LC)/atmospheric pressure chemical ionization–mass spectrometry (MS) on an Agilent 1260 Infinity series LC-MS according to [Hopmans et al. \(2016\)](#). Selective ion monitoring of the [M + H]⁺ was used to detect and quantify the different GDGTs, according to [Huguet et al. \(2006\)](#), except that a similar response factor was assumed for the GDGTs and the internal standard. The tetraether index of 86 carbon atoms (TEX₈₆) was calculated following [Kim et al. \(2010\)](#) [Eq. 3]:

$$TEX_{86} = \frac{GDGT-2 + GDGT-3 + Cren'}{GDGT-1 + GDGT-2 + GDGT-3 + Cren'} \quad (3)$$

One temperature calibration was tested for the isoGDGTs ([Eq. 4] [Tierney et al., 2010](#)), and two for the brGDGTs

TABLE 1 Long-chain diol concentrations and proxy.

Station code	Lake	Latitude (°N)	Longitude (°W)	Diol (Y/N)	C ₂₈ 1.14	C ₂₈ 1.13	C ₃₀ 1.15	C ₃₀ 1.14	C ₃₀ 1.13	C ₃₂ 1.15	LDI
ALZ10-04	Alzasca	8.5859	46.26545	Y	1.93	0.85	16.42	nd	nd	nd	0.95
BAL1502	Baldegg	8.261673	47.19715	N	nd	nd	nd	nd	nd	nd	
BIE14-67A	Biel	7.162719	47.08674	Y	nd	2.14	1.86	nd	nd	nd	0.47
BRE14-02	Brenet	6.325012	46.67339	Y	1.64	7.24	6.40	nd	nd	2.52	0.47
BRI1501	Brienz	7.951089	46.71752	N	nd	nd	nd	nd	nd	nd	
Burg 17	Burgäschli	7.668139	47.16906	N	nd	nd	nd	nd	nd	nd	
CON19-1	Constance	9.424173	47.60765	Y	0.00	3.74	3.49	nd	nd	nd	0.48
CAD09-02SC	Cadagno	8.711517	46.550083	N	nd	nd	nd	nd	nd	nd	
DAV20-1	Davos	9.852379	46.81804	Y	0.39	1.43	0.49	nd	nd	nd	0.25
EGL19-02	Egel	8.816667	47.25778	N	nd	nd	nd	nd	nd	nd	
ENG20-3	Engstlen	8.357076	46.77372	N	nd	nd	nd	nd	nd	nd	
GEN19-01	Geneva	6.588988	46.453	Y	6.34	19.13	2.75	0.92	2.05	1.80	0.11
GRE17_23	Greifen	7.166738	45.86832	Y	6.83	13.07	6.19	nd	10.34	6.35	0.21
GSB20-3	Great St Bernard	8.671848	47.35398	Y	nd	3.26	1.84	nd	nd	2.59	0.36
Hin	Hinterburg	8.067222	46.718056	N	nd	nd	nd	nd	nd	nd	
HUT19-02	Iffig	7.406306	46.386467	N	nd	nd	nd	nd	nd	nd	
HAL19-1	Halwil	8.216563	47.27509	Y	2.60	12.18	5.02	nd	1.10	1.38	0.27
HUT19-02	Hüttwil	8.84374	47.61026	Y	3.13	3.78	2.68	nd	0.95	2.25	0.36
INK15-1	Inkwiler	7.66315	47.19843	N	nd	nd	nd	nd	nd	nd	
JOU17-03	Joux	6.290169	46.64016	Y	1.07	5.57	4.98	nd	nd	1.10	0.47
K	Klontaler	8.98649	47.027839	N	nd	nd	nd	nd	nd	nd	
ROU14-02	Lac des Rousses	6.090158	46.504364	N	nd	nd	nd	nd	nd	nd	
4WS19-1	Lucern	8.350969	47.02162	Y	1.68	17.43	10.04	0.35	1.44	0.68	0.35
LUG14-7A	Lugano	8.956786	45.9416	Y	nd	1.48	nd	nd	nd	2.77	nd
LUN20-2	Lungern	8.160384	46.80015	Y	nd	4.14	3.18	nd	nd	nd	0.43
LUT19-02	Lutzel	8.773296	47.26013	Y	1.81	2.08	1.12	nd	nd	0.82	0.35
MAG19-1	Maggiore	8.715791	46.10145	Y	16.88	5.62	1.58	nd	nd	4.63	0.22
MAU15-01	Mauen	8.075466	47.17139	N	nd	nd	nd	nd	nd	nd	
MORG20-2	Morgins	6.84645	46.24887	Y	nd	nd	nd	nd	nd	nd	
MOS_07_B	Moos	7.476583	47.02325	Y	0.00	3.00	3.80	nd	nd	nd	0.56

(Continued on the following page)

TABLE 1 (Continued) Long-chain diol concentrations and proxy.

Station code	Lake	Latitude (°N)	Longitude (°W)	Diol (Y/N)	C ₂₈ 1.14	C ₂₈ 1.13	C ₃₀ 1.15	C ₃₀ 1.14	C ₃₀ 1.13	C ₃₂ 1.15	LDI
LM13-SC1	Murten	7.067306	46.92872	Y	0.42	2.32	2.80	nd	0.20	0.67	0.53
NEU19-01	Neuchatel	6.842856	46.90428	Y	0.92	6.61	4.39	nd	1.20	1.13	0.36
NOI20-3	Schwarz	7.283619	46.66787	N	nd	nd	nd	nd	nd	nd	
OES20-3	Oeschinen	7.723383	46.49902	N	nd	nd	nd	nd	nd	nd	
ROT15-01	Rot	8.314196	47.06973	Y	7.20	12.10	9.58	nd	nd	4.76	0.44
SAE19-3D	Saengeliweiher	7.764822	47.19823	Y	0.00	0.00	4.40	nd	nd	nd	1.00
SAR19-1	Sarnen	8.209465	46.86849	N	nd	nd	nd	nd	nd	nd	
NOI20-3	Schwarz	7.28361877	46.66787425	N	nd	nd	nd	nd	nd	nd	
SEL20-1	Seelisberg	8.572139	46.95832	Y	11.41	36.80	15.46	nd	2.51	2.84	0.28
SEM19-1	Sempach	8.1592	47.14058	Y	7.90	5.34	2.58	nd	3.27	4.20	0.23
SILS20-2	Sil	9.730362	46.4212	Y	0.00	1.01	1.23	nd	nd	nd	0.55
SILV20-3	Silvaplana	9.787737	46.44549	N	nd	nd	nd	nd	nd	nd	
STMO20-1	St Moritz	9.848104	46.49517	N	nd	nd	nd	nd	nd	nd	
SOP15-03	Soppen	8.08101	47.08945	Y	7.82	9.45	5.81	nd	nd	nd	0.38
TAI14-3	Taillères	6.575285	46.96707	Y	0.00	0.00	14.73	nd	6.58	10.97	0.69
TAN20-1	Taney	6.841237	46.34565	Y	12.36	5.31	3.67	nd	4.34	2.83	0.28
THU15-01	Thun	7.708551	46.69351	N	nd	nd	nd	nd	nd	nd	
Tr	Trub	8.396805	46.793405	N	nd	nd	nd	nd	nd	nd	
TUR20	Türlen	8.499799	47.27104	Y	8.17	16.00	9.34	nd	1.98	4.75	0.34
WAL-14-08	Walen	9.222889	47.12347	N	nd	nd	nd	nd	nd	nd	
ZUG18-7_1	Zug	8.486854	47.11993	Y	2.35	2.83	2.63	0.88	4.15	nd	0.27
ZUR-14-01	Zurich	8.597868	47.2791	Y	2.05	3.62	2.62	nd	nd	0.45	0.42

Nd, not detected.

using the MBT'_{5ME} and reconstructing mean annual air temperature and mean temperature for the month above freezing MAF ([Eq. 5] MAAT (J. M. Russell et al., 2018); [Eq. 6] (Bauersachs et al., 2024). They have all been applied to lakes in previous studies, for [Eq.6] the isomer ratio IR_{6ME} [Eq.8] was calculated to validate the temperature reconstruction.

$$LT = 38.874 \times TEX_{86} - 3.4992 \quad (4)$$

$$MBT'_{5ME} = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIIa + IIIb + IIIc + IIa + IIb + IIc} \quad (5)$$

