



UNIVERSITY OF
BIRMINGHAM

DEVELOPMENT AND CALIBRATION OF 3-HYDROXY
FATTY ACID-BASED PALAEOCLIMATIC PROXIES IN
PEDOGENIC AND LACUSTRINE ENVIRONMENTS

by

ALICE HARDMAN

A thesis submitted to the University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of Earth Sciences
School of Geography, Earth and Environmental Sciences
College of Life and Environmental Sciences
University of Birmingham

April 2024

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

Terrestrial environments are where the majority of climate change impacts will affect human societies, and the Intergovernmental Panel on Climate Change (IPCC) have called for high-resolution palaeoclimatic data to improve future temperature predictions in (often arid) regions with uncertainty in hydroclimatic regime, such as the Southwestern U.S. Environmental proxies derived from organic biomarkers are important tools for reconstruction of past climatic conditions, however their development and application between terrestrial environments is complex due to environmental heterogeneity. C_{10} - C_{18} 3-OH FAs are bacterial lipids investigated as a novel palaeoclimate biomarker group in recent years in soils, lake and marine sediments. However proxy- $\delta^{18}O$ relationships are highly variable between regional soil transects, and studies of lacustrine 3-OH FA distributions are limited. This thesis aims to investigate environmental drivers of 3-OH FA variability in pedogenic and lacustrine environments at a regional scale, to assess the performance of RAN_{17} as a temperature proxy in soils that span moisture and pH gradients, and compare 3-OH FA temperature proxy development and performance between and within soils and lakes, to improve regional calibration of terrestrial 3-OH FA-based environmental proxies for robust applicability.

In Chapter 3, relationships between 3-OH FA temperature and pH proxies (RAN_{15} , RAN_{17} , and $RIAN$) and contemporary environmental conditions are assessed in soils from the East Coast and Southwestern U.S., and the 3-OH FA distributions of these soils investigated for novel environmental proxy development. 3-OH FA proxy performance was compared with $\delta^{18}O$ -based MBT and CBT for the soils. pH and $\delta^{18}O$ are first-order drivers of 3-OH FA variability in U.S. soils, best described by the novel pH proxy $-\log(RIN_{USA})$. We find RAN_{17} is the primary 3-OH FA temperature response in U.S. soils with $pH < 7.2$, which may be associated with pH and moisture stress invoked on 3-OH FA bacterial soils in arid environments. In Chapter 4, relationships between 3-OH FA temperature proxies (RIN_{17} and RAN_{13}) are investigated in 26 lake

sediments from the SW U.S. and compared with existing data from Chinese lake sediments and the SW U.S. soils. 3-OH FA proxy performance was compared with brGDGT-based MBT'_{SME}.

$[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ is the 3-OH FA response to MAAT in U.S. lake sediments and is most responsive to T_{MAM} , differing from the MAAT responses of RIN₁₇ in Chinese lakes and RAN₅ in SW soils. Different 3-OH FA temperature proxies are viable within and between lacustrine and pedogenic environments, suggesting that lacustrine temperature calibrations from 3-OH FAs have a regional calibre at present. In Chapter 5, the potential to streamline 3-OH FA extraction from environmental samples is investigated by testing a low temperature long duration (70 °C, 14 hrs; 'Low n' Slow') acid hydrolysis method on 3-OH FA recovery and proxy data from re-extracted SW U.S. lake sediments and soils. The Low n' Slow method is successful at streamlining the laboratory procedure for 3-OH FA extraction and extracting 3-OH FAs, but raises implications for palaeoclimatic interpretations from 3-OH FAs extracted by different hydrolysis temperatures. In Chapter 6, the Low n' Slow method is applied to a 105 yr lacustrine sequence from Esthwaite Water (UK). 3-OH FA proxies are able to track environmental conditions at a sub-decadal resolution, however most proxies have a unique and contradicting response to temperature, which is potentially related to a history of nutrient enrichment at Esthwaite Water.

This thesis reveals important insights into the environmental drivers of 3-OH FA production in lacustrine and pedogenic environments, and produces regional lake and soil 3-OH FA-based proxy calibrations applicable to the U.S. The proxies likely represent both homeoviscous adaptation of bacterial membranes to ambient environmental conditions and bacterial community changes. There is much more to learn about 3-OH FA variability in response to environmental conditions, particularly in lacustrine environments, and we recommend a global-scale investigation of 3-OH FAs in lakes across different environmental gradients in future studies.

"No fear, no loss, no tears,

The time is almost here.

Our dreams will all come true, I promise you,

'Cause I can see for miles and miles.

In time we'll be dancing in the streets all night"

-In Time, Robbie Robb

ACKNOWLEDGEMENTS

I am grateful to many people who have helped and supported me during this PhD. I thank my academic supervisors James Bendle and Kweku Yamoah for your invaluable support and guidance over the past 5-6 years. James, it was a pleasure to meet Caroline, George and Ryan during the U.S. fieldtrip, and with Huub we had some wild times out there – driving boats, eating amazing BBQ food, collecting stinking lake mud – that I won't forget. Kweku, thank you for training me in the lab and introducing me to new experimental techniques, and helping me become more confident. Thank you also to Sarah Greene for support and advice during this PhD.

Thank you to Helen Whealton for running and integrating samples for GDGTs at the University of Bristol NERC NEIF facility. Nick Davidson, Kweku Yamoah, Matt Allison, and David Tubbs are thanked for various technical support in the lab and support for fieldwork equipment. Thank you to Grace Duffy who helped with the 3-OH FA extractions for the U.S. soils and lake sediments and for assisting on fieldwork. Thank you to Huub Zwart for organising the U.S. fieldtrip and collecting the samples, and to Vinothan Sivapalan who assisted with the East Coast fieldwork.

I would like to thank Amy Thorpe for sharing data from your PhD, for sharing your knowledge of Esthwaite Water and arranging the Esthwaite fieldwork to collect the sediment core. It has been a pleasure to work with you. Thank you to my lab mates in the organic geochemistry group over the years (Kate, Jonathan, Bridget, Matt, Patt, Changhai, Gemma), I have had a great time working alongside you and attending conferences together. Thank you to my office mates in the Earth Sciences department and from the CENTA cohort for the fun memories during my time at UoB.

Thank you to the EMBO for sponsoring me to undertake a 3-month placement at Stockholm University to do palaeoenvironmental work on a Holocene peat sequence, and I especially thank Jenny Sjöström and Elin Norström for your guidance and supervision during my time at

Stockholm and for supporting my EMBO application. Working at Stockholm University was an incredible opportunity that I will never forget.

Thank you to the NERC CENTA DTP for funding this PhD, NERC for funding GDGT analysis at Bristol, and the Leverhulme Trust and BRIDGE fund for funding the wider U.S. project and fieldwork.

Finally, I thank my friends and family for supporting me during this PhD, and who have been there during the ups and downs. I would especially like to thank Mum and Chloe and our cats Laces, Noah, O'Malley, and Milo for your continuous belief, encouragement, support, and 3 am mouse-out-the-house chasing entertainment.

THESIS CONTRIBUTIONS

The work in this thesis has been part of a multi-disciplinary collaboration with Amy Thorpe, an Earth Sciences PhD student at the University of Birmingham based at [UKCEH](#). Amy Thorpe specialises in microbial ecology and all U.S. soils, lake sediments and samples from the Esthwaite Water lake sediment core were also extracted by Amy Thorpe for 16S RNA sequencing to investigate the composition of bacterial communities in the samples. The 16S RNA sequencing data for all samples is reported in detail in Amy Thorpe's submitted PhD thesis which is cited as 'Thorpe (2023)' where the bacterial community data is referred to in this thesis. pH for all U.S. samples was also measured by Amy Thorpe at [UKCEH](#).

Additionally, the 2018 U.S. field campaign during which all U.S. samples were collected for this thesis was organised and conducted by Huub Zwart.

The structural template for this thesis is modified from an Overleaf latex template credited to Copyright 2022 Alexander Lyttle, which was available for distribution and modification under the LaTeX Project Public License (LPPL) version 1.3 or later.

Chapter specific contributions are as follows:

In **Chapter 3**, U.S. East Coast soils were collected by Huub Zwart and Vinothan Sivapalan. SW U.S. soils were collected by Huub Zwart, Alice Hardman and James Bendle. pH was measured by Amy Thorpe at [UKCEH](#). 3-OH FAs were extracted by Alice Hardman with occasional assistance from Grace Duffy. LOI was measured by Alice Hardman. GDGT distributions were analysed at the University of Bristol by Helen Whealton funded by a NERC NEIF grant awarded to Alice Hardman.

In **Chapter 4**, SW U.S. lake sediments were collected by Huub Zwart, Alice Hardman and James

Bendle. pH was measured by Amy Thorpe at [UKCEH](#). 3-OH FAs were extracted by Alice Hardman with occasional assistance from Grace Duffy. LOI was measured by Alice Hardman. GDGT distributions were analysed at the University of Bristol by Helen Whealton funded by a NERC NEIF grant awarded to Alice Hardman.

In **Chapter 5**, fieldwork to collect Esthwaite Water sediment core was conducted by Amy Thorpe, Eleanor Mackay and Alice Hardman. 3-OH FAs were extracted by Alice Hardman.

CONTENTS

ABSTRACT	i
ACKNOWLEDGEMENTS	iv
THESIS CONTRIBUTIONS	vi
CONTENTS	viii
LIST OF FIGURES	xvii
LIST OF TABLES	xxix
ACRONYMS	xxxii
EQUATIONS	xxxv
1 INTRODUCTION	1
1.1 Need for development of novel terrestrial palaeoclimatic biomarker proxies	1
1.2 3-Hydroxy fatty acids	5
1.2.1 Potential as a novel terrestrial palaeoclimate biomarker	5
1.2.2 Biochemistry	10
1.3 Homeoviscous adaptation: lipodomic modification of bacterial cell membranes for maintenance of optimum fluidity	12
1.4 Thesis aims and outline	14
2 METHODS	18
2.1 Sample collection and processing	18
2.2 Measurement and acquisition of environmental variables	19

2.2.1	Climate data	19
2.2.2	pH measurement	21
2.2.3	LOI measurement	21
2.3	Organic geochemistry: 3-OH FAs	21
2.3.1	3-OH FA extraction	21
2.3.2	3-OH FA analysis	22
2.3.3	Calculation of 3-OH FA-based climate proxies	23
2.4	Organic geochemistry: brGDGTs	24
2.4.1	brGDGT extraction	24
2.4.2	brGDGT analysis	24
2.4.3	Calculation of terrestrial GDGT-based climate proxies	25
2.5	Statistical approaches	27

3 DISTRIBUTIONS OF 3-HYDROXY FATTY ACIDS REPORTED FROM TWO ENVIRONMENTALLY CONTRASTING SOIL TRANSECTS IN THE USA 29

3.1	Introduction	30
3.2	Study area	34
3.3	Results	37
3.3.1	Environmental diversity across the East Coast and Southwestern U.S.	37
3.3.2	Organic geochemistry	40
3.3.2.1	3-OH FA distributions in East Coast and SW U.S. soils	40
3.3.2.2	GDGT distributions in SW U.S. soils	45
3.3.3	3-OH FA and brGDGT-derived palaeoclimate indices	45
3.3.3.1	3-OH FAs	45
3.3.3.2	brGDGTs	48
3.3.4	Application of linear and multivariate analyses to explore relationships between environmental variables and bacterial lipids	50

3.3.4.1	Response of 3-OH FA compounds to environmental variability and identification of primary environmental drivers	50
3.3.4.2	Response of established 3-OH FA-based proxies to environmental variables in U.S. soils	55
3.4	Discussion	65
3.4.1	U.S. soil 3-OH FA distributions in the context of global 3-OH FA variability	65
3.4.2	Response of 3-OH FAs to temperature in soils	68
3.4.3	Assessing RAN ₁₇ regionality in U.S. soils: redefining the regional environmental constraints of a RAN ₁₇ -based temperature calibration	74
3.4.4	Performance of lipid biomarker-based temperature calibrations in U.S. soils	82
3.4.4.1	RAN ₁₇ calibrations on compiled U.S. soils	82
3.4.4.2	RAN ₁₇ calibrations on EC and SW soils	88
3.4.4.3	Performance of RAN ₁₇ and MBT calibrations	90
3.4.5	Response of 3-OH FAs to pH in soils	93
3.4.6	Functionality of 3-OH FA-based environmental proxies in arid environments	100
3.5	Conclusions	105
4	DEVELOPMENT OF 3-HYDROXY FATTY ACID-BASED PALAEOTHERMOMETRY IN LACUSTRINE ENVIRONMENTS ACROSS THE SOUTHWESTERN UNITED STATES	107
4.1	Introduction	108
4.2	Study area	111
4.3	Results	113
4.3.1	Environmental diversity across the Southwestern U.S.	113
4.3.2	Lipid biomarker distributions in SW U.S. lake sediments	117
4.3.2.1	3-OH FA distributions	117

4.3.2.2	Comparison of 3-OH FA distributions between SW U.S. lake sediments and soils	117
4.3.2.3	GDGT distributions	123
4.3.2.4	Comparison of GDGT distributions between SW U.S. lake sedi- ments and soils	124
4.3.3	Lipid biomarker-based proxy distributions in SW U.S. lake sediments and comparisons with soils	126
4.3.3.1	3-OH FAs	126
4.3.3.2	brGDGTs	128
4.3.4	Application of linear and multivariate analyses to identify environmental drivers of 3-OH FA variability in SW U.S. lake sediments and soils	131
4.3.4.1	Environmental drivers of 3-OH FA variability in SW U.S. lake sediments	131
4.3.4.2	Environmental drivers of 3-OH FA variability in SW U.S. lake sediments omitting unique cold lake catchments	133
4.3.4.3	Quantifying 3-OH FA temperature variability revealed by multi- variate and linear correlation analysis	138
4.3.4.4	Response of established 3-OH FA-based proxies to environmental variables in SW U.S. lake sediments and soils	141
4.3.5	Performance of GDGT-based environmental proxies in SW lake sediments and soils	149
4.4	Discussion	151
4.4.1	3-OH FA distributions across lacustrine catchments in the context of global variability	151
4.4.1.1	3-OH FA distributions in lake sediments	151
4.4.1.2	3-OH FA distributions between lake sediments and soils	156

4.4.2	Consistency of temperature response amongst lipid biomarker distributions derived from lake sediments	163
4.4.2.1	3-OH FA-based temperature calibrations	163
4.4.2.2	brGDGT-based temperature calibrations	168
4.4.2.3	Seasonality of lipid biomarker-based temperature calibrations amongst SW U.S. lakes	171
4.4.3	Different bacterial strategies for temperature response between and within lake sediments and soils	174
4.4.3.1	Lacustrine environments	174
4.4.3.2	Pedogenic environments	175
4.4.3.3	Lacustrine vs pedogenic environments	176
4.4.4	Limitations of the regional stature of 3-OH FA-based temperature calibra- tions developed from SW U.S. lacustrine environments	178
4.5	Conclusions	180

5 INITIAL DEVELOPMENT OF A LOW N' SLOW ACID HYDROLYSIS METHOD FOR 3-OH FA EXTRACTION 182

5.1	Introduction	183
5.2	Experimental	186
5.2.1	SW samples dataset 1: original optimised acid hydrolysis method for 3-OH FA extraction	187
5.2.2	SW samples dataset 2: 3-OH FA extraction method employing Low n' Slow acid hydrolysis	187
5.3	Results	188
5.3.1	Low n' Slow 3-OH FA extraction of SW U.S. lake sediments and soils . . .	188
5.3.1.1	3-OH FA distributions	188
5.3.2	3-OH FA concentrations	192

5.3.2.1	3-OH FA-based environmental proxy distributions	194
5.4	Discussion	198
5.4.1	Effect of lower hydrolysis temperature on 3-OH FA recovery and distributions	198
5.4.2	Implications of differing acid hydrolysis temperatures on palaeoclimatic interpretation of 3-OH FAs	202
5.4.3	Streamlining 3-OH FA extraction: potential to further optimise 3-OH FA extraction from environmental media	205
5.5	Conclusions	206
6	APPLICATION OF LOW N' SLOW ACID HYDROLYSIS FOR 3-OH FA EXTRACTION IN A LAKE SEDIMENT CORE FROM ESTHWAITE WATER	209
6.1	Introduction	210
6.2	Summary of 3-OH FA data	211
6.3	3-OH FA distributions in context of environmental variability at Esthwaite Water .	214
6.4	Brief summary	218
7	CONCLUSIONS	220
7.1	Main findings of this thesis	220
7.2	Recommendations for future work	223
	REFERENCES	225
	APPENDIX A: SUPPLEMENTARY INFORMATION FOR CHAPTER 3	251
	Table A1	252
	Table A2	253
	Table A3	254
	Table A4	255
	Figure A5	256

Figure A6	257
Figure A7	258
Figure A8	259
Figure A9	260
Figure A10	261
Figure A11	262
APPENDIX B: SUPPLIMENTARY INFORMATION FOR CHAPTER 4	263
Figure B1	264
Figure B2	265
Figure B3	266
Figure B4	267
Figure B5	268
Figure B6	269
Figure B7	270
Figure B8	271
Figure B9	272
Figure B10	273
Figure B11	274
Figure B12	275
Figure B13	276
Figure B14	277
Figure B15	278
APPENDIX C: SUPPLEMENTARY INFORMATION FOR CHAPTER 5	279
Figure C1	280
Figure C2	281

Figure C3	282
Figure C4	283
Figure C5	284
Figure C6	285
Figure C7	286
APPENDIX D: SUPPLEMENTARY INFORMATION FOR CHAPTER 6	287
Figure D1	288
APPENDIX E: DATA TABLES	289
Table E1	290
Table E2	291
Table E3	292
Table E4	293
Table E5	294
Table E6	295
Table E7	296
Table E8	297
Table E9	298
Table E10	299
Table E11	300
Table E12	301
Table E13	302
Table E14	303
Table E15	304
Table E16	305
Table E17	306

Table E18	307
Table E19	308

LIST OF FIGURES

1.1	Chemical structures of brGDGTs (I-III) and Crenarchaeol (IV) adapted from De Jonge et al. (2014b). brGDGTs denoted by 'I' are 'tetra-methylated'. brGDGTs denoted by 'II' are 'penta-methylated', and brGDGTs denoted by 'III' are 'hexa-methylated'. brGDGTs denoted by 'a' have no cyclopentane moieties, those denoted by 'b' have one cyclopentane moiety, and those denoted by 'c' have two cyclopentane moieties. Prime ' denotes 6-methyl brGDGT isomers.	3
1.2	Chemical structures of 3-OH FAs.	6
1.3	Global distribution of soils, marine and lake sediments investigated for 3-OH FAs. Soils have been investigated by Véquaud et al. (2021b) (1), Huguet et al. (2019) (2), Wang et al. (2016) (3), Véquaud et al. (2021a) (4), and remaining soils unlabelled feature in the global soil transect of Wang et al. (2021). All lake sediments are associated with the Chinese lake transect of Yang et al. (2021). Marine sediments from the marginal seas of China and North Pacific Ocean have been investigated by Yang et al. (2020) and Dong et al. (2023).	7
1.4	Schematic diagram of the occurrence of 3-OH FAs in Gram-negative bacterial cell membranes. In colour to the left is the structure of the cell membrane. LPS is a component of the outer membrane, for which the structure is detailed to the right. 3-OH FAs are abundant by ester and amide bonds in the Lipid A component of LPS (Lipid A structure from <i>E. coli</i> visualised at bottom) Figure from Zhang et al. (2023).	11
3.1	Maps indicating location of EC and SW sampling sites. Samples from EC comprise soils, samples from SW are co-sampled lake sediments and catchment soils. A conveys MAAT variability across the study sites, and B conveys MAP.	36

3.2	Comparison of environmental variables along EC and SW U.S. soil transects, comprising: Pearson's correlation matrices for A EC soils, B SW soils, and crossplots between C MAAT and MAP, D MAAT and pH, E MAAT and T_{Soil} , F MAP and pH. Circles in upper half of matrix relate to relative magnitude of Pearson's R . Green circles indicate positive correlations, orange circles indicate negative correlations, and significant correlations are indicated by asterisks (p : $\leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$). Values in lower half of matrix report Pearson's R itself.	39
3.3	3-OH FA distributions in EC and SW U.S. soils reported as % fractional abundance of A individual 3-OH FA compounds, and B and C various 3-OH FA structural groups	42
3.4	Non-metric multidimensional scaling (NMDS) plot of composite 3-OH FA distributions for EC and SW U.S. soils with NMDS1 axis limit set between -0.2 and 0.2 (centre). 3-OH FA compounds with significant variability amongst the soils are visualised, and numbered labels indicate the sample name of unique soils. Upper right: zoomed out plot showing full scale NMDS1 and unique 3-OH FA distribution of site USA 062. Lower right: zoomed in plot highlighting the significantly variable 3-OH FA compounds that are less influential than those visualised in central plot. Associated statistics are reported in Appendix A4.	44
3.5	Boxplots visualising 3-OH FA-based proxy distributions within and between EC and SW soil transects	47
3.6	Boxplots visualising GDGT-based proxy distributions within and between EC and SW soil transects. Data for the EC soils is from Dirghangi et al. (2013), data from the SW soils is novel from this study.	49
3.7	Pearson's correlation heatmap between 3-OH FA compounds and environmental variables. 3-OH FA data is compiled for all soils, and not separated by US soil transect. Significance is indicated by asterisks (p : $\leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$))	52

3.8	PCA of U.S. soil 3-OH FA distributions. Variability of 3-OH FA compounds and quantitative environmental drivers is visualised, and significant environmental drivers ($p<0.05$) are highlighted following permutation analysis	53
3.9	RDA employing 3-OH FA proxies as response variables and environmental drivers as explanatory variables in U.S. soils	56
3.10	Crossplots between established 3-OH FA-based pH proxies and selected environmental variables (pH, MAP), comprising: BR, RIAN (negative logarithm of BR), RIN, and -log(RIN) with pH (A,C,E,G) and MAP (B,D,F,H).	58
3.11	Crossplots between RIN_{USA} and -log(RIN_{USA}) and selected environmental variables (pH and map), comprising: RIN_{USA} with pH (A), MAP (B), and MAAT (C), and -log(RIN_{USA}) with same (D,E,F).	61
3.12	Crossplots of RAN_{15} and RAN_{17} with MAAT (A,C) and pH (B,D) amongst U.S. soils	64
3.13	NMDS plot of 3-OH FA distributions in globally distributed soils after Véquaud et al. (2021b) ¹ , Yang et al. (2021) ² , present study ³ , Wang et al. (2021) ⁴ , Huguet et al. (2019) ⁵ , Wang et al. (2016) ⁶ . 3-OH FAs that significantly ($p<0.05$) vary amongst the samples are visualised following permutation analysis.	68
3.14	Variability of A RAN_{15} and B RAN_{17} with MAAT in globally distributed soils. Soil data after Véquaud et al. (2021b) ¹ , Yang et al. (2021) ² , present study ³ , Wang et al. (2021) ⁴ , Huguet et al. (2019) ⁵ , Wang et al. (2016) ⁶	70
3.15	Scatterplots between residuals from the RAN_{17} -MAAT regression (Fig. 3.12C) of compiled U.S. soils and selected environmental variables (A ΔT and B seasonality)	73

3.16	Scatterplots conveying the influence of environmental variability on the RAN ₁₇ -MAAT regression in U.S. soils, specifically the influence of: climate regime (A), MAP (C), ΔT (E), pH (G), SMO (I), and seasonality (J). Extreme environmental variability has been omitted from the RAN ₁₇ -MAAT regression for the following variables: climate regime (B), MAP (D), ΔT (F), pH (H).	76
3.17	Investigation of 3-OH FA distribution similarity between soils. A Dendrogram visualising results of cluster analysis of 3-OH FA distributions in U.S. soils. B PCA plot with soils re-visualised as clustered groups (from A) instead of EC and SW transect grouping. C RAN ₁₇ -MAAT regression with cluster grouping from A. D RAN ₁₇ -MAAT regression with cluster group 5 omitted.	80
3.18	Refined RAN ₁₇ -MAAT regression from U.S. soils that only features soils with pH<7.2. Residuals are also visualised.	82
3.19	Testing the influence of a pH<7.2 threshold (found to improve RAN functionality in U.S. soils) on the RAN ₁₇ -MAAT regression of globally compiled soils. Summary statistics in blue (upper) summarise calibration quality of global soils with pH<7.2, and those in pink (lower) summarise the calibration quality of the entire dataset.	84
3.20	Crossplots between contemporary MAAT and ΔT for various RAN ₁₇ (and one MBT) calibrations for the compiled U.S. soils. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after Peterse et al. (2012), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.	86
3.21	Crossplots between contemporary MAAT and ΔT for various RAN ₁₇ (and one MBT) calibrations for the EC and SW soil transects. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after Peterse et al. (2012), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.	89

3.22	Scatterplots between visualising comparability of A contemporary MAAT and MBT estimated MAAT and B ran ₁₇ estimated MAAT and MBT estimated MAAT, for U.S. soils.	91
3.23	-log(RIN _{USA})-pH regression from U.S. soils, with residuals visualised.	96
3.24	Variability of A RIAN and B -log(RIN _{USA}) with pH in globally distributed soils. Soil data after Véquaud et al. (2021b) ¹ , Yang et al. (2021) ² , present study ³ , Wang et al. (2021) ⁴ , Huguet et al. (2019) ⁵ , Wang et al. (2016) ⁶	97
3.25	Crossplots of: A contemporary pH and the brGDGT-based CBT index, B -log(RIN _{USA}) estimated pH and CBT estimated pH from U.S. soils	99
3.26	Crossplot visualising 3-OH FA and brGDGT-based pH proxy performance by comparing contemporary pH with ΔpH (proxy estimated pH - contemporary pH). . . .	100
3.27	Crossplot matrix visualising the variability in fractional abundance of the Bacterial phyla identified as most responsive to pH by Thorpe (2023), specifically with pH (A) and MAAT (B), and same for soils with pH<7.2 (C,D). Comprises unpublished bacterial data from Thorpe (2023).	103
3.28	Crossplots visualising α -C ₁₇ variability in response to selected Bacterial phyla amongst A U.S. soils and B U.S soils with pH<7.2. Comprises unpublished bacterial data from Thorpe (2023).	104
4.1	SW U.S. lakes. All lakes are scaled to the 5 km scale bar. This scale bar excludes map of U.S. States	113
4.2	A Spearman's correlation matrix between SW U.S. environmental variables. Significance is indicated by asterisks (p : $\leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$), and crossplots between B MAAT and soil pH, C lake sediment pH and soil pH	116
4.3	3-OH FA distributions in SW U.S. lake sediments and soils reported as % fractional abundance for individual 3-OH FA compounds	119

4.4	NMDS plot of a Bray-Curtis dissimilarity matrix of SW U.S. lake sediment and soil 3-OH FA fractional abundance. 3-OH FA compounds that have significant variability amongst the samples ($p < 0.05$) are visualised following permutation analysis	120
4.5	Pooled 3-OH FA 3-OH FA fractional abundance in SW U.S. lake sediments and soils. The fractional distributions of individual 3-OH FA compounds are visualised as a jitter and percentage 3-OH FA representation summarised as pie charts for three fractional abundance thresholds: (1) $< 10\%$, (2) $10\text{--}20\%$ abundance, (3) $> 20\%$ abundance	121
4.6	brGDGT distributions in SW U.S. lake sediments and soils reported as $\%$ fractional abundance for individual brGDGT compounds	124
4.7	Boxplots visualising 3-OH FA-based proxy distributions between SW lake sediments and soils. Descriptive statistics are visualised.	127
4.8	Boxplots visualising brGDGT-based proxy distributions between SW lake sediments and soils. Descriptive statistics are visualised.	130
4.9	PCA of relationships between lake sediment 3-OH FA compounds and environmental variables. Environmental variables were added <i>a posteriori</i> following permutation analysis. Significant environmental variables ($p < 0.05$) are indicated.	132
4.10	PCAs of 3-OH FA fractional distributions from (Left) SW lake sediments and (Right) soils, with cold sites omitted. Environmental variables were added <i>a posteriori</i> following permutation analysis. Significant environmental variables ($p < 0.05$) are indicated.	135
4.11	Pearson's correlation matrix between 3-OH FA fractional abundance of SW lake sediments and environmental variables, omitting cold lakes. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$).	136

4.12	Pearson's correlation matrix between 3-OH FA fractional abundance of SW soils and environmental variables, omitting cold lakes. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$).	137
4.13	Scatterplots between MAAT and A $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ index in SW U.S. lake sediments, and B $[(i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18}]$ index in SW U.S. soils, all omitting two cold lakes. Shaded graph area represents 95 % functional bounds. Dashed lines represent 95 % observational bounds.	139
4.14	RDA using fractional abundance of established 3-OH FA proxies as response variables and environmental conditions as explanatory variables for A SW lake sediments and B SW soils	143
4.15	Scatterplots between RAN_{13} and A MAAT and B T_{MAM} in SW lake sediments	145
4.16	Scatterplot between RAN_5 in SW soils with $pH > 7.2$	148
4.17	Scatterplot matrix between MBT'_{5ME} and A MAAT, B respective lake sediment and soil pH, C T_{MAM} , and D T_{DJF} in SW lake sediments and soils.	150
4.18	NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of SW lake sediments and the Chinese lake sediments of Yang et al. (2021). 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised.	153
4.19	NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of global sediments investigated for 3-OH FAs. 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised. Soils are derived from Wang et al. (2021), Véquaud et al. (2021b), Véquaud et al. (2021a), Wang et al. (2016) and Huguet et al. (2019), and SW soils. Lakes sediments comprise SW lakes in this study and those of Yang et al. (2021). Marine sediments comprise those from the marginal seas of China reported by Yang et al. (2020) and Dong et al. (2023).	157

4.20	Distributions of 3-OH FA fractional abundance amongst global sediments for (Upper) soils, (Middle) lake sediments, and (Lower) marine sediments. Plots to the left visualise individual 3-OH FA compounds, and plots to the right visualise 3-OH FA structural groups.	158
4.21	RAN distributions amongst globally distributed samples as boxplots. The RAN confidence ranges are also shown, calculated as mean and standard deviation. . . .	159
4.22	Scatterplots showing performance of three 3-OH FA-based calibrations in reconstructing MAAT in U.S. and Chinese lake sediments. Plots comprise reconstructed temperature vs MAAT of: A RAN ₁₃ calibrated directly from SW lakes, B RIN ₁₇ calibrated from Chinese lakes of Yang et al. (2021) with modifications outlined in main text, and C [(<i>i</i> -C ₁₃ + <i>i</i> -C ₁₄ + <i>α</i> -C ₁₅) / <i>n</i> -C ₁₈] index calibrated directly from SW lakes. Below are plots of ΔT (reconstructed MAAT - MAAT vs MAAT. . . .	165
4.23	Calibration and results of MLR calibration developed from SW lake sediments. A MLR calibration visualised as MLR reconstructed MAAT for calibration samples vs MAAT, with residuals below. B MLR reconstructed MAAT for SW U.S. (present study) and Chinese (Yang et al. 2021) lake sediments vs MAAT. Below is a plot of ΔT (reconstructed MAAT - MAAT) vs MAAT.	167
4.24	Scatterplots of A MBT' _{5ME} MAAT reconstructions and MAAT for SW soils. Below is a plot of ΔT (reconstructed MAAT - MAAT vs MAAT. B comparison of estimated MAAT between the three previously considered 3-OH FA-based calibrations and the SW U.S. MBT' _{5ME} calibration, amongst SW lake sediments. Below is a plot of ΔT (3-OH FA reconstructed MAAT - brGDGT reconstructed MAAT) vs MAAT . . .	171
5.1	NMDS plot of a Bray-Curtis dissimilarity matrix of SW U.S. lake sediment and soil 3-OH FA fractional abundance for samples extracted by Low n' Slow hydrolysis and standard acid hydrolysis. 3-OH FA compounds that have significant variability amongst the samples ($p < 0.05$) are visualised following permutation analysis. . . .	190

5.2	Ternary plot of SW U.S. lake sediments and soils 3-OH FA distributions extracted by two different acid hydrolysis methods.	192
5.3	$\Delta_{\text{Conc.}}$ (L. sediment 3-OH FA conc. - soil 3-OH FA conc.) of total 3-OH FAs for SW (Upper) lake sediments and (Lower) soils between hydrolysis methods for counterpart samples.	194
5.4	Boxplots visualising 3-OH FA-based proxy distributions in SW U.S lake sediments and soils extracted by Low n' Slow acid hydrolysis.	195
6.1	3-OH FA distributions Esthwaite Water lake sediments reported as % fractional abundance of individual 3-OH FA compounds	212
6.2	NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of Esthwaite Water lake sediments. 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised. This is also applicable for environmental variables.	213
6.3	Variability of 3-OH FA proxies throughout the Esthwaite Water lacustrine sequence. Chronology of important historic events related to water quality are indicated. In RED are events promoting nutrient enrichment and in GREEN are efforts to mitigate nutrient enrichment.	216
6.4	Pearson's correlation heatmap between 3-OH FA proxies and environmental variables. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)	218
A5	Pearson's correlation matrix between the fractional abundance of 3-OH FA compounds and environmental variables grouped by U.S. soil transect	256
A6	Pearson's correlation matrix between established 3-OH FA-based environmental proxies from U.S. soils and environmental variables.	257
A7	Pearson's correlation matrix between 3-OH FA-based proxies and environmental variables grouped by U.S. soil transect	258

A8	Crossplot matrix of established 3-OH FA-based temperature and pH proxies and MAAT in U.S. soils. Black regression lines and r^2 summarise all U.S. soils compiled. Regression lines displayed for significant correlations.	259
A9	Boxplots visualising distributions in environmental variables between cluster groups 1-4 with cluster group 5 (after Fig.3.17) for U.S. soils	260
A10	Crossplots between contemporary MAAT and ΔT for various RAN ₁₇ (and one MBT) calibrations for U.S. soils grouped by a pH 7.2 threshold. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after Peterse et al. (2012), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.	261
A11	Crossplot matrix of established 3-OH FA-based temperature and pH proxies and soil pH in U.S. soils. Black regression lines and r^2 summarise all U.S. soils compiled. Regression lines displayed for significant correlations.	262
B1	Barplots visualising difference in mean fractional abundance of 3-OH FA structural groups and descriptive statistics between SW U.S lake sediments and soils.	264
B2	Barplots visualising difference in mean fractional abundance of brGDGT structural groups and descriptive statistics between SW U.S lake sediments and soils.	265
B3	Boxplots showing distribution of air temperature variables along the SW U.S. lake transect.	266
B4	Contribution of 3-OH FA compounds to PC1 and PC2 in A & B all U.S. lake sediments (Fig.4.9), C & D U.S. lake sediments with cold lakes omitted (Fig.4.10), E & F U.S. soils with cold lakes omitted (Fig.4.10)	267
B5	RDA for SW U.S. A lake sediments and B soils omitting cold sites. 3-OH FA fractional abundance were used as response variables, and environmental parameters as explanatory variables.	268

B6	Crossplots between PC1 to 4 from PCA and lake sediment pH from PCA of SW lake sediments with cold sites omitted.	269
B7	Crossplots between PC1 to 4 from PCA and lake sediment pH omitting soils with $\text{pH} < 7$ for SW lake sediments with cold sites omitted.	270
B8	Crossplots between selected 3-OH FA compounds associated with PC1 and PC2 in SW lake sediments and lake sediment pH for SW lake sediments with cold sites omitted.	271
B9	Crossplots between PC1 to 4 from PCA of soils with cold sites omitted vs MAAT.	272
B10	Crossplot visualising relationship between soil pH and $[(i\text{-C}_{13} + i\text{-C}_{14} + a\text{-C}_{15}) / n\text{-C}_{18}]$ in SW U.S. soils.	273
B11	Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW lakes with cold sites omitted. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)	274
B12	Scatterplots between established 3-OH FA-based proxies and MAAT for SW U.S. lake sediments and soils with two cold lakes omitted.	275
B13	Scatterplots between established 3-OH FA-based proxies and MAAT for SW U.S. lake sediments and soils including two cold lakes.	276
B14	Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW soils with cold sites omitted. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)	277
B15	Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW lakes with $\text{pH} > 7.2$. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)	278
C1	3-OH FA distributions in SW U.S. lake sediments and soils extracted by Low n' Slow acid hydrolysis reported as % fractional abundance	280

C3	Comparison of 3-OH FA concentrations between SW lake sediments and soils extracted by Low n' Slow and original acid hydrolysis methods.	282
C4	3-OH FA distributions in SW U.S. soils (Upper) and lake sediments (Lower) extracted by Low n' Slow and original acid hydrolysis reported as % fractional abundance.	283
C5	Crossplot matrix of established 3-OH FA-based temperature and pH proxies and MAAT in U.S. lakes extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.	284
C6	Crossplot matrix of established 3-OH FA-based temperature and pH proxies and T_{MAM} in U.S. lakes extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.	285
C7	Crossplot matrix of established 3-OH FA-based temperature and pH proxies and MAAT in U.S. soils extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.	286
D1	Environmental variables from UKCEH monitoring records of Esthwaite Water. Alkalinity is measured as $\mu\text{g L}^{-1}$ CaCO_3 and pH. SRP is soluble reactive phosphate ($\text{PO}_4\text{-P}$), and Winter SRP is SRP averaged for winter months which was calculated by Thorpe et al. (2022). Data is integrated between water samples collected from lake surface waters between 0 and 5 m depth in water column. Data after Maberly et al. (2017) and (Feuchtmayr et al. 2021).	288

LIST OF TABLES

1.1	Statistics conveying accuracy of published temperature and pH calibrations developed from 3-OH FA-based proxies	9
A1	Statistics for comparison of mean and variance of environmental variables between EC and SW soil transects	252
A2	Statistics for comparison of 3-OH FA structural group fractional abundance distributions between EC and SW soil transects	253
A3	Statistics for comparison of 3-OH FA fractional abundance distributions between EC and SW soil transects	254
A4	Permutation test results for NMDS of 3-OH FA distributions amongst U.S. soils . .	255
C2	Statistics for comparison of mean and variance of 3-OH FA fractional abundances between Low n' Slow extracted lake sediments and soils	281
E1	U.S. East Coast soils 3-OH FA data (fractional abundance and proxies)	290
E2	Southwestern U.S. lake sediments and soils (extracted by original acid hydrolysis method) 3-OH FA data (fractional abundance and proxies)	291
E3	Southwestern U.S. lake sediments and soils (extracted by Low n' Slow hydrolysis method) 3-OH FA data (fractional abundance and proxies)	292
E4	3-OH FA proxy error calculation for Southwestern U.S. lake sediments and soils (extracted by both methods)	293
E5	Loss on Ignition (LOI) calculation	294
E6	3-OH FA standards used for SW U.S. lake sediments and soils GC-MS runs, and calculation of response factor in preparation for calculation of 3-OH FA concentrations in samples.	295
E7	SW U.S. lake sediments and soils (extracted by both hydrolysis methods) 3-OH FA concentrations	296

E8	Standard error calculation for 3-OH FA concentrations of SW U.S. lake sediments and soils extracted by both hydrolysis methods	297
E9	GDGT data from SW U.S. lake sediments and soils	298
E10	U.S. environmental conditions	299
E11	3-OH FA data for Esthwaite Water lake core	300
E12	Environmental data from Esthwaite Water. Data from Maberly et al. (2017), and data from 2014 to 2017 from Feuchtmayr et al. (2021).	301
E13	Calibration data for Chapter 3	302
E14	Calibration data for Chapter 4	303
E15	Statistics associated with PCA and RDA in Chapter 3. A Results of permutation analyses used to fit environmental variables to PCA (Figure 3.9). B Constrained variance of 3-OH FA proxy distributions amongst U.S. soils explained by environmental variables with significance indicated (Figure 3.10). C RDA results following forward selection of significant environmental descriptors on 3-OH FA proxy constrained variance (Figure 3.10).	304
E16	Results of permutation tests used to fit 3-OH FA fractional abundance distributions to NMDS of global soils in Figure 3.14.	305
E17	Statistics associated with PCA and RDA in Chapter 4. A Results of permutation analyses used to fit environmental variables to PCA for SW lake sediments (Figure 4.10 A), and B for soils (Figure 4.10 B). C Constrained variance of 3-OH FA fractional abundance distributions explained by environmental variables with significance indicated for SW lake sediments (Appendix B4 A), and D for SW soils (Appendix B4). E Constrained variance of 3-OH FA proxy distributions explained by environmental variables with significance indicated for SW lake sediments (Figure 4.15 A), and F for SW soils (Figure 4.15 B).	306

E18	Permutation tests for Chapter 4 NMDS ordination. Left: Fig.4.4. Middle: Fig.4.19. Right: Fig.4.20.	307
E19	Permutation tests for Chapter 5 NMDS ordination. From left to right: (i) 3-OH FA fractional abundance permutation Fig. 5.1. (ii) Fig.5.4, (iii) Appendix C8, (iv) Fig 5.8. (v) environmental variables permutation for Esthwaite Water lacustrine sequence, Fig.5.8	308

ACRONYMS

-COOH carboxylic acid functional group. 22

T_{DJF} mean winter air temperature (months of December, January, February). xxiii, 19, 34, 50, 55, 61, 64, 65, 71, 131, 134, 136, 144, 146, 147, 150, 151, 173, 177

T_{JJA} mean summer air temperature (months of June, July, August). 19, 34, 37, 40, 61, 65, 71, 134, 136, 141, 144, 146, 147

T_{MAF} mean temperature of months above freezing ($>0^{\circ}\text{C}$). 20, 109, 179, 181

T_{MAM} mean spring air temperature (months of March, April, May). ii, xxiii, xxviii, 19, 50, 51, 61, 65, 71, 134, 144–146, 150, 151, 172–174, 177, 180, 197, 198, 203, 207, 221, 285

T_{SON} mean autumn air temperature (months of September, October, November). 19, 61, 71, 146

T_{Soil} mean annual *in situ* soil temperature. xviii, 33, 34, 37–40, 54, 61, 71, 103, 113, 136, 137

T_{Water} lake surface water temperature. 10, 113, 114

ASV amplicon sequence variant. Sequences of DNA from marker genes used to group bacterial species following 16S RNA sequencing. 161

BFAs branched fatty alcohols. 4, 5, 30

branched ($\sum [\textit{iso-} + \textit{anteiso-3-OH FAs}]$). 41

BSTFA N,O-bis-(trimethylsilyl)trifluoroacetamide. 22

BWT bottom water temperature. 10, 222, 224

CA/SA catchment (watershed) area to lake surface area ratio. 114

CBT Cyclisation of branched tetraethers index. 2

DCM dichloromethane. [21](#), [22](#), [24](#)

DO dissolved oxygen. [110](#)

FAME fatty acid methyl ester. [22](#), [185](#), [199](#)

GC gas chromatography. [22](#)

GC-MS gas chromatography-mass spectrometry. [16](#), [22](#), [183](#), [187](#)

HCl hydrochloric acid. [21](#)

HGs heterocyte glycolipids. [4](#), [30](#)

IPCC Intergovernmental Panel on Climate Change. [i](#), [4](#)

LPS lipopolysaccharide. [10](#), [16](#)

LSWT Lake surface water temperature. [20](#)

MAAT mean annual air temperature. [i](#), [ii](#), [xvii–xxi](#), [xxiii](#), [xxiv](#), [xxvi–xxviii](#), [6](#), [7](#), [12](#), [14](#), [15](#), [19](#),
[20](#), [31](#), [33](#), [34](#), [36](#), [38](#), [39](#), [50](#), [51](#), [54](#), [59–66](#), [68–92](#), [95](#), [101–103](#), [105](#), [106](#), [108–114](#), [116](#),
[131–134](#), [136–142](#), [144–147](#), [149](#), [150](#), [154](#), [163–169](#), [171–179](#), [181](#), [196–198](#), [202](#), [203](#),
[220–222](#), [259](#), [261](#), [272](#), [275](#), [276](#), [284](#), [286](#)

MAP mean annual precipitation. [i](#), [xvii–xx](#), [14](#), [19](#), [33–41](#), [51](#), [54–56](#), [58–62](#), [66](#), [67](#), [72](#), [74–79](#),
[81](#), [94](#), [95](#), [100](#), [101](#), [105](#), [111](#), [114](#), [131](#), [134](#), [138](#), [141](#), [146–148](#), [151](#), [154](#), [160](#), [179](#)

MBT Methylation of branched tetraethers index. [2](#)

MeOH methanol. [22](#), [24](#)

MLR multiple linear regression. [32](#)

Na₂SO₄ sodium sulphate. 21

NAO North Atlantic Oscillation. 34

NMDS Non-metric multidimensional scaling analysis. 118

OH-FAME hydroxy fatty acid methyl ester. 22, 185

RMSE root mean square error. 28

SiO₂ silica gel (silicon dioxide). 22, 24

SMO soil moisture. xx, 18, 33–35, 40, 54, 65, 72, 75, 76, 78, 79, 81, 94, 100, 101

SRP soluble reactive phosphate (PO₄-P). 210

SST sea surface temperature. 3, 9, 23, 57, 108, 110, 164, 174, 176

TLE total lipid extract. 21, 22, 24

UKCEH UK Centre for Ecology and Hydrology. vi, vii, xxviii, 17, 21, 288

VPD vapour pressure deficit. 19, 35, 37, 40, 54, 55, 62, 65, 72, 81, 100, 114, 154

EQUATIONS

3-Hydroxy fatty acids (3-OH FAs):

$$\mathbf{RAN}_{15} = a\text{-}C_{15} / n\text{-}C_{15}$$

$$\mathbf{RAN}_{17} = a\text{-}C_{17} / n\text{-}C_{17}$$

$$\mathbf{Branching\ Ratio\ (BR)} = \sum [iso- + anteiso-] / \sum [normal-]$$

$$\mathbf{RIAN} = -\log(BR)$$

$$\mathbf{RAN}_{13} = a\text{-}C_{13} / n\text{-}C_{13}$$

$$\mathbf{RIN} = \sum [iso-] / \sum [normal-]$$

$$\mathbf{-log(RIN)} = -\log(\sum [iso-] / \sum [normal-])$$

$$\mathbf{RIN}_{USA} = \sum [i\text{-}C_{13} + i\text{-}C_{16} + i\text{-}C_{17}] / \sum [n\text{-}C_{14} + n\text{-}C_{18}]$$

$$\mathbf{-log(RIN}_{USA}) = -\log(\sum [i\text{-}C_{13} + i\text{-}C_{16} + i\text{-}C_{17}] / \sum [n\text{-}C_{14} + n\text{-}C_{18}])$$

$$\mathbf{RAN}_S = (i\text{-}C_{12} + a\text{-}C_{13} + i\text{-}C_{14}) / (i\text{-}C_{12} + a\text{-}C_{13} + n\text{-}C_{13} + i\text{-}C_{14} + n\text{-}C_{14})$$

$$\mathbf{RIN}_{17} = i\text{-}C_{17} / n\text{-}C_{17}$$

$$\mathbf{Lake\ 3\text{-}OH\ FA\ temperature\ response} = (i\text{-}C_{12} + a\text{-}C_{13} + a\text{-}C_{17}) / n\text{-}C_{18}$$

$$\mathbf{Soil\ 3\text{-}OH\ FA\ temperature\ response} = (i\text{-}C_{13} + i\text{-}C_{14} + a\text{-}C_{15}) / n\text{-}C_{18}$$

Glycerol dialkyl glycerol tetraethers (GDGTs):

$$\mathbf{R_{i/b}} = \sum [isoGDGTs] / \sum [brGDGTs]$$

$$\mathbf{BIT} = \frac{(Ia + IIa + IIIa + IIa' + IIIa')}{(Ia + IIa + IIIa + IIa' + IIIa' + cren)}$$

$$\mathbf{MBT} = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic) + (IIa^* + IIb^* + IIc^*) + (IIIa^* + IIIb^* + IIIc^*)}$$

$$\mathbf{CBT} = -\log_{10} \left(\frac{(Ib + IIb + IIb')}{(Ia + IIa + IIa')} \right)$$

$$\mathbf{MBT'}_{5ME} = (Ia + Ib + Ic) / (Ia + Ib + Ic + IIa + IIb + IIc + IIIa)$$

$$\mathbf{MBT'}_{6ME} = (Ia + Ib + Ic) / (Ia + Ib + Ic + IIa' + IIb' + IIc' + IIIa')$$

$$\mathbf{CBT'} = -\log_{10} \left(\frac{(Ic + IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc')}{(Ia + IIa + IIIa)} \right)$$

$$\mathbf{IR} = (IIa' + IIIa') / (IIa + IIIa + IIa' + IIIa')$$

$$\mathbf{DC'} = (Ib + IIb + IIb') / (Ia + IIa + IIa' + Ib + IIb + IIb')$$

CHAPTER 1: INTRODUCTION

1.1 Need for development of novel terrestrial palaeoclimatic biomarker proxies

As single-celled organisms, the physiological responses of microbes to external environmental conditions occur by lipodomic (lipid composition) adjustment of the cell membrane (Kaneda 1991, Guerzoni et al. 2001, Ernst et al. 2016)- the barrier between the external and internal environment of a microbe. Therefore certain lipid groups can track growth conditions such as temperature and pH, which are used as organic biomarkers (i.e. molecular fossils (Eglinton & Calvin 1967)) to reconstruct past environmental conditions (Tierney 2012, Schouten et al. 2013, Eglinton & Eglinton 2008, McClymont et al. 2023). Organic biomarkers are important tools to reconstruct palaeoclimatic conditions as they are often produced ubiquitously in a range of environments, have a long residence time and preservation potential in sedimentary archives due to the hydrophobic nature of lipids, and can be directly associated with a specific biochemical response mechanism to environmental forcing if the biological producers are known. Not all organic biomarkers are microbially derived, they can have any biological origin (McClymont et al. 2023). Biomarker proxies for the reconstruction of specific environmental conditions such as temperature and pH have been developed for application in marine (Kim et al. 2008, Müller et al. 1998) and terrestrial environments (Weijers et al. 2007), however proxy development is often more complex in the terrestrial realm due to environmental, topographic, and climatic heterogeneity causing complex biogeographic distributions in biota that produce organic biomarkers such as plants and microorganisms, resulting in a limited number of biomarker proxies available for temperature reconstruction in terrestrial environments.

Branched Glycerol Diakyl Glycerol Tetraethers (brGDGTs) are the most investigated biomarker group for temperature proxy calibration and application in terrestrial environments. brGDGTs were first detected in peat (Sinninghe Damsté et al. 2000), and have since been detected

in a range of environmental media derived from (but not limited to) pedogenic, lacustrine, and riverine environments (Weijers et al. 2006, Pearson et al. 2011, De Jonge et al. 2014a). brGDGTs comprise several suites of compounds that differ in degree of methylation, proportion of cyclopentane moieties and isomerism (Fig. 1.1). Degree of methylation and degree of cyclisation of brGDGTs are respectively correlated with temperature and pH, which were quantified by Weijers et al. (2007) as the MBT and CBT proxies. MBT has a long history of pH-dependence and has undergone revision and re-calibration (Peterse et al. 2012, De Jonge et al. 2014a) to deconvolve a brGDGT temperature response. MBT'_{5ME} omits 6-methyl brGDGT isomers based on new analytical methods which promote separation of 5- and 6-methyl brGDGT isomers which previously coeluted (De Jonge et al. 2014a, Hopmans et al. 2016). 6-methyl brGDGT compounds were found to have a pH response in soils (De Jonge et al. 2014b) and tropical lakes (Russell et al. 2018), however 6-methyl brGDGTs do vary in response to temperature instead of 5-methyl brGDGTs in some lacustrine environments (Dang et al. 2018), so the applicability of MBT'_{5ME} can be complex.

MBT'_{5ME} has been regionally and globally calibrated with contemporary air temperature in pedogenic (De Jonge et al. 2014a, Dearing Crampton-Flood et al. 2020) and lacustrine environments (Russell et al. 2018, Bauersachs et al. 2024, Zhuang et al. 2022, Raberg et al. 2021, Martínez-Sosa et al. 2021), however uncertainty concerning the use of brGDGTs remains. The performance of MBT and MBT'_{5ME} is variable, particularly in lacustrine environments (Dang et al. 2018), and arid pedogenic environments (Peterse et al. 2012, Dirghangi et al. 2013, Dearing Crampton-Flood et al. 2020) where soil moisture becomes an environmental stress on microbial communities. Although efforts have been made to deconvolve the pH dependence of MBT, the biological producers of brGDGTs are largely unknown and under investigation. *iso*-diabolic acids which are thought to be precursors of brGDGTs have been identified in strains of the bacterial phyla Acidobacteria (Sinninghe Damsté et al. 2018), however brGDGTs have characteristics of both bacterial and archaeal synthesis (Weijers et al. 2006). Furthermore, temperature and

pH-induced changes in bacterial community structure (involving Acidobacteria) are capable of influencing correlations between MBT'_{5ME} and temperature (De Jonge et al. 2019, van Bree et al. 2020) which raises implications for application and interpretation of this proxy in temporal records where bacterial community changes might occur.

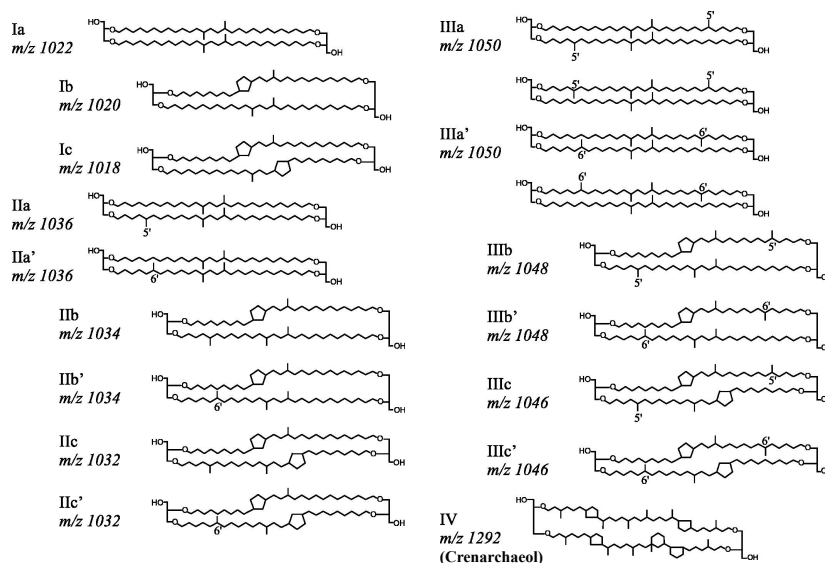


Figure 1.1: Chemical structures of brGDGTs (I-III) and Crenarchaeol (IV) adapted from De Jonge et al. (2014b). brGDGTs denoted by 'I' are 'tetra-methylated'. brGDGTs denoted by 'II' are 'penta-methylated', and brGDGTs denoted by 'III' are 'hexa-methylated'. brGDGTs denoted by 'a' have no cyclopentane moieties, those denoted by 'b' have one cyclopentane moiety, and those denoted by 'c' have two cyclopentane moieties. Prime ' denotes 6-methyl brGDGT isomers.

MBT'_{5ME} is the only extensively calibrated temperature proxy for application in terrestrial environments. Long-chain alkenones on which the UK'₃₇ SST proxy is based has limited applicability in lacustrine environments despite robust application in the marine realm as their lacustrine occurrence is restricted (Toney et al. 2010). Furthermore, applicability of isoprenoidal

(i)GDGT-based TEX₈₆ in lacustrine environments is generally limited to deep tropical lakes where Thaumarchaeota reside (Tierney et al. 2010), in contrast to their ubiquitous distribution in the marine realm. The influence of temperature on average chain length (ACL) of C₂₁-C₂₉ leaf wax *n*-alkanes has also been investigated across continental soil transects (Bush & McInerney 2015), however significant ACL-temperature correlations are regionally limited and are confounded by the effect of aridity amongst other implications. As a result, ACL has not been calibrated as a terrestrial temperature proxy, but has demonstrated reconstructive potential in certain marine environments (Seki et al. 2021). With terrestrial environments anticipated to experience enhanced future warming relative to oceans, and uncertainty over the expected magnitude and variability of climate warming and drought in terrestrial environments (IPCC 2021), the IPCC have issued bulletins in recent years highlighting the importance to gather more high-resolution terrestrial palaeoenvironmental data to narrow the range of uncertainty of future temperature predictions particularly in localised areas where hydroclimate modes tend to converge and which are often associated with aridity and drought (i.e. SW U.S., Southern Africa); such areas currently have a high degree of uncertainty in climate prediction models. (Jansen et al. 2007). In light of this, several novel biomarker groups have been explored for temperature proxy development; the HDI₂₆ proxy based on Heterocyte Glycolipids (HGs) derived from N₂-fixing heterocytous cyanobacteria was successfully applied to a ca.37 ka sediment core from Lake Tanganyika (East Africa) demonstrating potential as a lacustrine temperature proxy (Bauersachs et al. 2015, 2021), and temperature dependence of branched fatty alcohol (BFAs) distributions has been investigated in soils and peats from Mt. Shennongjia (China) using the BNA₁₅ index (Huang et al. 2013), however BNA₁₅ has not been investigated in detail beyond this local study. The potential for 3-hydroxy fatty acids (3-OH FAs) to record climate information in terrestrial and marine environments has also been investigated across numerous studies in recent years, which is summarised from here on.

1.2 3-Hydroxy fatty acids

1.2.1 Potential as a novel terrestrial palaeoclimate biomarker

α , β , and ω hydroxy fatty acids are widely abundant in nature, occurring in animal and plant tissues and microorganisms (refer to [Zhang et al. \(2023\)](#) and references therein for detailed review). Despite this diverse occurrence, C₁₀-C₁₈ 3-OH-FAs (hydroxyl group in β position) are diagnostic biomarkers for Gram-negative bacteria, being a component of Lipid A in the bacterial outer cell membrane ([Rietschel 1976](#), [Klok et al. 1988](#), [Wang et al. 2016](#)). 3-OH FAs differ in carbon number, structure, and in some cases saturation; 3-OH FA compounds with even C numbers have *iso*- and *normal*- homologues whilst compounds with odd C numbers also have an *anteiso*- homologue. *iso*-, *anteiso*-, and *normal*- compounds differ in degree of methyl branching; *iso*- structures comprise a methyl branch on the penultimate carbon, *anteiso*- structures comprise a methyl branch on the ante-penultimate carbon, and *normal*- structures comprise a straight carbon chain with no methyl branching (Fig.1.2). Methyl branching in the *iso*- and *anteiso*- positions is not unique to 3-OH FAs, and is widely observed amongst lipids in nature, including fatty acids (FAs) and BFAs ([Huang et al. 2013](#)). Mono-unsaturated 3-OH FAs are much less abundant, but have been reported for 3-OH FAs with even C numbers in lake sediments from China.

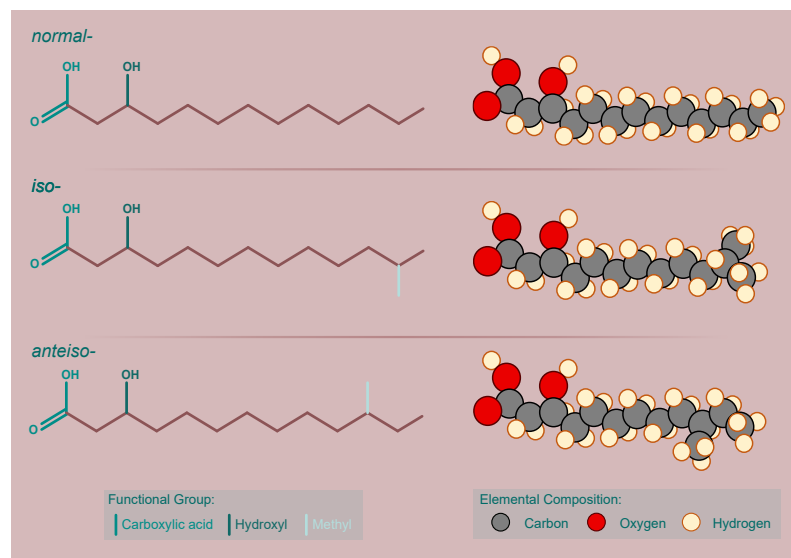


Figure 1.2: Chemical structures of 3-OH FAs.

There is surmounting evidence that different 3-OH FA structures are bio-synthesised in response to environmental conditions. Wang et al. (2016) identified a correlation between the ratio of *anteiso*- to *normal*- C₁₅ and C₁₇ 3-OH FAs and temperature across a contemporary Mt. Shennongjia (China) soil transect (Fig. 1.3), which are summarised as the RAN₁₅ and RAN₁₇ proxies and have moderate correlation coefficients with MAAT and respective calibration RMSEs of 2.60 and 2.70 °C (Table 1.1). The potential of RAN₁₅ and RAN₁₇ as temperature proxies has since been investigated across regional and global soil transects (Fig. 1.3), and their response to temperature has been highly variable between soil transects regionally, leading to variability in reconstructive potential of regional calibrations based on these proxies. Correlations between RAN₁₅ and MAAT amongst regional soil transects vary in coefficient (R^2) from 0.00 (Mts. Lautaret-Bauges, French Alps; (Véquaud et al. 2021a)) to 0.80 (Mt. Rungwe, Tanzania; (Huguet et al. 2019)), and from 0.00 (Mts. Lautaret-Bauges) to 0.81 (Peruvian Andes; (Véquaud et al. 2021b)) for RAN₁₇ (Table 1.1). In turn, RAN₁₅ and RAN₁₇ calibrations developed from regional soil transects with significant ($p < 0.05$) proxy-MAAT correlations ($R^2 \geq 0.49$ for both RAN₁₅ and

RAN₁₇) have variable RMSE between 1.98 (Mt. Shigyla, Tibet; (Véquaud et al. 2021b)) and 3.57 °C (Peruvian Andes) for RAN₁₅, and between 2.07 (Mt. Shigyla) and 3.19 °C (Mt. Rungwe) for RAN₁₇ (Table 1.1). The statistics of global temperature calibrations based on RAN₁₅ and RAN₁₇ are largely dependent on the regional soil transects used as a calibration dataset; RAN₁₇ is associated with a higher correlation coefficient with MAAT and lower RMSE in the compilation of altitudinal soil transects by (Véquaud et al. 2021b), however RAN₁₅ had an improved capability of tracking contemporary MAAT in the global soil transect of (Wang et al. 2021) due to a higher correlation coefficient and lower RMSE (Table 1.1).

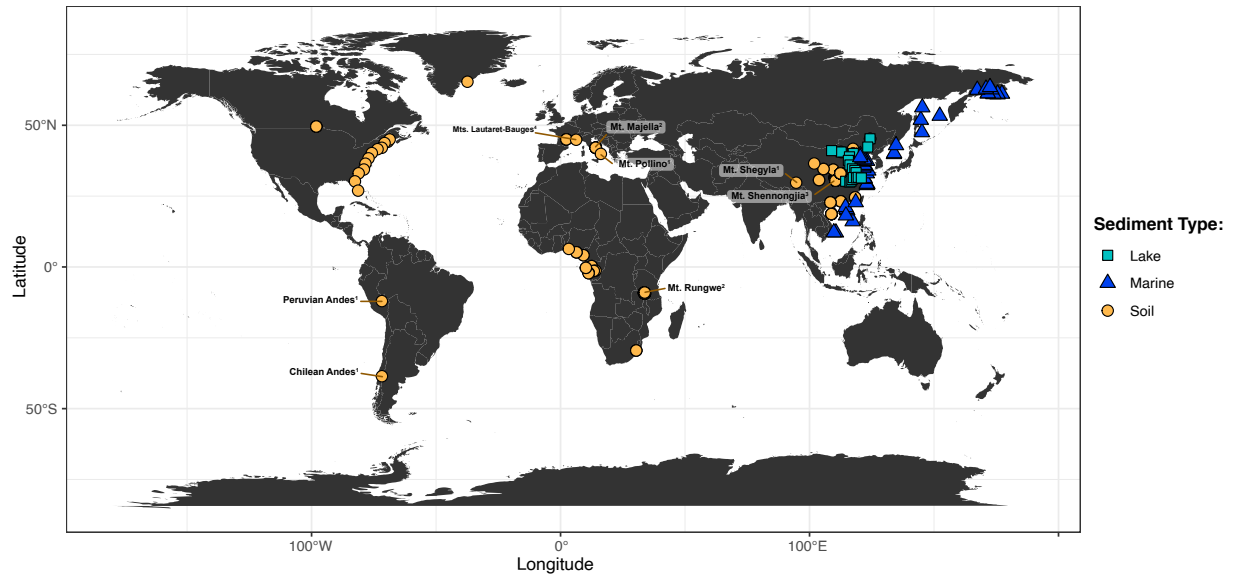


Figure 1.3: Global distribution of soils, marine and lake sediments investigated for 3-OH FAs. Soils have been investigated by Véquaud et al. (2021b) (1), Huguet et al. (2019) (2), Wang et al. (2016) (3), Véquaud et al. (2021a) (4), and remaining soils unlabelled feature in the global soil transect of Wang et al. (2021). All lake sediments are associated with the Chinese lake transect of Yang et al. (2021). Marine sediments from the marginal seas of China and North Pacific Ocean have been investigated by Yang et al. (2020) and Dong et al. (2023).

The ratio of $\Sigma(\text{iso-} + \text{anteiso-})$ to $\Sigma(\text{normal-})$ 3-OH FAs, summarised as the RIAN proxy, strongly ($R^2 = 0.70$) correlated with pH in soils from Mt. Shennongjia (Wang et al. 2016) with a calibration RMSE of pH 0.54, and also has a variable response to pH between regional soil transects. Correlation coefficients between RIAN and soil pH range between 0.04 (Peruvian Andes) and 0.70 (Mt. Shennongjia) and RMSE of significant correlations ($R^2 \geq 0.43$) between pH 0.40 (Mt. Shigela) and 0.83 (Mts. Lautaret-Bauges) (Table 1.1). The statistical integrity of global RIAN calibrations are dependent on the regional transects from which the calibration dataset is composed; the global RIAN-based pH calibration of Véquaud et al. (2021b) has an R^2 of 0.34, whereas that of (Wang et al. 2021) has a higher correlation coefficient of 0.66 with a lower RMSE (Table 1.1).

RAN₁₅, RAN₁₇, and RIAN are responsive to temperature and pH in many soils globally (Wang et al. 2021), but have a variable performance regionally, which has been attributed to limiting environmental conditions across some transects (i.e. small pH range (<2 pH units) amongst Mt. Rungwe, Mt. Majella, and Peruvian Andes soils which is a limiting factor for RIAN performance (Véquaud et al. 2021b)), as well as diversity in Gram-negative bacterial community structure between and within the soil transects (Véquaud et al. 2021b, Wang et al. 2021, Yang et al. 2020, 2021). Nevertheless, the regional RAN₁₅ soil calibration of Wang et al. (2016) indicated potential to reconstruct Holocene temperatures in a 9.0 KaBP stalagmite record from Heshang Cave (China) (Wang et al. 2018).

3-OH FA Proxy	Environmental Variable	Transect	Sample Type	R^2	p -value	RMSE	n	Calibration Study
RAN ₁₆	MAAT	Mt. Shennongjia, China	soil	0.51	<0.001	2.60°C	26	Wang, et al. (2016)
		Mt. Majella, Italy	soil	0.51	0.01	2.90°C	11	Huguet, et al. (2019)
		Mt. Rungwe, Tanzania	soil	0.52	<0.001	1.60°C	28	Huguet, et al. (2019)
		Peruvian Andes	soil	0.71	<0.001	3.57°C	14	Vequaud, et al. (2021b)
		Chilean Andes	soil	0.03	>0.05	-	10	Vequaud, et al. (2021b)
		Mt. Pollino, Italy	soil	0.54	<0.01	2.64°C	13	Vequaud, et al. (2021b)
		Mt. Shegyla	soil	0.57	<0.001	1.98°C	17	Vequaud, et al. (2021b)
		Mts. Lautaret-Bauges	soil	0	>0.05	-	49	Vequaud, et al. (2021a)
		globally distributed soils	soil	0.48	<0.001	4.90°C	186	Wang, et al. (2021)
		globally distributed soils	soil	0.37	<0.001	5.46°C	168	Vequaud, et al. (2021b)
RAN ₁₇	MAAT	Mt. Shennongjia, China	soil	0.48	<0.001	2.70°C	26	Wang, et al. (2016)
		Mt. Majella, Italy	soil	0.28	>0.05	3.80°C	11	Huguet, et al. (2019)
		Mt. Rungwe, Tanzania	soil	0.54	<0.001	2.50°C	28	Huguet, et al. (2019)
		Mts. Shennongjia, Rungwe, and Majella combined	soil	0.60	<0.001	5.10°C	65	Huguet, et al. (2019)
		Peruvian Andes	soil	0.81	<0.001	2.87°C	14	Vequaud, et al. (2021b)
		Chilean Andes	soil	0.04	>0.05	-	10	Vequaud, et al. (2021b)
		Mt. Pollino, Italy	soil	0.56	<0.01	2.57°C	13	Vequaud, et al. (2021b)
		Mt. Shegyla	soil	0.54	<0.001	2.07°C	17	Vequaud, et al. (2021b)
		Mts. Lautaret-Bauges	soil	0	>0.05	-	49	Vequaud, et al. (2021a)
		globally distributed soils	soil	0.41	<0.001	5.28°C	186	Wang, et al. (2021)
RIAN	soil pH	globally distributed soils	soil	0.41	<0.001	5.28°C	168	Vequaud, et al. (2021b)
		Mt. Shennongjia, China	soil	0.70	<0.001	pH 0.54	26	Wang, et al. (2016)
		Mt. Majella, Italy	soil	0.15	>0.05	-	11	Huguet, et al. (2019)
		Mt. Rungwe, Tanzania	soil	0.11	>0.05	-	28	Huguet, et al. (2019)
		Mts. Shennongjia, Rungwe, and Majella combined	soil	0.62	<0.001	pH 0.56	63	Huguet, et al. (2019)
		Peruvian Andes	soil	0.04	>0.05	-	14	Vequaud, et al. (2021b)
		Chilean Andes	soil	0.66	<0.01	pH 0.42	10	Vequaud, et al. (2021b)
		Mt. Pollino, Italy	soil	0.29	>0.05	-	13	Vequaud, et al. (2021b)
		Mt. Shegyla	soil	0.46	<0.01	pH 0.40	17	Vequaud, et al. (2021b)
		Mts. Lautaret-Bauges	soil	0.43	<0.001	pH 0.83	49	Vequaud, et al. (2021a)
RAN ₁₃	SST	Chinese marginal seas	marine sediment	0.92	<0.001	2.55°C	45	Yang, et al. (2020)
	MAAT	China	lake sediment	0.50	<0.01	3.00°C	24	Yang, et al. (2021)
RAN ₅	SST	Chinese marginal seas	marine sediment	0.92	<0.001	2.29°C	85	Dong, et al. (2023)
RIN ₁₇	MAAT	China	lake sediment	0.65	<0.01	2.60°C	24	Yang, et al. (2021)

Table 1.1: Statistics conveying accuracy of published temperature and pH calibrations developed from 3-OH FA-based proxies

3-OH FA distributions have also been investigated in marine and lake sediments (Fig. 1.3). Yang et al. (2020) identified a strong negative exponential correlation between the ratio of *anteiso-* to *normal*-C₁₃ 3-OH FAs and SST in 45 surface marine sediments from the marginal seas of China, which is expressed as RAN₁₃ (Table 1.1). RAN₁₃ in 40 additional marine sediments from the marginal seas of China investigated by Dong et al. (2023) were consistent with the calibration dataset of Yang et al. (2020) (Dong et al. 2023). Dong et al. (2023) additionally proposed RAN₅ as a SST proxy which describes the ratio of $\sum(i\text{-C}_{12}, a\text{-C}_{13}, i\text{-C}_{14})$ over $\sum(i\text{-C}_{12}, a\text{-C}_{13}, n\text{-C}_{13}, i\text{-C}_{14}, n\text{-C}_{14})$, after RAN₅ produced a strong negative linear correlation with SST in 85 marine sediments from the marginal seas of China, in a combined dataset of the marine sediments

investigated by [Yang et al. \(2020\)](#) and the novel sediments investigated by [Dong et al. \(2023\)](#). The ratio of *anteiso*- to *iso*-C₁₅ 3-OH FAs was also proposed as a proxy for [BWT](#) in the marginal seas of Eastern China by [Pan et al. \(2023\)](#). 3-OH FAs have been explored to a lesser degree in lacustrine environments than in pedogenic or marine environments, having been investigated across a lake transect from China where the ratio of *iso*- to *normal*-C₁₇ 3-OH FAs expressed as RIN₁₇ alongside RAN₁₃ correlated with T_{water} ([Yang et al. 2021](#)). 3-OH FA proxies to date have been fairly specific to sediment type, however RAN₁₅ developed from soils did track [BWT](#) in Eastern China marginal seas ([Pan et al. 2023](#)), and marine-derived RAN₁₃ was able to track T_{water} in Chinese lakes ([Yang et al. 2021](#)). The variability of 3-OH FA distributions and proxies between and within sediment types has been attributed to diversity in Gram-negative bacterial communities ([Wang et al. 2021](#), [Yang et al. 2020, 2021](#)) however data from 16S RNA sequencing to investigate composition of the bacterial communities producing 3-OH FAs is limited to a small number of lake and marine sediments and soils ([Yang et al. 2020, 2021](#)). Growth experiments on the Gram-negative bacterial phyla Bacteroidetes revealed that the bio-synthesis of any *anteiso*-3-OH FA compound can be responsive to temperature ([Hellequin et al. 2023](#)).

1.2.2 Biochemistry

C₁₀-C₁₈ 3-OH-FAs are produced predominantly but not exclusively by Gram-negative bacteria ([Wollenweber & Rietschel 1990](#), [Sjogren et al. 2003](#), [Zhang et al. 2023](#)), and occur within the Lipid A component of Lipopolysaccharide (LPS) which are large amphipathic glycoconjugates (consorted macrostructure of lipids and carbohydrates) that compose the outer monolayer of the outer cell membrane ([Wollenweber & Rietschel 1990](#), [Raetz et al. 2007](#)) (Fig.1.4). LPS comprises three components: (i) O-antigen, (ii) core polysaccharide which together with the O-antigen controls virulence (infective potential) of the bacterium, and (iii) Lipid A, an endotoxin that activates TLR4/MD2 animal immune receptors, inducing an inflammatory response and septic shock ([Raetz et al. 2007](#)). Lipid A comprises a 1,4-bisphosphorylated- β -d-glucosaminyl-(1 $\rightarrow$

6)-d-glucosamine disaccharide backbone with up to seven hydroxy fatty acids attached by either an ester or amide bond (Troudi et al. 2021). Lipid A composition differs between different bacterial strains, particularly in number of acyl chains and type (C number and structure) (Kato et al. 1998, Troudi et al. 2021). Lipid A in *E. coli* is hexa-acylated for example, comprising four *n*-C₁₄ hydroxy fatty acids directly bound to the glucosamine disaccharide moiety and two secondary hydroxy fatty acids ester bound to primary acyl chains (Raetz et al. 2007, Troudi et al. 2021). The primary role of 3-OH FAs bound within Lipid A is to anchor the O-antigen and core oligosaccharide to the lipid bilayer of the outer membrane, and adjust fluidity of the outer membrane.

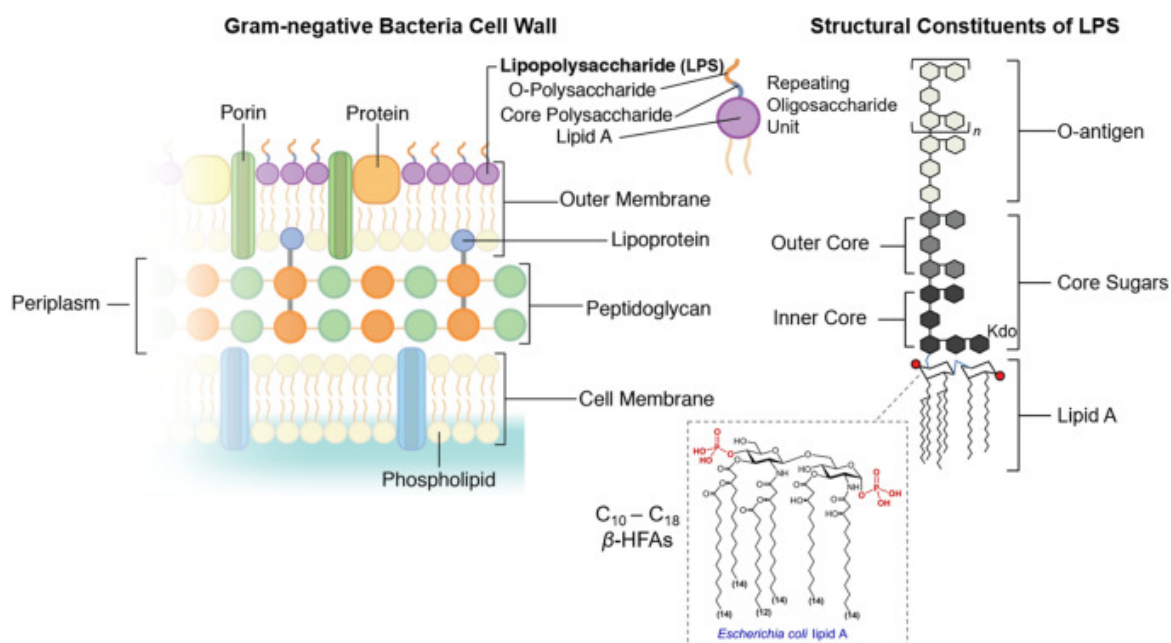


Figure 1.4: Schematic diagram of the occurrence of 3-OH FAs in Gram-negative bacterial cell membranes. In colour to the left is the structure of the cell membrane. LPS is a component of the outer membrane, for which the structure is detailed to the right. 3-OH FAs are bound by ester and amide bonds in the Lipid A component of LPS (Lipid A structure from *E. coli* visualised at bottom) Figure from Zhang et al. (2023).

1.3 Homeoviscous adaptation: lipodomic modification of bacterial cell membranes for maintenance of optimum fluidity

Homeoviscous adaptation is a cumulative term describing any lipodomic adjustment of a cell membrane to regulate fluidity (Guerzoni et al. 2001, Ernst et al. 2016). Biological processes across a cell membrane are optimised when the phospholipid bilayer is in a liquid-crystalline phase ($L\alpha$) (Garcia-Manyes et al. 2005), and maintenance of this phase is dependent on phase transition temperature (T_M) and average melting point of the fatty acids composing the bilayer (Kaneda 1991). T_M is the temperature at which the crystalline state of the lipid bilayer enters a more rigid gel phase ($L\beta$), and varies between lipid residue (Kaneda 1991, Garcia-Manyes et al. 2005). At T_M , half of the membrane lipids melt and the membrane exists in equal proportions of $L\alpha$ and $L\beta$ (Siliakus et al. 2017), and below T_M the lipid bilayer enters $L\beta$ becoming too rigid and increasing in thickness, limiting osmotic processes and functionality of protein channels that transfer ions across the membrane; most cell membranes can only function in $L\alpha$ (Russell & Fukunaga 1990), and so membrane modifications are required to maintain this phase.

$L\alpha$ of a lipid bilayer is jeopardised by ambient conditions in the external environment of a microbe (i.e. temperature), and lipodomic modification involving saturation, aromatic structure, branching, or acyl chain length adapt membrane fluidity accordingly (Siliakus et al. 2017, Kaneda 1991, Russell & Fukunaga 1990). At lower growth temperatures, microbes have a tendency to incorporate lipids with lower melting points into the lipid bilayer (i.e. higher degree of unsaturation, higher degree of branching) to promote a more disordered gel phase or to prevent a gel phase from occurring (Siliakus et al. 2017). Branched FAs have lower melting points and lower phase transition temperatures than their straight-chain counterparts, particularly *anteiso*-FAs, and particularly *a*-C₁₅ which has a melting point of 23.0 °C and T_M of -16.5 °C compared to the melting point of 34.2 °C and T_M of 52.5 °C of *n*-C₁₅ (Kaneda 1991, Zhu et al. 2005). As the melting point of *a*-C₁₅ occurs at what is considered a warm MAAT, T_M is likely

more influential in controlling membrane fluidity than lipid melting point. *anteiso*-FAs have lower melting points and T_M compared to their *iso*- homologues as the positioning of the methyl branch in the *anteiso*- position is more disruptive to intermolecular forces and dipoles between membrane lipids, and the crinkled carbon chain structure promoted by branching reduces the packing structure of membrane lipids and reduces the thickness of the cell membrane (Poger et al. 2014), increasing membrane fluidity and permeability, and alleviating the effects of $L\beta$.

Membrane modification in light of homeoviscous adaptation can also occur at different rates depending on lipid group. Cold shock in microbial cell membranes is induced when the membrane becomes rigid and increases in thickness. An increase in relative abundance of unsaturated FAs occurs rapidly by the deployment of desaturase enzymes which add a double carbon bond to saturated FAs already incorporated in the membrane (Siliakus et al. 2017), therefore modifying membrane lipids in-situ. However branched FAs are produced *de novo* (i.e. genetic code needs to adapt to environmental conditions to promote synthesis of branched lipids) so modification of cell membrane fluidity by degree of branching occurs by growth of new cells and not by adding branches to existing straight-chain lipids in the membrane (Siliakus et al. 2017). The initial homeoviscous adaptation response in microbial cell membranes is therefore to promote unsaturation in-situ, and then synthesise low-melting point lipids (promote unsaturation and branching) *de novo* with bacterial growth under persistent environmental conditions (Zhu et al. 2005, Russell & Fukunaga 1990).

In addition to a bacterial cold-shock response, 3-OH FAs in Gram-negative bacterial cell membranes can be adjusted in response to pH. Branched 3-OH FAs comprising *iso*- and *anteiso*- structures promote porosity when incorporated into a lipid bilayer, rendering the cell membrane more permeable to H^+ cations which is a favourable adaptation to low H^+ concentrations in the external environment (i.e. for bacteria residing in more alkaline soils) (Wang et al. 2021). This lipodomic adjustment of membrane permeability to maintain a H^+ gradient is termed

chemiosmotic coupling (Wang et al. 2021, Denich et al. 2003, Russell & Fukunaga 1990). As a result, the 3-OH FA branching ratio (BR) has been observed to correlate with soil pH in soils globally (Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021b, Wang et al. 2021). pH is also a primary environmental control on the structure of bacterial communities producing 3-OH FAs as different bacterial phyla have variable tolerance of acidic environmental conditions. Acidic soils can promote lower bacterial diversity composed of acidophilic specialists such as Acidobacteria, compared to neutral soils where bacterial community structure is likely to be more diverse (Wu et al. 2017).