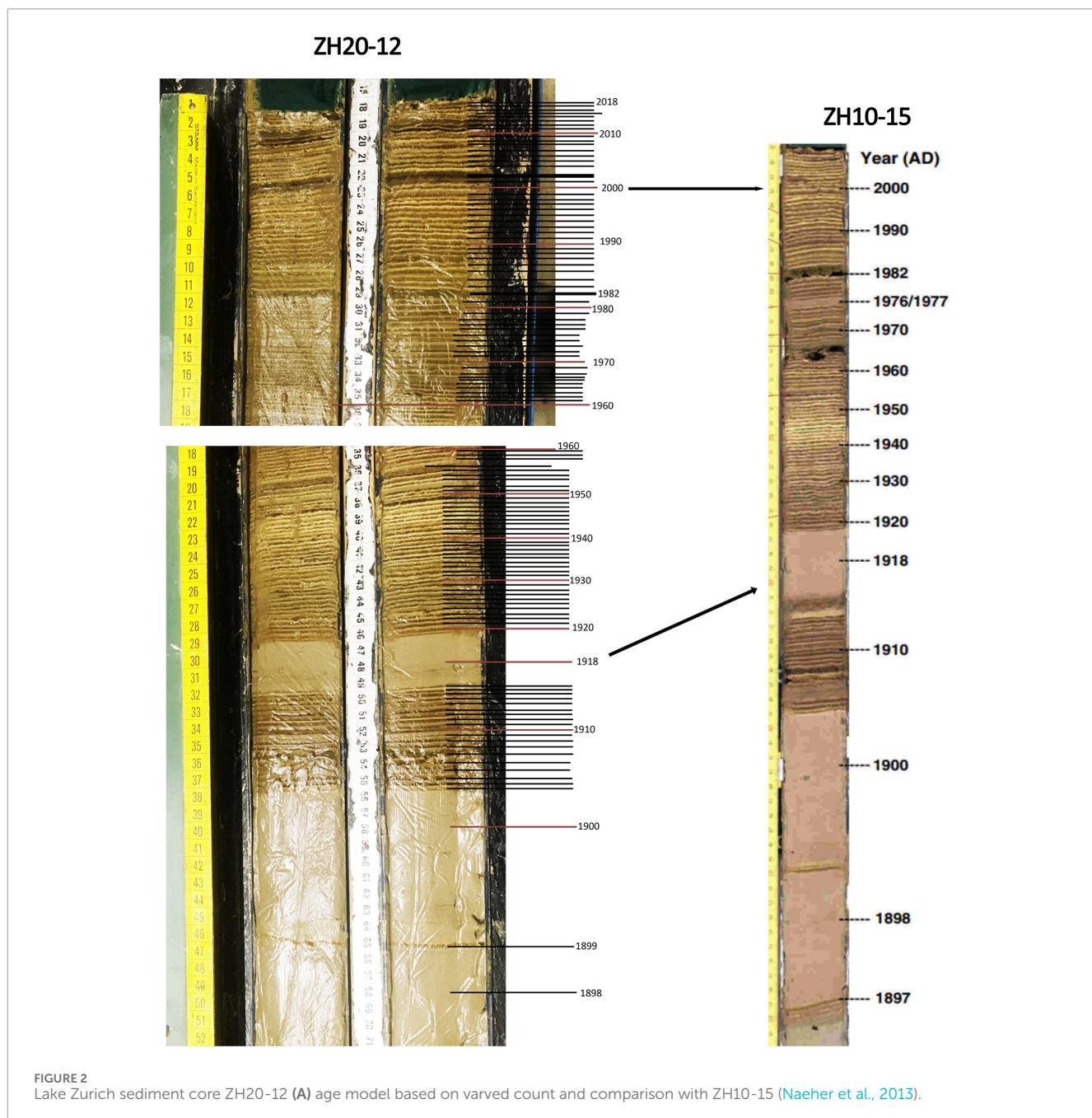
$$MAAT = -1.21 + 32.42 \times MBT'_{5ME} \quad (6)$$

$$MAF = 4.81 + 15.64 \times MBT'_{5ME} \quad (7)$$

$$IR_{6ME} = \frac{IIc' + IIb' + IIa' + IIIc' + IIIb' + IIIa'}{Ic' + IIb' + IIa' + IIIc' + IIIb' + IIIa' + IIa + IIb + IIc + IIIa + IIIb + IIIc} \quad (8)$$

2.2.3 Age model Lake Zurich core

The upper (>28 cm) laminations (seasonal forming varves) alternate between a light, slightly brownish layer and a dark, brown-black layer, likely representing spring/summer and fall/winter, respectively (Naehrer et al., 2013). The age model was based on varve counting and was crosschecked with previously published age models (Kelts, 1978; Naehrer et al., 2013; Figure 2). The youngest turbidite records the year 1918. The average sedimentation rate in



the varved section is approximately 0.29 cm yr^{-1} , very similar to the 0.28 cm yr^{-1} calculated by (Naeher et al., 2013).

2.2.4 Environmental parameters and statistical analysis

All background data are described by Martin et al., 2024 (submitted in the same issue, Supplementary Table S1). All analysis have been done on Rstudio (4.3.1, R Core Team, 2023), all graphics were done using ggplot2 (Wickham, 2016), principal component analysis were done using the *factoextra* package and *prcomp* function (Kassambara and Mundt, 2020). Random forest analysis was done following Martin et al., 2024 (submitted in the same issue). Briefly, in order to elucidate which environmental

parameters control the presence or absence of long-chain diols (all isomers summed), a random forest model was trained on the dataset using the *randomForest* package (Liaw and Wiener, 2002). As indicated by Martin et al., 2024 (submitted in the same issue), highly correlated parameters ($|r| > 0.7$) were removed such as elevation (as it correlates to air temperature), concentration in chlorine (Cl, correlating with sodium concentration, Na) concentration in calcium (Ca) (as it correlates with conductivity), total nitrogen (which correlates with concentration in nitrate) (Supplementary Table S2). Maximum depth was also removed due to its correlation to lake area. In addition, the model was tuned to select the best model parameters using 7-fold cross-validation. Accuracy, the number of correct predictions divided by total

number of predictions, was used as the metric to evaluate the performance of the model. The best performance was obtained for a $n_{tree} = 7000$ (Supplementary Figure S1A), and $mtry = 3$ (Supplementary Figure S1B). Accuracy, together with precision, the number of true positives divided by the total number of positive predictions, and specificity, which measures how well a machine learning model can detect true negative (true negative divided by true negative and false negative), were used to appreciate the model. Seven-fold cross-validation was performed to evaluate the model performance and the accuracy and importance were reported for each fold. Accumulated local effects (ALE) plots allow isolating the relationship of a given explanatory variable with the predicted outcome of the model (Molnar, 2020). They show the evolution of the prediction of the model across the range of values of each variable. ALE were obtained using the *FeatureEffects()* function from the R package *iml* (Molnar et al., 2018).

3 Results

3.1 Long-chain diol concentration and distribution

Long-chain diols were detected in 30 out of 52 lake surface sediments (57% of the lakes, Figure 1A). In the surface sediments where long-chain diols were detected, the total concentration varies from 2.2 to 78.8 ng g⁻¹ (Figure 1B). The main diols detected are the C₃₀ 1,15-diol (in average 30% ± 22%, with concentration varying from 0 to 16.4 ng g⁻¹, Table 1) the C₂₈ 1,13-diol (33% ± 15%, 0%–57%), the C₂₈ 1,14-diol (12% ± 12%, 0%–53%), and the C₃₂ 1,15-diol (10% ± 14%, with concentration varying from 0 to 11 ng g⁻¹, Table 1). The LDI [Eq. 1] varies from 0.11 to 0.69 (Table 1). To deconvolute the relation between long-chain diol isomers and environmental parameters in the lakes, a principal component analysis was performed (Figure 3A). On PC1 (explaining 38.6% of the variance), C₃₀ 1,13-diol loads positively, opposite to C₃₀ 1,15-diol. On PC2 (explaining 28.7% of the variance) C₃₀ 1,13- and C₃₀ 1,14-diols loads positively, opposite to C₃₂ 1,15-diol. On PC3 (explaining 22.4% of the variance), C₃₀ 1,15-diol loads positively.

In the core from Lake Zurich, long-chain diols were detected throughout the upper part of the core (Figure 4A). Their distribution in the core can be divided into three zones: top and bottom (0–7 cm and 18–25 cm) with dominant C₃₀ 1,15- and C₂₈ 1,13-diols (40% ± 11% and 40% ± 7%, respectively) and the middle of the core (8–17 cm) with the C₃₂ 1,15-diol dominating (38% ± 11%) (Figure 4B; Supplementary Table S3). Long-chain diol concentrations in the core are highly variable (0.4–126 ng g⁻¹) and show a very large increase between 2009–2013 (Figure 4A). LDI varies from 0.28 to 0.65, higher in the first 3 cm (0.49–0.65) compared with the rest of the core (Figure 5C).