1.4 Thesis aims and outline

This thesis uses surface soils and lake sediment samples from environmentally diverse transects across the U.S. to further investigate variability of 3-OH FA distributions in response to environmental forcing in pedogenic and lacustrine environments. The transects in question comprise topsoils between Maine and Florida along the U.S. East Coast, a transect previously investigated for *n*-alkane (Tipple & Pagani 2013) and GDGT (Dirghangi et al. 2013) distributions, and co-sampled lake sediments and catchment topsoils from 26 lake catchments distributed across the SW U.S. between Arizona and Texas, a lake transect previously investigated for δD values of *n*-alkanoic acids by Hou et al. (2008). The continental scope of these U.S. transects enables investigation of 3-OH FA distributions across large gradients in MAAT (6.08-23.0 °C), soil (pH 3.56-8.77) and lake sediment (pH 5.76-8.80) pH, and MAP (246-1474 mm). The SW U.S. is also an arid region; aridity is known to be a confounding environmental influence on the brGDGT-based MBT temperature proxy (Dirghangi et al. 2013, Peterse et al. 2012), however the influence of aridity on the 3-OH FA-based RAN₁₅ and RAN₁₇ temperature proxies has not been explored previously. Therefore, this U.S. study area offers this opportunity. Co-sampled lake sediments and catchment soils also offer the opportunity to directly compare 3-OH FA distributions between lacustrine and pedogenic environments, which is important to address

whether the same 3-OH FA-based proxies and calibrations are applicable between the environments, or whether in-situ lacustrine 3-OH FA production is occurring.

The outline and aims of this thesis are as follows:

CHAPTER 3: DISTRIBUTIONS OF 3-HYDROXY FATTY ACIDS REPORTED FROM TWO ENVIRONMENTALLY CONTRASTING SOIL TRANSECTS IN THE USA

This chapter aims to **assess the potential for temperature proxy development and calibration from 3-OH FAs in pedogenic environments at the continental scale**. Two studies investigating 3-OH FA variability in globally distributed soils reported a high degree of regional variability in correlations between the RAN_{15} and RAN_{17} proxies and [MAAT](#) and $RIAN$ and soil pH, and recommend the use of regional 3-OH FA-based RAN_{15} and RAN_{17} calibrations. However the environmental constraints on the regional applicability of 3-OH FA proxies have not been defined. Regional transects are currently defined by geographic area, for example, but different geographic areas can have similar environmental characteristics.

We therefore investigate the environmental response of RAN_{15} , RAN_{17} , and $RIAN$ amongst other previously developed and newly developed 3-OH FA-based proxies along pH, moisture and temperature gradients across the U.S. Aridity is known to be a confounding environmental variable on the performance of MBT, MBT', and MBT'_{5ME}, and we additionally investigate whether this also occurs for proxies based on 3-OH FAs by investigating 3-OH FA distributions from U.S. East Coast and SW U.S. soil transects. brGDGT distributions and environmental proxies (MBT, CBT, BIT, and $R_{i/b}$) have been reported by [Dirghangi et al. \(2013\)](#) for the for U.S East Coast soils. Therefore we present novel brGDGT data from the SW U.S. soils and compare with the existing East Coast data.

CHAPTER 4: DEVELOPMENT OF 3-HYDROXY FATTY ACID-BASED PALAEOTHERMOMETRY IN LACUSTRINE ENVIRONMENTS ACROSS THE SOUTHWESTERN UNITED STATES

This chapter aims to **investigate 3-OH FA distributions and responses to environmental variability within and between lacustrine and pedogenic environments**. Lacustrine environments are important archives for terrestrial palaeoclimatic information, however 3-OH FAs have only been investigated across one regional lake transect in China, therefore the variability of lacustrine 3-OH FA distributions and proxy applicability in lakes outside of China remain unknown. Lacustrine in-situ 3-OH FA production relative to catchment soils is investigated for the first time for 3-OH FAs (but is a common approach for brGDGTs), and proxy response to environmental conditions are compared between regional lake transects (3-OH FA data of [Yang et al. \(2021\)](#) from Chinese lakes) and between SW U.S. lake sediments and soils. 3-OH FA distributions and environmental proxies proposed from previous studies (RAN₁₅, RAN₁₇, RAN₁₃, RAN₅, RIN₁₇, RIN, BR, RIAN) are reported for the SW U.S. lakes, and 3-OH FA data from the SW soil transect is reassessed in this chapter based on results from Chapter 3. brGDGT distributions are reported from the SW U.S. lake sediments, and brGDGT-based environmental proxies (MBT'_{5ME}, CBT', MBT, IR, R_{i/b}, BIT) reported for both SW U.S. lake sediments and soils.

CHAPTER 5: INITIAL DEVELOPMENT OF A LOW N' SLOW ACID HYDROLYSIS METHOD FOR 3-OH FA EXTRACTION

This chapter aims to **assess the performance of a low temperature acid hydrolysis method in 3-OH FA extraction**. Acid hydrolysis is an essential acid-catalysed thermal pre-treatment required for 3-OH FA analysis as 3-OH FAs occur naturally bound in bacterial [LPS](#) and cannot be detected by [GC-MS](#) in this state. However the harsh thermal conditions of the hydrolysis method currently used (130 °C) have the potential to cause lipid degradation during extraction and requires specialised equipment that both potentially limit the applicability of 3-OH FA analysis in organic geochemistry laboratories.

The performance of two acid hydrolysis extraction methods that differ in temperature and duration on 3-OH FA yield and distribution is compared by extracting the SW U.S. lake sediments

and soils twice, firstly by the original 130 °C hydrolysis method (3 hrs, 3M HCl) ([Wang et al. 2012](#)) the data for which is reported in chapters 3 and 4, and secondly by a lower temperature 'Low n' Slow' method using 70 °C hydrolysis (14 hrs, 3M HCl) (conditions after [Ichihara & Fukubayashi \(2010\)](#) and a wider body of method development literature cited in Chapter 5) the data for which is exclusively reported in this chapter. The temperature and duration conditions for Low n' Slow hydrolysis were selected from an array of method studies focused on fatty acid extraction, and are tested for the first time for 3-OH FA extraction in the context of organic geochemistry in this thesis. The original 130 °C hydrolysis method optimised by [Wang et al. \(2012\)](#) was used for 3-OH FA extraction from U.S. lake sediments and soils in chapters 3 and 4 for direct comparability to previously published 3-OH FA data and to allow for development and refinement of the Low n' Slow hydrolysis method in future work, such as optimisation of the hydrolysis conditions temperature, duration, and acid concentration.

CHAPTER 6: APPLICATION OF LOW N' SLOW ACID HYDROLYSIS FOR 3-OH FA EXTRACTION IN A LAKE SEDIMENT CORE FROM ESTHWAITE WATER

This chapter aims to **conduct a preliminary investigation of how 3-OH FAs track variability in environmental conditions in a sediment core from a lake environment with a history of nutrient enrichment**. 3-OH FAs have been investigated across environmental gradients of many contemporary surface transects in pedogenic and lacustrine environments, but not in a down-core lacustrine sequence. Esthwaite Water in the English Lake District is a unique site as it has been continuously monitored for water quality and environmental conditions since 1945 by the Centre for Ecology and Hydrology ([UKCEH](#)) ([Maberly et al. 2017](#), [Thorpe et al. 2022](#)). 3-OH FA distributions throughout the lacustrine sequence are investigated, compared with temporal variability of environmental conditions, and the resolution of 3-OH FA proxy variability noted.

CHAPTER 2: METHODS

2.1 Sample collection and processing

When collecting soil samples, any overlying leaf litter was firstly removed to expose topsoil, and a trowel was used to dig and mix soil 0-10 cm within the subsurface. An approximately 300 g of sample of the mixed soil was collected. Along the SW U.S. lake transect, an individual soil sample was collected from the periphery of each lake site, from an area where as natural vegetation as possible was deemed to persist and where vegetation was observed to be representative of that within the lake catchment. One lake sediment sample (ca. 200 g) was collected from each lake using an Eckman grab. One lake sediment and catchment soil sample was collected per lake catchment due to the regional scope of the SW U.S. transect and the calibre of this study as a multi-lake calibration. However lake sediment 3-OH FA distributions are expected to vary within a lake due to factors such as detrital input and turbidity regime, however an investigation of intra-lake 3-OH FA variability is best suited to a study limited to sampling of sediment transects that span individual lakes, and should be conducted as a separate study in the future. At the SW U.S lakes, the lake sediment sample was collected from the depocentre of the lake where possible for consistency. Lake sediment was mixed upon collection and sampled from the centre of the grab sample to limit contamination. The Eckman was also used to measure lake depth.

All soils and lake sediments were wrapped in pre-combusted foil and stored in Whirl Pak sample bags at -20 °C until arrival at the laboratory. Samples were freeze dried and **SMO** was measured for soils by subtracting dry sample weight from frozen sample weight. The samples were then homogenised by being passed through a 1 mm aperture sieve to remove vegetation and stones, and ground by pestle and mortar.

2.2 Measurement and acquisition of environmental variables

2.2.1 Climate data

Lake surface water temperature and pH were point measured in-situ at the time of sampling using a multi-meter probe for comparison with SW U.S. lake 3-OH FA data, and to inform of the surface water conditions at the time of sampling. Elevation was obtained at each sample site using a portable Garmin GPS handset. Long-term climate data (monthly T max, T min, monthly VPD, and precipitation) was obtained from the PRISM climate group database of Oregon State University, a database which spatially interpolates meteorological data between climate stations across the contiguous U.S. (PRISM Climate Group 2014). PRISM data was accessed to compile contemporary instrumental climate data at all sample sites for 3-OH FA proxy calibration. Monthly T max and T min values were averaged between July 1998 and July 2018 (where data was available) to calculate 30-year normals of MAAT. Seasonal 30-year temperature normals were additionally calculated by averaging T max and T min values across the same temporal range exclusively for summer months (T_{JJA}), autumn months (T_{SON}), winter months (T_{DJF}), and spring months (T_{MAM}). MAP and VPD were calculated using the same method for temperature normals, and monthly precipitation values were averaged between 1998 and 2018. 30-year normals were calculated for all climatic parameters based on the assumption that 3-OH FAs in all samples have likely accumulated over a number of decades (Tipple & Pagani 2013, Weijers et al. 2007), and a long-term average accounts for the mixed nature of the samples, provides a holistic summary of climate at our SW U.S. sample sites, and potential variability of sedimentation rate between the lakes.

The 30-year air temperature normals derived from PRISM were used for temperature calibration of 3-OH FA proxies in this thesis, including for the SW U.S. lake sediments, based on the assumption that lake surface water temperature is influenced by air temperature. This is an assumption commonly made in temperature calibration studies in lacustrine environments (Raberg

et al. 2021, Tierney et al. 2010, Bauersachs et al. 2024, Yang et al. 2021) due to the absence of long-term monitoring of surface water temperature in most lakes. Lake surface water temperature is routinely recorded as a point measurement at the time of sampling, however MAAT captures annual temperature variability at a lake catchment, and at a multi-decadal scale when calculated as a long-term normal. The use of MAAT in lacustrine temperature proxy calibrations does have implications; consistency between lake surface water and air temperatures are known to vary with latitude and altitude, with high-latitude lakes being particularly susceptible to decoupling between the two temperature regimes during periods of ice cover. As a result, average air temperature of months above freezing (T_{MAF}) has been used in recent lacustrine calibrations incorporating mid-high latitude lakes (Raberg et al. 2021, Bauersachs et al. 2024), however is not applicable to the SW U.S. lakes which do not freeze. Additionally, use of MAAT for lacustrine temperature calibrations biases calibrations to 3-OH FA production in the pelagic zone, however the sources of 3-OH FAs in lacustrine environments are currently uncertain, and possibly originate from both the water column and benthic zone (Thorpe 2023). Overall, direct comparisons between lake surface water temperature (LSWT) and MAAT are limited to a small number of lakes (Armitage 2023) and the use of MAAT to represent lake surface water temperature is considered a viable assumption. Worldclim 2.1 (Fick & Hijmans 2017) hosts a wealth of additional climate data applicable to the SW U.S. and was used to generate study area maps for the U.S. in Chapter 3. Köppen-Geiger climate classifications were obtained from Beck et al. (2023).

Data for environmental parameters aside from climate were also acquired for the SW U.S. from a number of GIS platforms. The soil types present at each sample site were obtained from the Harmonized World Soil Database V2 (Nachtergaele et al. 2023), and information regarding bedrock geology was obtained from the USGS Mineral Resources Program (Horton 2017). Vegetation cover for each sample site was accessed from the USGS Gap Analysis Project (Davidson & McKerrow 2016) and Küchler (1964).

2.2.2 pH measurement

pH is an informative measurement as a primary constraint on nutrient availability, alkalinity, and microbial activity (Miller & Kissel 2010), and was measured for each lake sediment and soil sample at UKCEH Wallingford using a standard UKCEH protocol (Emmett et al. 2008). 5 g of homogenised sample was mixed with 12.5 ml of ultra-pure (Milli-Q, Millipore) water in a 1:2.5 sample-water ratio. The solution was mixed and left to stand for 30 minutes. A calibrated pH probe was then placed in the supernatant, and a measurement was recorded once the pH reading stabilised using an Orion Ross ultra pH/ATC triode and Orion Star pH meter (Thermo Fisher Scientific, MA, U.S.).

2.2.3 LOI measurement

Loss on ignition (LOI) (%) was measured for SW U.S. soils and lake sediments as a measure of total organic carbon (TOC) to normalise 3-OH FA concentrations in these samples. 1 g sample aliquots were combusted in a muffle furnace at 450 °C for 8 hours and LOI calculated as pre-combusted sample weight - combusted sample weight. 10 samples were run in triplicate and the mean of the standard deviations of LOI (%) calculated as a measure of standard error. This error was 0.008 g and added to the analytical error of 3-OH FA concentrations for final reported error of 3-OH FA concentrations (mg g⁻¹ LOI).

2.3 Organic geochemistry: 3-OH FAs

2.3.1 3-OH FA extraction

The U.S. samples were subjected to an optimised 3-OH FA extraction method developed by Wang et al. (2012). 20 g aliquots of sample were mixed with 30 ml 3M HCl and heated at 130 °C for 3 hours. Upon cooling, the supernatant was extracted four times with DCM and eluted through Na₂SO₄ chromatography minicolumns to yield TLEs. The TLEs were dried using a gentle stream

of N₂ and split 1:1 using 500 µl DCM:MeOH (9:1 v/v). 50% TLE then underwent targeted derivatisation for 3-OH FA extraction.

Fatty acids in the TLEs were methylated with 500 µl 1.2M HCl-MeOH. During this initial derivatisation step, samples were heated at 70 °C for 1.5 hrs in sealed vials, and the -COOH functional groups substituted for -CO(CH₃) to yield FAMEs and OH-FAMEs (Liu 1994). After cooling, samples were extracted four times with *n*-hexane, and dried using N₂. Non-OH-FAMEs and OH-FAMEs were then separated using a SiO₂ column chromatography method described by Jenske & Vetter (2008) and first used for 3-OH FA extraction by Wang et al. (2016). Non-OH-FAMEs were eluted in the first fraction using *n*-hexane:ethyl acetate (98:2, v/v), and OH-FAMEs in the second fraction using 100% ethyl acetate. The isolated OH-FAMEs underwent a final derivatisation step whereby 40 µl BSTFA was added to each sample and heated at 70 °C for 1.5 hrs in sealed GC-vials to initiate silylation of the hydroxy group on C₃, which is substituted for PO(CH₃)₃ (Orata 2012) to generate TMSi-FAMEs. Samples were dried with a gentle stream of N₂, and dissolved using 50 µl *n*-hexane in preparation for GC-MS analysis.

2.3.2 3-OH FA analysis

The TMSi-FAME derivatives of 3-OH FAs were analysed by GC-MS. Instrumentation comprised an Agilent 7890B gas chromatograph and a 5977A mass spectrometer, with an Agilent J&W DB-5 fused silica capillary column (60 m x 250 µm x 0.25 µm). Helium (99.999% purity) was used as the carrier gas, at a flow rate of 1.0 ml/min. The GC oven was programmed to heat from 60 °C to 200 °C at a ramp speed of 10 °C/min, and then up to 300 °C at a ramp speed of 2 °C/min and held at 300 °C for 15 minutes. The duration of analysis for each sample was approximately 80 minutes. This analysis method follows a similar method to that of Wang et al. (2016) who also analysed TMSi-FAME derivatives in Mt. Shennongjia soils.

A 1 µL sample aliquot was injected via an autosampler and ionised using 70 eV ionisation

energy, using *n*-hexane as an injection medium. 3-OH FA peaks were integrated manually using Agilent Masshunter software. Approximately 25% of samples were injected and integrated in duplicate to determine analytical error. An in-house standard mixture was composed using pure 3-Hydroxydodecanoic acid (*n*-C₁₂), 3-Hydroxytetradecanoic acid (*n*-C₁₄), and 3-Hydroxyhexadecanoic acid (*n*-C₁₆) acquired from Stratech Scientific Ltd., the final standard having a concentration of 100 µg mL⁻¹. One 3-OH FA standard was run in a sequence for every 5 samples and integrated in the same way as sample aliquots. 3-OH FAs were then identified using a diagnostic fragment ion *m/z* 175 ([CH₃]₃SiO=CHCH₂CO₂CH₃)⁺ from their TMSi-FAME derivatives, which is produced by cleavage between C₃ and C₄ (Wang et al. 2016).

3-OH FA concentrations for SW U.S. lake sediments and soils were calculated from the response area of averaged *n*-C₁₄ peaks in the 3-OH FA standards for each sample run, and corrected to 3-OH FA concentration per gram of sample and normalised to LOI. Error was calculated by calculating mean standard deviation between samples run in duplicate for each sample run. 3-OH FA proxy error was calculated by mean standard deviation for each proxy between samples run in duplicate.

2.3.3 Calculation of 3-OH FA-based climate proxies

A number of established 3-OH FA-based palaeoclimate proxies are calculated from our U.S. soil and lake sediment distributions. RAN₁₅, RAN₁₇, and RIAN are soil-based proxies for reconstruction of temperature (RAN) and soil pH (RIAN) developed by Wang et al. (2016) and are summarised by Equations 2.1-2.4. Additional proxies calculated comprise RAN₁₃, RIN and RAN₅ for SST reconstruction respectively developed by Yang et al. (2020) and Dong et al. (2023) using Equations 2.5-2.7, and RIN₁₇ for temperature reconstruction in lake sediments after Yang et al. (2021) (Eq. 2.8).

$$RAN_{15} = a-C_{15} / n-C_{15} \quad (2.1)$$

$$RAN_{17} = a-C_{17} / n-C_{17} \quad (2.2)$$

$$\text{Branching Ratio (BR)} = \sum [iso- + anteiso-] / \sum [normal-] \quad (2.3)$$

$$RIAN = -\log(BR) \quad (2.4)$$

$$RAN_{13} = a-C_{13} / n-C_{13} \quad (2.5)$$

$$RIN = \sum [iso-] / \sum [normal-] \quad (2.6)$$

$$RAN_S = (i-C_{12} + a-C_{13} + i-C_{14}) / (i-C_{12} + a-C_{13} + n-C_{13} + i-C_{14} + n-C_{14}) \quad (2.7)$$

$$RIN_{17} = i-C_{17} / n-C_{17} \quad (2.8)$$

2.4 Organic geochemistry: brGDGTs

2.4.1 brGDGT extraction

GDGTs in U.S. East Coast (EC) soils have previously been analysed by [Dirghangi et al. \(2013\)](#), therefore GDGT extractions in this study are only performed on SW soils. GDGT extraction was performed on the remaining 50% [TLE](#) aliquots of SW soils, the other 50% having been used for 3-OH FA extraction. [TLEs](#) were sequentially eluted through [SiO₂](#) chromatography minicolumns using *n*-hexane, [DCM](#), and [MeOH](#) to respectively separate apolar (F1), polar (F2), and highly polar (F3) lipid extracts, with brGDGTs collected in F3. The dried sample extracts were dissolved in 200 µl *n*-hexane:*i*-propoanol (IPA) (99:1, v/v) and filtered using 0.45 µm aperture PTFE filters.

2.4.2 brGDGT analysis

The SW soils were analysed for their core lipid GDGT distributions at the National Environmental Research Council (NERC) National Environmental Isotope Facility (NEIF), University of Bristol,

using a ThermoFisher Scientific Accela Quantum Access triple quadrupole MS in selected ion monitoring (SIM) mode. Normal phase separation was achieved using two ultra-high performance silica columns (Acquity UPLC BEH HILIC columns, 50 mm x ID 2.1 mm x 1.7 μm , 130 Å). Waters were fitted with a 2.1 mm x 5 mm guard cartridge after [Hopmans et al. \(2016\)](#). The HPLC pump was operated at a flow rate of 200 $\mu\text{L min}^{-1}$. GDGT determinations were based on single injections.

A 15 μL sample aliquot was injected via an autosampler, and analyte separation performed under a gradient elution. *n*-hexane:IPA (98.2:1.8 v/v) was eluted isocratically for 25 minutes, followed by an increase in solvent polarity to 3.5 % IPA in 25 minutes, and an increase to 10 % IPA in 30 minutes ([Hopmans et al. 2016](#)). A 45-minute washout period was applied between injections, whereby the column was flushed by injecting 25 μL *n*-hexane:IPA (99:1 v/v).

GDGT peaks were integrated manually using Xcalibur software (version 4.1). In-house generated standard solutions were measured daily to assess system performance. One peat standard was run in a sequence for every 10 samples and integrated in the same way as sample aliquots. SIM was used to monitor abundance of the $[\text{M}+\text{H}]^+$ ion of different GDGTs instead of full scan acquisition in order to improve the signal-to-noise ratio. SIM parameters were set to detect the protonated molecules of GDGTs using m/z 1302, 1300, 1298, 1296, 1294, 1292 for isoGDGTs, and m/z 1050, 1048, 1046, 1036, 1034, 1032, 1022, 1020, 1018 for brGDGTs, after [Schoon et al. \(2013\)](#).

2.4.3 Calculation of terrestrial GDGT-based climate proxies

In Chapter 3, GDGT-based indices were calculated for comparability with those reported by [Dirghangi et al. \(2013\)](#). $R_{i/b}$ (Eq. 2.9) conveys the fractional proportion of isoGDGTs relative to brGDGTs. BIT was developed by [Hopmans et al. \(2004\)](#) as a proxy for soil input into lacustrine settings (Eq. 2.10). MBT and CBT (Eqs. 2.11-2.15) are respective indices for reconstruction of

temperature and pH after [Peterse et al. \(2012\)](#) and are refined versions of the original MBT-CBT temperature index developed by ([Weijers et al. 2007](#)).

$$R_{i/b} = \sum [isoGDGTs] / \sum [brGDGTs] \quad (2.9)$$

$$BIT = \frac{(Ia + IIa + IIIa + IIa' + IIIa')}{(Ia + IIa + IIIa + IIa' + IIIa' + cren)} \quad (2.10)$$

$$MBT = \frac{(Ia + Ib + Ic)}{(Ia + Ib + Ic) + (IIa^* + IIb^* + IIc^*) + (IIIa^* + IIIb^* + IIIc^*)} \quad (2.11)$$

$$CBT = -\log_{10} \left(\frac{(Ib + IIb + IIb')}{(Ia + IIa + IIa')} \right) \quad (2.12)$$

In Chapter 4, a more recent suite of brGDGT-based proxies was calculated following [De Jonge et al. \(2014a\)](#) (MBT'_{5ME}, MBT'_{5ME}, CBT'), [De Jonge et al. \(2014b\)](#) (Isomer Ratio; IR), and [De Jonge et al. \(2024\)](#) (degree of cyclisation, DC').

$$MBT'_{5ME} = (Ia + Ib + Ic) / (Ia + Ib + Ic + IIa + IIb + IIc + IIIa) \quad (2.13)$$

$$MBT'_{6ME} = (Ia + Ib + Ic) / (Ia + Ib + Ic + IIa' + IIb' + IIc' + IIIa') \quad (2.14)$$

$$CBT' = -\log_{10} \left(\frac{(Ic + IIa' + IIb' + IIc' + IIIa' + IIIb' + IIIc')}{(Ia + IIa + IIIa)} \right) \quad (2.15)$$

$$IR = (IIa' + IIIa') / (IIa + IIIa + IIa' + IIIa') \quad (2.16)$$

$$DC' = (Ib + IIb + IIb') / (Ia + IIa + IIa' + Ib + IIb + IIb') \quad (2.17)$$

$$(2.18)$$

*5-methy and 6-methyl counterparts

2.5 Statistical approaches

Descriptive statistics in the form of students *t*-tests and *f*-tests were used to compare mean and variance (respectively) of 3-OH FA compound and proxy distributions between the EC and SW soil transects, and are reported as a *t(df)* or *f* statistic. Non-parametric versions (Mann-Whitney U test and Fligner-Kileen test) were applied where relevant, and are reported as a *w* or *f* statistic. The Shapiro-wilk test for normality was applied as a prerequisite to all descriptive statistical analyses. Co-variance between environmental variables and influence of environmental variables on both 3-OH FA fractional abundance and lipid biomarker-derived proxies is described using linear correlation coefficients (Pearson's *r*-value) for correlations with $p \leq 0.05$. Several ordination-based multivariate analyses are also employed in this study: Non-metric multidimensional scaling (NMDS) of Bray-Curtis distance matrices is used to visualise composite 3-OH FA distributions amongst U.S. soils, and globally distributed soils, and an ANOSIM test was performed to compare

similarity of the 3-OH FA distributions using different categorical environmental grouping classifications (i.e. geographical soil transect, climate regime, vegetation type), Principal component analysis (PCA) is used to determine trends of meaningful variance within the soil 3-OH FA distributions, and RDA is used to assess which environmental drivers best describe variance in 3-OH FA-based environmental proxies. The `envfit` function of the R Vegan package was used to fit 3-OH FA compounds onto the NMDS and environmental variables onto the PCA *a posteriori* using 999 permutations. Calibration accuracy is reported using the squared Pearson's correlation coefficient and [RMSE](#). R Studio version 4.3.2 ([R Core Team 2023](#)) was used for all data analysis, and the R package Vegan ([Oksanen et al. 2022](#)) was used for ordination analysis.

**CHAPTER 3: DISTRIBUTIONS OF 3-HYDROXY FATTY ACIDS REPORTED FROM
TWO ENVIRONMENTALLY CONTRASTING SOIL TRANSECTS IN THE
USA**

3.1 Introduction

Lipid biomarkers derived from sediments and soils, notably brGDGTs and isoGDGTs, are widely used to reconstruct past climates in a range of marine and terrestrial environments (Dang et al. 2016, Schouten et al. 2013, McClymont et al. 2023, De Jonge et al. 2014a). Currently, the brGDGT-based MBT'_{5ME} index is the only widely employed organic biomarker-based temperature proxy in terrestrial environments (De Jonge et al. 2014b). However brGDGTs have a number of limitations: the MBT temperature proxy has a history of pH dependence (Weijers et al. 2007, Peterse et al. 2012), the performance of MBT'_{5ME} as a temperature proxy can be influenced by bacterial community changes (De Jonge et al. 2019), and can severely underestimate temperature in soils that form in arid environments (Dirghangi et al. 2013, Peterse et al. 2012). New proxies for the reconstruction of past environmental conditions are therefore under development from other lipid biomarker groups. Examples include RAN₁₅, RAN₁₇, and RIAN from 3-OH FAs (Wang et al. 2016), BNA₁₅ from BFAs (Huang et al. 2013, Li et al. 2022), and HG₂₈ and HG₃₀ from HGs (Bauersachs et al. 2015, Klages et al. 2020), all which are under various stages of calibration against contemporary environmental conditions. 3-OH FAs are the only of these biomarker groups for which temperature responses in terrestrial (soils and lakes) and marine environments have been quantified, (Wang et al. 2016, Yang et al. 2020) and to have globally calibrated temperature proxies (RAN₁₅, RAN₁₇, RIAN) in soils (Wang et al. 2021, Véquaud et al. 2021b). The quality of palaeoenvironmental reconstructions from any proxy is dependent upon the contemporary environmental variability from which they were calibrated - the environmental constraints on a temperature proxy are often not entirely limited to temperature. As terrestrial environments are heterogeneous, the biological producers of lipid biomarkers are capable of having complex physiological adaptations to environmental forcing (Kaneda 1991, Siliakus et al. 2017), and changes in bacterial community structure which also has the potential to confound proxy-environment relationships (De Jonge et al. 2019). Consequently, it is necessary to

investigate the environmental variability that has potential to confound a proxy reconstruction, and develop new proxies or revise existing calibrations accordingly.

C₁₀ to C₁₈ 3-OH FAs are a group of lipids produced almost exclusively in the outer cell membranes of Gram-negative bacteria, and are ubiquitously produced in nature and preserved in terrestrial and marine sediments. 3-OH FA-based temperature proxy calibrations have been developed in a range of sediments at a range of scales (regional to global) (Wang et al. 2016, Huguet et al. 2019, Wang et al. 2021, Véquaud et al. 2021b, Yang et al. 2020, 2021, Dong et al. 2023), however soil-derived 3-OH FAs have been investigated most extensively for proxy development, and comprise a collection of 280 soils from four regional montane transects (Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021a) and two global calibrations (Véquaud et al. 2021b, Wang et al. 2021). From pedogenic environments, RAN₁₅ and RAN₁₇ have consistently conveyed variability in response to temperature, and RIAN with soil pH (Wang et al. 2016, Huguet et al. 2019, Wang et al. 2021). The RAN proxies describe a relative increase in *n*-C₁₅ & 17 relative to their *anteiso*- counterparts at higher MAAT which is regarded as a thermally mitigated homeoviscous adaptation response in the cell membrane lipid bilayer of Gram-negative bacteria (Kaneda 1991). The RIAN proxy employs the whole suite of C₁₀ to C₁₈ 3-OH FAs and is the negative logarithm of the branching ratio (BR), and generally describes a relative increase in branched (*iso*- + *anteiso*-) relative to *normal*-3-OH FAs with increasing (more alkaline) soil pH (Wang et al. 2016).

Although these 3-OH FA proxies have strong linear relationships with contemporary MAAT and pH across some mountain transects (Table 1.1) (Wang et al. 2016, Véquaud et al. 2021b) and moderate relationships globally (RAN₁₅ $R^2=0.48$, RAN₁₇ $R^2=0.41$, RIAN $R^2=0.66$) (Wang et al. 2021), the response of each proxy to environmental conditions has been extremely variable between soil transects (Huguet et al. 2019, Véquaud et al. 2021a,b). Scatter (particularly observed by Véquaud et al. (2021b) in a compilation of mountain soil transects) is a profound feature of all

proposed proxy-environment relationships for 3-OH FAs. A RAN₁₅ calibration from Mt. Shennongjia (China), was able to reconstruct a 9000 yr temperature record in a stalagmite from Heshang Cave (China) (Wang et al. 2018). As other organic biomarkers such as brGDGTs are preserved in sediments on multi-millennial time scales (De Jonge et al. 2021) and have been applied in palaeoclimate reconstructions up to Oligocene-Miocene (DeCelles et al. 2018), Eocene (Inglis et al. 2017, Naafs et al. 2018), and Cretaceous (Super et al. 2018), there is potential for 3-OH FAs to be applied as a biomarker on long-term geological timescales although their preservation beyond the chronology of the Heshang Cave stalagmite has not been investigated to date. However the applicability of regional 3-OH FA calibrations outside of their calibration transect is uncertain.

Polynomial temperature and pH calibrations have been explored as an alternative to linear regression with the aim of improving the reconstructive potential of environmental calibrations based on 3-OH FAs. Methods considered include MLR (Véquaud et al. 2021b) and machine learning-assisted models such as Gaussian process emulator (Wang et al. 2021) and random forest (Véquaud et al. 2021b). Despite producing higher correlation coefficients and lower RMSEs (Table 1.1) (Wang et al. 2021), the linear RAN₁₅ regression from Mt. Shennongjia was better at reconstructing temperature in the Heshang Cave stalagmite record than polynomial calibrations (Véquaud et al. 2021b). The development of regional 3-OH FA proxy calibrations has therefore been recommended to account to regional variability of RAN₁₅, RAN₁₇, and RIAN (Wang et al. 2021, Véquaud et al. 2021b). Regional calibrations are useful for application in unique environments where conditions diverge from normal global ranges (i.e. aridity, sodic, saline), whereas global calibrations are useful to incorporate diverse environmental variability into one calibration and apply one calibration in different regions, and so both calibration types have applicability, and a variety of calibration scales are often considered as a pre-requisite to palaeoclimate reconstruction.

However, regional transects used for calibration development are often characterised by geographic area, whereas in nature different regional transects can have similar environmental characteristics which are transient between regions; the definition of a region and therefore applicability of a regional proxy calibration is unknown. Furthermore, environmental variables other than temperature and pH that might cause scatter in RAN_{15} and RAN_{17} -temperature relationships have not been investigated in detail. The brGDGT-based MBT'_{5ME} temperature proxy has had a long history of pH dependence (Weijers et al. 2007, Peterse et al. 2012, De Jonge et al. 2014a), and pH-induced shifts in bacterial community structure are capable of influencing MBT'_{5ME} performance as a temperature proxy (De Jonge et al. 2019). SMO and MAP are known to confound the performance of MBT, MBT' and MBT'_{5ME} in arid environments where low moisture becomes an environmental stress for bacterial growth (Dirghangi et al. 2013, Xie et al. 2012, Dang et al. 2016). As 3-OH FAs and brGDGTs share at least one bacterial producer in Acidobacteria, the environmental variables confounding MBT'_{5ME} temperature reconstructions could also influence 3-OH FAs. Additionally, pedoclimate is more heterogeneous than the atmospheric annual-averaged environmental parameters often assessed in calibrations such as MAAT and MAP, and T_{Soil} is more likely to be the best representation of bacterial growth temperature. Direct measurements of soil temperature are limited and unconventional, and it has become standard procedure to use MAAT as an analogue (Weijers et al. 2007, Peterse et al. 2012). However, in recent years high resolution satellite-derived GIS T_{Soil} data has become available enabling the role of T_{Soil} relative to MAAT to be investigated in pedogenic temperature proxy calibrations.

There is a complex array of environmental conditions that have the potential to confound 3-OH FA proxy relationships with temperature and pH, which are known to limit applicability of brGDGT proxies, but have not been explored for 3-OH FAs. Additionally the environmental constraints of regional calibrations remain unexplored. We therefore investigate 3-OH FA variability within and between soil transects from the U.S. East Coast (EC) and SW U.S. (SW) -

two environmentally contrasting soil transects in pH, aridity, and **MAAT** - to assess the influence of **MAP**, **SMO**, soil pH, T_{Soil} , and **MAAT** seasonality on 3-OH FA-based temperature and pH proxies and calibrations to determine if confounding scatter could be refined from regional or global 3-OH FA based proxy regressions to improve calibrations.

3.2 Study area

62 soil samples were collected from two U.S. environmental transects; 36 along an EC latitudinal transect spanning Maine to Florida, and 26 from a SW U.S. longitudinal transect between Arizona and Texas (Fig. 3.1). The EC transect follows the soil transect of [Dirghangi et al. \(2013\)](#) who previously investigated GDGT distributions in the EC soils allowing for comparison of 3-OH FA and brGDGT data. This study expands the EC transect of [Dirghangi et al. \(2013\)](#) by including 5 soils from Florida to increase the maximum temperature range of the soil transect to ca. 23 °C. The SW transect comprises lake catchment soils following the transect of [Hou et al. \(2008\)](#) who investigated δD from *n*-alkanoic acids, which is useful for comparison with future 3-OH FA-based isotopic datasets, as 3-OH FAs are amenable to C and H isotope analysis. Both transects are subjected to a continental climate; the EC receives moisture from the **NAO** and westerly trade winds from the continental interior and is fairly flat (-4 to 249 m asl), whilst the SW is generally arid and has basin and range topography which denotes high spatial variability in elevation (35 to 2758 m asl) ([Hou et al. 2008](#)).

MAAT ranges between 6.1 and 23.0 °C (mean = 15.5 ± 4.77 °C (*SD*)) amongst the compiled soils, which is 6.5-23.0 °C (mean = 14.1 ± 4.9 °C (*SD*)) for EC soils, and 6.1-21.9 °C (mean = 17.4 ± 3.9 °C (*SD*)) for SW soils (Fig. 3.1A). Seasonal temperature for the soils is also highly variable with $\Delta T_{\text{DJF-JJA}}$ ranging between 8.0 and 19.6 °C. In this study, high values for seasonality are indicative of a larger offset between **mean summer air temperature (months of June, July, August)** (T_{JJA}) and **mean winter air temperature (months of December, January, February)** (T_{DJF}).

These high values tend to occur in warmer regions of the study areas such as Arizona and Florida. The soils are subject to **MAP** of 245-1474 mm (mean = 930 ± 389 mm (*SD*)) (Fig. 3.1B), however precipitation is variable between the transects as **MAP** is 1087-1357 mm (mean = 1214 ± 67 mm (*SD*)) for EC soils, and 245-1474 mm (mean = 537 ± 294 mm (*SD*)) for SW soils. SW soils are therefore generally arid, and EC soils are humid with little variability, however the SW transect covers a wider **MAP** range where most soils experience an arid to semi-arid climate (**MAP**<500 mm), and a minority are subject to humid conditions (Fig. 3.1B). **VPD** (known to significantly co-vary with % humidity) ranges between 4.7-23.0 % (mean = 10.5 ± 4.70 % (*SD*)), which is also highly variable between EC (mean = 7.55 ± 1.82 % (*SD*)) and SW (mean = 14.7 ± 4.30 % (*SD*)) soils. % **SMO** for compiled soils is 0.20-53.5 % (mean = 14.9 ± 12.3 % (*SD*)), which is extremely variable between the EC (mean = 19.9 ± 12.9 % (*SD*)) and SW (mean = 8.00 ± 7.32 % (*SD*)) soils. EC soils collected from latitudes below ca.39°N (south of Virginia) have comparable **SMO** to SW soils, however **SMO** is extremely variable amongst EC soils and is exceptionally high (up to 50 %) in NE EC soils. The climatic circumstances likely responsible for this are detailed by [Hardman \(2019\)](#). The lowest **SMO** values are observed in approximately 1/3 of SW soils generally from arid sites in New Mexico and Arizona.

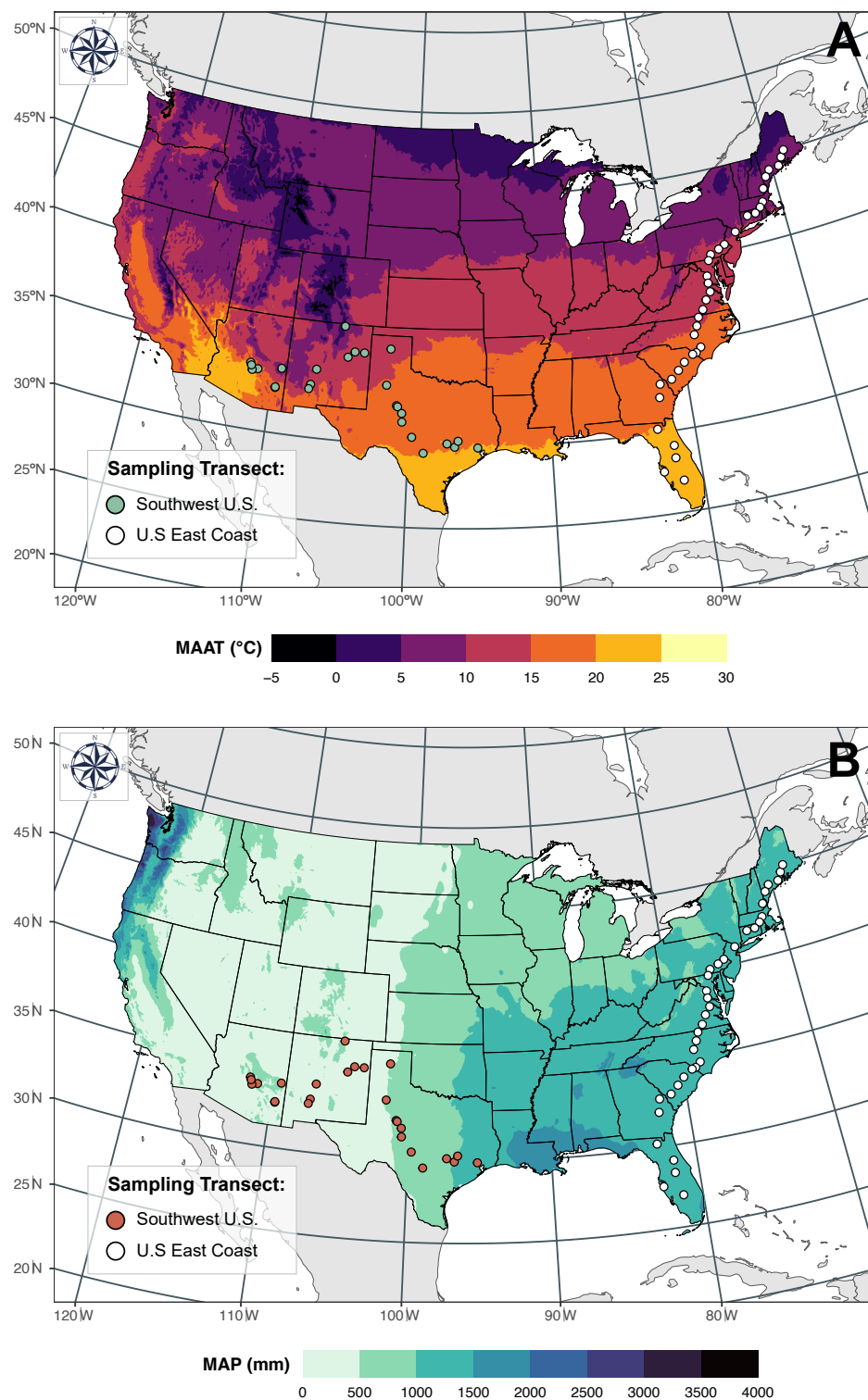


Figure 3.1: Maps indicating location of EC and SW sampling sites. Samples from EC comprise soils, samples from SW are co-sampled lake sediments and catchment soils. **A** conveys **MAAT** variability across the study sites, and **B** conveys **MAP**.

Along the EC, Köppen climate regimes transition from snow in the NE through to warm temperate, and equatorial in Florida (Beck et al. 2023) which primarily support coniferous to deciduous woodland vegetation, whilst wetland vegetation of the Everglades is supported by an equatorial climate. The SW comprises warm temperate, arid, and snow climate regimes (Beck et al. 2023) which support grassland plains across eastern Texas, progressing to shrubsteppe and desert scrub with increasing aridity towards the West. Vegetation across the U.S transects is supported by an array of soil types; podzols (sandy soils leached of organic matter and nutrients (Sauer et al. 2007)) and acrisols are predominant along the EC and associated with humid climate regimes, particularly podzols with boreal vegetation. Luvisols, kastanozems, and yermosols are predominant across the SW and associated with humid (luvisols), semiarid (kastanozems), and arid (yermosols) climate regimes. Additionally, pH for the compiled soils comprises a mixture of acidic and alkaline soils 3.56-8.77 (mean = 5.87 ± 1.58 (*SD*)). The acidic soils are generally associated with the EC (mean = 4.78 ± 0.83 (*SD*)) and more neutral to alkaline soils are associated with the SW (mean = 7.34 ± 1.07 (*SD*)). The environmental conditions of the U.S. therefore reveal a humid environment along the EC with acidic soils and low VPD and a generally semi-arid to arid environment across the SW where the majority of soils are neutral to alkaline in nature and VPD is high. Vegetation and soil type appear to be climatically influenced across the U.S. study area.

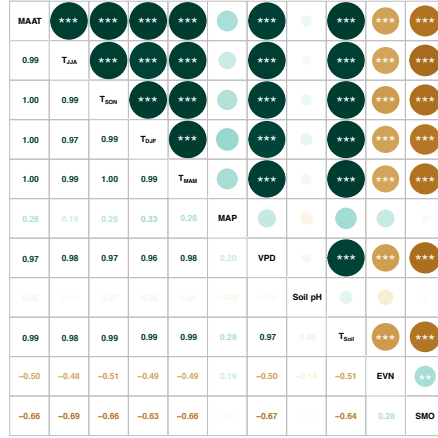
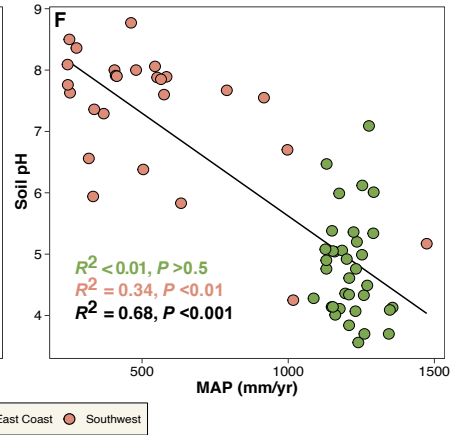
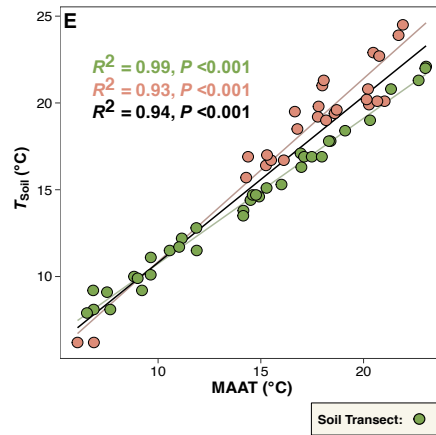
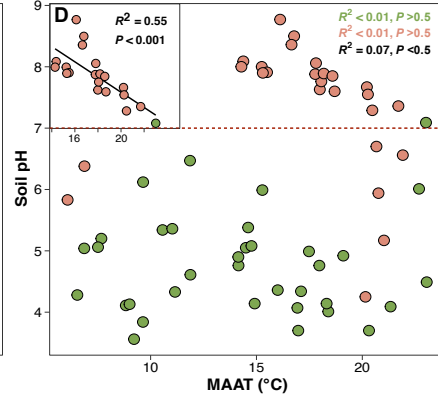
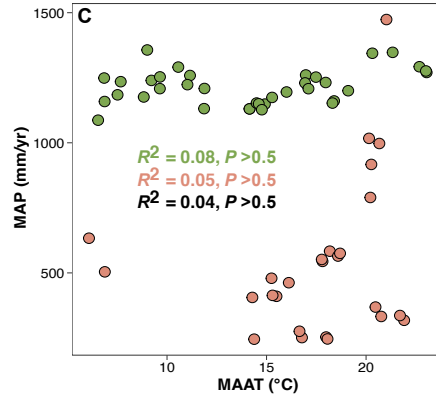
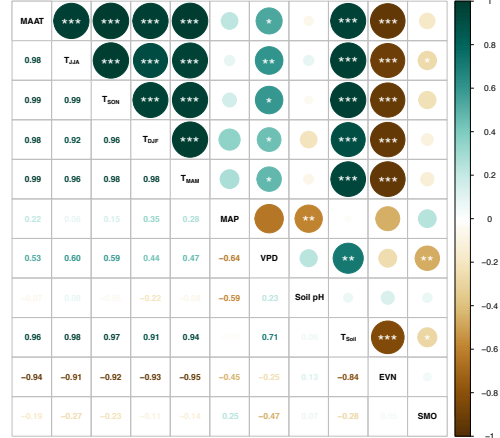
3.3 Results

3.3.1 Environmental diversity across the East Coast and Southwestern U.S.

The EC and SW have contrasting environments, with substantial variability in climate regime. All measured environmental variables significantly differ between the transects, most predominantly for moisture-related variables, T_{JJA} , and T_{Soil} ($p < 0.001$) (Appendix A1). Significantly higher ($p < 0.001$) and more variable ($p < 0.05$) VPD in the SW (Appendix A1) is consistent with the low MAP observed here, in addition to significantly lower MAP, and higher T_{Soil} , pH, and air

temperatures which cumulatively indicate the SW is a hotter, arid, and more evaporative environment relative to the humid EC that is subject to oceanic moisture and cold continental winds.

Pearson's correlation matrices performed separately upon environmental variables reported from the EC and SW (Fig. 3.2) are used to assess how the variables correspond to each other along each transect, and whether this response is similar between transects. As expected, MAAT and seasonal air temperatures are highly correlated along each transect, despite both having strong variability in seasonal temperature regime between sample sites (Fig. 3.2A & B). T_{Soil} , too, is highly correlated with air temperatures along both transects, although ΔT ($T_{\text{Soil}} - \text{MAAT}$) is generally higher across the SW compared to the EC (Fig. 3.2E). There is no significant covariance between MAAT and MAP along either transect (Fig. 3.2C) despite positive r values indicating MAP generally increases with MAAT. As individual transects and combined, soil pH has no significant relationship with MAAT, however a crossplot between these variables reveals a potential pH threshold for covariability with MAAT (Fig. 3.2D); alkaline soils (pH>7) almost exclusively originating from the SW negatively covary with MAAT to a highly significant degree ($r^2=0.55$, $p<0.001$), whilst no such trend is observed amongst acidic soils (pH<7) for either transect, indicating pH 7 may be an important threshold for pH and MAAT covariability.

A U.S. East Coast**B SW U.S.**

Soil Transect: ● East Coast ● Southwest

Figure 3.2: Comparison of environmental variables along EC and SW U.S. soil transects, comprising: Pearson's correlation matrices for **A** EC soils, **B** SW soils, and crossplots between **C** MAAT and MAP, **D** MAAT and pH, **E** MAAT and T_{Soil} , **F** MAP and pH. Circles in upper half of matrix relate to relative magnitude of Pearson's R . Green circles indicate positive correlations, orange circles indicate negative correlations, and significant correlations are indicated by asterisks (p : ≤ 0.5 = *, ≤ 0.1 = **, ≤ 0.001 = ***). Values in lower half of matrix report Pearson's R itself.

Notable differences between the two transects do however arise amongst the moisture-related environmental variables. Along the EC, soil moisture correlates with all environmental variables, excluding **MAP** and soil pH, to a highly significant degree, only retaining significant relationships with **VPD**, T_{JJA} , and T_{Soil} across the SW (Fig. 3.2A & B). The strong negative relationship between **VPD** and **SMO** across both transects is not surprising and is widely documented as a general representation of evaporative moisture availability; as **SMO** decreases, **VPD** has a tendency to increase alongside relative humidity (Dubey & Ghosh 2023). The significance of T_{JJA} suggests the pedospheric to atmospheric moisture balance may be more seasonally mitigated in the SW compared to the EC. Furthermore, a significant negative relationship is observed between SW soil pH and **MAP**, which does not occur amongst EC soils, a relationship that becomes stronger when both soil transects are compiled and soils become more acidic with increasing **MAP** ($r^2=0.68$, $p<0.001$) (Fig. 3.2F), a known phenomenon induced by the leaching of base cations (Ca^{2+} and Mg^{2+}) from soils by percolating rainfall, rendering soils more acidic in humid environments with higher **MAP** (Fenn et al. 2006). Overall, environmental variables across the EC and SW covary in ways which are and are not comparable. The physical and hydroclimatic variables of the SW are likely more seasonally influenced than along the EC, however combining the soil transects does not appear to hide any overarching trends in environmental covariance.

3.3.2 Organic geochemistry

3.3.2.1 3-OH FA distributions in East Coast and SW U.S. soils

A broad range of 3-OH FAs between C_{10} and C_{18} were detected in the EC and SW soils (Fig. 3.3A). In most soils, $n-C_{14}$ (EC: mean= 26.1 ± 5.6 % (*SD*); SW: mean= 21.6 ± 6.2 % (*SD*)) is of highest abundance, with $n-C_{18}$, $n-C_{12}$, and $n-C_{16}$ being most abundant in some samples (Fig. 3.3A); this variability in dominant 3-OH FA compound has been observed amongst globally distributed soils (Wang et al. 2021). $i-C_{17}$ is the dominant 3-OH FA compound in some SW soils (EC: mean= 6.2 ± 1.6 % (*SD*); SW: mean= 14.5 ± 4.6 % (*SD*); Fig. 3.3 A). 3-OH FA distributions

across both soil transects follow a similar structural breakdown of fractional abundance in which a small number of compounds (generally *normal*- structures with even C numbers) occupy higher fractional abundances (>10 %) on average (Fig. 3.3 A), and the majority of 3-OH FA compounds are present in fractional abundances <5 % total 3-OH FAs (Fig. 3.3 A). *iso*-3-OH FAs with odd C numbers are the second most abundant structural group for both transects, and *iso*-3-OH FAs with even C numbers and *normal*-3-OH FAs with odd C numbers are produced in almost identical proportions (Fig. 3.3 C). *anteiso*-3-OH FAs are only detected amongst odd C numbers which is likely related to the biochemistry of Gram-negative bacterial 3-OH FA synthesis.

Although 3-OH FA distributions from the EC and SW soil transects follow a similar overall structural composition, some 3-OH FA compounds are consistently produced at different fractional abundances or with different variability, leading to a significant difference in composite 3-OH FA distributions between the soil transects. *normal*-3-OH FAs are significantly more abundant amongst EC soils compared to those in SW soils ($p<0.001$), and *branched* 3-OH FAs are more abundant amongst SW soils in respect to their EC counterparts ($p<0.001$) (Fig. 3.3B). The enrichment of branched 3-OH FAs in SW soils is the direct result of *iso*-3-OH FAs between the transects ($p<0.001$), not *anteiso*-3-OH FAs which have no significant difference in fractional abundance ($p=0.17$) (Fig. 3.3B). All structural group statistics are reported in Appendix A2. *i*-C₁₇ ($p<0.001$) and *i*-C₁₃ ($p<0.001$) are the predominant compounds promoting *iso*- enrichment in SW soils (Fig. 3.3A), which is consistent with our observations of *i*-C₁₇ as the dominant 3-OH FA compound in several SW soils. *n*-C₁₄ ($p<0.01$), *n*-C₁₈ ($p<0.05$), and *n*-C₁₂ ($p<0.05$) are, in turn, the primary drivers of *normal*-3-OH FA enrichment in EC soils. All 3-OH FA comparison statistics are reported in Appendix A3 Therefore, *anteiso*-3-OH FAs are comparable between the soil transects, however they differ in proportion of *normal*- and *iso*-3-OH FAs; *i*-C₁₇ enrichment in SW soils is particularly notable as this has generally been observed in low MAP soils.

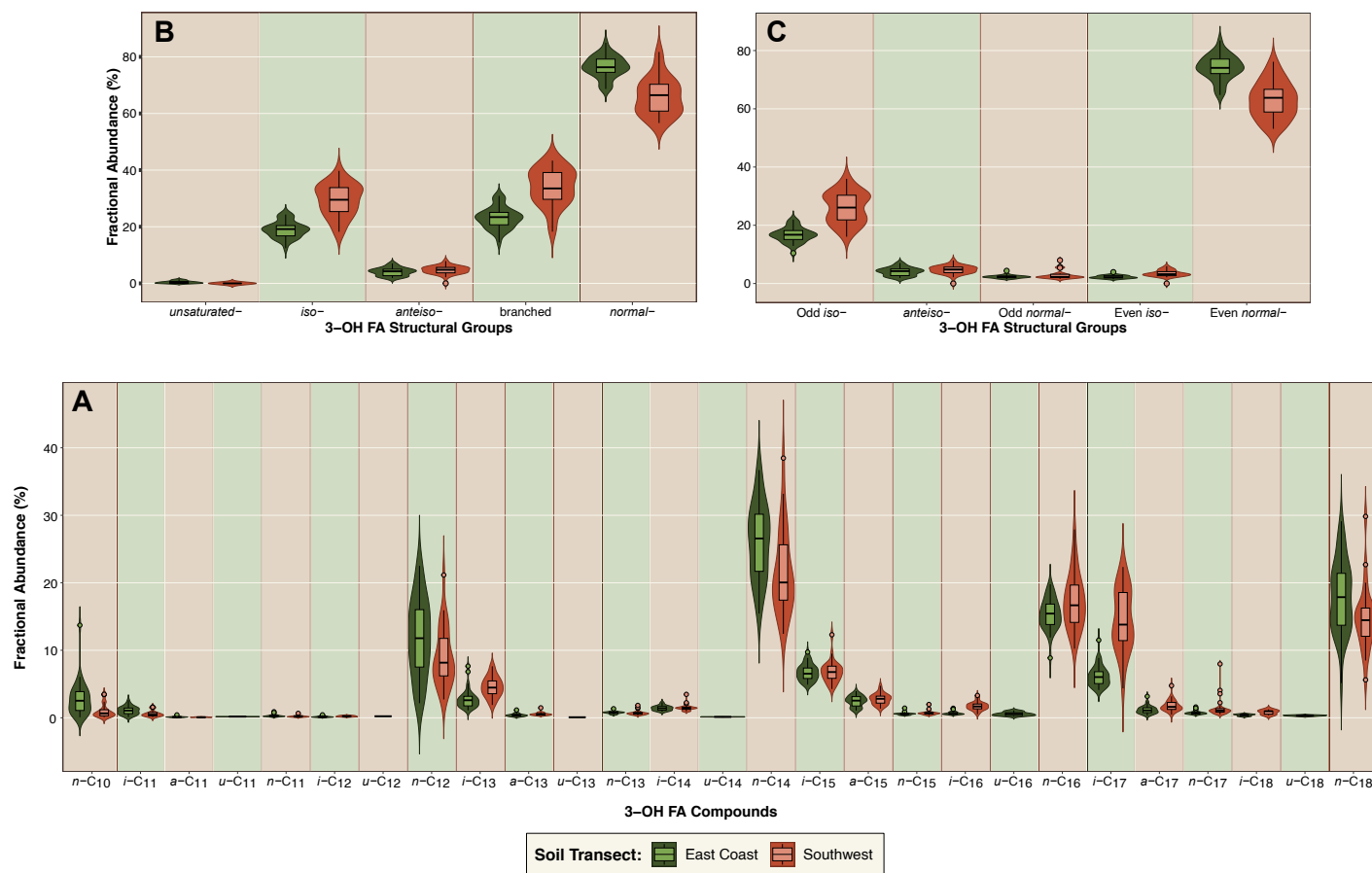


Figure 3.3: 3-OH FA distributions in EC and SW U.S. soils reported as % fractional abundance of **A** individual 3-OH FA compounds, and **B** and **C** various 3-OH FA structural groups

NMDS visualisation of a Bray-Curtis distance matrix of 3-OH FA fractional compositions in U.S. soils produced two mostly distinct confidence ellipses between the EC and SW transects that partially overlap (Fig. 3.4). Although the 3-OH FA distributions are significantly differentiated by soil transect as a grouping variable as indicated by an ANOSIM test ($r^2=0.45$, $p<0.001$), soils USA002 (Maine), USA024 (North Carolina), USA037 and USA040 (Florida), USA043 (Mary's Lake, NM), USA047 (Dankworth Pond, AZ), and USA066 (Lake Sommerville, TX) fall within the overlapped area of the confidence ellipses for each transect (Fig. 3.4), suggesting that these sites have similar 3-OH FA distributions that cannot be distinguished by transect. Conversely, soils from sites USA016 (Pennsylvania), USA031 (South Carolina), USA048 (Roper Lake, AZ), USA051 (Horseshoe Lake, AZ), USA053 (Crescent Lake, AZ), and USA062 (Lake Junction, TX) fall outside the confidence ellipses for either soil transect; Soil from the catchment of Lake Junction (USA062) has the most unique 3-OH FA distribution of all soils, being uniquely enriched in n -C₁₇ (Fig. 3.4), and upon further inspection has a limited composition of 3-OH FAs attributed to the absence of short chain (C₁₀ and C₁₁) and C₁₅ 3-OH FAs.

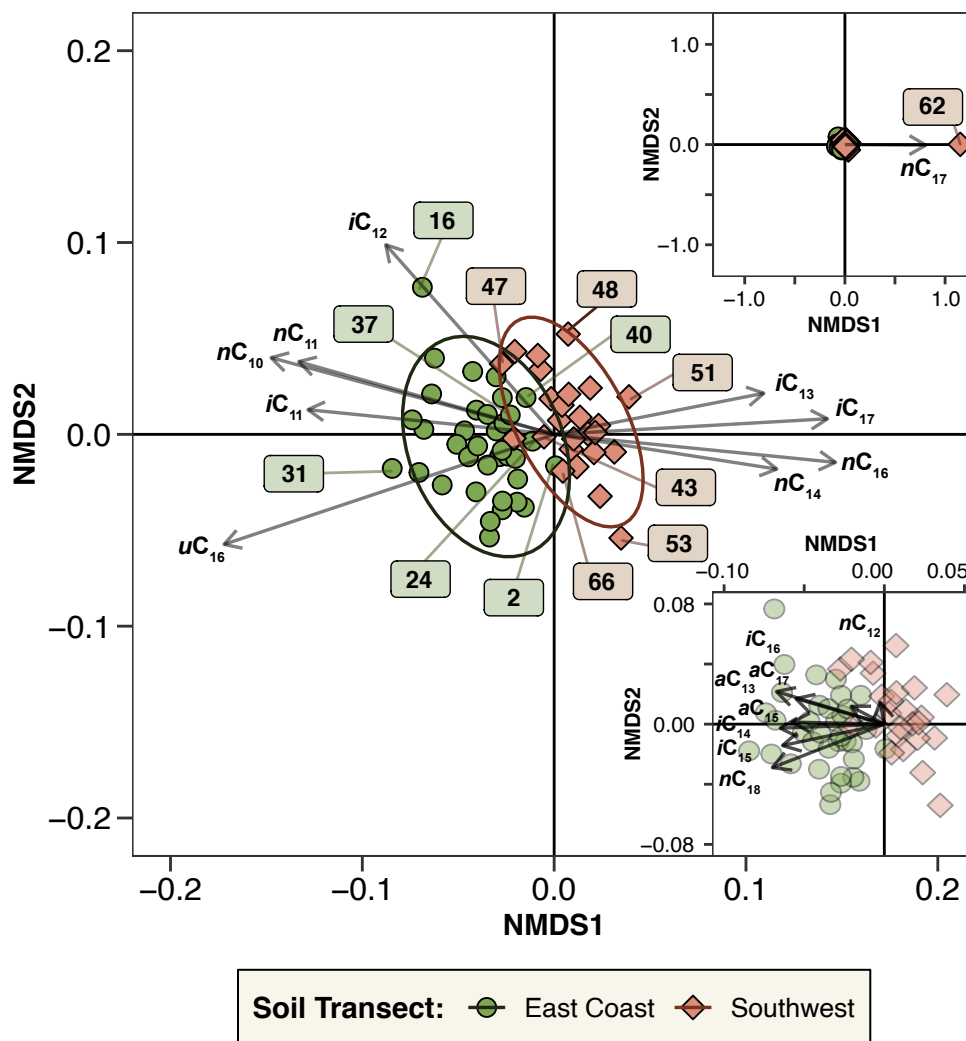


Figure 3.4: Non-metric multidimensional scaling (NMDS) plot of composite 3-OH FA distributions for EC and SW U.S. soils with NMDS1 axis limit set between -0.2 and 0.2 (centre). 3-OH FA compounds with significant variability amongst the soils are visualised, and numbered labels indicate the sample name of unique soils. **Upper right:** zoomed out plot showing full scale NMDS1 and unique 3-OH FA distribution of site USA 062. **Lower right:** zoomed in plot highlighting the significantly variable 3-OH FA compounds that are less influential than those visualised in central plot. Associated statistics are reported in Appendix A4.

3.3.2.2 GDGT distributions in SW U.S. soils

GDGT distributions in U.S. EC soils have been reported by [Dirghangi et al. \(2013\)](#) as % isoGDGTs, % brGDGTs, $R_{i/b}$, BIT, MBT, and CBT, and were not reanalysed along the EC for the present study. GDGTs were extracted and analysed from the SW soils for which there is no preexisting data. For the SW soils, a summary of GDGT data is as follows:

A range of br- and isoGDGTs were detected in the SW soils. The proportion of brGDGTs ranged from 24.4 to 100 % total GDGTs (mean = 63.3 ± 21.7 % (*SD*)), and isoGDGTs from 0 to 75.6 % (mean = 36.7 ± 21.7 % (*SD*)), brGDGTs are therefore generally more abundant in the soils than isoGDGTs. Of the brGDGT distributions, brGDGT-IIa is predominant (41.5 % total mean brGDGT composition), leading to the dominance of compounds with no cyclopentyl moieties (denoted with 'a'; 82.7 % total brGDGTs) together with brGDGTs-Ia (24.8 %) and IIIa (12.8 %). brGDGTs with one ('b') or two ('c') cyclopentyl moieties are far less abundant, particularly 'c' compounds (0.87 % total brGDGT composition) whereby brGDGTs-IIIc and IIIc' are below detection limit in all SW soils. Compounds with one methyl branch (denoted as 'II') are most abundant when brGDGT distributions are classed by degree of methylation, attributed to brGDGT-IIa and the less abundant IIb (8.9 %). 5-methyl and 6-methyl GDGTs were separated following the analytical method of [De Jonge et al. \(2014b\)](#), which was developed after the publication of GDGT data from EC soils ([Dirghangi et al. 2013](#)). SW soil brGDGTs are predominantly composed of 5-methyl compounds (96.3 % total brGDGTs), in comparison to 6-methyl compounds (3.7 %).

3.3.3 3-OH FA and brGDGT-derived palaeoclimate indices

3.3.3.1 3-OH FAs

Established 3-OH FA-based proxies for temperature (RAN_{15} , RAN_{17} , RAN_{13} , RAN_S , RIN_{17} , RIN) and pH (BR, RIAN) were calculated for the EC and SW soil transects and are reported in

appendices [E1](#) and [E2](#). All 3-OH FA-based proxies produced a range of values across each soil transect (Appendices [E1](#) and [E2](#)), which indicates that all proxies are responsive to either environmental or biological forcing across each transect. Between the two soil transects, RAN_{15} ($t=-0.23(54)$, $p>0.05$) and RAN_{17} ($w=391$, $p>0.05$) are the only proxies that have no significant difference in mean distribution. Alongside equal variance for both proxies ($RAN_{15}:f=0.48, p>0.05$; $RAN_{17}:f=0.66, p>0.05$), it can be concluded there is no significant difference in distribution of these proxies between the soil transects (Fig. [3.5](#)). Values are significantly higher in SW soils relative to EC soils for all remaining proxies, excluding RIAN which - as the negative logarithm of BR - mirrors BR in reverse and therefore generates lower proxy values in SW soils (Fig. [3.5](#)). RIAN also has the most substantial difference in values between the EC and SW soils ($t=7.51(42)$, $p<0.001$). Proxies that significantly differ in mean value between the soil transects do not necessarily significantly differ in variance; RAN_5 has no significant difference in variability ($f=0.82, p>0.05$) between the soil transects despite values being enriched in SW soils ($t=-3.06(59)$, $p<0.01$), and the same is observed for the variance ($f=1.38, p>0.05$) and median ($w=154, p<0.001$) of RAN_{13} values, which indicates that these proxies are more variable between the soil transects than within each soil transect.

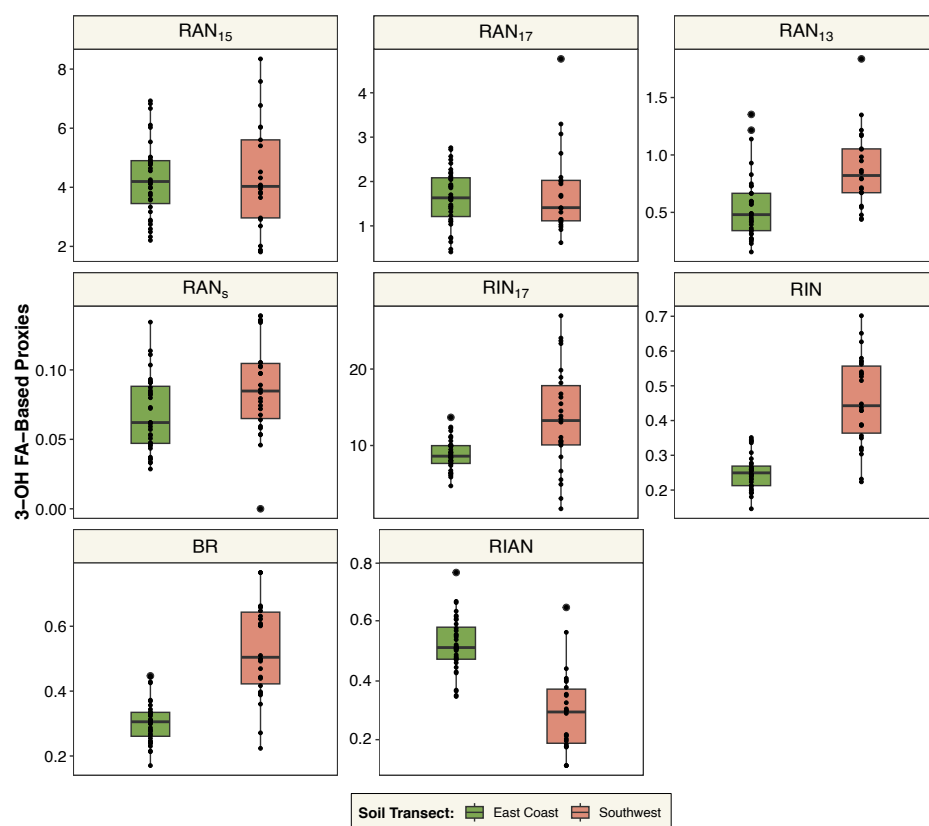


Figure 3.5: Boxplots visualising 3-OH FA-based proxy distributions within and between EC and SW soil transects

It is also notable that some soils did not produce RAN_{15} , RAN_{17} , RAN_{13} , and RAN_s values, but all soils produced values for more cumulative proxies such as RIN, BR, and RIAN which employ a wide range of 3-OH FAs in their ratios. Although only a very small number of soils could not produce the specified RAN proxies and our onward capability to use either soil transect to explore the environmental potential of these proxies in further detail is not affected, such an observation suggests that proxies which are dependent on a cumulative sum of different 3-OH FA compounds are potentially more reliable than proxies which depend on a specific 3-OH FA compound as a numerator and denominator. Proxies that conform to the latter are limited by their dependence on the preservation of specific 3-OH FA compounds, which may be environmentally

mediated as more soils could not produce proxy values in SW soils than EC soils.

3.3.3.2 brGDGTs

The brGDGT-based MBT and CBT proxies for the respective reconstruction of temperature and pH are calculated from the GDGT distributions of our SW soils in addition to the $R_{i/b}$ and BIT indices (Fig. 3.6), which are more descriptive of the GDGT distributions themselves; $R_{i/b}$ is a ratio of archaeal-derived isoprenoidal GDGTs to bacterially-derived branched GDGTs, whilst BIT is a measure of crenarchaeol abundance in a sample. The MBT'_{5ME} and CBT' proxies are revised versions of MBT and CBT that are widely employed in current palaeoclimate research, accounting for the separation of 5-methyl and 6-methyl brGDGTs using the analytical method of [De Jonge et al. \(2014b\)](#). However, the original MBT and CBT proxies developed by [Weijers et al. \(2007\)](#) are specifically calculated for direct comparability to the EC soil GDGT data of [Dirghangi et al. \(2013\)](#) in which 5-methyl and 6-methyl brGDGTs are not separated.

We previously reported brGDGTs are generally more enriched in SW soils relative to isoGDGTs, which is further supported as $R_{i/b}$ values range between 0 and 3.1 (mean = 0.81 ± 0.76 (*SD*)) and BIT values between 0.3 and 1.0 (mean = 0.69 ± 0.19 (*SD*)), which suggests that although brGDGTs preside over isoGDGTs in most soils ($R_{i/b} < 1$), they are not completely dominant ($R_{i/b} < 0.5$; ([Naafs et al. 2017b](#))). Soils from the catchments of Roper Lake, AZ, and Lake Houston, TX, are highly enriched in isoGDGTs ($R_{i/b} > 2$). When compared with the respective indices reported by [Dirghangi et al. \(2013\)](#) from EC soils, the two soil transects have completely distinct $R_{i/b}$ and BIT distributions that differ in both mean and variance (Fig. 3.6); $R_{i/b}$ values are higher ($w=34.5$, $p<0.001$) and more variable ($f=34.76$, $p<0.001$) in SW soils whereas values in EC soils are substantially lower and have restricted variability (mean = 0.008, *SD* = 0.01). EC soil GDGT distributions are therefore almost completely dominated by bacterial-derived brGDGTs, whilst SW soils have more of a mixed archaeal and bacterial-derived GDGT composition of variable proportion. The BIT index supports the $R_{i/b}$ observations; EC soils follow a somewhat

quintessential soil GDGT distribution, whereby BIT values are concentrated close to a value of 1 (mean=0.998 $\pm$ 0.007 (*SD*)) - the maximum limit for BIT - which has been shown as generally characteristic for soils globally (Hopmans et al. 2004). Lower ($w=780$, $p<0.001$) and highly variable ($f=37.70$, $p<0.001$) BIT values for SW soils indicates crenarchaeol has a substantially higher abundance in SW soils than in EC soils, and is extremely variable intrinsically across the SW soil transect.

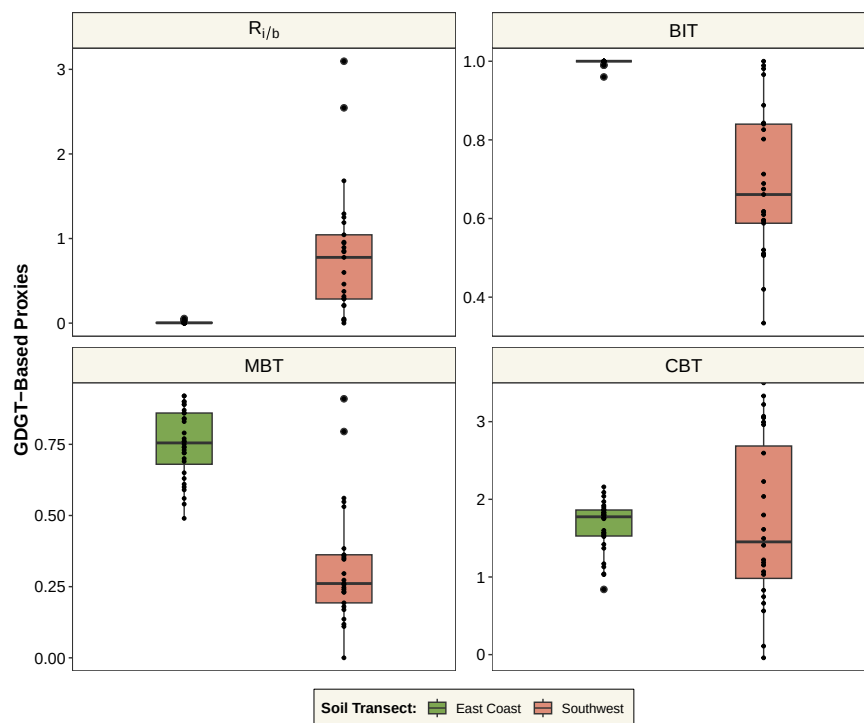


Figure 3.6: Boxplots visualising GDGT-based proxy distributions within and between EC and SW soil transects. Data for the EC soils is from Dirghangi et al. (2013), data from the SW soils is novel from this study.

The MBT and CBT proxies also produce different distributions between the EC and SW soil transects (Fig. 3.6); MBT values range between 0.49 and 0.92 (mean=0.75 $\pm$ 0.12 (*SD*)) for EC soils, and between 0.11 and 0.91 (mean=0.34 $\pm$ 0.20 (*SD*)) for SW soils, two distributions which

differ significantly in average ($w=712$, $p<0.001$) but retain equal variance ($f=0.38$, $p>0.05$). CBT values range between 0.84 and 2.16 (mean= 1.66 ± 0.33 (SD)) for EC soils, and between -0.04 and 3.33 (mean= 1.68 ± 1.03 (SD)) for SW soils, significantly differing in variance ($f=18.26$, $p<0.001$) but not mean ($w=421$, $p>0.05$). The two proxies thus have opposed responses; CBT values are more consistent between the transects, but have a higher degree of intrinsic variability between SW soils, and MBT retains intrinsic variability between transects (Fig. 3.6).

3.3.4 Application of linear and multivariate analyses to explore relationships between environmental variables and bacterial lipids

3.3.4.1 Response of 3-OH FA compounds to environmental variability and identification of primary environmental drivers

3-OH FA fractional abundance for the U.S. soils were subjected to Pearson's correlation analysis in conjunction with environmental variables. The matrix reveals that almost entirely different subsets of 3-OH FA compounds are capable of having a direct response to temperature and soil pH. n -C₁₀, a -C₁₅, i -C₁₄, n -C₁₆ and a -C₁₇ have significant variation in fractional abundance in response to MAAT (Fig. 3.7); n -C₁₆ is the only of these compounds to have a positive relationship with temperature. In turn, i -C₁₃, n -C₁₃, n -C₁₄, a -C₁₅, i -C₁₆, i -C₁₇, a -C₁₇, n -C₁₇, and n -C₁₈ have significant variation in response to soil pH (Fig. 3.7), whereby n -C₁₄ and n -C₁₈ are the only compounds for which this relationship is negative. a -C₁₅ and a -C₁₇ are the only mutual compounds between these 3-OH FA subsets, and both have a negative relationship with MAAT and a positive one with soil pH. Furthermore, a -C₁₅ has a stronger association with MAAT, and a -C₁₇ with soil pH. n -C₁₀ is responsive to all environmental variables except for pH. a -C₁₅ has highly significant correlations with all seasonal temperatures, but a -C₁₇ only has a significant relationship with T_{DJF} and T_{MAM} suggesting production of a -C₁₇ may be more seasonal and therefore it does not have as strong a relationship with MAAT as a -C₁₅. As a result, the cumulative fractional abundance of *anteiso*-3-OH FAs has a significant response to both MAAT and pH, most significantly for T_{DJF}

and T_{MAM} , whilst all structural groups (*iso*-, *anteiso*- and *normal*-) significantly respond to pH.

When Pearson's correlation analysis was applied to the EC and SW soils as separate data pools, 3-OH FAs between each transect had profoundly different variability in response to the same environmental variables (Appendix A5). $n\text{-C}_{14}$ and $n\text{-C}_{16}$ have a significant response to $MAAT$ in EC soils which is masked in the composite Pearson's matrix (Fig. 3.7), whereas $a\text{-C}_{15}$ is the only compound to significantly respond to $MAAT$ in SW soils. Furthermore, the cumulative fractional abundances of *iso*- and *normal*-3-OH FAs have no significant relationship with soil pH in EC soils unlike SW soils. There is almost no direct temperature response among SW soil 3-OH FAs, and no significant response to MAP in EC soil 3-OH FAs, despite having significant relationships with these parameters when soils from both regions are compiled together. Across both transects however, $i\text{-C}_{13}$ and $i\text{-C}_{16}$ have the same response to pH, and $a\text{-C}_{15}$ to $MAAT$; although the two transects represent two distinctly opposed environments in regard to MAP and soil pH, some 3-OH FA compounds have been identified which portray the same response to environmental forcing.

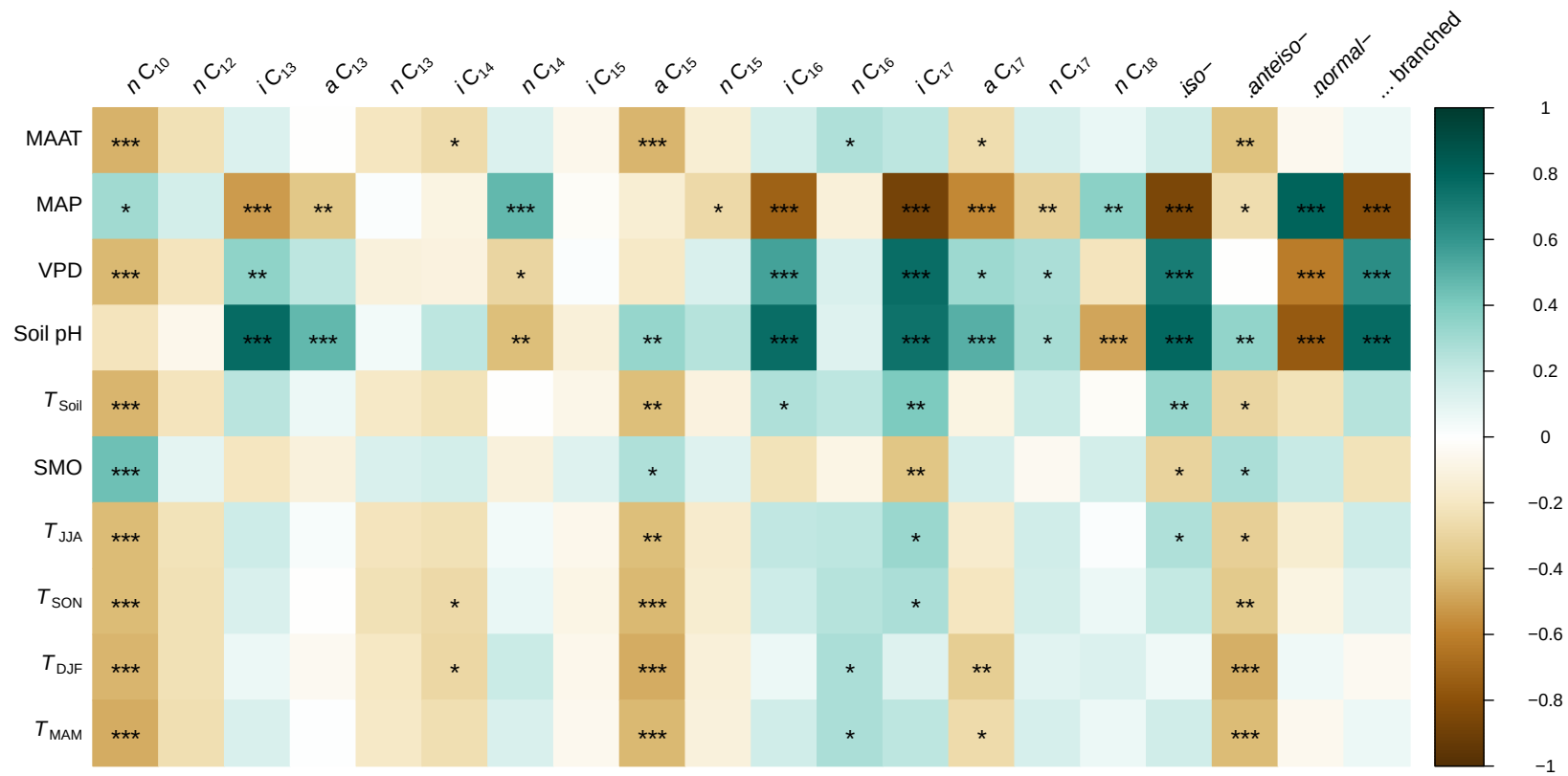


Figure 3.7: Pearson's correlation heatmap between 3-OH FA compounds and environmental variables. 3-OH FA data is compiled for all soils, and not separated by US soil transect. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$))

PCA was applied to explore and summarise meaningful trends of variance within the 3-OH FA distributions of our U.S. soils. 3-OH FAs that consistently fall below detection limit in the soils - *i*-C₁₁, *a*-C₁₁, *u*-C₁₁, *n*-C₁₁, *i*-C₁₂, *u*-C₁₂, *u*-C₁₃, *u*-C₁₄, *u*-C₁₆, *i*-C₁₈, and *u*-C₁₁ - were removed from the PCA dataset. By omitting these compounds, EC and SW soil 3-OH FA distributions become less distinguished in comparison to the previous NMDS plot (Fig. 3.8), as regionally-influential compounds - such as unsaturated structures that were predominantly yielded from EC soils - were removed.