3.2 GDGT concentration and distribution in the core

In the core, 17 brGDGTs were detected, to be noted is the presence of brGDGT-IIIa' identified by Weber et al. (2015). The tetra- and pentamethylated brGDGTs were dominant (av. 41.2%

± 4.3% and 40.9% ± 4.0%, respectively), with brGDGT-IIIa being the dominant one throughout the core (22% ± 5%), followed by brGDGT-IIIa' (14% ± 4%), IIa (15% ± 4%) and IIa' (15% ± 6%). The total brGDGT concentration varies from 23 to 860 ng g⁻¹ (Supplementary Table S4). Their distribution changed very little throughout the core, except during 1990–1993 when there was an increased abundance of brGDGT-IIa' compared to brGDGT-IIa. MBT_{SME} values vary very little and average at 0.30 ± 0.05. IR_{GME} averages at 0.38 ± 0.10, and is twice above 0.5 in 2002 and 2009 (0.59 and 0.55, respectively, Supplementary Table S4).

Six isoGDGTs were quantified: isoGDGT-0 to 3, crenarchaeol and crenarchaeol isomer. Crenarchaeol (40% ± 10%) and isoGDGT-0 (40% ± 10%) are the dominant isoGDGTs in the core. Total concentration varies from 209 to 16,600 ng g⁻¹ (Supplementary Table S4). TEX₈₆ is invariant through the core at 0.31 ± 0.02 (Supplementary Table S4).

3.3 Statistical analysis

The random forest model trained on this dataset (using 12 environmental parameters) is better at predicting the presence (sensitivity 68%) than the absence (specificity 44%) of long-chain diols (Table 2). The average accuracy for the 7-fold cross validation analysis is 57% ± 4% (ave. ± standard error). The importance of each variable can be quantified by the mean decrease in accuracy (MDA, representing the loss of accuracy associated with the removal of a given explanatory variable) and the mean decrease in Gini (Gini, which measures the loss of purity of the nodes caused by the exclusion of a given variable). The environmental parameters with the highest MDA values are mean annual air temperature (MAAT) followed by the concentration in potassium (K), sodium (Na) and the area of the lake. The parameters with the highest Gini values are MAAT, K, Na, conductivity, and the lake area (Figure 3A). These results are stable across the folds of cross-validation.

ALE plots (that show the evolution of the probability of presence of long-chain diols) reveal a threshold for some parameters such as MAAT, K, and lake area with a higher chance of finding long-chain diols after a certain value (Figure 3B). Other parameters do not present neither threshold nor gradient but rather optimum/a: Na, conductivity and NO₃⁻.

4 Discussion

4.1 Environmental controls on the presence of long-chain diols in lake surface sediments

Only 57% of the lakes ($n = 52$) had long-chain diols (Figure 1A). The dominant long-chain diols in the Swiss lakes are the C₂₈ 1,13- and C₃₀ 1,15-diols (37% ± 16%, 32% ± 21%, respectively). This distribution is similar to that of Lake Geneva suspended particulate matter during the late summer productive season (Lattaud et al., 2021). Although the accuracy of the model is relatively low (61%), the results of the random forest model point to the dominant influence of temperature on the presence of long-chain diols

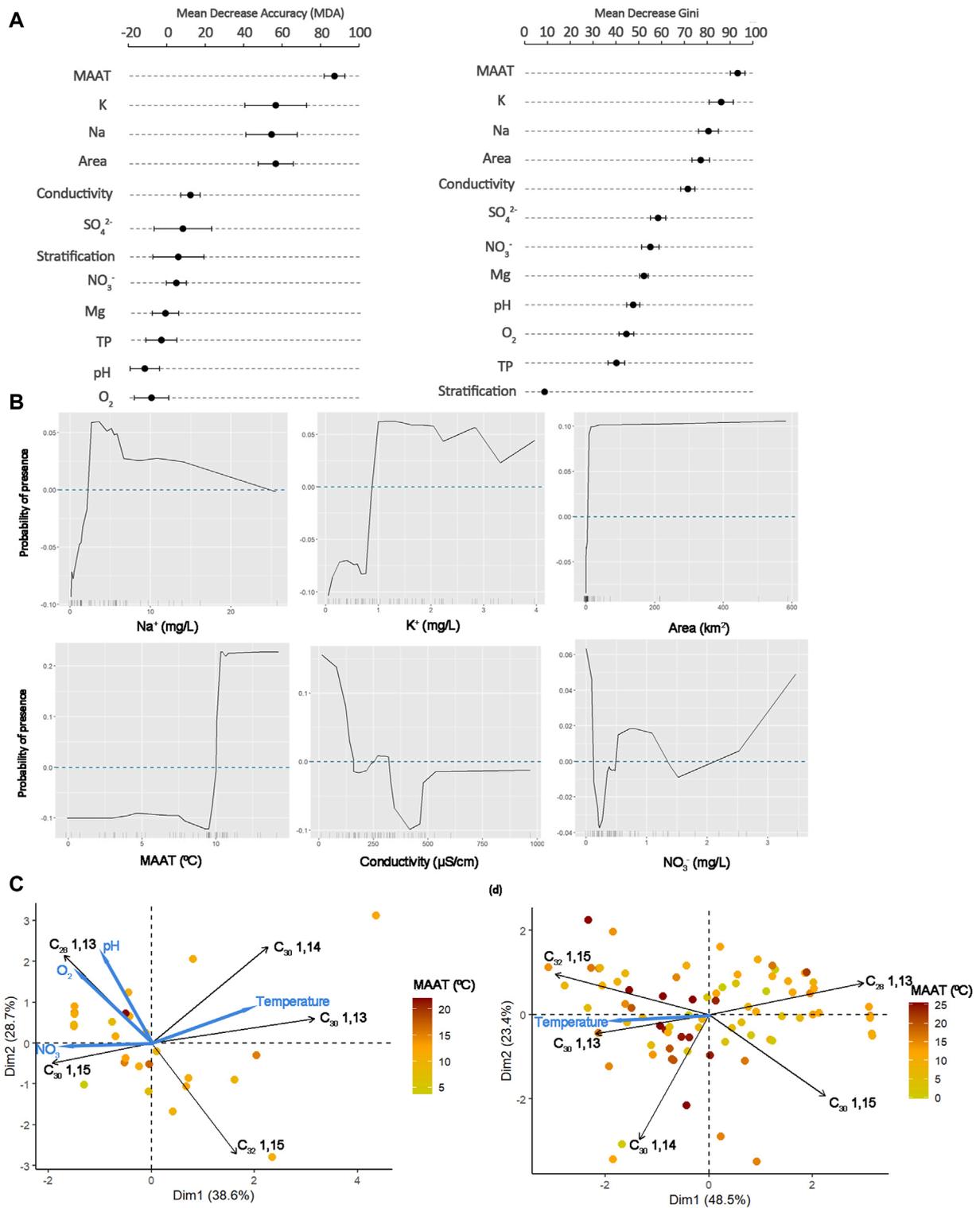
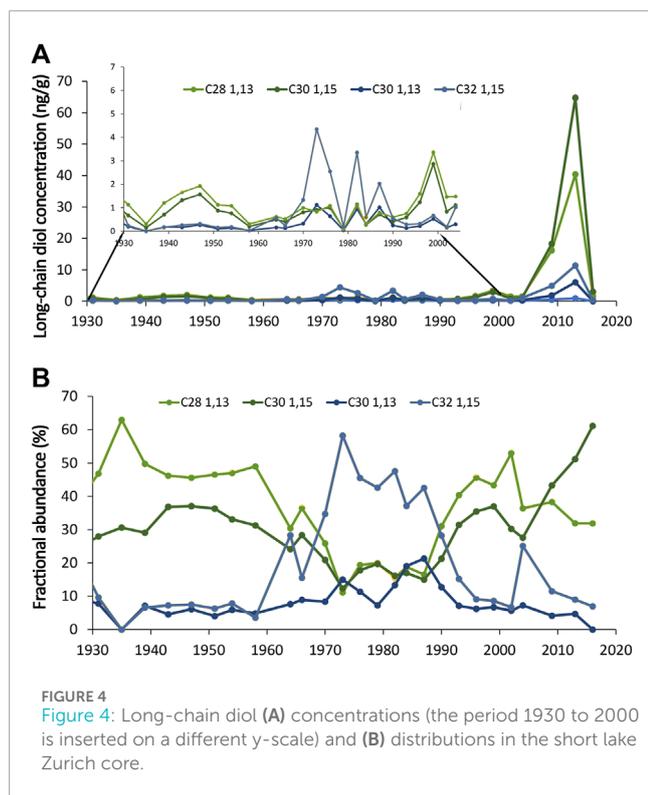


FIGURE 3
(A) Results from the random forest model (variable importance), mean decrease accuracy and Gini from the 7-fold cross-validation (average and standard error) **(B)** ALE plots for the most important parameters. Principal component analysis of the fractional abundance of long-chain diols and selected environmental parameters in **(C)** Swiss lakes and **(D)** Swiss lakes and global lakes (Rampen et al., 2014). MAAT denotes mean annual air temperatures.



in the Swiss lakes (Figure 3A). In particular, the probability of their occurrence increases above 10°C (Figure 3B). In the Swiss lakes, the frequency distribution plot confirms that long-chain diols are less frequent in lakes below 10°C than in those above (Supplementary Figure S2A). However, the low frequency of long-chain diols at temperatures below 10°C seems to be specific to Swiss lakes as cultures of pure eustigmatophytes have been found to grow at a large temperature range (8–20°C; Rampen et al., 2014) and the global lake dataset (although not reporting the absence of long-chain diols in lakes) shows the presence of long-chain diols below 10°C.

In the Swiss lakes, Na⁺ and K⁺ average concentrations in the lakes where long-chain diols are present are significantly higher than in the lakes without (*t*-test $p = 0.03$ and $p = 0.02$, respectively; Kolmogorov-Smirnov test $p = 0.85$). K⁺ is an important ion for photosynthesising organisms such as green unicellular algae, hence it is not surprising that its presence is determinant for the presence of long-chain diols in lakes (Figure 3A). The probability of presence of long-chain diols is increased for concentrations above a threshold of 1 mg L⁻¹ (Figure 3B). In addition, Na⁺ is also an important ion for long-chain diol presence, between concentration of 2.5 and 7 mg L⁻¹ (Figures 3A,B). Usually, organisms maintain a low level of Na⁺ and a high level of K⁺ in their cells especially when dealing with environments with high Na⁺ concentration and high alkalinity (Li et al., 2023). In addition, Na⁺ increase can create a stress factor for green algae which biomass generally decreases (Shetty et al., 2019; Sinetova et al., 2021). The presence of long-chain diols in relation to conductivity shows higher probability at low conductivities probably reflecting the optima at low concentrations of K⁺, Na⁺ and other ions.

The area of the lakes (which is correlated with the depth of the lake, $r = 0.71$) has a high MDA and appears as an important parameter for long-chain diol occurrence (Figure 3A). Area and depth define a lake's mixing model and its stratification state. In turn, they will define nutrient distribution in the lake's water column.

The presence of several ranges of nitrate concentrations where the probability of presence of long-chain diol increases might be related to the adaptability of eustigmatophytes to different nutrient states. Considering only freshwater algae, long-chain diols with 28, 30 and 32 carbon atoms have so far only been found in pure culture of unicellular algae such as *Eustigmatos*, *Vischeria* and *Nannochloropsis*, part of the Eustigmatophycean genera (Volkman et al., 1992; Rampen et al., 2014; Kryvenda et al., 2018). A previous study in Lake Geneva (Switzerland) highlighted the potential for long-chain diols to be produced by yet uncultured eustigmatophytes (Lattaud et al., 2021). Not much is known about eustigmatophytes in natural freshwater environments, only that Eustigmatophyceae can grow on a wide variety of nitrogen source (Santos, 1996) and can be photo and heterotroph (Santos, 1996; Marudhupandi et al., 2016). Moreover, *Nannochloropsis* spp. can be present in high abundance and can represent the majority of the primary production of inland lakes and ponds of different trophic state (Krienitz et al., 2000; Fietz et al., 2005; Fawley and Fawley, 2007; Fawley et al., 2014). However, nutrient levels do not seem to be important parameters for the presence of long-chain diols in the model (Figure 3A).

In summary, long-chain diol presence seems to be mainly controlled by the temperature of the lake and its ion composition. However, as this model was trained on a limited dataset, more lakes with their biogeochemical parameters should be added to increase its reliability and accuracy.

4.2 Environmental controls on the distribution of long-chain diol isomers in lake surface sediments

A Principal Component Analysis (PCA) was performed on the long-chain diol isomer distribution to assess their variation, selected environmental variables were added to identify the controls of isomer distribution (Figure 3C). On PC1, C₃₀ 1,13-diol loads positively opposite to C₃₀ 1,15-diol. This relation is similar to that observed in marine environments where C₃₀ 1,15-diol fractional abundance is positively correlated to temperature (e.g., de Bar et al., 2020). However, in this lake dataset, although temperature also loads on PC1, it is to the opposite of C₃₀ 1,15-diol. This is a major difference between lake and marine realm (Rampen et al., 2012) and might impact the applicability of the temperature proxy LDI as it is now defined. However, it is to be noted that the PC1 and PC2 only explained 67.3% of the variance. Temperature effect on isomer distribution was shown in pure culture before (Rampen et al., 2014), and linked to membrane adaptability to temperature change, but not in a classical way with increased unsaturation at lower temperatures (e.g., Russell and Fukunaga, 1990; Suutari and Laakso, 1994), eustigmatophytes may have adopted other methods to adapt their lipid composition to temperature (Rampen et al., 2014). In the PCA including global lakes (lakes from Rampen et al., 2014), the relation between the isomers is the same (Figure 3D). On PC1 (explaining

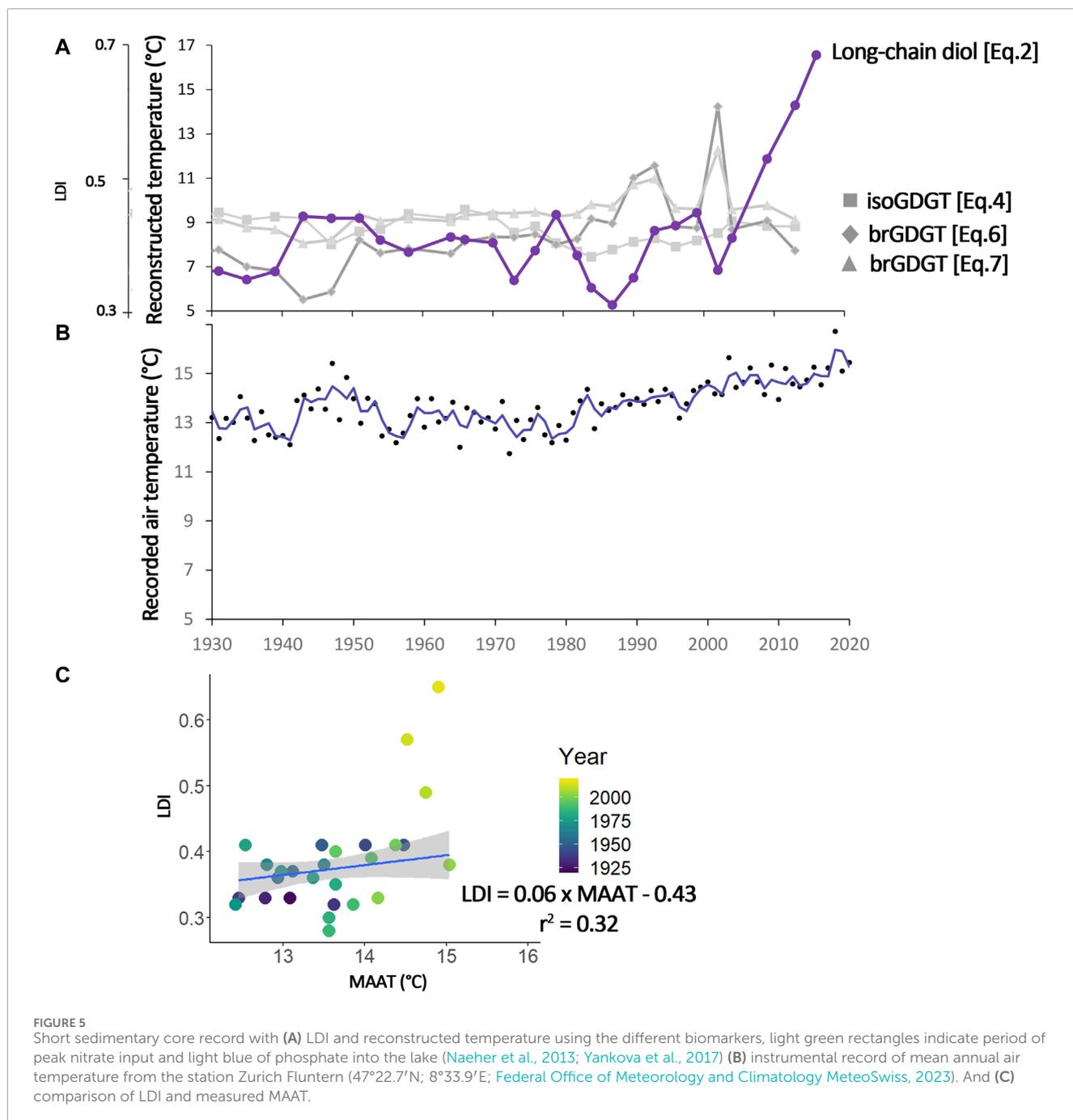


TABLE 2 Confusion matrix from the random forest model (accuracy = 61.5%).