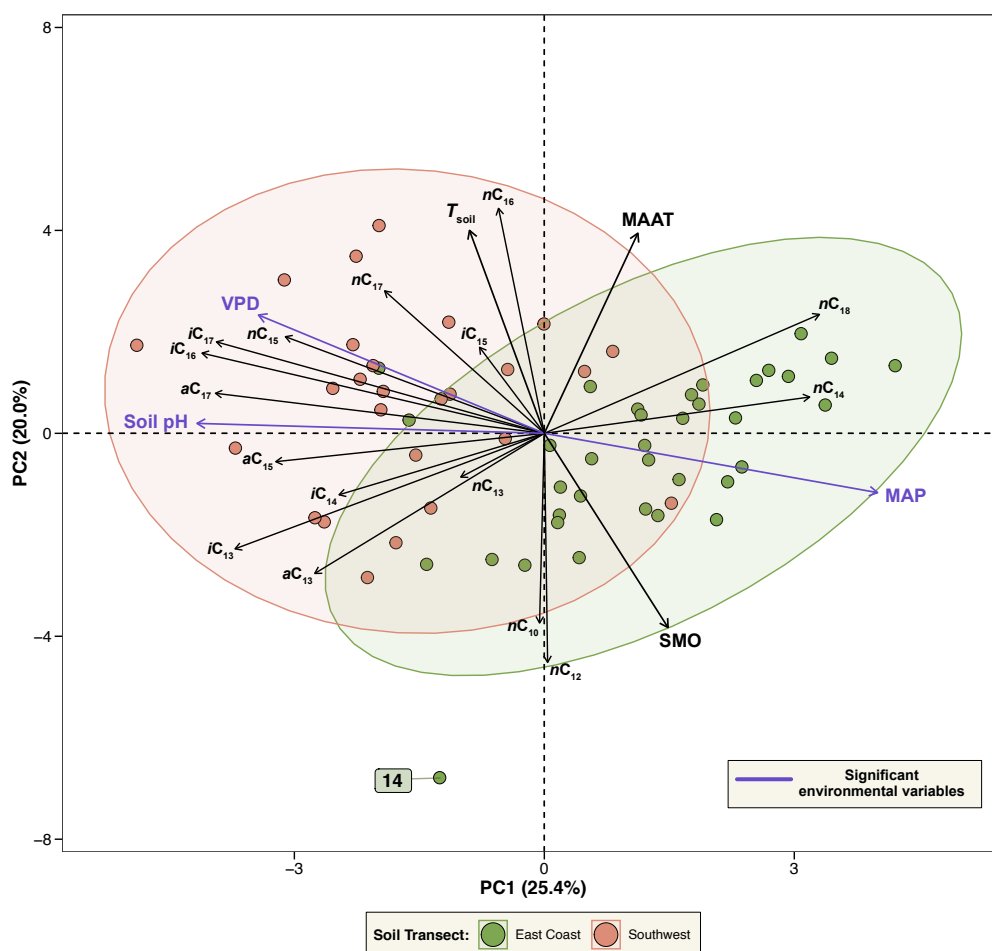


Figure 3.8: PCA of U.S. soil 3-OH FA distributions. Variability of 3-OH FA compounds and quantitative environmental drivers is visualised, and significant environmental drivers ($p < 0.05$) are highlighted following permutation analysis

Combined, PC1 and PC2 explain 40.4 % of variance in the U.S. soils 3-OH FA dataset. PC1 explains the largest proportion of this variance (25.5 %) which is associated with the partial separation of the soil transects; the upper range of EC soil PC1 values plot in the right half of the PCA plot, and the lower range of SW U.S. soil PC1 values in the left half (Fig. 3.8). *i*-C₁₆, *a*-C₁₇, *i*-C₁₇, *i*-C₁₃, *n*-C₁₈, *a*-C₁₅, *n*-C₁₄, and *n*-C₁₅ are the primary contributing 3-OH FA compounds to PC1. *n*-C₁₄ and *n*-C₁₈ are the only of these compounds to be positively associated with EC soils, plotting in an opposed ordination to other 3-OH FAs contributing to PC1 (Fig. 3.8). PC2 also explains a large proportion of 3-OH FA variance (20.0 %), almost as much as PC1, for which the primary contributing 3-OH FAs are *n*-C₁₂, *n*-C₁₆, *n*-C₁₀, *n*-C₁₇ and *n*-C₁₃, all normal structures compared to the mixture of structural 3-OH FAs associated with PC1.

Quantitative environmental variables (MAAT, MAP, SMO, soil pH, VPD, T_{Soil}) were added to the PCA plot following permutation analysis (Fig. 3.8), which reveals three variables that significantly explain 3-OH FA variance: soil pH ($r^2=0.66, p<0.001$), MAP ($r^2=0.59, p<0.001$), and VPD ($r^2=0.29, p<0.001$). Soil pH is strongly orientated along PC1, as are MAP and VPD to a lesser extent (Fig. 3.8). MAP plots in the positive range of PC1 values, almost directly opposed to VPD and soil pH which is consistent with the negative linear relationship between soil pH and MAP previously revealed (Fig. 3.2E), and the general distinction of acidic and alkaline soil pH between the transects which is inferred along PC1. 3-OH FA variance induced by moisture-related variables is therefore associated with PC1, and is observed across both soil transects combined. No significant variance in the compiled soil 3-OH FA dataset is caused by either MAAT or T_{Soil} even though significant correlations between a number of 3-OH FA compounds and these temperatures have been observed (Fig. 3.7). This result is likely because 3-OH FA temperature response is restricted to EC soils, and therefore regionally biased, whereas the 3-OH FA pH response is emphasised by the diverse pH range of both transects combined. This calls into question how a regional transect is defined, emphasising the need to understand regional environmental characteristics of contemporary soil transects used for lipid biomarker proxy calibration.

3.3.4.2 Response of established 3-OH FA-based proxies to environmental variables in U.S. soils

3-OH FA-based proxies for temperature and pH have distinct responses to environmental variables across U.S. soils. RDA performed on the proxy distributions in conjunction with environmental variables (Fig. 3.9) reveals that the measured quantitative environmental variables for U.S. soils explain 54.6 % constrained variance amongst the proxies, therefore it is likely that some proxies summarise 3-OH FA response to at least some environmental variables. Furthermore, proxy response to temperature and pH is likely to be independent as pH-related variables are associated with RDA1, and RDA2 for temperature-related variables (Fig. 3.9). Performing forward selection to identify statistically important explanatory variables of the constrained variance revealed MAP, VPD, soil pH, and T_{DJF} explain 46.2 % of constrained variance. The forward-selected variables are consistent with those that significantly described 3-OH FA variance during PCA analysis (MAP, soil pH, VPD; Fig. 3.8), and it can be established that existing 3-OH FA-based proxies best describe soil pH and moisture-related variables amongst all U.S. soils when their geographic transects are not used as a grouping variable.

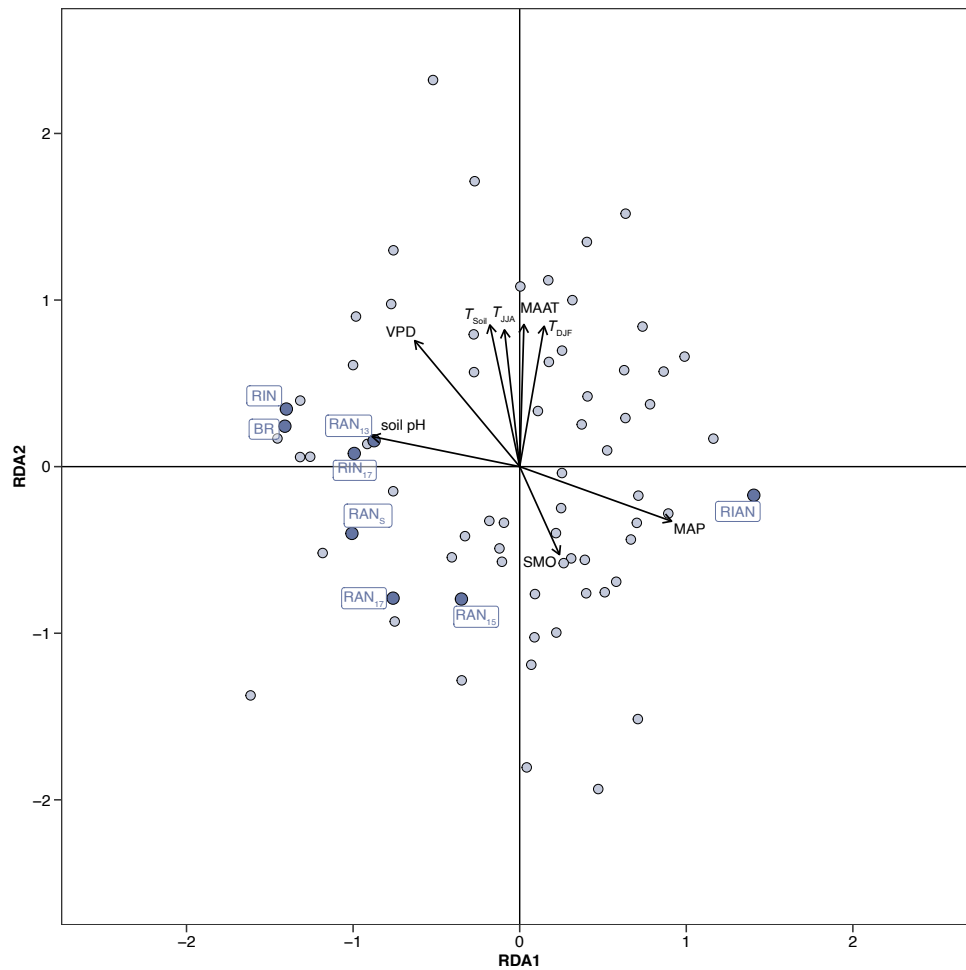


Figure 3.9: RDA employing 3-OH FA proxies as response variables and environmental drivers as explanatory variables in U.S. soils

As soil pH and **MAP** strongly co-vary in U.S. soils and are primary environmental drivers of variance within the 3-OH FA composition of these soils, 3-OH FA proxy response to pH is focused on first. RIAN, RIN, BR, RAN₁₃ and RIN₁₇ are all associated with soil pH and **MAP** along RDA1 (Fig. 3.9), however a Pearson's correlation matrix reveals that all 3-OH FA-based proxies excluding RAN₁₅ and RAN₁₇ have highly significant correlations with soil pH and (excluding only RAN₁₅) **MAP** (Appendix A6). Soil pH and **MAP** have highly significant correlations amongst almost all proxies because 3-OH FAs are cumulatively responding to these variables; 3-OH FA

compounds employed in the RAN indices and RIN_{17} are also employed in the more cumulative proxies RIN, BR, and RIAN, and so the three latter are more constructive proxies for conveying a U.S. soil pH response. An exponential regression might best describe RIN and BR (Fig. 3.10), and therefore RIAN - the linear conversion of BR - might be a more appropriate representation of BR. A linear conversion for RIN has not been required to fit RIN variance to an environmental driver in previous studies as significant RIN-environmental correlations (i.e. RIN and SST) have been linear (Yang et al. 2020), and $-\log(RIN)$ produces a strong positive linear correlation with soil pH ($r^2=0.61$, $p<0.001$; Fig. 3.10).

Fitting BR and RIN with an exponential curve neither improves or deteriorates their regressions with pH from that of a linear model (Fig. 3.10). Similarly, the difference in correlation between RIAN and $-\log(RIN)$ with pH is negligible compared to that of BR and RIN (Fig. 3.10). However all iterations remain under consideration for proxy calibration as the residuals appear slightly tighter in the regressions of RIAN and $-\log(RIN)$ with pH (Fig. 3.10). Of the established 3-OH FA proxies, BR and RIN (and their linear-converted iterations) summarise the most 3-OH FA variance in response to soil pH, which is likely best described by RIN as the cumulative proportion of *iso*- and *normal*-3-OH FAs are highly responsive to pH ($i=r^2=0.62$, $p<0.001$; $n=r^2=0.58$, $p<0.001$), whereas a weak correlation is observed between *anteiso*-3-OH FAs and pH ($r^2=0.13$, $p<0.01$) (Fig. 3.10).

$$-\log(RIN) = -\log(RIN) \quad (3.1)$$

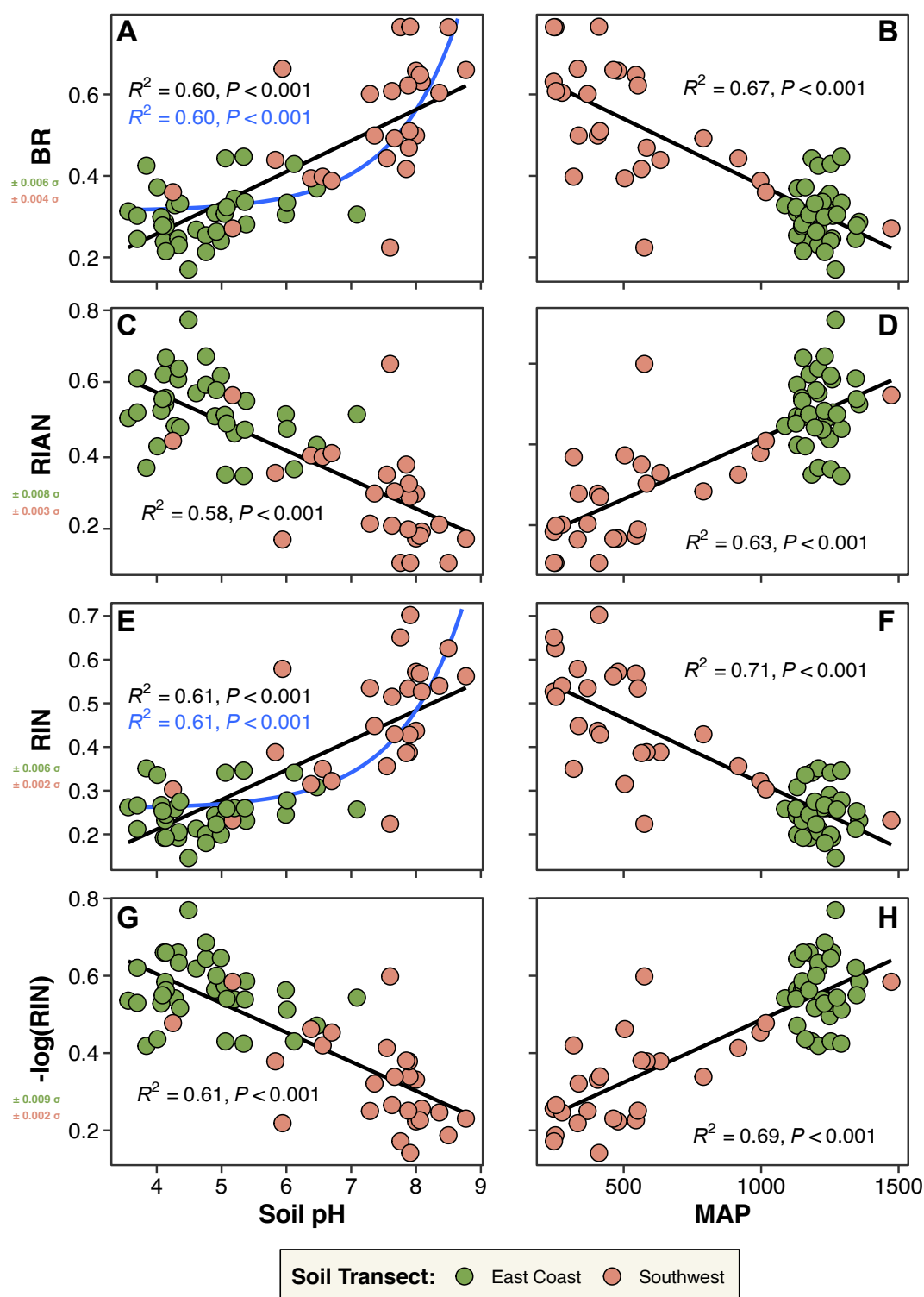


Figure 3.10: Crossplots between established 3-OH FA-based pH proxies and selected environmental variables (pH, MAP), comprising: BR, RIAN (negative logarithm of BR), RIN, and $-\log(\text{RIN})$ with pH (A,C,E,G) and MAP (B,D,F,H).

As soil pH and MAP strongly co-vary amongst the U.S. soils, BR and RIN also have strong negative linear relationships with MAP, which is $r^2=0.71$, $p<0.001$ for RIN and $r^2=0.67$, $p<0.001$ for BR which supersede the strength of their relationships with pH. Additionally, both regressions are stronger for BR and RIN compared to their linear-corrected counterparts with MAP because a linear model best describes BR and RIN and so a linear-correction is not required for their relationships with MAP as it potentially is for their relationships with soil pH. As observed for soil pH, the stronger MAP correlation for RIN relative to BR is attributed to the cumulative response of *iso*- ($r^2=0.72$, $p<0.001$) and *normal*-3-OH FAs ($r^2=0.65$, $p<0.001$) 3-OH FAs to MAP, and not *anteiso*-3-OH FAs ($r^2=0.07$, $p=0.04$). Overall, the strong response of RIN and BR (and -log(RIN) and RIAN by association) to soil pH and MAP is associated more so with SW soils ($r^2=0.30-0.44$) than EC soils ($r^2=<0.01-0.1$) likely because the SW soils cover wider gradients in MAP and pH. In any case, these proxies have a pronounced response to pH and MAP when the U.S. soils are compiled together compared to when considered as individual transects, however the covariance of soil pH and MAP is a potential limitation to their functionality as pH proxies as it is difficult to deconvolve. None of the following proxies - RIN, BR, RIAN - or descriptive indices - $\sum[\textit{iso}$ -3-OH FAs], $\sum[\textit{normal}$ -3-OH FAs] - have a significant correlation with MAAT and are therefore independent of temperature.

Most 3-OH FA compounds associated with soil pH along PC1 are *iso*- and *normal*-3-OH FA compounds employed in RIN, however not all 3-OH FAs employed in RIN are associated with PC1. Upon further inspection, *i*-C₁₃ ($r^2=0.61$, $p<0.001$), *i*-C₁₆ ($r^2=0.50$, $p<0.001$), and *i*-C₁₇ ($r^2=0.55$, $p<0.001$) significantly increase in fractional abundance in response to soil pH amongst U.S. soils, and *n*-C₁₄ ($r^2=0.16$, $p=0.001$) and *n*-C₁₈ ($r^2=0.26$, $p<0.001$) undergo a significant decline (Fig. 3.7). When $\sum(\textit{i}$ -C₁₃, *i*-C₁₆, *i*-C₁₇) are resolved by $\sum(\textit{n}-C₁₄, *n*-C₁₈) (Eq.3.23), the resulting RIN_{USA} index produces a strong and highly significant positive correlation with pH ($r^2=0.67$, $p<0.001$), however does not have an even distribution of residuals, and produces a high degree of scatter when soil pH>7.5 when fitted with a linear regression (Fig. 3.11). When a$

negative logarithm function is applied to the RIN_{USA} (Eq.3.3), as was applied to BR and RIN to produce RIAN and $-\log(RIN)$, the resulting correlation with soil pH supersedes that of the RIN_{USA} (and any established 3-OH FA pH proxy, Fig. 3.10) ($r^2=0.72$, $p<0.001$) (Fig. 3.11). Although RIN_{USA} and its linear corrected rendition $-\log(RIN_{USA})$ also have significant regressions with MAP (Fig. 3.11 B & E), we interpret these proxies primarily as a reflection of soil pH amongst U.S. soils as both generate variance amongst EC and SW soils in response to pH, but only vary in response to MAP in SW, not EC soils. However, EC and SW soils have weak but significant correlations with MAAT as individual transects for the proxies in question (Fig. 3.11 C & F), but not when compiled, and potential implications of this should be considered upon calibration of this proxy. Regardless of this minor MAAT dependence at a regional level, $-\log(RIN_{USA})$ (Eq.3.3) is the strongest pH proxy for calibration for U.S. soils.

$$RIN_{USA} = \sum [i-C_{13} + i-C_{16} + i-C_{17}] / \sum [n-C_{14} + n-C_{18}] \quad (3.2)$$

$$-\log(RIN_{USA}) = -\log(RIN_{USA}) \quad (3.3)$$

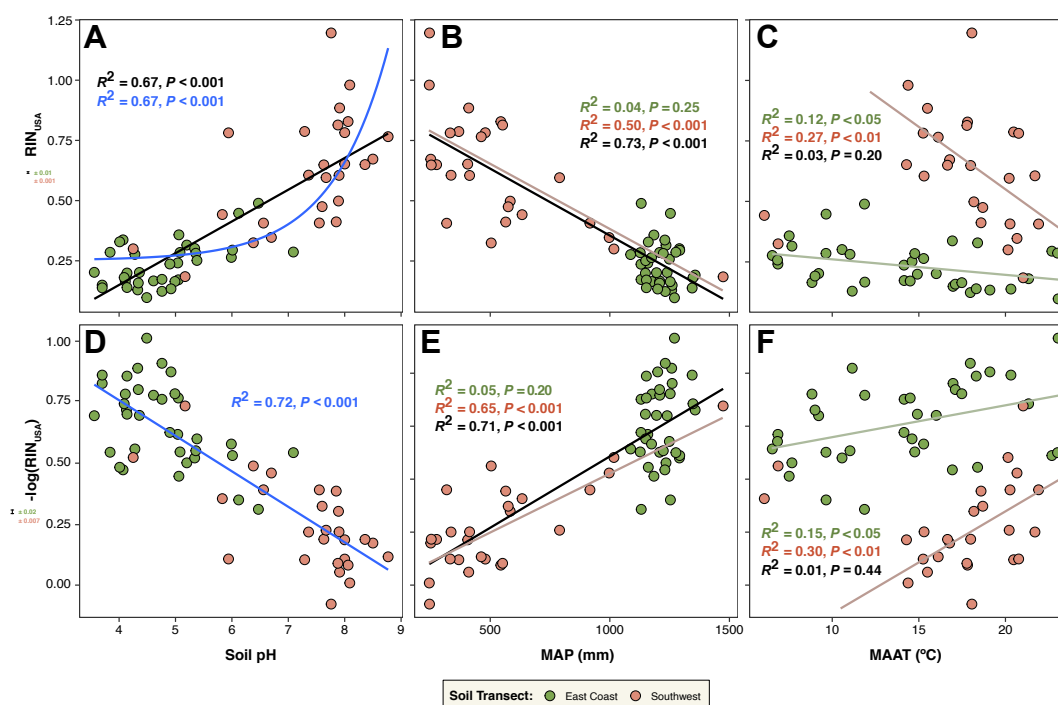


Figure 3.11: Crossplots between RIN_{USA} and -log(RIN_{USA}) and selected environmental variables (pH and MAP_{map}), comprising: RIN_{USA} with pH (A), MAP (B), and MAAT (C), and -log(RIN_{USA}) with same (D,E,F).

Deconvolving a 3-OH FA response to temperature from the U.S. soils is more complex than for pH and MAP as (i) no established 3-OH FA proxies are obviously associated with temperature variables (MAAT, T_{JJA} , T_{SON} , T_{DJF} , T_{MAM} , T_{Soil}) along RDA2 (Fig. 3.9) and (ii), of the 3-OH FA compounds associated with temperature variables along PC2 in our PCA, only $n\text{-C}_{10}$ significantly decreases with MAAT ($r^2=0.16, p<0.001$), and so developing a novel temperature proxy from the U.S. soil 3-OH FA distributions does not seem a viable option as it was for pH. However, the

Pearson's matrix between U.S. soil 3-OH FA proxies and environmental variables reveals RAN₁₅ and RAN₁₇ - temperature proxies developed from soils (Wang et al. 2016) - as the only proxies to significantly correlate with all temperature variables (Appendix A6). Both RAN₁₅ and RAN₁₇ have significant negative linear correlations with temperature variables, and although this is more pronounced for RAN₁₇ ($r^2=0.29$, $p<0.001$) compared to RAN₁₅ ($r^2=0.13$, $p<0.01$), both correlations are weak, which results from scatter and is particularly pronounced for RAN₁₅ (Fig. 3.12).

MAAT did not significantly explain any variance amongst compiled 3-OH FAs as indicated by PCA analysis (Fig. 3.8), nor was selected as a primary environmental driver of proxy variance by forward selection during RDA analysis (Fig. 3.9). However, a number of 3-OH FA compounds were significantly associated with temperature variables along PCA2, only *n*-C₁₀ produces a significant correlation with MAAT and Pearson's correlation matrices between 3-OH FA-based proxies and environmental variables performed separately for each U.S. soil transect reveals that temperature signatures amongst the proxies are highly regional (Appendix A7). RAN₁₃ and RIN₁₇ are the only proxies not to significantly respond to any environmental parameter (excluding RAN₁₃ and soil pH) amongst EC soils. Furthermore, the responsive proxies generally have significant relationships with all environmental variables excluding MAP and soil pH (Appendix A7), which is likely related to the fairly low pH and MAP gradients of EC soils; the EC is humid and comprises acidic soils. The reverse occurs amongst SW soils; most 3-OH FA proxies (excluding RAN₁₅ and RAN₁₃ which have no significant response to any environmental variable) have significant correlations with MAP, VPD, and soil pH and not with temperature (Appendix A7). Environmental drivers appear to generate a temperature response amongst EC soils, and a pH and moisture response amongst SW soils, however when compiled, an overarching pH response amongst U.S. soils is revealed, and the temperature response of EC 3-OH FAs undermined. 3-OH FA response to temperature appears largely regional between the U.S. soil transects, which is likely why MAAT has not been identified as a primary driver of 3-OH FA variance in multivariate

analyses compiling all U.S. soils together.

3-OH FAs have a significant response to temperature in EC soils, however as most 3-OH FA proxies have significant responses to **MAAT**, meaningful proxy responses need to be identified. RAN_{17} produces the most significant linear correlation with **MAAT** amongst EC soils, which is negative ($r^2=0.54$, $p<0.001$), and is subceded in magnitude by RAN_{15} ($r^2=0.33$, $p<0.001$) (Fig. 3.12). The cumulative proportion of *anteiso*-3-OH FAs also strongly correlates with **MAAT** in a negative linear correlation ($r^2=0.51$, $p<0.001$), however the predominant contributing 3-OH FA compounds to the cumulative proportion of *anteiso*-3-OH FAs are *a*-C₁₅ and *a*-C₁₇ which are more productively employed in the RAN_{15} and RAN_{17} indices which have been deployed widely in globally distributed soils (Wang et al. 2021). The cumulative proportion of *normal*-3-OH FAs also significantly correlates with **MAAT** in EC soils ($r^2=0.25$, $p<0.01$), which likely summarises the association of a number of *normal*- 3-OH FA compounds with temperature variables along PC2 (Fig. 3.8). As **MAAT** also generates significant RAN_{17} variability in SW soils ($r^2=0.33$, $p<0.01$), not RAN_{15} ($r^2=0.03$, $p=0.49$), is known to respond to **MAAT** in many soils globally, and responds to temperature independently of soil pH (Fig. 3.12), RAN_{17} indicates the best functionality as a temperature proxy in U.S. soils, at least for those of the EC transect.

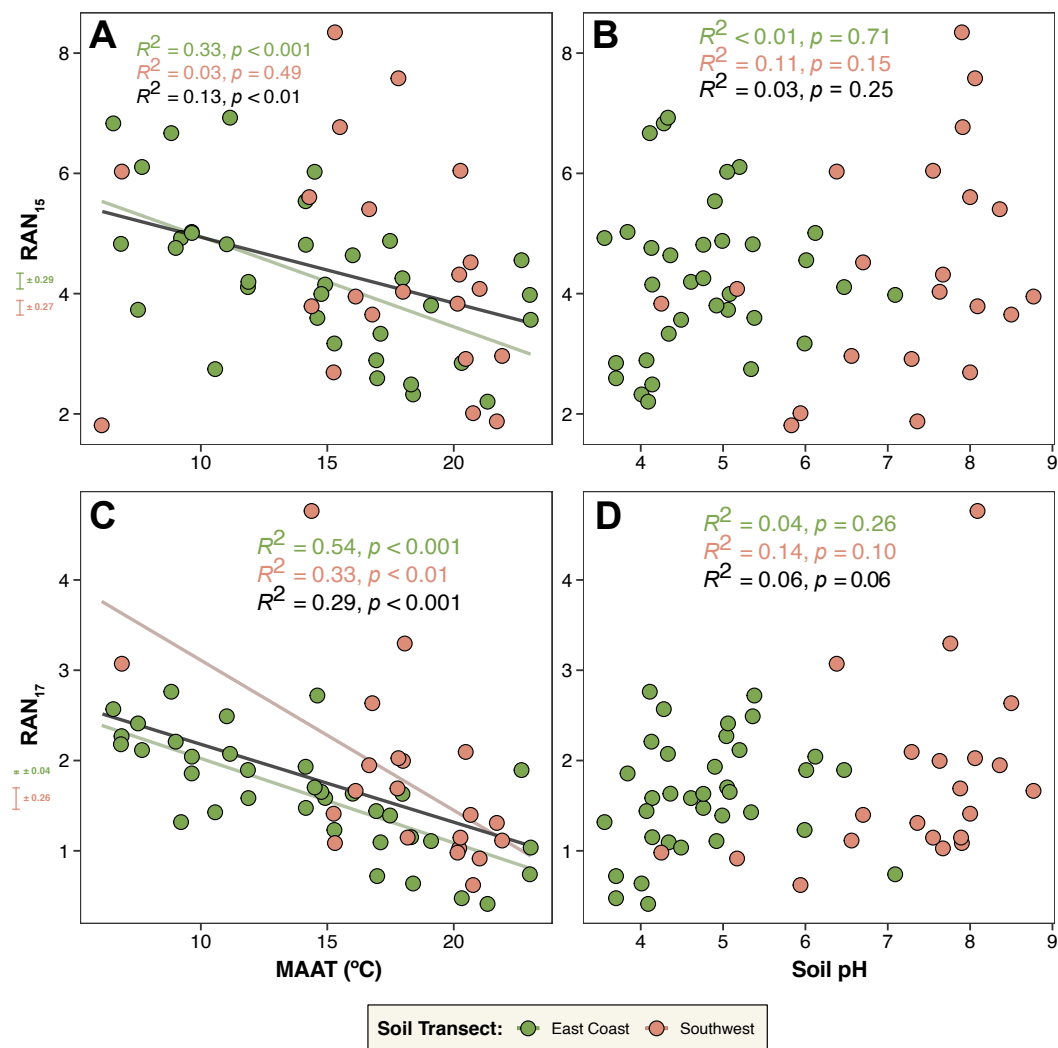


Figure 3.12: Crossplots of RAN₁₅ and RAN₁₇ with MAAT (A,C) and pH (B,D) amongst U.S. soils

There are other notable and potentially confounding environmental influences on the most viable pH and temperature proxies identified in U.S. soils. Firstly, there is evidence of seasonal bias in air temperature on RAN₁₅ and RAN₁₇ which is associated with SW soils. Forward selection of environmental variables in RDA extracted T_{DJF} amongst those most responsible for constrained significance amongst 3-OH FA proxies, and indeed, T_{DJF} generates a slightly stronger response from RAN₁₅ and RAN₁₇, albeit correlations for both proxies remain weak. In SW soils, RAN₁₇

most significantly correlates with T_{DJF} ($r^2=0.36$, $p<0.01$) and least with T_{JJA} ($r^2=0.28$, $p<0.001$), but no seasonal variability is observed amongst EC soils (Appendix A7). Amongst compiled U.S. soils, RAN_{15} instead has variability amongst seasonal air temperatures having strongest linear relationships with $MAAT$, T_{DJF} , and T_{MAM} (Appendix A7). Secondly, soil pH only significantly explains RIN and BR variance in SW soils, and appear respond to temperature, SMO , and VPD in EC soils. These pH proxies only have a widespread pH response when all U.S. soils are combined.

We additionally observe that two cold outlier soils in the SW soil transect consistently follow the 3-OH FA proxy variability of EC soils, and that omitting these soils could reveal significant proxy relationships with environmental variables in SW soils. When the SW cold outliers are re-classed as EC soils, the RAN_{17} - $MAAT$ relationship is rendered weaker in SW soils, and RAN_S has highly significant variability with $MAAT$ ($r^2=0.49$, $p<0.001$) instead of RAN_{17} (Appendix A8).

Therefore 3-OH FA distributions and established 3-OH FA-based proxies for temperature and pH reconstruction have highly variable responses to environmental drivers between U.S. soil transects, and when compiled. This calls into question whether 3-OH FA distributions in U.S. soils are better suited for the development of regional U.S. temperature and pH calibrations that employ all soils, or whether more regional calibrations employing soils from the individual geographical transects are more appropriate for this dataset. We consider the 3-OH FA calibration potential and integrity of our U.S. soils in our subsequent discussion.

3.4 Discussion

3.4.1 U.S. soil 3-OH FA distributions in the context of global 3-OH FA variability

3-OH FA distributions have been reported from a number of globally distributed soils which are commonly grouped into regional transects (Wang et al. 2021, Véquaud et al. 2021b), and the identification of 3-OH FAs in the U.S. soils further supports the ubiquitous occurrence of 3-OH

FAs in the pedosphere. 3-OH FA distributions have generally been investigated across montane transects, initially from Mt. Shennongjia, China from which the first 3-OH FA-based proxies for temperature and pH reconstruction were developed (Wang et al. 2016), followed by contrasting tropical (Mt. Rungwe, Tanzania) and temperate (Mt. Majella, Italy) transects (Huguet et al. 2019) in addition to a study of environmental variability across the French Alps (Véquaud et al. 2021a). Global compilations of 3-OH FA distributions in soils have been reported by Wang et al. (2021) featuring a range of montane and lowland soils, and Véquaud et al. (2021b) who investigated a number of montane soil transects from the Chilean and Peruvian Andes, Mt. Pollino (Italy), and Mt. Shegyla (Tibet) in addition to existing transects. Combined, these globally distributed soils originate from a diverse range of environments encompassing a MAAT range of -0.4 to 27 °C, an acidic to alkaline pH range (3.40-9.20), an elevation range of -4.0 to 4479 m asl, and a MAP range of at least 245 to 3313 mm (where reported). Soils from the aridic climate regime of the SW U.S. are particularly important for their role in increasing 3-OH FA representation from pedogenic environments that receive <500 mm rainfall per year, and comprise the only hyperarid soils (n=2) from which 3-OH FAs have been reported.

3-OH FA distributions from all compiled soils are highly variable, and all 3-OH FA compounds vary significantly amongst global soil 3-OH FA distributions (Fig. 3.13), which is not surprising considering the vast environmental gradients from which they are derived. However, regardless of this variability, most soils fall within the 95 % threshold for soil 3-OH FA variability (marked by a confidence ellipsis), and samples outside this threshold are not associated with a particular soil transect, rather a small number of samples from multiple transects (Fig. 3.13). Multiple lines of variability can be identified from such samples; those that protrude into the upper left NMDS quadrant are associated with a high abundance of normal 3-OH FA compounds with positive C numbers (n -C₁₄, n -C₁₆, i -C₁₈) which predominately occur in soils from the EC and global soil transect of Wang et al. (2021) (Fig. 3.13). Samples that protrude into the lower left NMDS quadrant are associated with enrichment in longer-chain 3-OH FA compounds (n -C₁₇,

i-C₁₈, *i*-C₁₆, *i*-C₁₇) which occurs in soils from the SW, global soil transect of Wang et al. (2021), and Mt. Majella (Fig. 3.13). *i*-C₁₇ enrichment in SW soils is fairly uncharacteristic relative to globally distributed soils (Wang et al. 2021, Véquaud et al. 2021b), although some soils have been observed to have *i*-C₁₇ as a dominant compound from the global soil calibration of Wang et al. (2021) (n=4), Mt. Pollino (n=1) (Véquaud et al. 2021b), French Alps (n=1) (Véquaud et al. 2021a), and in a soil from the catchment of Lake Chagan (China) (n=1) (Yang et al. 2021). Interestingly all soils with *i*-C₁₇ dominance have a MAP<560 mm and pH>5.94; they originate from arid environments. In the upper right NMDS quadrant, samples are associated with enrichment of shorter-chain 3-OH FAs (*i*-C₁₁, *n*-C₁₁, *n*-C₁₀, *n*-C₁₂) occurring in soils from the global soil transect, EC, Mt. Rungwe and Chilean Andes (Fig. 3.13). The least variability is associated with the lower right NMDS quadrant where *i*-C₁₂ is particularly enriched in one SW U.S. soil. Of all soils, two have particularly unique 3-OH FA distributions, one of which is soil from the catchment of Lake Junction (USA062) from the SW soil transect.

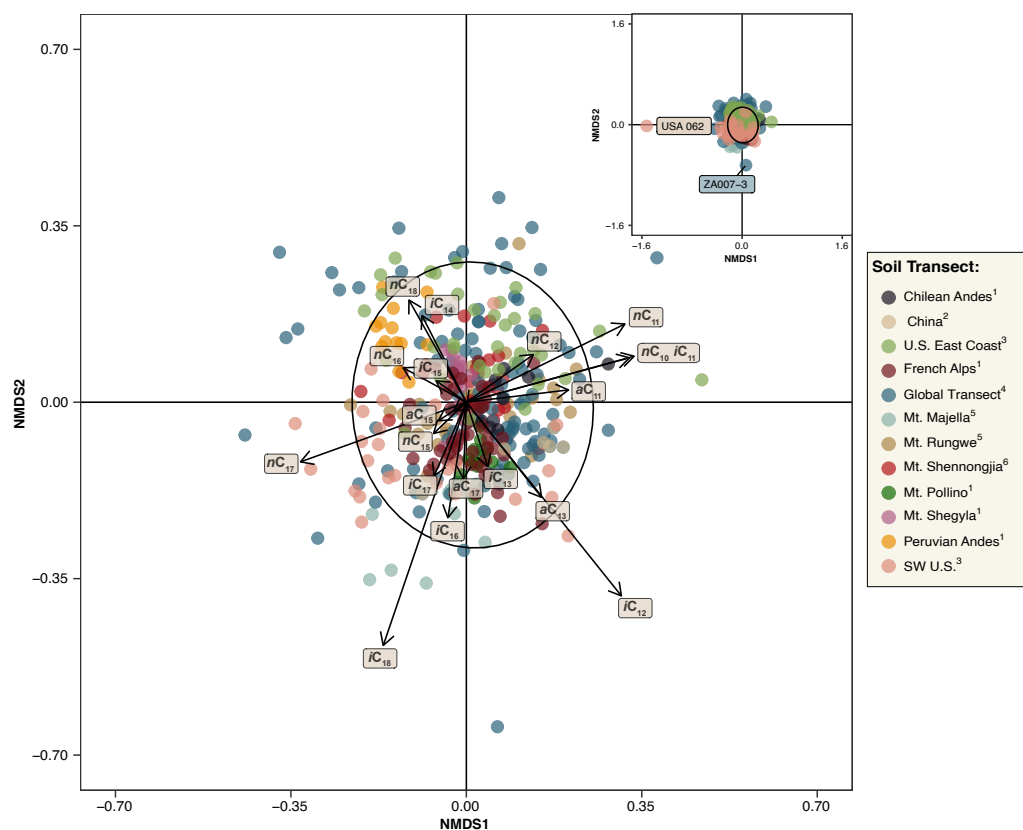


Figure 3.13: NMDS plot of 3-OH FA distributions in globally distributed soils after Véquaud et al. (2021b)¹, Yang et al. (2021)², present study³, Wang et al. (2021)⁴, Huguet et al. (2019)⁵, Wang et al. (2016)⁶. 3-OH FAs that significantly ($p < 0.05$) vary amongst the samples are visualised following permutation analysis.