	N	Y	
N	15	8	0.348
Y	12	17	0.414

48.5% of the variance) temperature is again loading close to C_{30} 1,13-diol confirming the relationship between long-chain diol isomer and temperature in lakes.

Aside from temperature, the concentration in nitrate (NO_3^-) loads negatively on PC1, close to the C_{30} 1,15-diol. Nitrate concentration might be an additional control on the distribution of long-chain diol isomers. Nutrient concentration (for example, nitrate) could control the presence of specific long-chain diol producers (which would produce more C_{30} 1,15-diol). In pure eustigmatophyte culture, nitrite depletion has been found to decrease the amount of C_{32} 1,15-diol (Balzano et al., 2017) highlighting the possibility for eustigmatophyte to change their long-chain diol composition depending on nutrient state. Increased amounts of C_{30} 1,15-diol are observed in Lake Murten and Moos where large amounts of nitrate are present ($>2 \text{ mg L}^{-1}$).

On PC2 (Figure 3A), C_{28} 1,13-diol, pH and oxygen concentration load positively, opposite to the C_{32} 1,15-diol. pH and oxygen concentration can be linked to the presence of a chemocline and potential stratification in the lakes. Previous studies have found that stratification is likely enhancing the production of some long-chain diol isomers, in particular C_{32} 1,15-diol (Lattaud et al., 2021). In pure culture of eustigmatophytes, oxidative stress in the form of hydrogen peroxide adjunct resulted in a general decrease of long-chain diol concentration but without changing the long-chain diol isomer distribution (Balzano et al., 2017). Hence, the increased proportion of C_{32} 1,15-diol under stratification change might be linked to species change rather than species adaptation.

Hence, temperature seems to control not only the long-chain diol presence but also the distribution of the C_{30} 1,15- and C_{30} 1,13-diol isomers, making them potentially great paleothermometers. Several other environmental factors such as variation in nitrate and oxygen concentrations, and pH seem to influence the other isomers, i.e., C_{28} 1,13- and C_{32} 1,15-diols. The latter parameters do not show any impact on the probability of presence of long-chain diols. We tested these relations in the upper section (27 cm long) of a short sedimentary core from Lake Zurich covering instrumental data.

4.3 Environmental controls on sedimentary long-chain diols

Following the results from the surface sediments, this study looked at the presence of long-chain diols and their link with mean annual air temperature using the previously known relationship in marine settings (e.g., Rampen et al., 2012; de Bar et al., 2020). Long-chain diol distribution, LDI, and LDI-reconstructed temperatures are then compared with instrumental air temperatures. The region of Zurich (Switzerland) has been monitored for temperature since 1864 (Federal Office of Meteorology and Climatology MeteoSwiss, 2023; Supplementary Table S5). The last 40 years show a gradual increase in mean annual air temperatures (reaching 15.4°C in 2020, Figure 5B), that reflects anthropogenic climatic change also recorded in the rest of the globe. During the period covered by the record (1930–2020), LDI and measured temperatures show a moderate but significant correlation ($r = 0.56$, $p < 0.05$, $n = 27$, Figure 5C) indicating the potential for LDI to track changes in air temperature. In particular, the warm period between 1940 and 1955, is recorded by the LDI (Figure 5A). However, during some specific periods, LDI and LDI-reconstructed temperatures are not correlating with recorded temperatures (see deviation from correlation in Figure 5C). For example, during the most recent years (2008 onward, $n = 4$) when LDI-reconstructed temperatures are increasing much faster than measured temperatures. This is also a period of higher long-chain diol concentrations with enhanced proportions of the C_{30} 1,15- and C_{28} 1,13-diols (Figure 4A). These long-chain diols have been found to be indicative of lower temperature and increased nitrate concentration for the C_{30} 1,15-diol and oxygen and pH increase for the C_{28} 1,13-diol. However, Lake Zurich surface waters are warmer and less oxygenated since 1950 (Naeher et al., 2013; Yankova et al., 2017). Nitrate concentrations are higher in the surface waters since 2005 (Yankova et al., 2017), which might explain the high proportion of C_{30} 1,15-diol. In addition, there is a decrease in phosphate concentration and a reduction of spring

mixing of the water column in Lake Zurich resulting in a decrease of diatom blooms since 2000 (Yankova et al., 2017). This could free niches for the long-chain diol producers, specifically for C_{30} 1,15- and C_{28} 1,13-diol producers disregarding environmental conditions. In conclusion, this recent deviation from recorded air temperature is likely coming from a change in nitrate concentration of the surface waters and likely decrease of competition.

Another period deviates from the recorded temperature, between 1970 and 1983, when the LDI-reconstructed temperatures are lower than expected (Figures 5A,B). During this period, there is a change in the distribution of long-chain diol isomers with an enhanced proportion and concentration of C_{32} 1,15- and C_{30} 1,13-diol in the sediment (from av. $12\% \pm 6\%$ and $5\% \pm 2\%$ before 1970 to av. $38\% \pm 11\%$ and $12\% \pm 5\%$ between 1970–1983, respectively, Figure 4B). A change in nutrient concentration (in particular nitrate) driven by recorded historical maximum nutrient input from agriculture occurred during that period (e.g., Naeher et al., 2013; Yankova et al., 2017; Fiskal et al., 2019). Nitrate inputs can change the proportion of C_{30} 1,15-diol over C_{30} 1,13-diol, affecting the LDI-temperature reconstructions. The enhanced nutrient input drove an enhanced stratification in the lake and the apparition of bottom water anoxia (Naeher et al., 2013) but also less oxygen in surface waters (Yankova et al., 2017). This change in oxygen concentration can explain the change in fractional abundance and absolute concentration of C_{32} 1,15-diol. Long-chain diol isomers seem to degrade at a similar rate in natural marine environment and in aerobic and anaerobic experiments (Grossi et al., 2001; Rodrigo-Gámiz et al., 2016; Reiche et al., 2018). Hence differential degradation cannot explain the different proportion of the isomers. Long-chain diol distribution change could be explained by a community shift driven by a change in nutrients in the lake, similar to what has been observed in Lake Geneva during the two long-chain diol productive seasons. In Lake Geneva, observed change in long-chain diol distribution can likely be explained by a shift in producers (Lattaud et al., 2021).

In comparison to LDI-reconstructed temperatures, temperatures reconstructed with other biomarkers such as branched (br) or isoprenoid (iso) GDGTs do not record the early 1940–1955 warming (Figure 5A). Positively, the br and iso GDGT-reconstructed temperatures do not react to changes in nutrient concentration. However, all br and iso GDGT-based temperature reconstructions suggest colder than measured mean air temperatures (Figure 5A). Finally, in the latest years of the record (after 2000), GDGT-reconstructed temperatures are not following the measured air temperature, similarly to LDI-reconstructed temperature. The post-2000 years are a period of enhanced concentrations in br and isoGDGTs. High concentrations and fractional abundance in crenarchaeol (0.4 ± 0.1 throughout the core) might indicate the presence of Thaumarchaeota living below the thermocline (Sinninghe Damsté et al., 2022). This is in agreement with low and stable TEX_{86} (0.31 ± 0.02) reconstructing below-thermocline temperatures ([Eq. 4] $8.65 \pm 0.49^\circ\text{C}$) similar to the temperature measured below the thermocline, which remained stable throughout the considered period (e.g., Yankova et al., 2017). At least part of the brGDGTs are produced *in situ* in the lake as indicated by the presence of brGDGT-IIIa⁺ (Weber et al., 2015), in addition to potential soil-derived brGDGT inputs (e.g., Weijers et al., 2009). The 6-methyl isomer ratio (IR_{6ME}), which

can be used as a quality control for the use of MBT_{5ME} for temperature reconstruction in European lakes, remains below cutoff for most of the core, indicating that temperature should be reconstructed accurately (Bauersachs et al., 2024). However, large variation in oxygen concentration can affect brGDGT-reconstructed temperatures and might explain why temperature reconstructions are not following measured temperatures in this record (Halamka et al., 2021; O'Beirne et al., 2023; Wu et al., 2021). In addition, changes in brGDGT source proportion lined with growing agriculture around the lake, enhancing soil inputs in the lake, could bias the temperature reconstruction (Martin et al., 2019).

In summary, reconstructing lake temperature using long-chain diols can be possible, although clear knowledge of nutrients, and/or oxygen concentration change is needed. In addition, it is necessary to either refine or develop a new LDI index that would exclude long-chain diols affected by other biogeochemical parameters.

5 Conclusion

Long-chain diols were studied in 52 Swiss lake surface sediments and one short lake sediment core (Lake Zurich). The environmental parameters influencing their presence were tested using random forest modelling, highlighting the control of temperature, sodium, and potassium, with a threshold of 10°C for temperature (which might be a regional specificity). The isomer distribution is linked to temperature, nitrate concentration, oxygen and pH. In the short core, long-chain diols show a good correlation with recorded temperature. However, periods of enhanced agriculture-driven nutrients and enhanced de-oxygenation modify this relation. In conclusion, long-chain diols are a great potential paleothermometer when nutrient levels are stable, while they hold the key to obtain nutrient and oxygenation records when those vary greatly. More studies of presence/absence of long-chain diols in lakes in parallel with biogeochemical lake parameter measurements are needed to fully develop strong long-chain diol proxies.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: <https://github.com/JulieLattaud/RandomForestDiol>.

Author contributions

JL: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Writing–original draft. CM: Methodology, Resources,

Writing–review and editing. RL: Resources, Writing–review and editing. BZ: Formal Analysis, Writing–review and editing. ND: Project administration, Resources, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. JL was funded by a Rubicon grant (019.183EN.002) from NWO, Netherlands Organization for scientific research, and a Career Seed Grant (SEED-11 20-1) from ETH Zurich. Open access funding by ETH Zurich.

Acknowledgments

Core ZH20 was obtained with the help of Adrian Gilli, laboratory support at ETHZ was delivered by Daniel Montluçon and Timothy I. Eglinton. We thank the reviewers and editor for their time and comments that improve this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feart.2024.1409137/full#supplementary-material>

References

- Atwood, A. R., Volkman, J. K., and Sachs, J. P. (2014). Characterization of unusual sterols and long chain diols, triols, keto-ols and n-alkenols in El Junco Lake, Galápagos. *Org. Geochem.* 66, 80–89. doi:10.1016/j.orggeochem.2013.11.004
- Balzano, S., Lattaud, J., Villanueva, L., Rampen, S., Brussaard, C. P. D., van Bleijswijk, J., et al. (2018). A quest for the biological sources of the ubiquitous long chain alkyl diols in the marine realm. *Biogeosciences* (March), 1–40. doi:10.5194/bg-2018-97
- Balzano, S., Villanueva, L., de Bar, M., Sinnighe Damsté, J. S., and Schouten, S. (2017). Impact of culturing conditions on the abundance and composition of long chain alkyl diols in species of the genus *Nannochloropsis*. *Org. Geochem.* 108 (May), 9–17. doi:10.1016/j.orggeochem.2017.02.006
- Bauersachs, T., Schubert, C. J., Mayr, C., Gilli, A., and Schwark, L. (2024). Branched GDGT-based temperature calibrations from Central European lakes. *Sci. Total Environ.* 906, 167724. doi:10.1016/j.scitotenv.2023.167724