3.4.2 Response of 3-OH FAs to temperature in soils

RAN₁₅ and RAN₁₇ are established 3-OH FA-based temperature proxies that have demonstrated significant variance in response to MAAT in soils globally (Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021b, Wang et al. 2021) (Fig. 3.14). Of these proxies, RAN₁₇ was identified as the most viable temperature response in our U.S. soils, having a moderate but highly significant negative relationship with MAAT in EC soils, and a weak relationship with MAAT amongst all U.S. soils combined (Fig. 3.14). In almost all soil transects investigated for 3-OH FAs, both RAN indices generally decline in value with increasing air temperature, therefore indicating that α -C₁₅

and α -C₁₇ decrease in abundance relative to their *normal*- structural counterparts as a significant response to MAAT (Wang et al. 2016). When U.S. soils are compiled, neither α -C₁₇ or n -C₁₇ have any significant variability in response to temperature ($p=0.12-0.27$), however amongst EC soils α -C₁₇ alone significantly declines in abundance with increasing MAAT (Appendix A5), indicating that although RAN₁₇ has significant variation in response to temperature, α -C₁₇ is the exclusive temperature-dependent component in EC soils responsible for the functionality of this proxy. The general response of RAN₁₅ and RAN₁₇ to MAAT is widely regarded as a physiological response of Gram-negative bacterial 3-OH FA producers to growth temperature in their external environment (Wang et al. 2016, 2021). This temperature response is termed homeoviscous adaptation and condones a deliberate relative adjustment of these specific 3-OH FA compounds to regulate and maintain an optimum membrane fluidity (Berry & Foegeding 1997).

As single-celled organisms, microbes are known to bio-synthetically adapt their cell membrane chemistry to mitigate against the ambient nature of external environmental variables (Guerzoni et al. 2001). Homeoviscous adaptation is a cumulative term for any articulated adjustment of the lipodome for the regulation of cell membrane fluidity (Ernst et al. 2016), and can involve changes in lipid saturation, aromatic structure, branching, or acyl chain length. The response of RAN₁₅ and RAN₁₇ to MAAT in globally distributed soils suggests that Gram-negative bacteria biosynthesise α -C₁₅₋₁₇ relative to n -C₁₅₋₁₇ 3-OH FAs in response to growth temperature, therefore invoking a structural change in 3-OH FA branching. Changes in the relative proportion of *anteiso*- fatty acids in response to growth temperature has been documented from a wide body of research reporting bacterial culturing experiments; in geothermal hotsprings (Nordström & Laakso 1992) and Bacteroidetes (Hellequin et al. 2023). Relatively increasing the abundance of α -C₁₅ and α -C₁₇ is a thermally viable response to lower bacterial growth temperatures because *anteiso*- fatty acids have a lower melting point in respect to their *normal*- counterparts (Kaneda 1991). Incorporating fatty acids with lower melting points into the phospholipid bilayer of a cell membrane provides a thermal buffer to ambient growth temperature by lowering the order-disorder

transition temperature of the membrane, to maintain a liquid-crystalline state for optimum fluidity (Annous et al. 1999, Kaneda 1991). The RAN₁₅ and RAN₁₇ proxies therefore represent a viable response of 3-OH FA bio-synthesis to temperature.

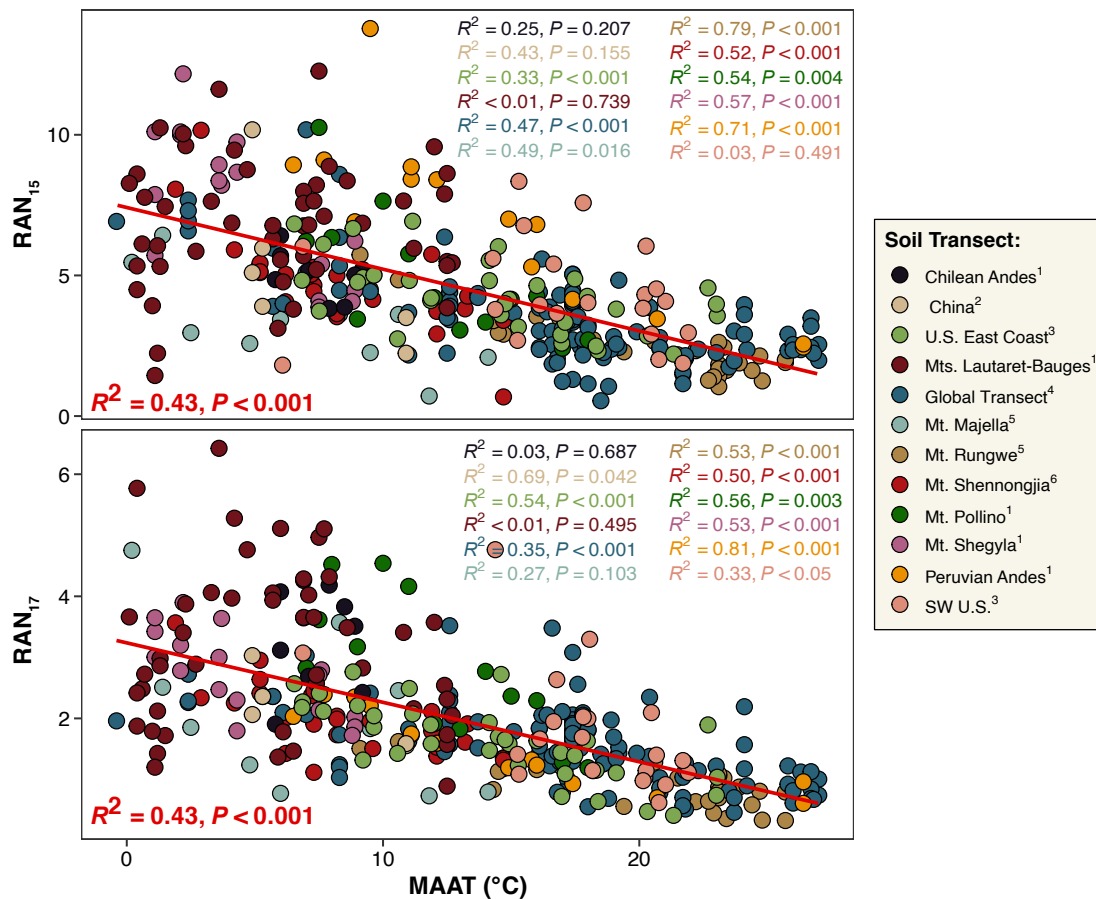


Figure 3.14: Variability of **A** RAN₁₅ and **B** RAN₁₇ with MAAT in globally distributed soils. Soil data after Véquaud et al. (2021b)¹, Yang et al. (2021)², present study³, Wang et al. (2021)⁴, Huguet et al. (2019)⁵, Wang et al. (2016)⁶.

Despite the weak to moderate response of RAN₁₇ to temperature in U.S. soils and evidence for this proxy representing a homeoviscous membrane adaptation response of 3-OH FA bacterial producers to temperature, strong regional variability of RAN₁₇ in addition to functionality of other

3-OH FA-based temperature proxies was observed between the EC and SW soil transects. Our study of U.S. soils is not the first time strong variability between and within 3-OH FA-based temperature proxies has been observed. Two global soil calibrations of RAN_{15} and RAN_{17} have been conducted in soils, and with all soil transects compiled (U.S. and global soils), RAN_{15} -MAAT regressions vary between $r^2=0$ (Mts Lautaret-Bauges, France) and $r^2=0.79$ (Mt. Rungwe, Tanzania), and those for RAN_{17} between ($r^2=0$ (Mts Lautaret-Bauges) and $r^2=0.81$ (Peruvian Andes) (Véquaude et al. 2021b) (Fig. 3.14). Although regressions between the regional transects are not definitively comparable due to differences in sample size, there is considerable variability in proxy performance amongst soils. The EC RAN_{17} -MAAT relationship is comparable to that of Mts Shennongjia, Rungwe, Shegyla, and Pollino ($r^2=0.49-0.56$) falling in the mid range of global soil RAN_{17} values, whereas both U.S. soil transects fall at the lower end of RAN_{15} functionality in response to temperature, and completely breaks down as a temperature proxy in SW soils, which is only observed amongst soils from the French Alps in the global compilation of soils investigated for 3-OH FAs (Véquaude et al. 2021a,b) (Fig. 3.14). Considerable variability in the performance of the RAN_{15} and RAN_{17} temperature proxies in U.S. soils (and more generally in global soils), is likely associated with environmentally and biologically-induced differences in 3-OH FA distribution between the transects which are all related to the diverse biogeography that these soils cover.

A further question to address when considering the performance of a temperature proxy is what growth temperature the Gram-negative bacterial producers of 3-OH FAs are most likely responding to, and contending temperatures aside from MAAT are its seasonal variations (T_{JJA} , T_{SON} , T_{DJF} , T_{MAM}) and T_{Soil} . A number of samples from both U.S. soil transects have an offset in ΔT (T_{Soil} -MAAT); in Arizona and Florida, soils have a tendency of positive offset from MAAT and becoming warmer, whilst soils in Maine have a negative ΔT , and are generally colder than MAAT. T_{Soil} is more susceptible to strong diurnal fluctuations compared to air temperature, particularly in hot regions with highly exposed soils such as deserts (Lozano-Parra et al. 2018).

For this reason, desert soil microbial communities are reported to be adapted to a wider T_{Optimum} range and to be composed of multiple communities of bacterial taxa, with some cold-adapted taxa more adapted to growth under T_{Min} , and more thermophilic taxa adapted to growth towards T_{Optimum} , compared to bacterial communities that might reside in temperate soils (van Gestel et al. 2013, Pettersson & Bååth 2003). Seasonality in air temperature also has the potential to influence the quality of a temperature proxy calibration, and cold season bias has led to growing season temperature (i.e. annual period when soils or lakes are not frozen) to be considered for some recent MBT'_{5ME} calibrations (Raberg et al. 2021, De Jonge et al. 2024). Although growing season temperature data is not available for U.S. soils, we are able to consider the potential influence of seasonal air temperature variability. As both U.S. soil transects cover a range of climate regimes, seasonality is variable across the U.S. soils, and is described by seasonality ($\Delta T_{\text{JJA-DJF}}$). We assess the influence of both ΔT and seasonality on the RAN₁₇-MAAT regression for compiled U.S. soils by plotting these parameters with the regression residuals, and find that seasonality does not significantly influence the residuals ($r^2=0.02$, $p=0.31$), ΔT does weakly but significantly influence them ($r^2=0.19$, $p<0.01$) (Fig. 3.15), the significance of which will be considered in our onward discussion. Additionally, the environmental variables reported in this study - MAAT, MAP, soil pH, VPD, seasonality, SMO, and ΔT - describe a small proportion of 3-OH FA variance in U.S. soils along PCs 1 and 2 at 45.4 %. These parameters are not fully representative of the pedoenvironmental conditions that influence bacterial diversity and growth in soils. Other variables such as cation (Mg^{2+} , Al^{3+} , Na^{+}) chemistry and soil grain size composition (% clay, silt, sand) that are not reported may account for part of the remaining variance. Ultimately, PCA revealed that soil pH and MAP are significant environmental drivers of 3-OH FA variance amongst those reported.

Overall, the RAN₁₇ proxy has been identified as most viable for temperature reconstruction amongst U.S. soils, but has a variable response to MAAT between the EC and SW transects. This proxy variability is not unique, having been observed for both RAN₁₅ and RAN₁₇ in soils globally.

The RAN₁₇-MAAT relationship appears more pronounced in EC than SW soils, however there is actually no significant difference in Pearson's correlation coefficient between U.S. soil transects for either proxy (RAN₁₅: $z=1.74$, $p=0.08$; RAN₁₇: $z=0.96$, $p=0.34$), which suggests that many samples between EC and SW soils have a comparable proxy response; although SW soils have a higher degree of RAN₁₇ scatter when regressed with MAAT, many samples actually follow the same trend of variability as the EC soils. These results indicate that RAN₁₇ variability in response to temperature may not be best described by classifying U.S. soils by their EC or sampling transect, and leads to the question do two separate regional calibrations need to be developed for RAN₁₇-based temperature reconstruction in U.S. soils, or can a compiled U.S. soil calibration be developed after refinement after we address what environmental variables might be causing scatter.

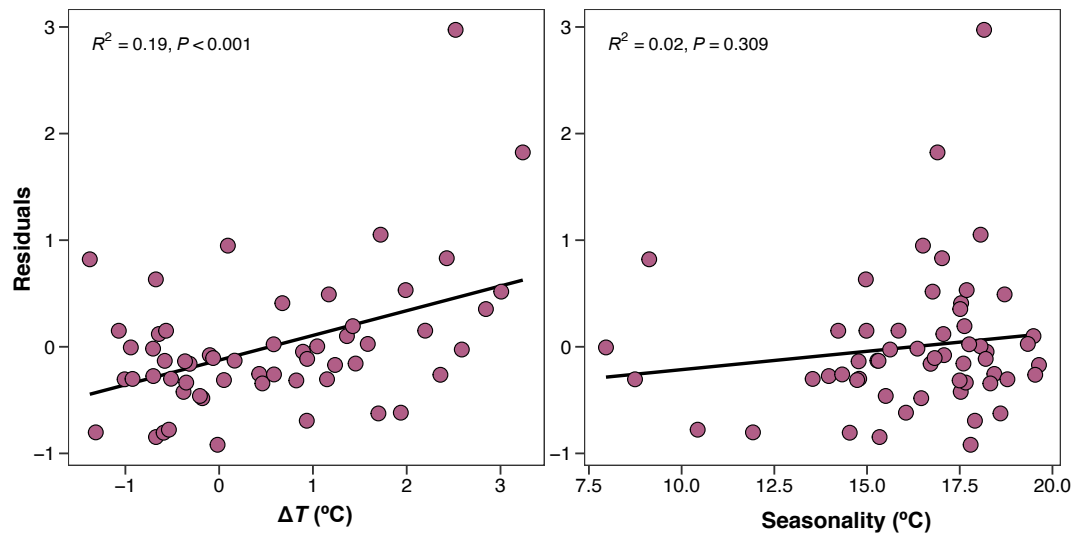


Figure 3.15: Scatterplots between residuals from the RAN₁₇-MAAT regression (Fig. 3.12C) of compiled U.S. soils and selected environmental variables (**A** ΔT and **B** seasonality)

3.4.3 Assessing RAN₁₇ regionality in U.S. soils: redefining the regional environmental constraints of a RAN₁₇-based temperature calibration

RAN₁₇ was identified as the most viable 3-OH FA-based temperature proxy in U.S. soils, however has differing variability between EC and SW transects which results in a weak RAN₁₇-MAAT when compiled. As RAN₁₇ regressions with MAAT are not completely distinguished between the two U.S. soil transects, the development of two separate regional RAN₁₇-based temperature calibrations for each transect may not be required. Instead, we adopt the approach of several recent modifications to the global MBT'_{5ME} soil calibration of [De Jonge et al. \(2014a\)](#) whereby the influence of non-MAAT environmental variability on the MBT'_{5ME} temperature regression has been investigated using regional environmentally specific soil transects, in order to remove confounding environmental scatter and improve proxy-based temperature reconstructions from global calibrations. To date, MBT'_{5ME} modifications have been performed due to confounding environmental scatter from soil pH ([De Jonge et al. 2021](#)), shifts in soil bacterial community composition ([De Jonge et al. 2019](#)), and MAP ([Dang et al. 2016](#), [Li et al. 2018](#)), as well as soil temperature ([Halffman et al. 2022](#)), and vegetation effects ([Liang et al. 2019](#)). Following similar procedures for the 3-OH FA-based proxies may generate a more robust RAN₁₇-based temperature calibration for the U.S. soils. We therefore compile the EC and SW soils, removing their geographical transect groupings, and explore other environmental variables that explain scatter in the RAN₁₇-MAAT relationship of these soils, and subsequently investigate.

Environmental variables are summarised by climate regimes - a climate regime is a description of the temperature and precipitation regimes of an environment, which can in turn mitigate vegetation ecozones and pedoclimate - and so we start our investigation by exploring the climate regimes that influence our U.S. soils. The EC and SW soil transects both encompass a range of climate regimes - snow, warm temperate, and equatorial for the EC, and arid, snow arid, and warm temperate for the SW ([Beck et al. 2023](#)); warm temperate is a mutual climate

classification between the transects, and all climate regimes along the EC are humid, whereas those of the SW substantially differ in precipitation regime. The most extreme residuals in the RAN_{17} -MAAT regression for U.S. soils arise from SW soils influenced by an arid climate regime (Fig. 3.16 A). These arid soils are associated with a range of extreme environmental variability, including low MAP and SMO, alkaline pH, high ΔT , and high seasonality (Fig. 3.16). Of these environmental conditions, low MAP and associated SMO are known to promote soil aeration and oxidation, which are unfavourable for the growth of anaerobic (and therefore Gram-negative) bacteria (Picek et al. 2000) and is hypothesised to have caused a breakdown in the brGDGT-based MBT temperature index and significantly underestimated MBT-reconstructed MAATs in arid soils of the western U.S. (Dirghangi et al. 2013). Removing soils from the arid climate regime substantially improves the U.S. soil RAN_{17} -MAAT regression ($r^2=0.60$) (Fig. 3.16B). However, defining scatter by categorical climate regimes is not robust, and a quantitative measure for this variability should be identified.

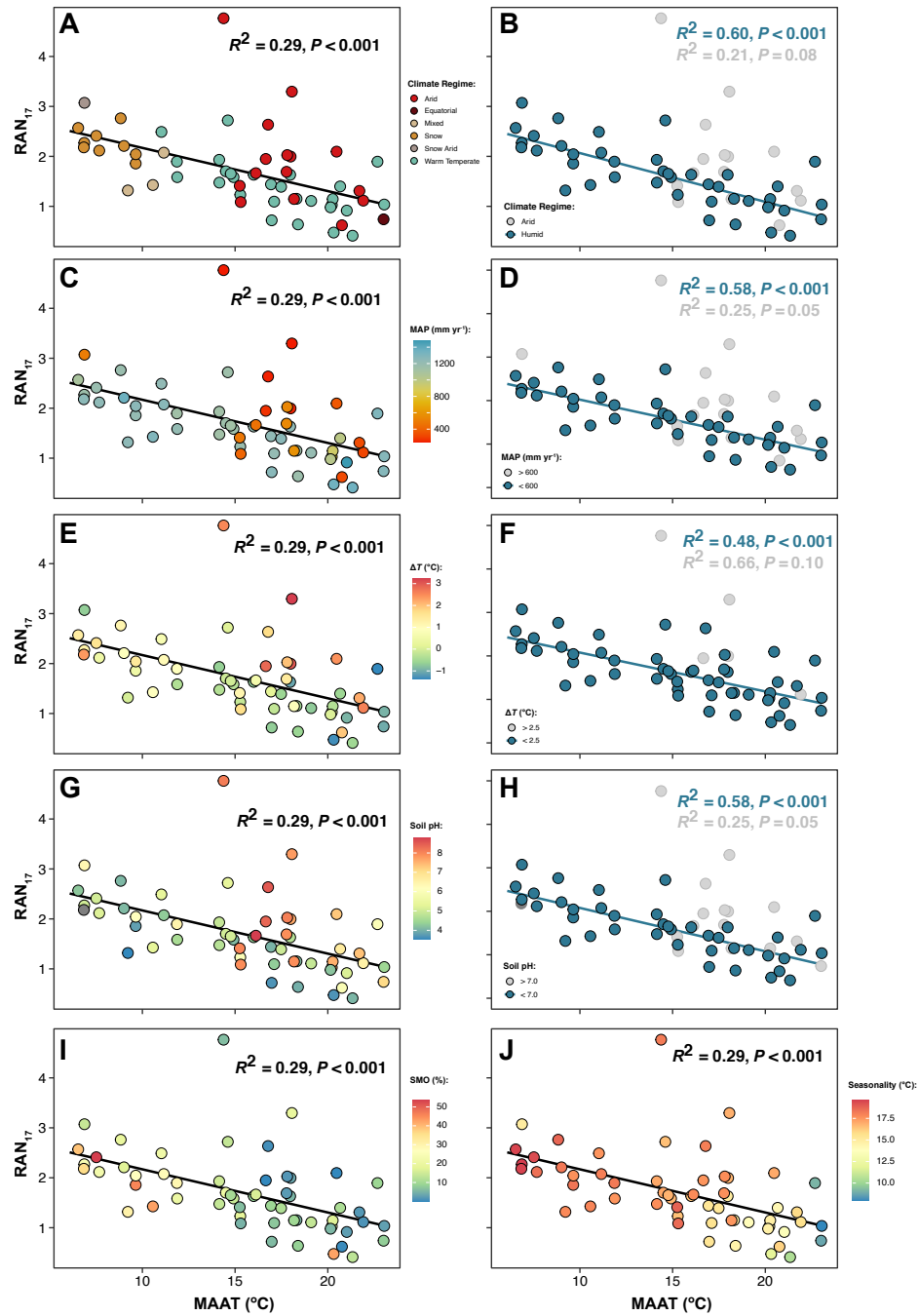


Figure 3.16: Scatterplots conveying the influence of environmental variability on the RAN_{17} - $MAAT$ regression in U.S. soils, specifically the influence of: climate regime (A), MAP (C), ΔT (E), pH (G), SMO (I), and seasonality (J). Extreme environmental variability has been omitted from the RAN_{17} - $MAAT$ regression for the following variables: climate regime (B), MAP (D), ΔT (F), pH (H).

Dirghangi et al. (2013) also report the MBT breakdown occurred at a MAP threshold of 800 mm/yr below which MBT temperature estimates were rendered unreliable, which extended to temperature calibrations of globally distributed soils (Peterse et al. 2012). As MAP is an important environmental control on aridity, we additionally assess the influence of different MAP thresholds on the RAN₁₇-MAAT relationship, first by excluding soils below the 800 mm MAP threshold from the regression. Doing so retains four warm temperate SW U.S. samples and also substantially improves the RAN₁₇-MAAT regression of U.S. soils ($r^2=0.57$). As the exact MAP threshold of MBT functionality and increase in $R_{i/b}$ has been observed to be variable (threshold of >600 mm reported by Xie et al. (2012)), we also apply a MAP threshold of 600 mm which retains 5 Warm Temperate SW U.S. soils and has a negligible improvement on the RAN₁₇-MAAT regression compared to the 800 mm MAP threshold ($r^2=0.58$) (Fig. 3.16C & D). We find that 600 mm is the optimum MAP threshold for RAN₁₇ functionality and lowering the MAP threshold further does not remove all confounding scatter. To further highlight the potentially confounding role of MAP on the RAN₁₇ temperature proxy, two SW soils were identified as hyperarid which is unique amongst the global soils explored for 3-OH FAs, and these two soils produce the most extreme residuals in the RAN₁₇-MAAT regression.

We have previously reported the variability of ΔT amongst U.S. soils in addition to the the significant influence of ΔT on RAN₁₇-MAAT residuals, both which indicate that pedoclimate and atmospheric climate may be disassociated to varying degrees amongst the soils, which also has the potential to confound any 3-OH FA response to temperature. Furthermore, ΔT is a known confounding variable in palaeoclimate reconstruction, that is capable of significantly offsetting temperature estimates in palaeosols (Gallagher et al. 2019). Several SW soils exhibited $\Delta T > 2.5$ °C and maximum observed ΔT was 3.2 °C. Although ΔT observed in the SW is not considered to be extreme in terms of global variability ($\Delta T > 4.5$ °C is regarded as a more extreme threshold (Molnar 2022, Gallagher et al. 2019)) we still assess it's effect on RAN₁₇ functionality to determine if RAN₁₇ might be responding to T_{Soil} in soils that produce scatter when regressed with MAAT.

Omitting soils exceeding a threshold of $\Delta T > 2.5$ °C isolates five SW U.S. soils from the RAN_{17} regression, each associated with the arid climate regime, and include the two hyperarid soils, however the temperature regression itself ($r^2=0.48$) is not improved as much as by omitting all arid soils or those below the 600 mm MAP threshold (Fig. 3.16E & F). This result may arise because ΔT is highly variable amongst warm U.S. soils - some warm soils have high ΔT and some do not - and is likely to be governed by environmental variables that influence soil physiology such as exposure from vegetation cover, SMO and soil type. SW soils are associated with shrubsteppe and desert scrub vegetation cover and low SMO whilst warm soils along the EC originate from Florida which has dense vegetation cover and moist soils associated with the Everglades, which are more likely to be primary environmental regulators on ΔT compared to MAAT.

The U.S. soils also cover a wide pH range, which we consider here as a potentially confounding environmental variable on the RAN_{17} -MAAT relationship because soil pH is a known driver of microbial respiration and changes in bacterial community structure in soils. Thorpe (2023) reports that bacterial communities amongst the U.S. soils are best structured by soil pH, which may introduce a biologically-mitigated pH bias on the RAN_{17} temperature proxy in these soils. De Jonge et al. (2021) identified a pH-induced change in brGDGT geochemical fingerprint at a threshold of pH 6.1, revealing that the brGDGT-based MBT'_{5ME} temperature proxy can become pH-dependent in the absence of a strong temperature gradient. Results such as this from brGDGT temperature proxy development are important because Acidobacteria is understood to be a biological source of brGDGTs and as a Gram-negative bacterial phyla, are also a microbial producer of 3-OH FAs. We also test pH 7.0 as a threshold in the RAN_{17} -MAAT regression as soil pH amongst the alkaline soils of the SW were observed to covary with MAAT. Grouping the U.S. RAN_{17} -MAAT regression using a pH=6.1 threshold after De Jonge et al. (2021) reveals a more robust temperature response in low pH soils (pH<6.1; $r^2=0.54$) in which six SW U.S. soils are retained, in comparison to that of more alkaline soils with pH>6.1 ($r^2=0.12$). However a pH=7.0 threshold optimises pH grouping of the RAN_{17} temperature response in U.S. soils (pH<7.0:

$r^2=0.58$; $\text{pH}>7.0$: $r^2=0.25$) (Fig. 3.16G & H). As the bacterial producers of 3-OH FAs are best structured by soil pH amongst U.S. soils (Thorpe 2023), the refined RAN_{17} temperature regression using acidic soils for which $\text{pH}<7.0$ is more likely to represent a more legitimate temperature response than omitting all soils that developed under an arid climate regime although the later produces a slightly improved regression; the alkaline pH of many SW soils is likely to be climatically induced, resulting from a combination of hot MAAT and low MAP - in short, the aridity of the region.

Our environmental investigation on confounding scatter in the RAN_{17} -MAAT regression has been fairly superficial thus far; potential scatter has been identified related to an arid climate regime, that is described by soils with $\text{MAP}>600$ mm, soil $\text{pH}>7.0$ and include soils with $\Delta T>2.5$ °C, a high seasonality and low SMO, however the thresholds at which these variables promote RAN_{17} scatter are not conclusive and are known to be transient between studies investigating their influence on performance of the brGDGT-based MBT_{5ME} temperature proxy and its earlier renditions. Legitimate scatter in the RAN_{17} proxy is likely to also be observed in the composite 3-OH FA distributions between soils. Therefore, with the aim of making definitive conclusions about what U.S. samples, if any, have the potential to produce legitimate 3-OH FA proxy scatter in response to temperature, we perform cluster analysis on the composite 3-OH FA distributions of all U.S. soils (Fig. 3.17A). Cluster analysis reveals a number of distinguished groups in 3-OH FA distribution at high resolution (Fig. 3.17A). Cluster groups 2 and 5 are most distinguished, and have respective compositions of primarily EC and SW soils (Fig. 3.17B). SW U.S. soils are split between cluster groups 3 and 5; these two groups have differing variability in 3-OH FA composition - group 3 soils are strongly associated with *i*-C₁₆, *i*-C₁₇, *a*-C₁₇, and *n*-C₁₅, variability in group 5 soils is not associated with any particular 3-OH FA compound.

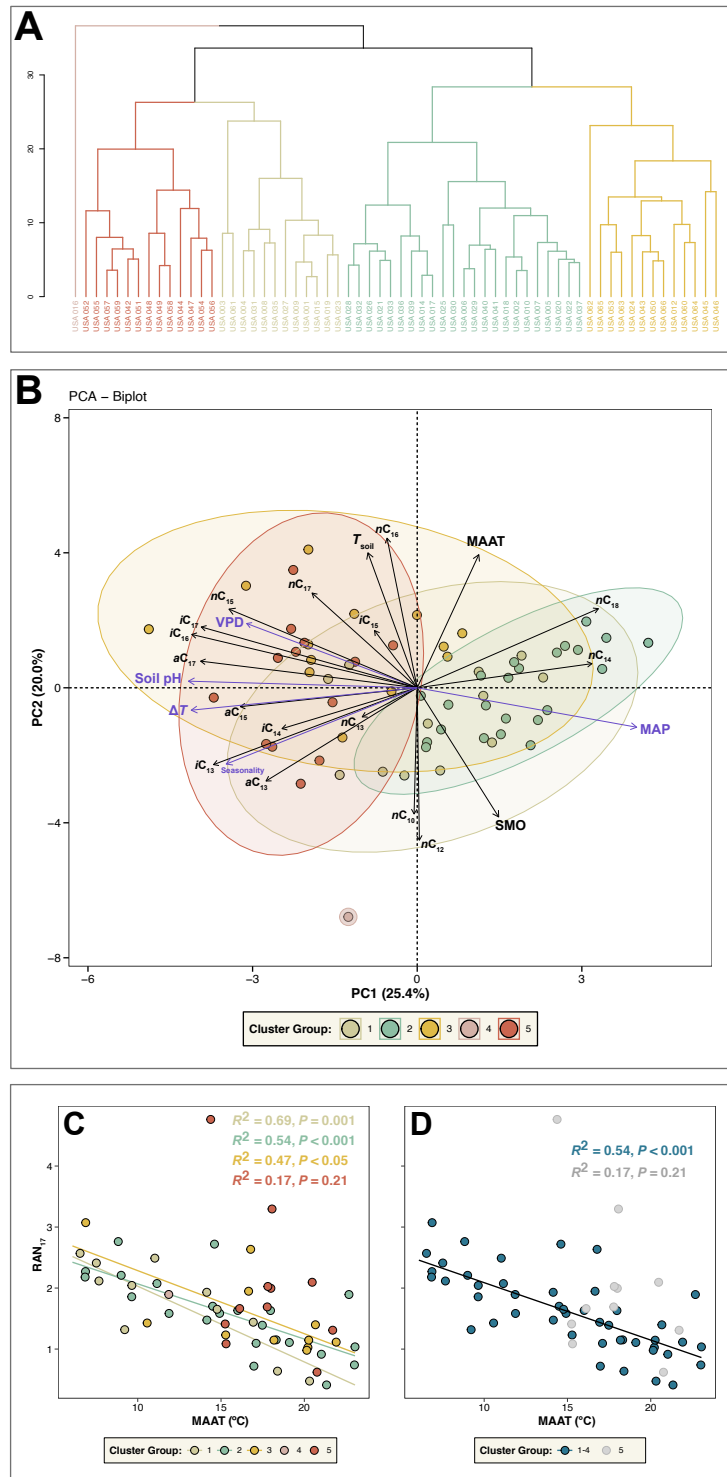


Figure 3.17: Investigation of 3-OH FA distribution similarity between soils. **A** Dendrogram visualising results of cluster analysis of 3-OH FA distributions in U.S. soils. **B** PCA plot with soils re-visualised as clustered groups (from A) instead of EC and SW transect grouping. **C** RAN₁₇-MAAT regression with cluster grouping from A. **D** RAN₁₇-MAAT regression with cluster group 5 omitted.

Additionally, group 3 SW soils are most associated with the group 2 EC soils, and group 5 SW soils are not, consolidating our hypothesis that a number of SW soils have comparable 3-OH FA distributions to EC soils. Of the clustered groups, the group 5 soils are the only that do not produce a significant correlation (excluding group 4 that comprises one soil) between RAN_{17} and $MAAT$ (Fig. 3.17C). When group 5 soils are omitted, the correlation between RAN_{17} and $MAAT$ substantially improves ($r^2=0.54$, $p<0.001$), amongst U.S. soils (Fig. 3.17D). Group 5 soils are primarily associated with high ΔT , soil pH, VPD and low MAP compared to other U.S. soils (Appendix A9). The MAP threshold on RAN_{17} that we previously considered can be disregarded as a number of low MAP soils are not associated with cluster group 5 - this is not a clustering based on MAP. Instead, VPD and soil pH are the primary environmental variables distinguishing group 5. VPD has a well established role in mitigating evapotranspiration in plants, and can become an environmental stress on plants in semiarid and arid environments. VPD is also influential on SMO, however as any direct control VPD might have on soil bacterial community structure is poorly understood, we do not use a VPD threshold to explain RAN_{17} scatter in response to $MAAT$. We thus conclude that a soil pH threshold of $pH<7.2$ (revised from our pre-cluster threshold of $pH=7.0$) to best describe scatter in the RAN_{17} - $MAAT$ regression; this threshold omits all (but one low pH outlier) group 5 soils from the RAN_{17} regression, in addition to a number of SW soils from other cluster groups, likely because the clustered groups are based on composite 3-OH FA distributions of the U.S. soils and not just compounds employed in RAN_{17} . Our final RAN_{17} - $MAAT$ linear regression - which excludes two high index outliers that fall outside 95 % confidence bounds - produces a strong negative correlation ($r^2=0.72$, $p<0.001$) (Fig. 3.18), and results in the following RAN_{17} calibration for reconstruction of $MAAT$ (Eq. (3.4)):

$$MAAT (^{\circ}C) = 25.1 - 6.66 \times RAN_{17} \quad (RMSE = 2.72^{\circ}C \quad n = 40) \quad (3.4)$$

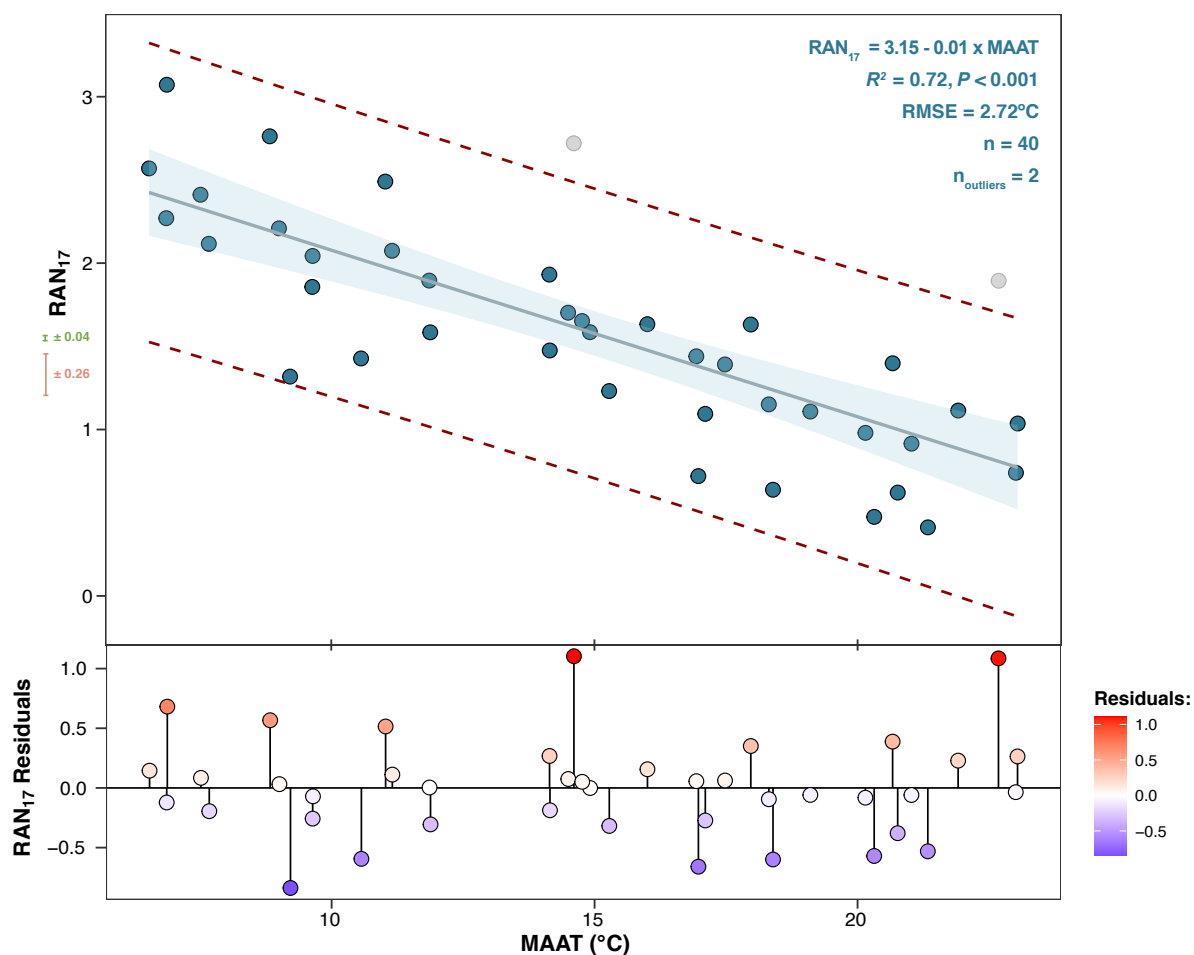


Figure 3.18: Refined RAN_{17} - $MAAT$ regression from U.S. soils that only features soils with $pH < 7.2$. Residuals are also visualised.

3.4.4 Performance of lipid biomarker-based temperature calibrations in U.S. soils

3.4.4.1 RAN_{17} calibrations on compiled U.S. soils

We have identified a collective of U.S. soils (those with $pH < 7.2$) in which RAN_{17} variability in response to $MAAT$ is optimised, and now assess the quality of several RAN_{17} -based temperature calibrations in addition to our optimised regional U.S. calibration, on the U.S. soils. Omitting soils with $pH > 7.2$ from the global soil dataset does not improve the RMSE of a global RAN_{17} calibration, despite improving the Pearson's correlation coefficient of the RAN_{17} - $MAAT$

regression itself (Fig. 3.19). This suggests that the pH<7.2 threshold observed to improve the RAN₁₇-MAAT regression in U.S. soils may be a regional effect and not globally influential, or that other environmental variables are promoting scatter in the global RAN₁₇ regression in addition to soil pH. Considerable scatter in the global RAN₁₇ regression indeed remains, particularly amongst colder soils >12 °C. Global soils from the calibrations of Wang et al. (2021) and Véquaud et al. (2021b) have not been combined into a single dataset before, and therefore the calibration for all soils in Figure 3.19 is novel and expressed as the equation:

$$MAAT (^{\circ}C) = 21.7 - 4.38 \times RAN_{17} \quad (RMSE = 5.37^{\circ}C \quad n = 337) \quad (3.5)$$

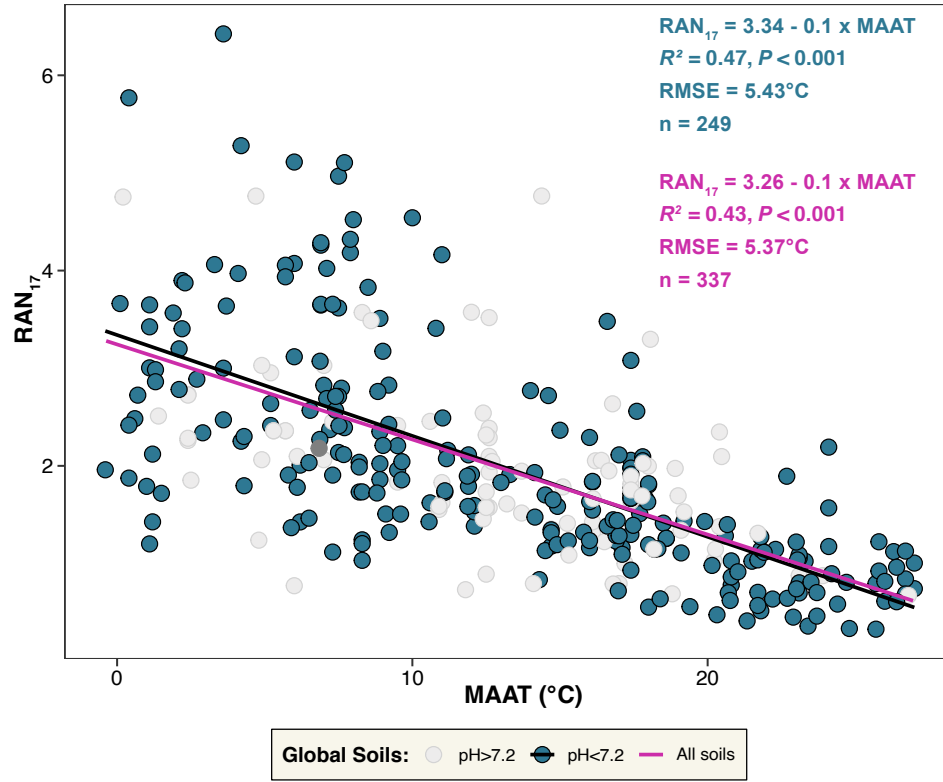


Figure 3.19: Testing the influence of a pH<7.2 threshold (found to improve RAN functionality in U.S. soils) on the RAN_{17} -MAAT regression of globally compiled soils. Summary statistics in blue (upper) summarise calibration quality of global soils with pH<7.2, and those in pink (lower) summarise the calibration quality of the entire dataset.