- Baxter, A. J., van Bree, L. G. J., Peterse, F., Hopmans, E. C., Villanueva, L., Verschuren, D., et al. (2021). Seasonal and multi-annual variation in the abundance of isoprenoid GDGT membrane lipids and their producers in the water column of a meromictic equatorial crater lake (Lake Chala, East Africa). *Quat. Sci. Rev.* 273, 107263. doi:10.1016/j.quascirev.2021.107263
- Bossard, P., Gammeter, S., Lehmann, C., Schanz, F., Bachofen, R., Bürgi, H.-R., et al. (2001). Limnological description of the lakes Zürich, Lucerne, and Cadagno. *Aquat. Sci.* 63 (3), 225–249. doi:10.1007/PL00011353
- Cranwell, P. A., Eglinton, G., and Robinson, N. (1987). Lipids of aquatic organisms as potential contributors to lacustrine sediments-II. *Org. Geochem.* 11 (6), 513–527. doi:10.1016/0146-6380(87)90007-6
- de Bar, M. W., Weiss, G., Yildiz, C., Rampen, S. W., Lattaud, J., Bale, N. J., et al. (2020). Global temperature calibration of the Long chain Diol Index in marine surface sediments. *Org. Geochem.* 142, 103983. doi:10.1016/j.orggeochem.2020.103983
- De Jonge, C., Hopmans, E. C., Zell, C. I., Kim, J.-H. H., Schouten, S., Sinninghe Damsté, J. S., et al. (2014). Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: implications for palaeoclimate reconstruction. *Geochimica Cosmochimica Acta* 141, 97–112. doi:10.1016/j.gca.2014.06.013
- De Jonge, C., Radujković, D., Sigurdsson, B. D., Weedon, J. T., Janssens, I., and Peterse, F. (2019). Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils. *Org. Geochem.* 137, 103897. doi:10.1016/j.orggeochem.2019.07.006
- Fawley, K. P., Eliáš, M., and Fawley, M. W. (2014). The diversity and phylogeny of the commercially important algal class Eustigmatophyceae, including the new clade Goniochloridales. *J. Appl. Phycol.* 26 (4), 1773–1782. doi:10.1007/s10811-013-0216-z
- Fawley, K. P., and Fawley, M. W. (2007). Observations on the diversity and ecology of freshwater Nannochloropsis (Eustigmatophyceae), with descriptions of new taxa. *Protist* 158 (3), 325–336. doi:10.1016/j.protis.2007.03.003
- Federal Office of Meteorology and Climatology MeteoSwiss (2023). Zurich/fluntern. Available at: www.meteosuisse.admin.ch/services-et-publications/applications/ext/climate-tables-homogenized.html.
- Fietz, S., Bleiß, W., Hepperle, D., Koppitz, H., Krienitz, L., and Nicklisch, A. (2005). First record of Nannochloropsis limnetica (eustigmatophyceae) in the autotrophic picoplankton from Lake Baikal. *J. Phycol.* 41 (4), 780–790. doi:10.1111/j.0022-3646.2005.04198.x
- Fiskal, A., Deng, L., Michel, A., Eickenbusch, P., Han, X., Lagostina, L., et al. (2019). Effects of eutrophication on sedimentary organic carbon cycling in five temperate lakes. *Biogeosciences* 16 (19), 3725–3746. doi:10.5194/bg-16-3725-2019
- Friedrich, J., Janssen, F., Aleynik, D., Bange, H. W., Boltacheva, N., Çagatay, M. N., et al. (2014). Investigating hypoxia in aquatic environments: diverse approaches to addressing a complex phenomenon. *Biogeosciences* 11 (4), 1215–1259. doi:10.5194/bg-11-1215-2014
- García-Alix, A., Toney, J. L., Jiménez-Moreno, G., Pérez-Martínez, C., Jiménez, L., Rodrigo-Gámiz, M., et al. (2020). Algal lipids reveal unprecedented warming rates in alpine areas of SW Europe during the industrial period. *Clim. Past* 16 (1), 245–263. doi:10.5194/cp-16-245-2020
- Grossi, V., Blokker, P., and Sinninghe Damsté, J. S. (2001). Anaerobic biodegradation of lipids of the marine microalga Nannochloropsis salina. *Org. Geochem.* 32 (6), 795–808. doi:10.1016/S0146-6380(01)00040-7
- Häggi, C., Schefuß, E., Sawakuchi, A. O., Chiessi, C. M., Mulitza, S., Bertassoli, D. J., et al. (2019). Modern and late Pleistocene particulate organic carbon transport by the Amazon River: insights from long-chain alkyl diols. *Geochimica Cosmochimica Acta* 262, 1–19. doi:10.1016/j.gca.2019.07.018
- Halamka, H., McFarlin, J. M., Younkin, A. D., Depoy, J., Dildar, N., and Kopf, S. (2021). Oxygen limitation can trigger the production of branched GDGTs in culture. *Geochim. Perspect. Lett.* 19, 36–39. doi:10.7185/geochemlet.2132
- Hopmans, E. C., Schouten, S., and Sinninghe Damsté, J. S. (2016). The effect of improved chromatography on GDGT-based palaeoproxies. *Org. Geochem.* 93, 1–6. doi:10.1016/j.orggeochem.2015.12.006
- Huguet, C., Kim, J.-H., Damsté, J. S. S., and Schouten, S. (2006). Reconstruction of sea surface temperature variations in the Arabian Sea over the last 23 kyr using organic proxies (TEX86 and U37K'). *Paleoceanography* 21 (3). doi:10.1029/2005pa001215
- Jankowski, T., Livingstone, D. M., Bühner, H., Forster, R., and Niederhauser, P. (2006). Consequences of the 2003 European heat wave for lake temperature profiles, thermal stability, and hypolimnetic oxygen depletion: implications for a warmer world. *Limnol. Oceanogr.* 51 (2), 815–819. doi:10.4319/lo.2006.51.2.0815
- Kassambara, A., and Mundt, F. (2020). *factoextra: Extract and Visualize the Results of Multivariate Data Analyses* (package version 1.0.7). Available at: <https://CRAN.R-project.org/package=factoextra>.
- Kelts, K. R. (1978). “Geological and sedimentary evolution of lakes zuerich and zug, Switzerland [Dr.Nat., eidgenössische Technische Hochschule zuerich (Switzerland)],” in *ProQuest dissertations and theses*. Available at: <https://www.proquest.com/docview/302900865/abstract/CDAD50028F6347B6PQ/1>.
- Kim, J.-H. H., van der Meer, J., Schouten, S., Helmke, P., Willmott, V., Sangiorgi, F., et al. (2010). New indices and calibrations derived from the distribution of crenarchaeal isoprenoid tetraether lipids: implications for past sea surface temperature reconstructions. *Geochimica Cosmochimica Acta* 74 (16), 4639–4654. doi:10.1016/j.gca.2010.05.027
- Krienitz, L., Hepperle, D., Stich, H.-B., and Weiler, W. (2000). Nannochloropsis limnetica (Eustigmatophyceae), a new species of picoplankton from freshwater. *Phycologia* 39 (3), 219–227. doi:10.2216/i0031-8884-39-3-219.1
- Kryvenda, A., Rybalka, N., Wolf, M., and Friedl, T. (2018). Species distinctions among closely related strains of Eustigmatophyceae (Stramenopiles) emphasizing ITS2 sequence-structure data: Eustigmatos and Vischeria. *Eur. J. Phycol.* 53 (4), 471–491. doi:10.1080/09670262.2018.1475015
- Lattaud, J., Balzano, S., van der Meer, M. T. J., Villanueva, L., Hopmans, E. C., Sinninghe Damsté, J. S., et al. (2021). Sources and seasonality of long-chain diols in a temperate lake (Lake Geneva). *Org. Geochem.* 156, 104223. doi:10.1016/j.orggeochem.2021.104223
- Lattaud, J., De Jonge, C., Elling, F. J., Pearson, A., and Eglinton, T. I. (2021b). Microbial lipid signatures in Arctic deltaic sediments – insights into methane cycling and climate variability. *Org. Geochem.* 157, 104242. doi:10.1016/j.orggeochem.2021.104242
- Lattaud, J., Kirkels, F., Peterse, F., Freymond, C. V., Eglinton, T. I., Hefter, J., et al. (2018). Long-chain diols in rivers: distribution and potential biological sources. *Biogeosciences* 15 (13), 4147–4161. doi:10.5194/bg-15-4147-2018
- Li, W., Zhang, Y., Ren, H., Wang, Z., OuYang, Y., Wang, S., et al. (2023). Identification of potassium transport proteins in algae and determination of their role under salt and saline-alkaline stress. *Algal Res.* 69, 102923. doi:10.1016/j.algal.2022.102923
- Liaw, A., and Wiener, M. (2002). Classification and regression by randomForest. *R. News* 2 (3), 18–22.
- Livingstone, D. M. (2003). Impact of secular climate change on the thermal structure of a large temperate central European lake. *Clim. Change* 57 (1), 205–225. doi:10.1023/A:1022119503144
- Martin, C., Ménot, G., Thouveny, N., Davtian, N., Andrieu-Ponel, V., Reille, M., et al. (2019). Impact of human activities and vegetation changes on the tetraether sources in Lake St Front (Massif Central, France). *Org. Geochem.* 135, 38–52. doi:10.1016/j.orggeochem.2019.06.005
- Martin, C., Richter, N., Lloren, R., Amaral-Zettler, L., and Dubois, N. (2024). Machine learning reveals that sodium concentration and temperature influence alkenone occurrence in Swiss and freshwater lakes globally. *Frontiers in Earth Science*.
- Marudhupandi, T., Sathishkumar, R., and Kumar, T. T. A. (2016). Heterotrophic cultivation of *Nannochloropsis salina* for enhancing biomass and lipid production. *Biotechnol. Rep.* 10, 8–16. doi:10.1016/j.btre.2016.02.001
- Molnar, C. (2020). Interpretable machine learning. Available at: <https://christophm.github.io/interpretable-ml-book/>.
- Molnar, C., Casalicchio, G., and Bischl, B. (2018). Iml: an R package for interpretable machine learning. *J. Open Source Softw.* 3 (26), 786. doi:10.21105/joss.00786
- Naeher, S., Gilli, A., North, R. P., Hamann, Y., and Schubert, C. J. (2013). Tracing bottom water oxygenation with sedimentary Mn/Fe ratios in Lake Zurich, Switzerland. *Chem. Geol.* 352, 125–133. doi:10.1016/j.chemgeo.2013.06.006
- Naeher, S., Suga, H., Ogawa, N. O., Schubert, C. J., Grice, K., and Ohkouchi, N. (2016). Compound-specific carbon and nitrogen isotopic compositions of chlorophyll a and its derivatives reveal the eutrophication history of Lake Zurich (Switzerland). *Chem. Geol.* 443, 210–219. doi:10.1016/j.chemgeo.2016.09.005
- North, R. P., North, R. L., Livingstone, D. M., Köster, O., and Kipfer, R. (2014). Long-term changes in hypoxia and soluble reactive phosphorus in the hypolimnion of a large temperate lake: consequences of a climate regime shift. *Glob. Change Biol.* 20 (3), 811–823. doi:10.1111/gcb.12371
- O’Beirne, M. D., Scott, W. P., and Werne, J. P. (2023). A critical assessment of lacustrine branched glycerol dialkyl glycerol tetraether (brGDGT) temperature calibration models. *Geochimica Cosmochimica Acta* 359, 100–118. doi:10.1016/j.gca.2023.08.019
- Peterse, F., van der Meer, J., Schouten, S., Weijers, J. W. H., Fierer, N., Jackson, R. B., et al. (2012). Revised calibration of the MBT-CBT paleotemperature proxy based on branched tetraether membrane lipids in surface soils. *Geochimica Cosmochimica Acta* 96, 215–229. doi:10.1016/j.gca.2012.08.011
- Raberg, J., Harning, D., Crump, S., de Wet, G., Blumm, A., Kopf, S., et al. (2021). Revised fractional abundances and warm-season temperatures substantially improve brGDGT calibrations in lake sediments. *Biogeosciences Discuss.*, 1–36. doi:10.5194/bg-2021-16
- Rampen, S. W., Datema, M., Rodrigo-Gámiz, M., Schouten, S., Reichart, G. J., and Sinninghe Damsté, J. S. (2014). Sources and proxy potential of long chain alkyl diols in lacustrine environments. *Geochimica Cosmochimica Acta* 144, 59–71. doi:10.1016/j.gca.2014.08.033

- Rampen, S. W., Schouten, S., Wakeham, S. G., and Sinninghe Damsté, J. S. (2007). Seasonal and spatial variation in the sources and fluxes of long chain diols and mid-chain hydroxy methyl alkanolates in the Arabian Sea. *Org. Geochem.* 38 (2), 165–179. doi:10.1016/j.orggeochem.2006.10.008
- Rampen, S. W., Willmott, V., Kim, J. H., Uliana, E., Mollenhauer, G., Schefuß, E., et al. (2012). Long chain 1,13- and 1,15-diols as a potential proxy for palaeotemperature reconstruction. *Geochimica Cosmochimica Acta* 84, 204–216. doi:10.1016/j.gca.2012.01.024
- R Core Team (2023). *R: a language and environment for statistical computing*. USA: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.
- Reiche, S., Rampen, S. W., Dorhout, D. J. C., Sinninghe Damsté, J. S., and Schouten, S. (2018). The impact of oxygen exposure on long-chain alkyl diols and the long chain diol index (LDI) – a long-term incubation study. *Org. Geochem.* 124, 238–246. doi:10.1016/j.orggeochem.2018.08.003
- Robinson, N., Eglinton, G., Cranwell, P. A., and Zeng, Y. B. (1989). Messel oil shale (western Germany): assessment of depositional palaeoenvironment from the content of biological marker compounds. *Chem. Geol.* 76 (1–2), 153–173. doi:10.1016/0009-2541(89)90134-4
- Rodrigo-Gámiz, M., Rampen, S. W., Schouten, S., and Sinninghe Damsté, J. S. (2016). The impact of oxic degradation on long chain alkyl diol distributions in Arabian Sea surface sediments. *Org. Geochem.* 100, 1–9. doi:10.1016/j.orggeochem.2016.07.003
- Romero-Viana, L., Kienel, U., and Sachse, D. (2012). Lipid biomarker signatures in a hypersaline lake on Isabel Island (Eastern Pacific) as a proxy for past rainfall anomaly (1942–2006AD). *Palaeogeogr. Palaeoclimatol. Palaeoecol.* 350–352, 49–61. doi:10.1016/j.palaeo.2012.06.011
- Russell, J. M., Hopmans, E. C., Loomis, S. E., Liang, J., and Sinninghe Damsté, J. S. (2018). Distributions of 5- and 6-methyl branched glycerol dialkyl glycerol tetraethers (brGDGTs) in East African lake sediment: effects of temperature, pH, and new lacustrine paleotemperature calibrations. *Org. Geochem.* 117, 56–69. doi:10.1016/j.orggeochem.2017.12.003
- Russell, N. J., and Fukunaga, N. (1990). A comparison of thermal adaptation of membrane lipids in psychrophilic and thermophilic bacteria. *FEMS Microbiol. Lett.* 75 (2), 171–182. doi:10.1016/0378-1097(90)90530-4
- Santos, L. M. A. (1996). The Eustigmatophyceae: actua knowledge and research perspective. *Nova Hedwig. Beih.*, 391–405. doi:10.1007/978-3-319-28149-0_39
- Shetty, P., Gitau, M. M., and Maróti, G. (2019). Salinity stress responses and adaptation mechanisms in eukaryotic green microalgae. *Cells* 8 (12), 1657. Article 12. doi:10.3390/cells8121657
- Shimokawara, M., Nishimura, M., Matsuda, T., Akiyama, N., and Kawai, T. (2010). Bound forms, compositional features, major sources and diagenesis of long chain, alkyl mid-chain diols in Lake Baikal sediments over the past 28,000 years. *Org. Geochem.* 41 (8), 753–766. doi:10.1016/j.orggeochem.2010.05.013
- Sinetova, M. A., Sidorov, R. A., Medvedeva, A. A., Starikov, A. Y., Markelova, A. G., Allakhverdiev, S. I., et al. (2021). Effect of salt stress on physiological parameters of microalgae *Vischeria punctata* strain IPPAS H-242, a superproducer of eicosapentaenoic acid. *J. Biotechnol.* 331, 63–73. doi:10.1016/j.jbiotec.2021.03.001
- Sinninghe Damsté, J. S., Weber, Y., Zopfi, J., Lehmann, M. F., and Niemann, H. (2022). Distributions and sources of isoprenoidal GDGTs in Lake Lugano and other central European (peri-)alpine lakes: lessons for their use as paleotemperature proxies. *Quat. Sci. Rev.* 277, 107352. doi:10.1016/j.quascirev.2021.107352
- Suutari, M., and Laakso, S. (1994). Microbial fatty acids and thermal adaptation. *Crit. Rev. Microbiol.* 20 (4), 285–328. doi:10.3109/10408419409113560
- Tierney, J. E., Russell, J. M., Eggermont, H., Hopmans, E. C., Verschuren, D., and Sinninghe Damsté, J. S. (2010). Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments. *Geochimica Cosmochimica Acta* 74 (17), 4902–4918. doi:10.1016/j.gca.2010.06.002
- Toney, J. L., García-Alix, A., Jiménez-Moreno, G., Anderson, R. S., Moossen, H., and Seki, O. (2020). New insights into Holocene hydrology and temperature from lipid biomarkers in western Mediterranean alpine wetlands. *Quat. Sci. Rev.* 240, 106395. doi:10.1016/j.quascirev.2020.106395
- van Bree, L. G. J., Peterse, F., van der Meer, M. T. J., Middelburg, J. J., Negash, A. M. D., Crop, W. D., et al. (2018). Seasonal variability in the abundance and stable carbon-isotopic composition of lipid biomarkers in suspended particulate matter from a stratified equatorial lake (Lake Chala, Kenya/Tanzania): implications for the sedimentary record. *Quat. Sci. Rev.* 192, 208–224. doi:10.1016/j.quascirev.2018.05.023
- Versteegh, G. J. M., Bosch, H.-J., and Leeuw, J. W. D. E. (1997). Potential palaeoenvironmental information of C24 to C36 mid-chain diols, keto-ols and mid-chain hydroxy fatty acids: a critical review. *Org. Geochem.* 27 (97), 1–13. doi:10.1016/S0146-6380(97)00063-6
- Villanueva, L., Besseling, M., Rodrigo-Gámiz, M., Rampen, S. W., Verschuren, D., and Sinninghe Damsté, J. S. (2014). Potential biological sources of long chain n-alkyl diols in a lacustrine system. *Org. Geochem.* 68, 27–30. doi:10.1016/j.orggeochem.2014.01.001
- Volkman, J. K., Barrett, S. M., and Blackburn, S. I. (1999). Eustigmatophyte microalgae are potential sources of C29 sterols, C22–C28 n-alcohols and C28–C32 n-alkyl diols in freshwater environments. *Org. Geochem.* 30 (5), 307–318. doi:10.1016/S0146-6380(99)00009-1
- Volkman, J. K., Barrett, S. M., Dunstan, G. A., and Jeffrey, S. W. (1992). C30–C32 alkyl diols and unsaturated alcohols in microalgae of the class Eustigmatophyceae. *Org. Geochem.* 18 (1), 131–138. doi:10.1016/0146-6380(92)90150-V
- Wang, Z., Zhou, Y., Xing, L., Bao, R., and Zhang, J. (2022). Seasonal variation of long-chain alkyl diols proxies in suspended particulate matter of the changjiang river estuary. *J. Geophys. Res. Biogeosciences* 127 (12), e2022JG007049. doi:10.1029/2022JG007049
- Weber, Y., De Jonge, C., Rijpstra, W. I. C., Hopmans, E. C., Stadnitskaia, A., Schubert, C. J., et al. (2015). Identification and carbon isotope composition of a novel branched GDGT isomer in lake sediments: evidence for lacustrine branched GDGT production. *Geochimica Cosmochimica Acta* 154, 118–129. doi:10.1016/j.gca.2015.01.032
- Weijers, J. W. H., Panoto, E., van Bleijswijk, J., Schouten, S., Rijpstra, W. I. C., Balk, M., et al. (2009). Constraints on the biological source(s) of the orphan branched tetraether membrane lipids. *Geomicrobiol. J.* 26 (6), 402–414. doi:10.1080/01490450902937293
- Wickham, H. (2016). “Data analysis,” in *Ggplot2: elegant graphics for data analysis*. Editor H. Wickham (Germany: Springer International Publishing), 189–201. doi:10.1007/978-3-319-24277-4_9
- Wu, J., Yang, H., Pancost, R. D., Naafs, B. D. A., Qian, S., Dang, X., et al. (2021). Variations in dissolved O₂ in a Chinese lake drive changes in microbial communities and impact sedimentary GDGT distributions. *Chem. Geol.* 579, 120348. doi:10.1016/j.chemgeo.2021.120348
- Xu, Y., Simoneit, B. R. T., and Jaffé, R. (2007). Occurrence of long-chain n-alkenols, diols, keto-ols and sec-alkanols in a sediment core from a hypereutrophic, freshwater lake. *Org. Geochem.* 38 (6), 870–883. doi:10.1016/j.orggeochem.2007.02.003
- Yankova, Y., Neuenschwander, S., Köster, O., and Posch, T. (2017). Abrupt stop of deep water turnover with lake warming: drastic consequences for algal primary producers. *Sci. Rep.* 7 (1), 13770. doi:10.1038/s41598-017-13159-9
- Zhang, Y., Su, Y., Liu, Z., Du, Y., Yu, J., and Jin, M. (2019). Fatty acid and alcohol compositions in lacustrine sediments as indicators of environment and ecosystem of lakes in Eastern China. *Ecol. Indic.* 97, 290–300. doi:10.1016/j.ecolind.2018.10.029
- Zhang, Y., Yaling, S., Zhengwen, L., Xiangchao, C., Jinlei, Y., Xiaodan, D., et al. (2015). Long-chain n-alkenes in recent sediment of Lake Lugu (SW China) and their ecological implications. *Limnologia* 52, 30–40. doi:10.1016/j.limno.2015.02.004
- Zhang, Z., Metzger, P., and Sachs, J. P. (2011). Co-occurrence of long chain diols, keto-ols, hydroxy acids and keto acids in recent sediments of Lake El Junco, Galápagos Islands. *Org. Geochem.* 42 (7), 823–837. doi:10.1016/j.orggeochem.2011.04.012