A number of RAN_{17} -based temperature calibrations have previously been developed from regional (Wang et al. 2016, Huguet et al. 2019) to global (Wang et al. 2021, Véquaud et al. 2021b) soil transects, which we also use to generate MAAT estimates in U.S. soils. Equations (3.6) to (3.8) are regional calibrations respectively developed from Mt. Shennongjia (Wang et al. 2016), Mt. Majella and Mt. Rungwe (Huguet et al. 2019) soils, and these three transects are combined to produce Equation (3.9) (Huguet et al. 2019). Equation (3.10) expresses a global RAN_{17} temperature calibration developed by Wang et al. (2021). Each of these calibrations are based on linear models, although we acknowledge that there has been an effort in recent years to develop

3-OH FA-based temperature calibrations using alternative methods such as multiple linear regression (Véquaud et al. 2021b) and random forest algorithms (Véquaud et al. 2021b, Wang et al. 2021) which employ a wider range of 3-OH FA compounds and have been reported to improve the predictive power of 3-OH FA-based temperature reconstruction. We test the performance of linear regressions only for direct comparison to the U.S. soils RAN₁₇ temperature calibration.

$$MAAT (^{\circ}C) = 26.36 - 9.09 \times RAN_{17} \quad RMSE = 2.7^{\circ}C \quad (3.6)$$

$$MAAT (^{\circ}C) = 30.12 - 8.24 \times RAN_{17} \quad RMSE = 2.5^{\circ}C \quad n = 28 \quad (3.7)$$

$$MAAT (^{\circ}C) = 10.80 - 1.84 \times RAN_{17} \quad RMSE = 3.8^{\circ}C \quad n = 11 \quad (3.8)$$

$$MAAT (^{\circ}C) = 25.74 - 7.38 \times RAN_{17} \quad RMSE = 5.1^{\circ}C \quad n = 65 \quad (3.9)$$

$$MAAT (^{\circ}C) = 37.68 - 14.49 \times RAN_{17} \quad RMSE = 5.2^{\circ}C \quad n = 185 \quad (3.10)$$

Estimated **MAAT** from the 7 RAN₁₇-based temperature calibrations reported in Equations (3.4) to (3.10) are visualised as scatterplots between ΔT (RAN₁₇-reconstructed temperature-**MAAT**) and contemporary **MAAT** (Fig. 3.20). The calibrations have a general tendency to underestimate **MAAT** in the U.S. soils as mean ΔT is negative for all estimates excluding those calculated from the RAN₁₇ calibration from Mt. Rungwe soils (Huguet et al. 2019) (Fig. 3.20). The RAN₁₇ calibration from Mt. Rungwe soils is the only to overestimate temperatures in U.S. soils (mean $\Delta T = 0.64$), whereas of the RAN₁₇ calibrations, that from Mt. Majella soils underestimated **MAAT** to the greatest degree (mean $\Delta T = -7.88$), and progressively underestimated with increasing contemporary **MAAT** (Fig. 3.20). RAN₁₇ calibrations from Mt. Rungwe and Mt. Majella soils likely have this opposed response because the two soil transects originate from environments with differing climate regimes; Mt. Rungwe is a tropical montane environment, and Mt. Majella is a temperate montane environment (Huguet et al. 2019). The Mt.

Rungwe calibration is more consistent at reconstructing temperature in U.S. soils probably because Mt. Rungwe soils cover a wider **MAAT** range and the general environmental conditions (warm and humid) are more comparable to the U.S. soils than the colder Mt. Majella soils.

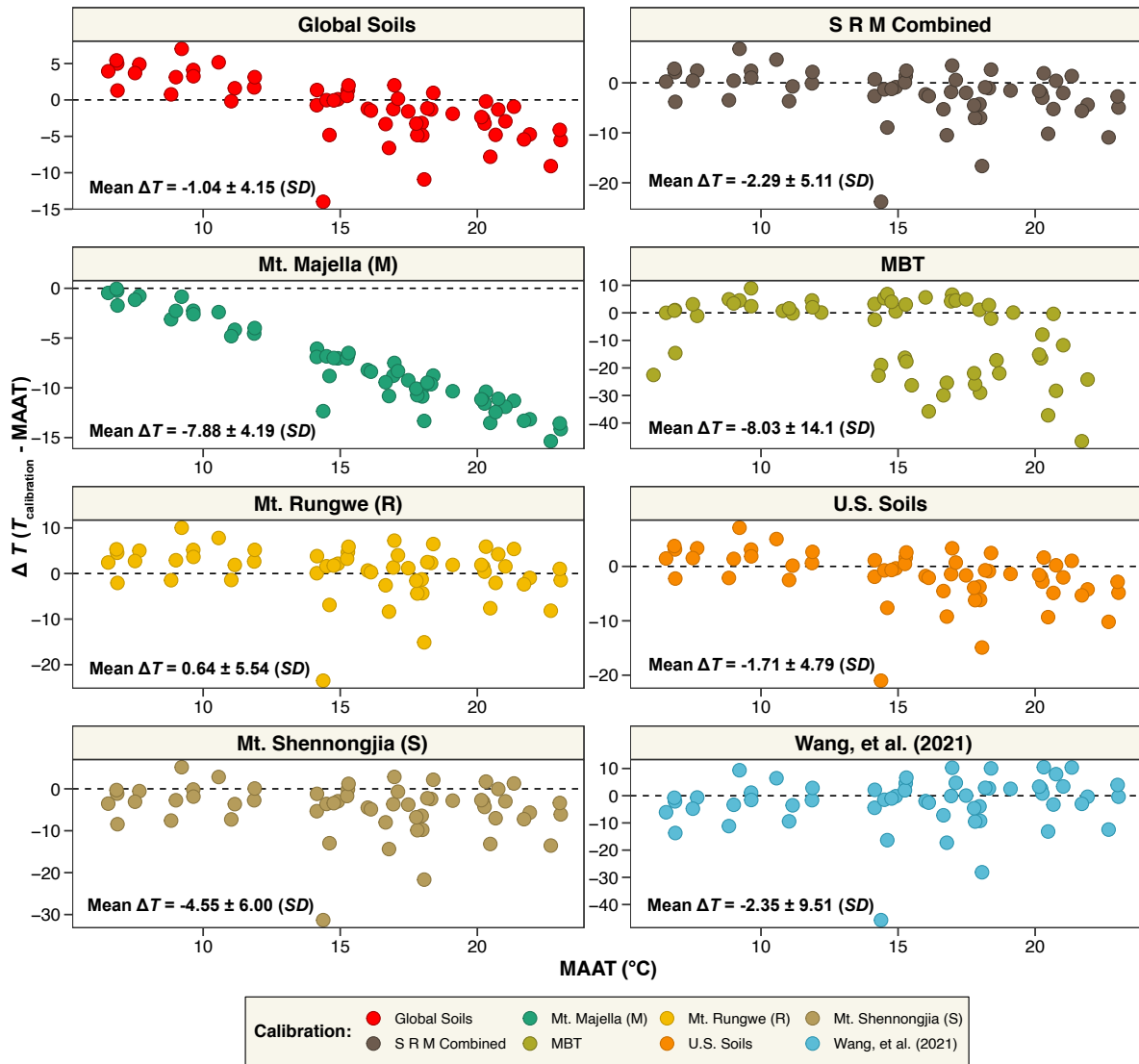


Figure 3.20: Crossplots between contemporary **MAAT** and ΔT for various RAN_{17} (and one MBT) calibrations for the compiled U.S. soils. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after [Peterse et al. \(2012\)](#), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.

The performance of the Mt. Majella, Mt. Shennongjia, and Mt. Rungwe RAN₁₇ calibrations is improved when the three transects are combined (mean $\Delta T = -2.29$) however reconstructed temperature remains highly variable (Fig. 3.20). Nevertheless, this compiled soils RAN₁₇ calibration is more comparable to mean ΔT produced by the two global soil calibrations considered. The calibrations from compiled global soils with pH<7.2 (Global Soils) and the global soils compiled by Wang et al. (2021) (Wang, et al. (2021)) generally have lower mean ΔT compared to regional calibrations despite having the latter having lower RMSEs. The 'Global Soils' RAN₁₇ calibration has a tendency to overestimate temperature in cold U.S. soils, and progressively underestimates temperature in warmer soils above ca.16 °C, which is likely due to the 'Global Soils' calibration having a slightly higher RMSE and incorporating more environmental variability by combining the soil datasets of Wang et al. (2021) and Véquaud et al. (2021b), which comprise cold soils that do not have a RAN₁₇ temperature response alone (Véquaud et al. 2021b), relative to the calibration of Wang et al. (2021). However this tailing of end member cold and warm soils is not as pronounced in the distribution of estimated temperatures from the regional RAN₁₇ U.S. soil calibration from the U.S. soils themselves, although ΔT becomes more variable and under-predicted above 15 °C as SW soils with more alkaline pH tend to occupy a higher MAAT threshold.

Although regional temperature calibrations can have a strong predictive power in the regions from which they were developed - the regional Mt. Shennongjia 3-OH FA-based RAN₁₅ calibration was capable of producing reliable Holocene temperature estimates from the HS4 stalagmite sourced from central China (Wang et al. 2018) - they rarely perform well in samples from other localities (Lei et al. 2023). Instead, global calibrations have a tendency to produce comparable means to contemporary MAAT but over and underestimate MAAT in cold and warm soils (Peterse et al. 2012), which has been widely observed amongst temperature estimates from global calibrations from both 3-OH FA and brGDGT proxies (Weijers et al. 2007, Dang et al. 2018, Véquaud et al. 2021b), and is not only restricted to soils (Lei et al. 2023).

3.4.4.2 RAN₁₇ calibrations on EC and SW soils

The RAN₁₇ calibrations are more capable of estimating [MAAT](#) in EC soils relative to SW soils (Fig. [3.21](#)), and in pH<7.2 soils relative to more alkaline pH>7.2 soils (Appendix [A10](#)). Almost all calibrations produce temperature estimates with lower mean ΔT in SW soils. [MAAT](#) from Lake Escondida soil (USA044) is consistently and extremely underestimated (Fig. [3.21](#)). The RAN₁₇ calibration developed directly from U.S. soils with pH<7.2 has lowest mean ΔT amongst estimates for EC rather than soils with pH>7.2 which is consistent with the U.S. EC being a more homogeneous environment, and pH is not the only environmental driver of 3-OH FA production that differs between the transects from our previous discussion. A benefit of the pH<7.2 threshold on the U.S. soils RAN₁₇ calibration is that mean ΔT remains very low (-0.33) amongst pH<7.2 reconstructions in U.S. soils with lowest observed standard deviation for this calibration, which incorporates some of the variable SW soils. The RAN₁₇ calibration from combined Mts. Shennongjia, Rungwe, and Majella soils has a similar performance with the global soil calibration for each U.S. transect and when compiled, which is likely because these soil transects were included in the soil dataset from which the global calibration was developed ([Wang et al. 2021](#)). The Mt. Rungwe RAN₁₇ calibration is the only to have a comparable performance between EC and SW soils, and variability is comparable to the U.S. soil calibration temperature estimates.

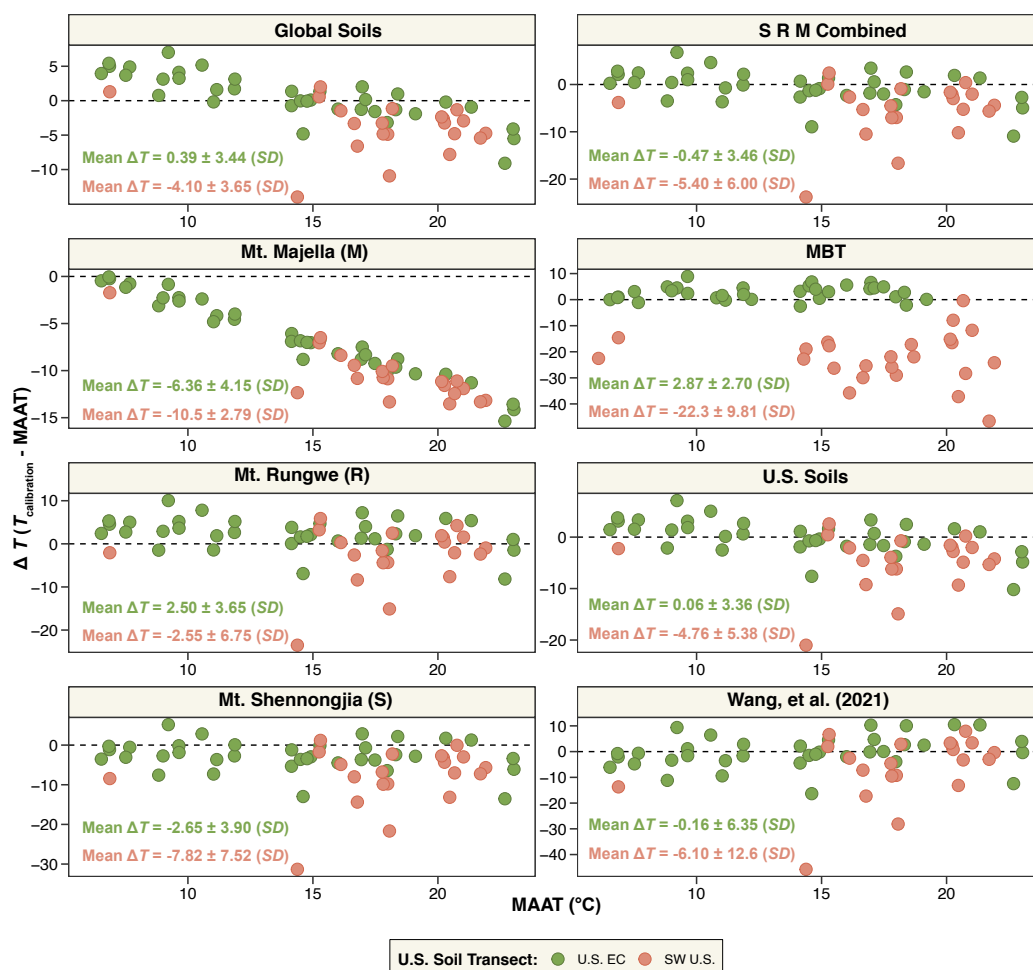


Figure 3.21: Crossplots between contemporary $MAAT$ and ΔT for various RAN_{17} (and one MBT) calibrations for the EC and SW soil transects. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after Peterse et al. (2012), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.

Interestingly, the 'Global Soils' calibration based on RAN_{17} from all soils investigated for 3-OH FAs produced lowest mean ΔT for U.S soils with $pH < 7.2$ (mean $\Delta T = 0.16$). Although the 'Global Soils' and global soil calibration of Wang et al. (2021) are both global calibrations, the former produced more comparable $MAAT$ estimates to contemporary $MAAT$ with lower variability and the RMSE is slightly higher for this calibration relative to that of Wang et al.

(2021), which is likely attributed to the substantially higher sample number of the former calibration ($n=337$) relative to the latter ($n=185$) providing a higher predictive power. The regional U.S. soil calibration is more functional in estimating MAAT in U.S. soils relative to the global soil calibration as the global calibration slightly overestimates MAAT in cold soils and underestimates MAAT in warm soils compared to the regional U.S. calibration (Fig. 3.21); ultimately both calibrations have drawbacks, despite the regional U.S. soils calibration being developed directly from the soils in question, and the global calibration has a lot of scatter particularly amongst cold soils (that was not removed by a $\text{pH}>7.2$ threshold) which is likely responsible for the poor estimation of end-member temperatures, and should be explored in future 3-OH FA calibration development studies.

3.4.4.3 Performance of RAN₁₇ and MBT calibrations

In addition to RAN₁₇, the performance of the brGDGT-based MBT temperature proxy was also assessed amongst the U.S. soils, to compare the performance of 3-OH FA-based and brGDGT-based temperature proxies. The global MBT calibration of Peterse et al. (2012) was used to generate MBT-based MAAT estimates from the SW soils; this calibration was also used by Dirghangi et al. (2013) for MBT in EC soils, and the compiled results are visualised in Figures 3.20 and 3.22.

Temperature reconstructions based on MBT are highly variable between EC and SW soils. MBT estimated temperature tracks contemporary MAAT amongst EC soils (Fig. 3.22A) and although over-estimated ($\Delta T=2.87\text{ }^{\circ}\text{C}$), ΔT is evenly distributed across the $6.5\text{--}23.0\text{ }^{\circ}\text{C}$ temperature range of the EC soil transect (Fig. 3.21). This is unlike global RAN₁₇ calibrations which over or underestimate cold and warm MAAT (case of Global Soils) or produce highly variable ΔT (case of Wang, et al. (2021)) (Fig. 3.21). For the SW soils, MBT MAAT is consistently underestimated ($\Delta T=-22.3\text{ }^{\circ}\text{C}$) and does not significantly track contemporary MAAT, although a positive trend amongst SW samples between ca.15 and $20\text{ }^{\circ}\text{C}$ suggests an alternative

temperature calibration might be appropriate for a subset of SW soils (Fig. 3.22A) . When the soils are compiled, $\Delta T = -8.03$ °C for MBT temperature estimates with extremely high variability ($SD = 14.1$ °C) (Fig. 3.20), confirming MBT has a different performance in temperature reconstruction between the EC and SW soil transects.

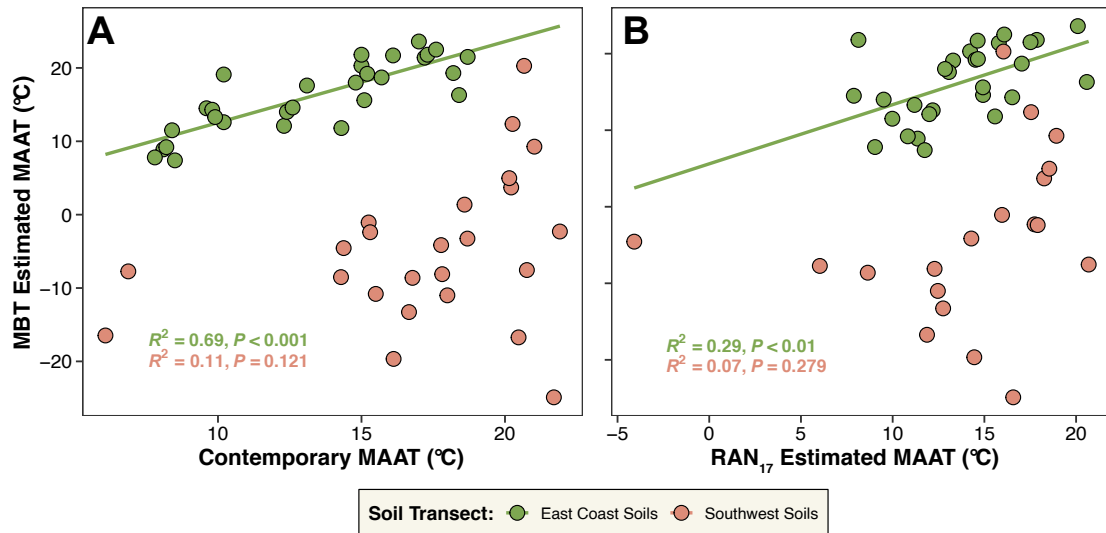


Figure 3.22: Scatterplots between visualising comparability of **A** contemporary MAAT and MBT estimated MAAT and **B** ran₁₇ estimated MAAT and MBT estimated MAAT, for U.S. soils.

This offset in ΔT between transects is consistent with previous investigations of MBT in arid soils. Dirghangi et al. (2013) observed this phenomenon across an arid soil transect from the western U.S. - soils that formed under similar environmental conditions to those of the SW U.S. - and has been observed in global soil investigations (Weijers et al. 2007, Peterse et al. 2012). The inability of MBT to function as a temperature proxy in arid Western U.S. soils was associated with uniquely high $R_{i/b}$ values compared to those in EC soils, suggesting a dominance of isoGDGT production in the arid soils. This is indicative of a shift from bacterial to archaeal GDGT production induced by moisture stress (Dirghangi et al. 2013). High $R_{i/b}$ values are also associated

with SW soils, albeit not as extreme as in western U.S. soils, and so a similar occurrence is plausible. However, although $R_{i/b}$ values are higher and more variable in SW relative to EC soils, archaeal production of isoGDGTs is not dominant for the majority of SW soils, rather brGDGT production is more mixed between bacterial and archaeal sources in SW relative to EC soils.

brGDGTs and 3-OH FAs are likely to share mutual bacterial producers in the phylum Acidobacteria as certain brGDGT compounds have been detected in small concentrations in cultured strains of Acidobacteria, and the suspected brGDGT precursor monomers were abundant (Sinninghe Damsté et al. 2018). However the complete offset in ΔT from MBT temperature estimates between soil transects was not observed for RAN_{17} calibrated using U.S. soils with $pH < 7.2$; $\Delta T = 0.08$ °C for EC soils so RAN_{17} temperature estimates are more comparable to contemporary MAAT and estimates for SW soils are less offset from MAAT ($\Delta T = -4.78$ °C) (Fig. 3.21). Furthermore, when the EC and SW transects are compiled, $\Delta T = -1.71$ °C for RAN_{17} estimates, indicating the performance of the regional U.S. RAN_{17} calibration is more consistent between EC and SW soils compared to MBT.

MBT was calibrated using a global calibration, whereas the RAN_{17} temperature estimates for comparison are derived from a regional calibration specific to the U.S soils and to the general pH range of the EC soils. However we have already noted that the global MBT calibration has a better performance in EC temperature reconstruction than the global RAN_{17} calibrations considered. As MBT and RAN_{17} are both capable for reconstructing temperature in EC soils with a degree of consistency between biomarker groups, but have incomparable performance in SW soils (Fig. 3.22B), MBT and RAN_{17} (and therefore brGDGTs and 3-OH FAs) likely have different responses to aridity. The low pH and humid conditions of the EC soils are suited for the growth of acidophilic bacterial specialists such as Acidobacteria (Thorpe 2023, Wang et al. 2021) which likely comprise taxa that produce both 3-OH FAs and brGDGTs which could account for the more consistent calibration performance amongst EC soils. However RAN_{17} could more viably

reconstruct temperature in SW soils from almost all calibrations considered in SW soils considered, which may be because (i) pH>7.2 soils that comprise a subset of SW soils were omitted from the regional RAN₁₇ U.S. calibration, or (ii) the environmental conditions of the SW U.S. may be more suited to 3-OH FA rather than brGDGT production. brGDGTs likely originate from bacterial and archaeal sources, whilst C₁₀-C₁₈ 3-OH FAs are predominantly bacterial in origin, and so it is possible that the MBT index is more susceptible to environmental changes promoting a shift from bacterial to archaeal brGDGT production, than 3-OH FAs.

3.4.5 Response of 3-OH FAs to pH in soils

Development of proxies for pH reconstruction is important as soil pH is a primary control on microbial activity in soils and nutrient availability to plants, and is an informative indicator of substantial environmental change. 3-OH FA distributions have been observed to have a more cumulative response to pH in soils, compared to temperature proxies such as RAN₁₅ and RAN₁₇. The 3-OH FA pH response in soils is widely reported as the RIAN index which employs the branching ratio Eq.(2.3) and has a negative relationship with pH in some regional soil transects (Wang et al. 2016, 2021, Huguet et al. 2019)). In U.S. soils, RIAN also produces a strong and highly significant negative linear correlation ($r^2=0.58$, $p<0.001$) (Fig. 3.10) across a pH gradient of 3.56 to 8.77. The negative correlation commonly observed between RIAN and pH reflects the positive relationship between BR and pH which indicates the fractional proportion of branched (*iso*- & *anteiso*-) 3-OH FAs increases relative to *normal*-3-OH FAs with increasingly alkaline pH, which is regarded as a bacterial membrane response to steep proton (H⁺) gradients (Wang et al. 2016, 2021). Increasing the proportion of branched 3-OH FA compounds under more alkaline pH, and therefore lower H⁺ concentration, promotes increased membrane fluidity and thus porosity, rendering the lipid bilayer more permeable to maintain chemiosmotic coupling - the maintenance of electrochemical ion gradient across the cell membrane (Lane 2017)- under ambient H⁺ concentrations (Russell 1995, Denich et al. 2003)

However, the performance of RIAN as a pH proxy is extremely variable between regional environmental transects, more so than for 3-OH FA-based temperature proxies (Véquaud et al. 2021b). A significant linear relationship between RIAN and pH has not been observed amongst many of the montane soil transects that have contributed to global soil calibrations, such as the Peruvian Andes, Mt. Rungwe, Mt. Pollino, and Mt. Majella (Véquaud et al. 2021b, Huguet et al. 2019). The limited RIAN response to pH in these soil transects confounded the development of a linear model-based global soil calibration by Véquaud et al. (2021b) who used a multiple linear regression-based calibration method instead of a linear RIAN model. Conversely, RIAN has exhibited some strong relationships with pH across soil transects from Mt. Shennongjia ($r^2=0.71$) Wang et al. (2016) and the Chilean Andes ($r^2=0.66$) (Véquaud et al. 2021b), in compiled Mt. Rungwe, Mt. Shennongjia, and Mt. Majella soils ($r^2=0.62$) (Huguet et al. 2019), and in the globally distributed soils investigated by Wang et al. (2021) ($r^2=0.66$, $n=186$); RIAN is capable of functionality at a range of scales, and tends to produce stronger pH correlations when samples from different regions are compiled together to cover wider pH ranges (Huguet et al. 2019, Wang et al. 2021). The U.S. soils also have substantial variation in response of RIAN to pH. RIAN has a weak correlation with pH in SW soils ($r^2=0.30$), and no significant functionality in EC soils (Appendix A11). Instead, RIAN has a weak but highly significant response to SMO in the EC soils ($r^2=0.30$) which is not observed in SW soils ($r^2=0.05$), where RIAN significantly responds to MAP, more so than pH. It is possible that more complex environmental circumstances mitigate RIAN than soil pH alone, not just amongst U.S. soils, but globally. As soil pH and MAP strongly covary for U.S. soils, the any confounding implications asserted on RIAN by moisture variability is difficult to deconvolve which is a limitation to a pH calibration developed from U.S. soils.

Nevertheless, the RIAN proxy does not best represent 3-OH FA variability in response to pH in U.S. soils. The wide pH and MAP ranges were identified as the primary environmental drivers of 3-OH FA variance in U.S. soils, and the most viable pH proxy was found to be $-\log(\text{RIN}_{\text{USA}})$ that employs $i\text{-C}_{13}$, $i\text{-C}_{16}$, $i\text{-C}_{17}$, $n\text{-C}_{14}$, and $n\text{-C}_{18}$ (Eq.(3.3)). The correlation between

$-\log(RIN_{USA})$ and pH is optimised by combining the U.S. soils ($r^2=0.72$, $p<0.001$) instead of grouping by regional EC and SW classification, and is presented with residuals in Figure 3.23. As the RIAN and $-\log(RIN_{USA})$ have optimum functionality amongst compiled U.S. soils, it is inferred these proxies require a wide pH range to generate variability, as observed for RIAN. Soil pH can be reconstructed from this $-\log(RIN_{USA})$ regression using the calibration outlined in Equation (3.11). Residuals of the modified RIN-pH regression have weak but significant correlations with ΔT ($r^2=0.22$, $p<0.001$), seasonality ($r^2=0.14$, $p<0.01$), and MAP which are all known to covary to an extent with pH. Residuals of RIAN also correlate with these variables, and so developing a 3-OH FA-based pH calibration from RIAN instead of $-\log(RIN_{USA})$ would not remove moisture dependence from a pH proxy in U.S. soils. Additionally, RIAN residuals are found to significantly correlate with MAAT ($r^2=0.07$, $p<0.05$) and employing $-\log(RIN_{USA})$ instead of RIAN removes this direct MAAT influence. These observations further highlight the complex influence of pedoclimate on 3-OH FA variability, and likely roots back to how a mixture of pedoclimatic drivers influence respiration and speciation amongst Gram-negative bacterial communities that are producing 3-OH FAs. It is also possible that the $-\log(RIN_{USA})$ proxy that is most responsive to pH in U.S. soils also pertains to the chemiosmotic coupling adaptation.

$$SoilpH = 8.43 - 5.25 \times -\log(RIN_{USA}) \quad (RMSE = 0.86, n = 59) \quad (3.11)$$

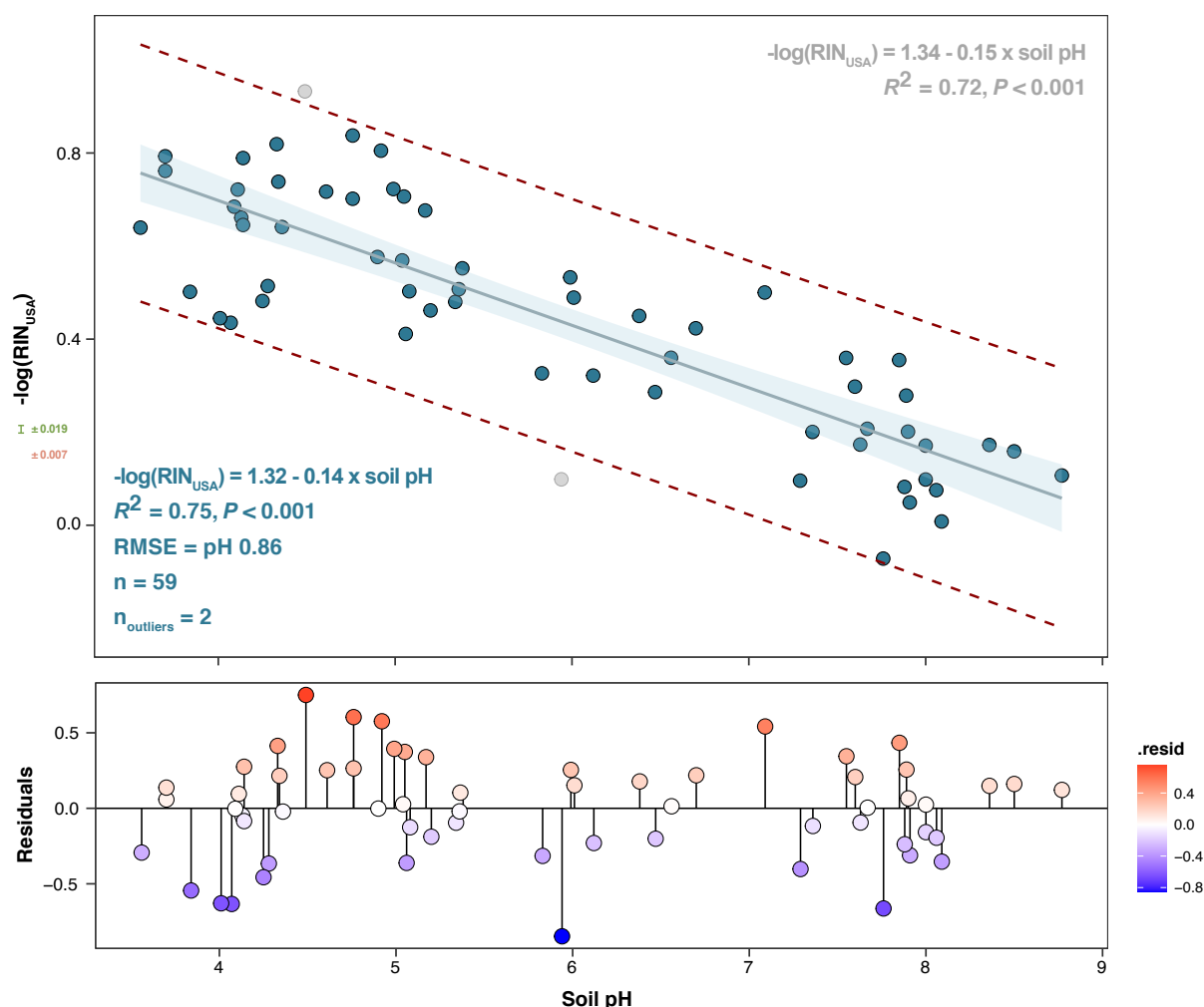


Figure 3.23: $-\log(\text{RIN}_{\text{USA}})$ -pH regression from U.S. soils, with residuals visualised.

A 3-OH FA-based pH proxy developed by a linear regression model other than RIAN has not been explored in soils, and $-\log(\text{RIN}_{\text{USA}})$ produces a stronger correlation with pH amongst global soils ($r^2=0.55$, $p<0.001$) relative to RIAN ($r^2=0.46$, $p<0.001$). $-\log(\text{RIN}_{\text{USA}})$ produces a better correlation with pH because scatter is substantially reduced amongst U.S. soils, which produce tighter $-\log(\text{RIN}_{\text{USA}})$ residuals in response to pH compared to RIAN. This result is expected because $-\log(\text{RIN}_{\text{USA}})$ was developed directly from first order 3-OH FA variability between U.S. soils. However $-\log(\text{RIN}_{\text{USA}})$ also produces a stronger correlation with pH globally because

$-\log(\text{RIN}_{\text{USA}})$ in Mt. Majella, Mt. Pollino, and Mt. Shegyla soils is more responsive to pH than RIAN (Fig. 3.24). RIAN remains the most descriptive proxy for pH-driven 3-OH FA variability in Mt. Shennongjia and Mts. Lautaret-Bauges (Fig. 3.24). pH variability is not improved by $-\log(\text{RIN}_{\text{USA}})$ (difference is negligible) in global soil transect soils from Wang et al. (2021), and neither RIAN or $-\log(\text{RIN}_{\text{USA}})$ describe regional 3-OH FA variance in response to pH in soil from the Chilean Andes, Mt. Rungwe, or the Peruvian Andes (Fig. 3.24). Therefore, like RIAN, $-\log(\text{RIN}_{\text{USA}})$ is also highly variable between regional soil transects, but has the potential to improve soil pH reconstruction in some soils.

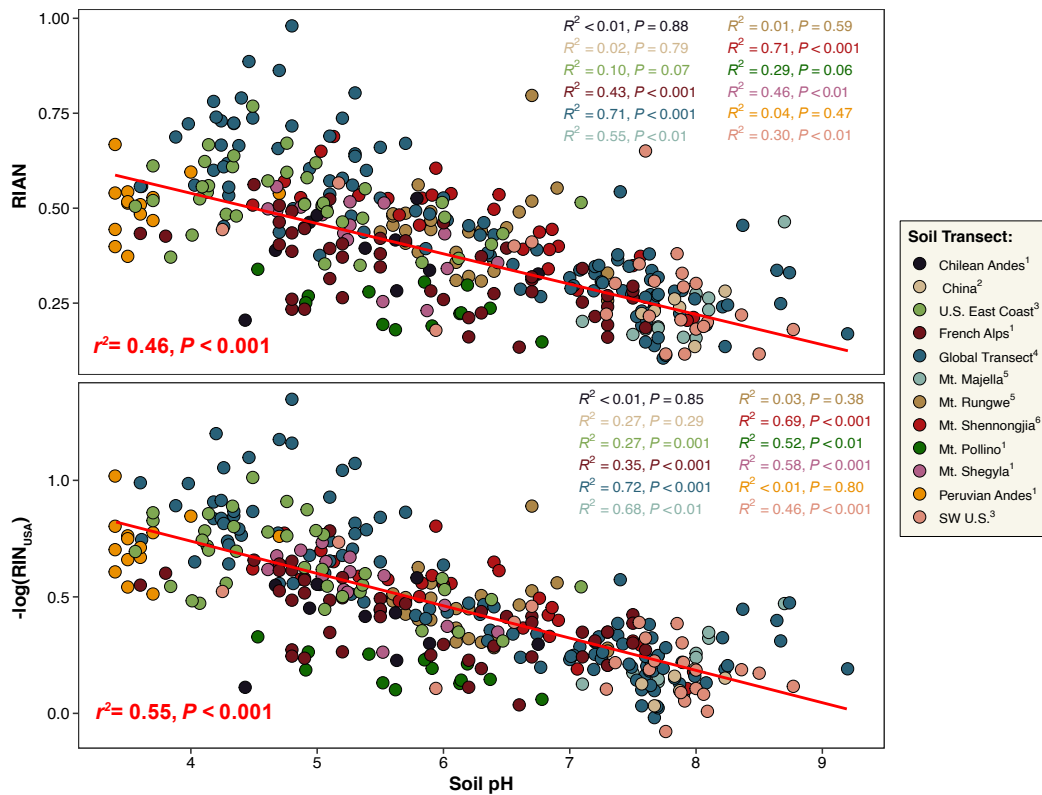


Figure 3.24: Variability of A RIAN and B $-\log(\text{RIN}_{\text{USA}})$ with pH in globally distributed soils. Soil data after Véquaud et al. (2021b)¹, Yang et al. (2021)², present study³, Wang et al. (2021)⁴, Huguet et al. (2019)⁵, Wang et al. (2016)⁶.

brGDGTs have also been observed to undergo structural changes in response to pH. Specifically, the abundance of brGDGT compounds with differing numbers of cyclopentane moieties can vary with pH, and this degree of cyclisation is expressed as the CBT pH proxy (Eq.2.15). CBT does not vary in response to pH amongst either EC or SW soils as individual transects; the significant correlation between CBT and pH in SW soils is the result of a high degree of scatter in some soils (Fig. 3.25A). However CBT does produce a weak but significant correlation with pH when the U.S. soil transects are combined ($r^2=0.12$, $p<0.05$) (Fig. 3.25A), suggesting that a wide pH range is required to induce CBT functionality, which was also observed from the optimised correlations of RIAN and $-\log(\text{RIN}_{\text{USA}})$ with pH over a wider pH range. However, estimated soil pH from the brGDGT-based CBT and 3-OH FA-based $-\log(\text{RIN}_{\text{USA}})$ proxies do not significantly correlate when regressed against each other ($r^2=0.02$, $p=0.34$) (Fig. 3.25B), which is not surprising as $-\log(\text{RIN}_{\text{USA}})$ has a higher degree of variance in response to pH ($r^2=0.75$) compared to CBT ($r^2=0.12$), and CBT estimated pH was calculated using the global CBT calibration of Weijers et al. (2007) whilst that for $-\log(\text{RIN}_{\text{USA}})$ was the regional U.S. calibration, and developed using PCA specifically from the U.S. soils. Furthermore, the structural responses and biological producers between the two biomarker groups are different; that of brGDGTs is based on cyclisation and brGDGTs have archaeal and bacterial producers, whilst a 3-OH FA pH response is based on methyl branching which is a temperature response amongst GDGTs. Overall, CBT and $-\log(\text{RIN}_{\text{USA}})$ do not have comparable functionality as pH proxies in U.S. soils likely because they are based on different calibrations that were developed from different soil datasets (CBT from a global calibre, regional calibre for $-\log(\text{RIN}_{\text{USA}})$), the lipodomic response to pH initiated by bacteria is not necessarily the same as that of Archaea and also intrinsically differ between the biomarker groups, and therefore differing proxy performance between 3-OH FAs and brGDGTs is expected.

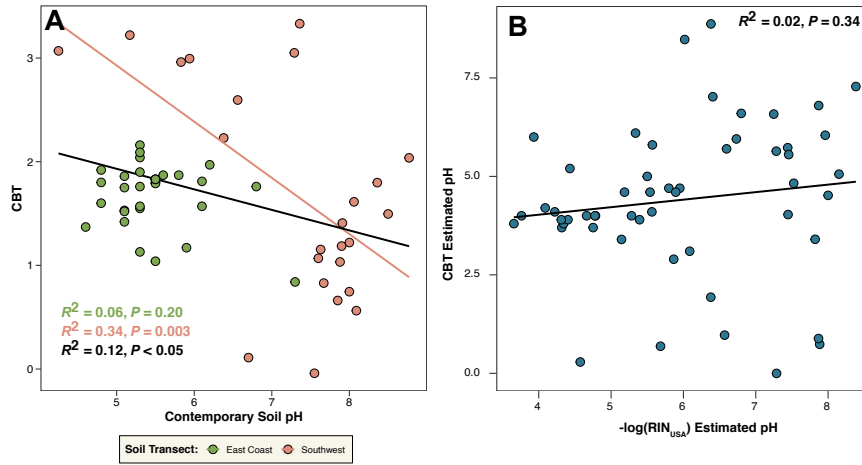


Figure 3.25: Crossplots of: **A** contemporary pH and the brGDGT-based CBT index, **B** -log(RIN_{USA}) estimated pH and CBT estimated pH from U.S. soils

The difference in performance of -log(RIN_{USA}) and CBT is particularly emphasised when considering ΔpH between proxy estimated and contemporary pH (Fig. 3.26). ΔpH distributions between CBT and -log(RIN_{USA}) are homogeneous in neither mean ($t = -7.09(74)$, $p < 0.001$) or variance ($f = 4.81$, $p < 0.001$), and pH is severely underestimated for all but 5 soils from CBT across the entire pH range of U.S. soils (mean $\Delta pH = -1.93$) compared to -log(RIN_{USA}) which produces ΔpH distributed around 0 (mean $\Delta pH = 0.01$) (Fig. 3.26). Figure 3.26 further reveals that -log(RIN_{USA}) has a tendency to overestimate pH in more acidic soils, and underestimate pH in more alkaline soils ($r^2 = 0.19$, $p < 0.001$), however the low correlation coefficient suggests this is a weak effect. Therefore, -log(RIN_{USA}) is capable of producing accurate pH estimates at least regionally for U.S. soils over a large pH gradient.

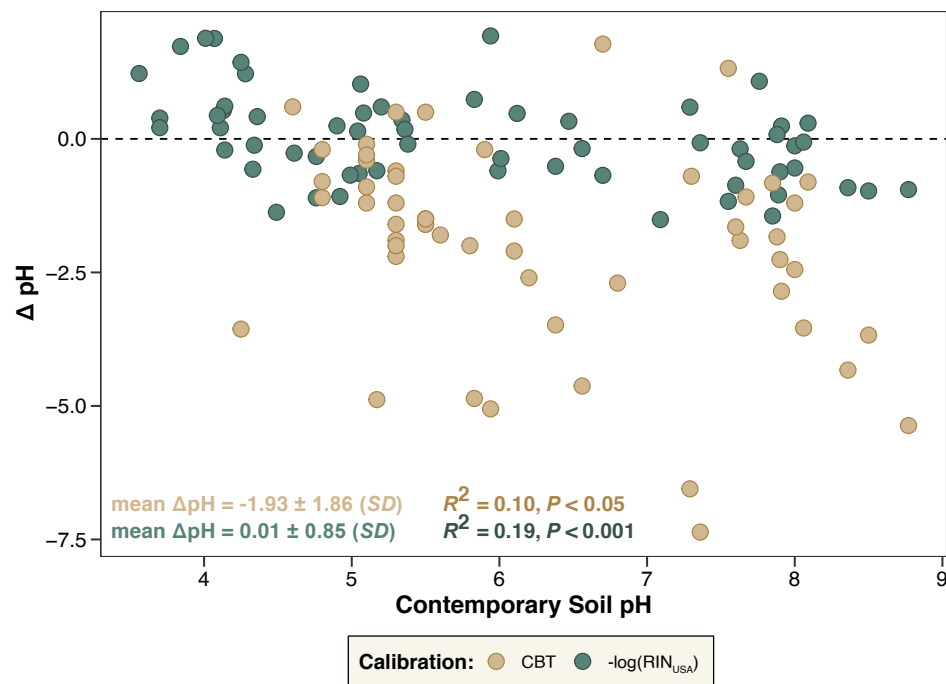


Figure 3.26: Crossplot visualising 3-OH FA and brGDGT-based pH proxy performance by comparing contemporary pH with ΔpH (proxy estimated pH - contemporary pH).

3.4.6 Functionality of 3-OH FA-based environmental proxies in arid environments

Omitting arid soils or those with an alkaline pH from a U.S. RAN₁₇ temperature calibration conveys that RAN₁₇ does not function as a temperature proxy in soils of this condition, at least regionally amongst U.S. soils. Some of the environmental variables that are associated with aridity (low [MAP](#), [SMO](#), seasonality, and high [VPD](#) and ΔT) are known to be influential on the performance of the brGDGT-based MBT'_{5ME} index (and earlier MBT and MBT' iterations) as a temperature proxy ([Peterse et al. 2012](#), [Dirghangi et al. 2013](#), [Dang et al. 2016](#)); the confounding influence of arid climate regimes on microbial lipid biomarker biosynthesis and their environmental proxies is not a new concept and is being investigated rigorously in the calibration of brGDGT temperature proxies.

In drier soils, moisture in the form of **SMO** or **MAP** can become a critical environmental stress on microbial respiration (Dang et al. 2016, Menges et al. 2014, Wildung et al. 1975), a notion which is supported amongst U.S. soils by the weak response of RAN_{17} to **MAP** and **SMO**. A transient **MAP** threshold below which MBT, MBT' or MBT'_{5ME} response to temperature becomes confounded has been reported from several brGDGT calibration studies. Dang et al. (2016) identified **MAP**=500 mm as a threshold between **MAP** (**MAP**<500 mm) and **MAAT** (**MAP**>500 mm) response in the MBT' temperature proxy. Dirghangi et al. (2013) reported a limiting **MAP** threshold of 700 mm below which MBT severely underestimated **MAAT**. Following the separation of 5- and 6-methyl brGDGTs (De Jonge et al. 2014a), a **MAP** and pH dependence of 6-methyl brGDGTs as reported, leading to the development of MBT'_{5ME} which omits the confounding 6-methyl compounds. The IR_{6ME} index was developed by Dang et al. (2016) to distinguish soils with a **MAP** response $IR_{6ME}>0.5$ from those with a **MAAT** response $IR_{6ME}<0.5$, and omitting soils with $IR_{6ME}>0.5$ improved the global MBT' regression with **MAAT** but removing soils with **MAP**-dependent brGDGT distributions. The IR_{6ME} proxy is now used as a prerequisite to assess potential **MAP** influence on MBT' and MBT'_{5ME} prior to application as a temperature proxy, however the ratio between 5-methyl and 6-methyl brGDGTs also has a complex association with soil pH (Yang et al. 2015).

The confounding influence of **MAP** and soil pH on brGDGT-based temperature proxies remains under investigation, and has led to the development of a range of MBT, MBT', MBT'_{5ME} temperature calibrations from regional to global scales. Global soil calibrations of biomarker-derived proxies cover a wider range of environmental variability than those of a regional calibre (Weijers et al. 2007, Peterse et al. 2012, De Jonge et al. 2014a), however have the restricted ability to reconstruct temperature in environmentally specialised environments such as arid regions, or those with alkaline pH. The MBT'_{5ME} proxy that has been globally calibrated with contemporary **MAAT** (De Jonge et al. 2014a) severely overestimated **MAAT** in the dry and cold soils of the Qinghai-Tibetan Plateau, leading to the development of a regional brGDGT-based

temperature proxy MBT_{5/6} that incorporates two temperature-dependent 6-methyl brGDGT compounds (Ding et al. 2015). Yang et al. (2014) used a stepwise selection method instead of linear regression to develop a brGDGT-based temperature proxy in arid soils of central China, where MBT values were low. Therefore, regional temperature calibrations may best suit applications in environmentally specialised arid regions, however rarely function outside of the region from which they were developed.

Moisture or pH stress can also lead to shifts in bacterial community structure which calls into question whether variability of RAN₁₇ and RIN_{Mod} in respective response to MAAT and soil pH are the result of phenotypic plasticity of 3-OH FA distributions, bacterial community change, or a mixture of both. To address the bacterial communities producing 3-OH FAs in U.S. soils, community composition at the Phylum level is similar between EC and SW U.S. soils, and is strongly and predominantly structured by soil pH (Thorpe 2023). The phyla Acidobacteria and Verrucomicrobia decline in abundance with increasing soil pH, whilst Bacterioidetes, and Gemmatimonadetes become more abundant with increasing soil pH (Thorpe 2023) (Fig. 3.27A), which has been widely observed amongst soil pH gradients (Bartram et al. 2014, Jones et al. 2009). Bacteroidetes ($r^2=0.51$, $p<0.001$) and Gemmatimonadetes ($r^2=0.45$, $p<0.001$) are the phyla most responsive to pH (Fig. 3.27A), and none of the four bacterial phyla in question have any significant correlation with MAAT amongst the U.S. soils (Fig. 3.27B). *a*-C₁₇ and *n*-C₁₇, the 3-OH FA compounds employed in the RAN₁₇ temperature proxy, both have extremely weak correlations with soil pH in these soils ($r^2=0.08$, $p=0.03$), and no significant relationship with MAAT ($r^2=0.08$, $p=0.03$). We have proposed that RAN₁₇ is more functional as a temperature proxy in U.S. soils with pH<7.2, and amongst these soils *a*-C₁₇ has a significant but weak negative correlation with MAAT ($r^2=0.23$, $p<0.01$) and not with pH. In this subset of soils, the four bacterial phyla that are predominantly associated with soil pH across the whole dataset, remain significantly associated with pH, however the response of Gemmatimonadota to pH becomes less substantial ($r^2=0.25$), and Bacteroidota has the most responsive variability to pH ($r^2=0.56$),

suggesting that Gemmatimonadota have a preference for higher pH pedoenvironment compared to Bacteroidota (Fig. 3.27C). Interestingly, in U.S. soils with pH<7.2 Acidobacteria is the only of the four bacterial phyla to have a significant positive relationship with MAAT ($r^2=0.26$, $p<0.001$) (Fig. 3.27D). Acidobacteria does not have any response to MAAT when soil pH>7.2, therefore it is possible that Acidobacteria are driving the functionality of the RAN₁₇ proxy in soils with pH<7.2.

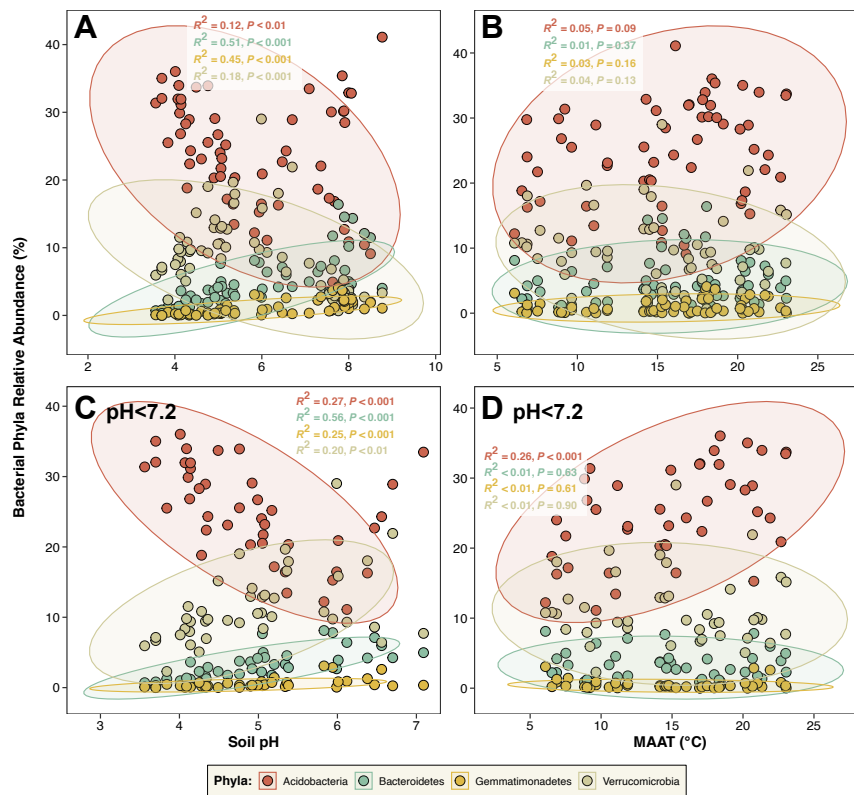


Figure 3.27: Crossplot matrix visualising the variability in fractional abundance of the Bacterial phyla identified as most responsive to pH by Thorpe (2023), specifically with pH (A) and MAAT (B), and same for soils with pH<7.2 (C,D). Comprises unpublished bacterial data from Thorpe (2023).

Acidobacteria have been shown to shift in bacterial community structure at 14 °C, albeit in geothermally warmed soils with a substantial T_{Soil} offset from MAAT, which is known to

regionally confound the functionality of MBT_{5ME} as a temperature proxy (De Jonge et al. 2019). However, a recent Bacteroidetes culture experiment using growth temperature as an independent variable revealed that Bacteroidetes are capable of increasing the relative proportion of *iso*- or *anteiso*- relative to *normal*-3-OH FAs under lower growth temperature (Hellequin et al. 2023). An accumulation of low melting-point fatty acids (branched, unsaturated, short chain) has also been reported from a number of bacterial culture experiments (Nordström & Laakso 1992, Kaneda 1991, Prakash et al. 2015). Ultimately, the bacterial producers of 3-OH FAs are capable of undergoing lipidomic restructure in response to air temperature, however there is also evidence that the RAN₁₇ temperature proxy is associated with Acidobacteria in soils where pH<7.2 (Fig. 3.28A & B) which should be explored further using 16S RNA sequencing and further culture experiments. -log(RIN_{USA}) is likely the result of combined community variability in response to a large pH gradient; *n*-C₁₄, *n*-C₁₈, and *i*-C₁₄ are associated with several bacterial phyla but specific controls are uncertain.

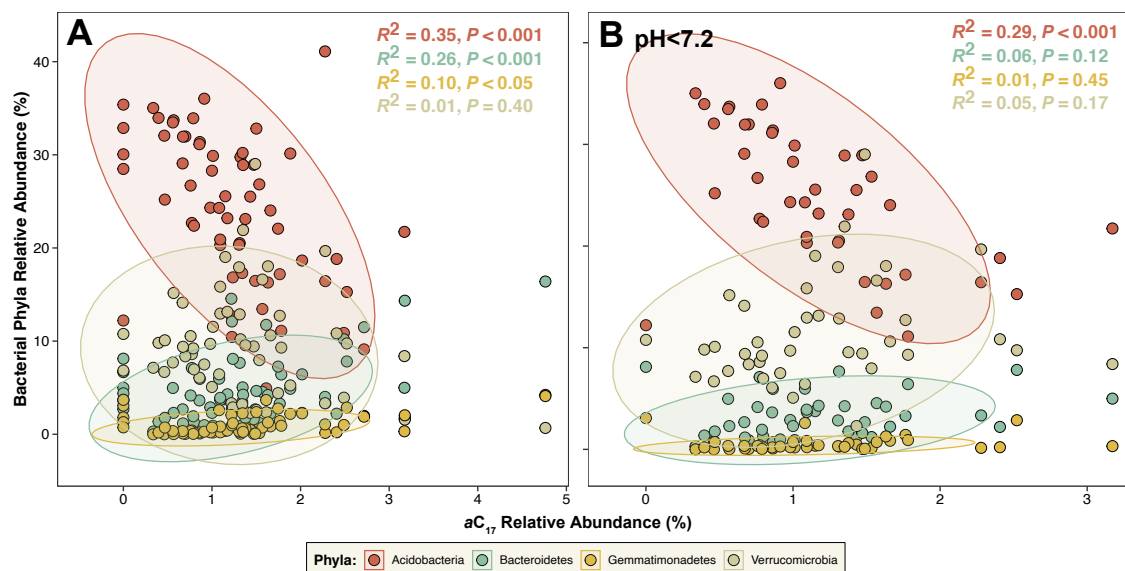


Figure 3.28: Crossplots visualising *a*-C₁₇ variability in response to selected Bacterial phyla amongst **A** U.S. soils and **B** U.S. soils with pH<7.2. Comprises unpublished bacterial data from Thorpe (2023).

3.5 Conclusions

3-OH FA distributions from two contemporary U.S. soil transects that span environmental gradients in soil pH, MAP, and MAAT were investigated. pH and MAP were the primary environmental drivers of 3-OH FA variability in the soils, however RAN₁₇ was most responsive to MAAT which initially appeared restricted to humid EC soils. An investigation of the environmental drivers of RAN₁₇ revealed the RAN₁₇ -MAAT regression was optimised in pH<7.2 soils. Omitting pH>7.2 soils from a globally compiled soil dataset did not improve RMSE of a global temperature calibration, so we developed (i) a regional RAN₁₇ linear regression model from pH<7.2 U.S. soils and (ii) a global RAN₁₇ linear regression model compiling the soil datasets of Wang et al. (2021), Véquaud et al. (2021b), and U.S. soils of the present study. The regional U.S. calibration demonstrated best potential at producing MAAT estimates in U.S. soils. Our U.S.-based RAN₁₇ calibration is a regional calibration suited to humid environments with soil pH<7.2, and has potential to reconstruct temperature in these environments. To test for applicability of this regional calibration outside of the U.S. and in temporal records, a range of regional and global-scale RAN₁₇ calibrations should be applied and compared, in addition to the well-established MBT'_{5ME} proxy.

-log(RIN_{USA}) best represented a 3-OH FA pH response instead of RIAN. Although pH strongly co-varied with MAP in U.S. soils, -log(RIN_{USA}) produced variability across both EC and SW soil transects and had optimum functionality when the soil transects were combined into one dataset, over a wider pH range. -log(RIN_{USA}) produced an improved regression with MAAT compared to RIAN. We developed a regional U.S. calibration based on -log(RIN_{USA}) using a linear regression model which produced pH estimates distributed around $\Delta T=0$, compared to CBT which consistently underestimated pH.

Arid environments are complex for biomarker proxy development, as low soil moisture and often associated alkaline pH invoke stress on microbial producers. The Gram-negative bacterial

producers of 3-OH FAs are known to undergo changes in community structure in response to temperature and pH (particularly Acidobacteria), however culture experiments have also revealed that Gram-negative bacterial strains can invoke a lipodomic plastic response producing fatty acids with lower melting points (i.e. branched 3-OH FAs) in response to ambient growth temperature. The $\text{pH} < 7.2$ threshold for RAN_{17} functionality may be associated with Acidobacteria diversity. The 3-OH FAs employed in $-\log(\text{RIN}_{\text{USA}})$ are associated with all bacterial phyla that respond to pH. Overall, the two proxies calibrated in this study are capable of producing temperature and pH estimates for U.S. soils despite the strong environmental variability from which the soils are derived. This study has started investigation of environmental drivers that can confound the functionality 3-OH FA-based proxies, however the pH 7.2 threshold did not remove RAN_{17} scatter in cold soils despite improving the correlation coefficient of RAN_{17} -[MAAT](#) regression amongst globally compiled soils, which should be investigated in future studies.

Ultimately, we have developed a regional RAN_{17} calibration for temperature reconstruction from U.S. soils that has a soil pH threshold of < 7.2 (calibration dataset comprises humid and acidic soils) to accommodate for bacterial community changes induced by acid-alkaline pH changes and humid-arid moisture changes. The performance of our U.S. RAN_{17} calibration was tested in arid and humid soils alongside the brGDGT-based MBT proxy, and was able to reconstruct temperature in arid soils where MBT-estimated temperatures were severely under-predicted, despite the acidic pH range of the calibration dataset. Our $-\log(\text{RIN}_{\text{USA}})$ soil pH calibration is based on a novel proxy designed in this study, and has demonstrated a stronger potential than RIAN to reconstruct pH from soils globally. The performance of these new 3-OH FA-based temperature and pH calibrations should be investigated amongst temporal sediments in future studies.

**CHAPTER 4: DEVELOPMENT OF 3-HYDROXY FATTY ACID-BASED PALAEOTHER-
MOMETRY IN LACUSTRINE ENVIRONMENTS ACROSS THE SOUTH-
WESTERN UNITED STATES**

4.1 Introduction

Lake sediments are terrestrial palaeoclimate archives of high regard due to their ability to accumulate continuous incremental biomolecular records, precise chronologies, and their global distribution (Wang et al. 2018, Raberg et al. 2021). Lakes are also capable of undergoing sedimentation on the scale of tens to hundreds of thousands of years, and to a million years in rare cases. However, exposed palaeo lake sediments (e.g. Oligocene-Miocene Kailas Formation, Tibet (DeCelles et al. 2018)) provide lake sediments on deep geological time scales for palaeoclimate reconstruction. As a result, the distributions of a number of microbial lipid palaeoclimate biomarkers have been investigated in lakes at a range of scales (intra-lacustrine to global) for the development of lacustrine-specific biomarker-based temperature proxies and calibrations. $U^{K'}_{37}$ and TEX_{86} are widely applied palaeothermometer proxies in the marine realm for reconstruction of SST (Müller et al. 1998, Kim et al. 2008), respectively derived from long-chain alkenones and isoprenoidal (i)GDGTs. However the application of both in lacustrine environments is limited; $U^{K'}_{37}$ has not been widely calibrated because long-chain alkenones are rarely abundant in lakes (Toney et al. 2010), and TEX_{86} application is generally restricted to deep tropical lakes (Tierney et al. 2010) where the primary source of iGDGTs is thought to be aquatic Thaumarchaeota archaea as is the case in marine environments.

Branched (br)GDGTs are ubiquitous in lacustrine environments, however, and the brGDGT-based MBT'_{5ME} temperature proxy has been calibrated regionally (Russell et al. 2018, Zhao et al. 2023, Bauersachs et al. 2024) and globally (Martínez-Sosa et al. 2021, Raberg et al. 2021) in lacustrine environments due to the ubiquitous production of brGDGTs in environmental media. MBT'_{5ME} describes a decrease in brGDGT methylation at higher MAAT which is thought to represent a physiological homeoviscous adaptation response of brGDGT microbial producers to growth temperature, which occurs in cell membranes. Prior to the development of MBT'_{5ME} , its previous iterations MBT and MBT' had a long history of pH-dependence (Weijers et al. 2007,

Peterse et al. 2012), which was deconvolved by the development of a novel analytical method promoting the separation of brGDGT isomers with a methyl group in the C₅ and C₆ positions which previously coeluted (De Jonge et al. 2014a, Hopmans et al. 2016). 6-methyl brGDGTs were identified as the pH-dependent confounding component of MBT and MBT', with the brGDGT temperature response occurring in 5-methyl compounds (De Jonge et al. 2014b). MBT'_{5ME} has demonstrated significant variability in response to MAAT in both pedogenic and lacustrine environments, however lacustrine-specific temperature calibrations are required because soil-based calibrations have a tendency to underestimate temperature in lakes (Tierney et al. 2010), which is likely related to in-situ lacustrine production of brGDGTs by a different community of source organisms to those in soils (Tierney et al. 2010, Loomis et al. 2011).

Identification of brGDGT source organisms remains an issue for the application of MBT'_{5ME} in any environment; although brGDGTs are thought to be both Archaeal and bacterial in origin (Weijers et al. 2006) and have been associated with a certain strains of Acidobacteria which produce lipids with a similar structure to brGDGTs (Sinninghe Damsté et al. 2011), temperature and pH-induced changes in Acidobacterial community in soils are known to have a confounding influence on MBT'_{5ME} temperature reconstructions (De Jonge et al. 2019). Additional uncertainty concerning MBT'_{5ME} calibration in lacustrine environments arises from how the growth temperature brGDGT producers are responding to is best quantified in calibrations; MAAT is the conventional temperature that biomarker-based temperature proxy transfer functions are based on, and is assumed to be representative of lake surface water temperature. However recent studies have reported warm season temperature bias in lacustrine MBT'_{5ME} temperature reconstructions in mid to high latitude lakes (Shanahan et al. 2013, Sun et al. 2011, Bauersachs et al. 2024), leading to the investigation of average temperature of months above freezing (T_{MAF}) in recent MBT'_{5ME} calibrations (Bauersachs et al. 2024, Raberg et al. 2021, Martínez-Sosa et al. 2021) that is considered to be more representative of growth temperature for brGDGT microbial producers. T_{MAF} accounts for warm season bias with increasing latitude because T_{MAF} is equivalent to

[MAAT](#) in tropical environments, and represents an increasingly limited temporal window with increasing latitude. Furthermore, MBT'_{5ME} is not capable of reconstructing temperature in all lacustrine environments ([Dang et al. 2018](#)), and the influence of heterogeneous environmental variability amongst lakes (salinity, [DO](#), turbidity, lake depth, thermal regime, conductivity, nutrient concentrations) remains poorly understood.

The uncertainties associated with brGDGT-based temperature proxies have promoted the exploration of temperature proxy development from new lipid biomarker groups, and one such group undergoing extensive development is 3-OH FAs. 3-OH FAs are bacterial membrane lipids that have been investigated in a number of sedimentary environments in recent years as a novel organic biomarker group for the development of palaeoenvironmental proxies. C₁₀ to C₁₈ 3-OH FAs are known to originate from the outer cell membranes of Gram-negative bacteria ([Alexander & Zähringer 2002](#), [Klok et al. 1988](#), [Wollenweber et al. 1982](#)), and thus the biological role of 3-OH FAs and the biochemical circumstances of their origin are well understood. Investigation of 3-OH FA distributions in soils across a montane soil transect from Mt. Shennongjia (China) revealed the ratio of *anteiso*- to *normal*- C₁₅ and C₁₇ 3-OH FAs - expressed as the respective proxies RAN₁₅ and RAN₁₇ - produced significant variability in response to [MAAT](#). 3-OH FA distributions have since been extensively investigated in soils across regional ([Huguet et al. 2019](#), [Véquaude et al. 2021a](#)) and global ([Wang et al. 2021](#), [Véquaude et al. 2021b](#)) transects, across which RAN₁₅ and RAN₁₇ have had a highly variable response to [MAAT](#) likely due to regional variability in bacterial community structure, leading to the recommendation of regional calibrations of RAN₁₅ and RAN₁₇. 3-OH FAs have additionally been investigated in marine settings which led to the development of RAN₁₃ ([Yang et al. 2020](#)) and RAN_S ([Dong et al. 2023](#)) as 3-OH FA-based proxies for [SST](#) reconstruction in the marginal seas of China. Additionally, 3-OH FA distributions have been reported from a regional lake transect across central China from which calibrations based on RAN₁₃ and RAN₁₇ were developed for [MAAT](#) reconstruction ([Yang et al. 2021](#)), and thus 3-OH FA distributions vary spatially between sediment types.

However, lacustrine 3-OH FA distributions are relatively unexplored despite the important stature of lake sediments for terrestrial palaeoclimate reconstruction. The applicability of RAN₁₃ and RIN₁₇ outside the regional variability of the Chinese lake transect of [Yang et al. \(2021\)](#) is unknown, as is the influence of the complex environmental heterogeneity of lacustrine environments on 3-OH FA distributions. The extent to which confounding implications from lacustrine MBT'_{5ME} calibrations, such as seasonal bias and in-situ aquatic production, is also uncertain. We therefore investigate 3-OH FA distributions across a SW U.S. lake transect, from which co-sampled catchment soils were also collected, to determine whether any 3-OH FA temperature variability identified amongst U.S. lake sediments is consistent with that of the Chinese lake sediments, and also how consistent 3-OH FA distributions and temperature variability is between lake sediments and their catchment soil counterparts presented in Chapter 3. This study will also compare temperature variability between brGDGTs and 3-OH FAs.

4.2 Study area

Surface sediments and catchment soils were co-sampled from 26 lake catchments distributed between Phoenix, Arizona, and Houston, Texas (Fig. 4.1), broadly following the sampling transect of [Hou et al. \(2008\)](#) who sampled surface sediments and lake water from 32 SW lakes to investigate *D/H* ratios of sedimentary leaf waxes in lacustrine settings, particularly across this environmentally diverse region. The SW U.S. spans diverse gradients in [MAAT](#) (6.1-21.9 °C) and [MAP](#) (245-1474 mm), elevation attributed to the basin and range topography of Western Texas through to Arizona, and a lake sediment pH of 5.90-8.80. The SW U.S. transect expands the temperature range of lacustrine environments investigated for 3-OH FAs, from the ca. 5.0-17.0 °C [MAAT](#) range of the Chinese lake transect investigated for 3-OH FAs by [Yang et al. \(2021\)](#). The atmospheric and pedoclimatic conditions across the SW U.S. soil transect were previously reported in Chapter 3, and portrayed a generally hot semiarid to arid environment promoting evaporation of moisture from soils into the atmosphere instead of precipitation, and capturing a range of climate

regimes (Warm Temperate, Arid, Cold Arid) across basin and range topography. Additionally, the SW U.S. lakes comprised warm temperatures between 19.2-33.4 °C at the time of sampling (Summer), were generally alkaline at the time of sampling (surface water pH 7.70-9.78), and vary in depth between 0.2 and 40 m, and surface area between 0.02 and 51.6 km² (Appendix [E10](#)).

Refer to Chapter 2 for a detailed overview of methods used to generate and analyse data for this present chapter.

We calibrate all temperature proxies throughout this Chapter to [MAAT](#) for the SW U.S. lakes based on the assumption that [MAAT](#) is coupled with lake surface water temperature, which is a routine assumption for lacustrine temperature proxies derived from organic biomarkers ([Yang et al. 2021](#), [Bauersachs et al. 2024](#), [Raberg et al. 2021](#)). It is currently unknown whether lacustrine 3-OH FAs are produced in the water column, in the sediment, or a mixture of both. Using MAAT as a calibration temperature will likely bias temperature estimates to 3-OH FA production in surface waters.

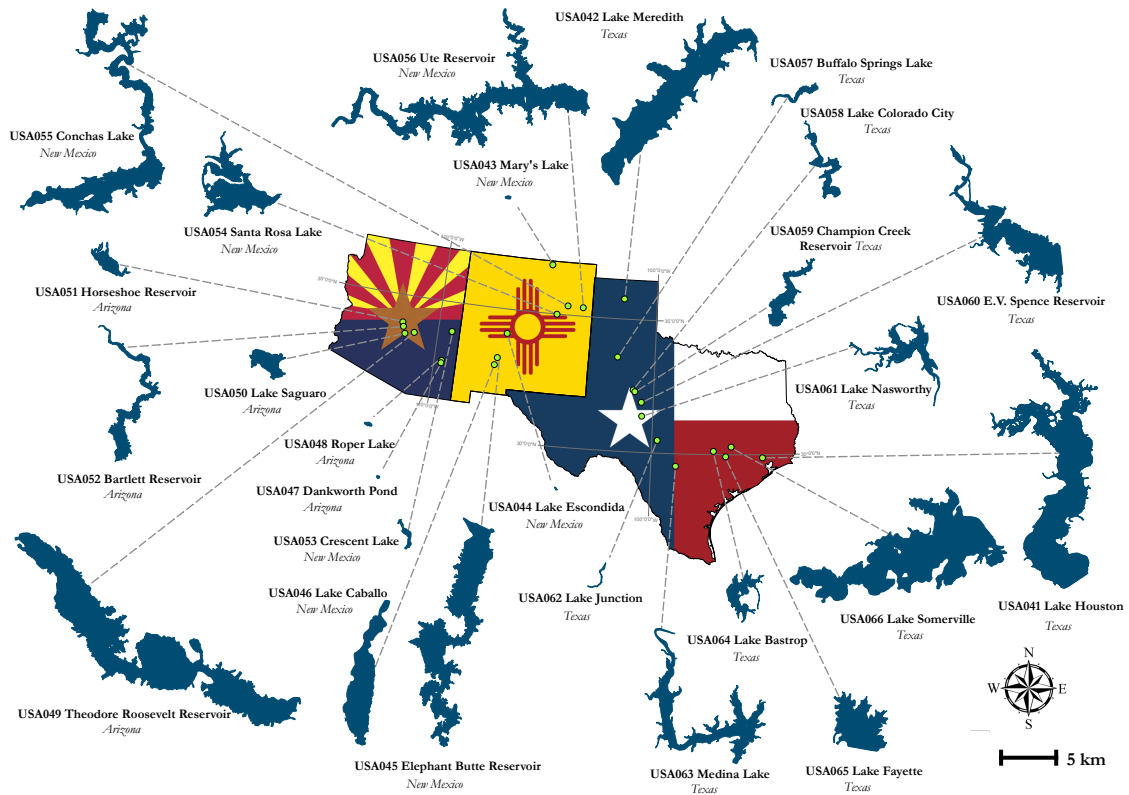


Figure 4.1: SW U.S. lakes. All lakes are scaled to the 5 km scale bar. This scale bar excludes map of U.S. States

4.3 Results

4.3.1 Environmental diversity across the Southwestern U.S.

Surface water temperatures (T_{Water}) amongst the SW lakes range between 19.2 °C (Crescent Lake, AZ) and 33.4 °C (Lake Fayette, TX) and have highly significant covariance with $MAAT$, seasonal air temperatures, and T_{Soil} (Fig. 4.2A) despite being point measurements recorded in summer months that are more transient across an annum than reported in this study. This point measurement nature may explain in part why T_{Water} does not significantly covary with seasonality as was observed for air and soil temperatures. Lake sediment pH ranges between pH 5.76 and 8.80, whilst lake water pH ranges between pH 7.70 and 9.78, both distributions occupying values generally >6 pH comprising neutral to alkaline sediments on the global scale (Slessarev et al.

2016). Lake Houston and Crescent Lake have uniquely low lake sediment pH (5.76 and 5.90, respectively) compared to other SW lakes across the transect. All lakes were alkaline at the time of sampling, exceptionally so for Dankworth Pond compared to other SW lakes (pH 9.78). Soil pH of SW soils is known to covary with MAAT when pH>7 (Fig. 4.2B), however lake sediment pH does not covary with MAAT or any other environmental variable, despite having significant positive covariance with soil pH (Fig. 4.2A & C). Conversely, water pH does not significantly covary with either lake sediment, soil pH, or MAAT, only positively correlating with T_{Water} (Fig. 4.2A), suggesting that T_{Water} is the primary environmental control (amongst measured variables) on water pH. Overall, soil pH appears to be strongly influenced by MAAT across the SW transect, but lake sediment pH does not.

Of soil, lake sediment, and lake surface water pH, water pH has the most alkaline distribution (mean=8.47,SD=0.45) compared to sediment and soil pH, whilst soil pH has more variability (mean=7.34,SD=1.07) compared to lake sediment pH (mean=7.51,SD=0.65). Sediment pH is therefore uncoupled from lake surface water pH suggesting the environmental controls on soil, sediment, and aqueous pH differ. Several physical properties of the SW lakes were also recorded, namely sampling depth, CA/SA, and % lake surface area within catchment. Lake depth does not significantly covary with any measured environmental variable (Fig. 4.2A), and is either random across the transect or not climatically controlled. CA/SA and % lake surface area significantly and negatively covary because lake surface area is incorporated into CA/SA, and indicates that higher CA/SA values are associated with smaller lake surface areas relative to overall catchment area. CA/SA has significant variance with mean annual precipitation (MAP), vapour pressure deficit (VPD), seasonality, and ΔT (Fig. 4.2A). CA/SA comprises four extremely high outliers (>5000) caused by extremely large catchment areas relative to lake surface area for Lake Junction, Buffalo Springs Lake, Dankworth Pond, and Horseshoe Reservoir. The initial two lakes are situated in residential areas of Texas, whilst the latter two are not. Without these outliers, CA/SA generally increases at low MAP, potentially because most SW lakes are reservoirs that have been engineered

to accumulate water from a large land surface area in this arid region where water is known to be scarce ([Kendrick 1993](#)).

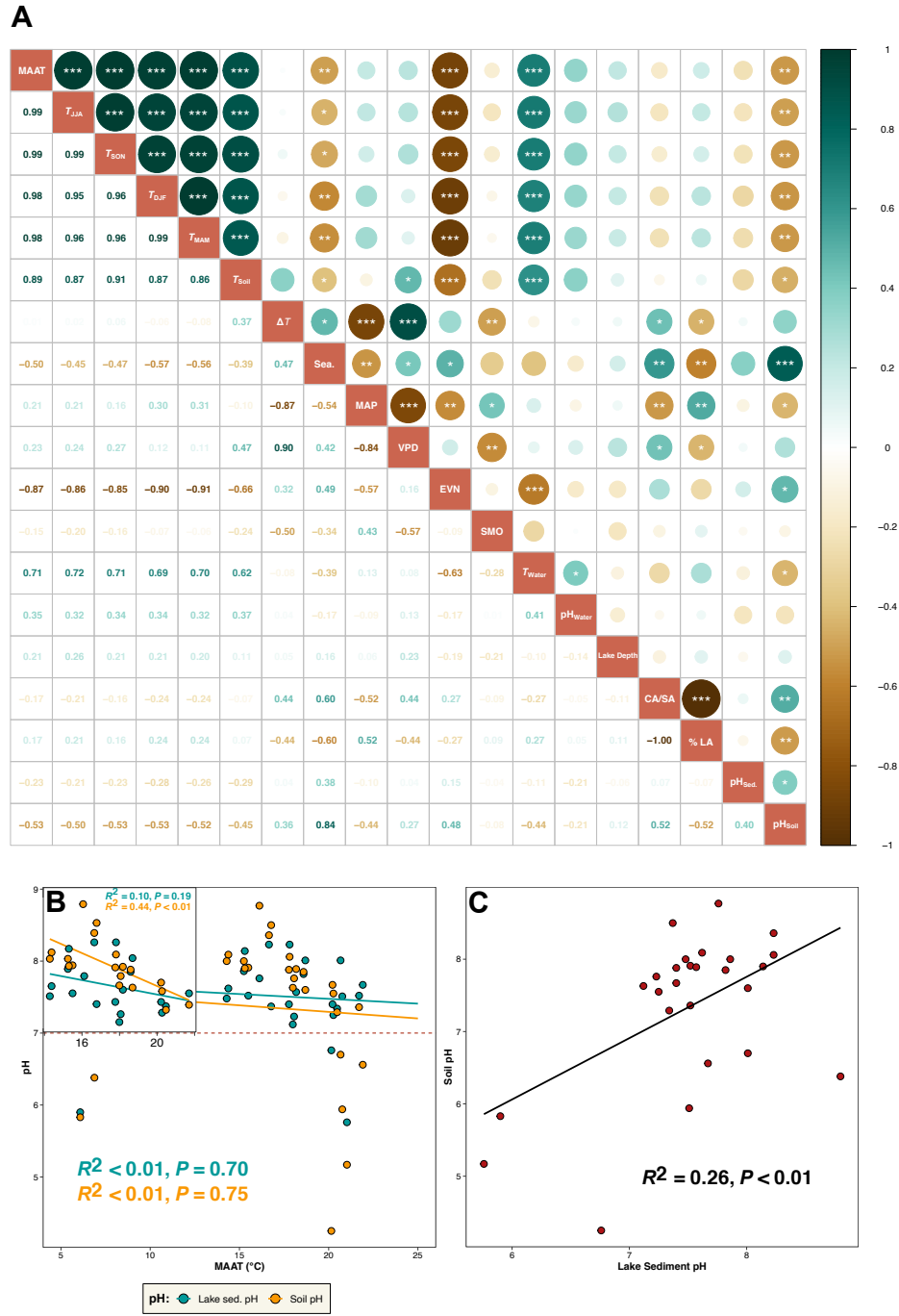


Figure 4.2: **A** Spearman's correlation matrix between SW U.S. environmental variables. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$), and crossplots between **B** MAAT and soil pH, **C** lake sediment pH and soil pH

4.3.2 Lipid biomarker distributions in SW U.S. lake sediments

4.3.2.1 3-OH FA distributions

It is well established that 3-OH FAs with carbon numbers between C₁₀ and C₁₈ are produced almost exclusively by Gram-negative bacteria (Klok et al. 1988, Wollenweber et al. 1982); this range of 3-OH FA compounds is therefore the focus of our onward discussion. A wide range of 3-OH FAs were detected in our SW lake sediments, comprising *iso*-, *anteiso*-, *normal*-, and *unsaturated*- structures (Fig. 4.3), which generally follow 3-OH FA distributions observed in sediments globally (Wang et al. 2016, Yang et al. 2021, 2020, Wang et al. 2021) whereby 3-OH FAs with odd C numbers have three structural isomers, *iso*-, *anteiso*-, and *normal*-, ranking which are highest to lowest abundance, respectively, and those with even C numbers excluding C₁₀ have *normal*- and *iso*- isomers, ranking same from highest to lowest abundance. Mono-unsaturated 3-OH FAs predominantly but not exclusively associated with even C numbers have been reported from a small number of lake sediments (Yang et al. 2021), and do occur inconsistently within the SW lake sediments. *normal*-3-OH FAs with even C numbers are predominant in the lake sediments (mean=64.1 %, SD=57.8 %). *n*-C₁₄ and *n*-C₁₆ are almost mutually dominant in SW lake sediment 3-OH FA distributions (Fig. 4.3), in addition to *n*-C₁₂ in one sample. Lake Junction (USA062) and Medina Lake (USA063) have unique 3-OH FA distributions compared to other SW lake sediments, whereby only *normal*-3-OH FAs with even C numbers were detected in addition to *i*-C₁₇ for both and *i*-C₁₃ for Medina Lake.

4.3.2.2 Comparison of 3-OH FA distributions between SW U.S. lake sediments and soils

3-OH FAs in lake sediments are relatively unexplored in comparison to 3-OH FAs in topsoils which have been analysed at a global scale (Wang et al. 2021, Véquaud et al. 2021b). To assess how 3-OH FA distributions differ between lake sediments and soils along our SW transect, the 3-OH FA distributions were compared using NMDS analysis, and the structure of pooled

fractional abundances for the two sedimentary groups was compared following the method of [Martínez-Sosa et al. \(2021\)](#), in addition to distributions of individual 3-OH FA compounds between the two groups. NMDS analysis reveals that lake sediments and soils produce distinguished clusters of 3-OH FA distributions across the SW U.S., which is confirmed by ANOSIM analysis whereby the ANOSIM R statistic of 0.55 ($p < 0.001$) indicates that sediment type significantly describes a moderate level of dissimilarity amongst the 3-OH FA distributions. The NMDS also confirms that sediments from lakes Junction and Medina have unique 3-OH FA distributions compared to other SW lake sediments as they produce extreme outliers outside the 95 % confidence ellipsis of lake sediment 3-OH FA variability (Fig. 4.4). A unique 3-OH FA composition in soil from the catchment of Lake Junction was reported in Chapter 3, therefore Lake Junction has extremely unique 3-OH FA distributions in both lake sediment and catchment soil, which are both comparable to each other having high NMDS1 values (Fig. 4.4), suggesting that either (i) unique 3-OH FA distributions are produced across the Lake Junction catchment and not restricted to either sediment type, or (ii) soils at Lake Junction may have a unique bacterial community producing 3-OH FAs, and soil input into the lake is overprinting any 3-OH FA distribution produced by an in-situ lake bacterial community.

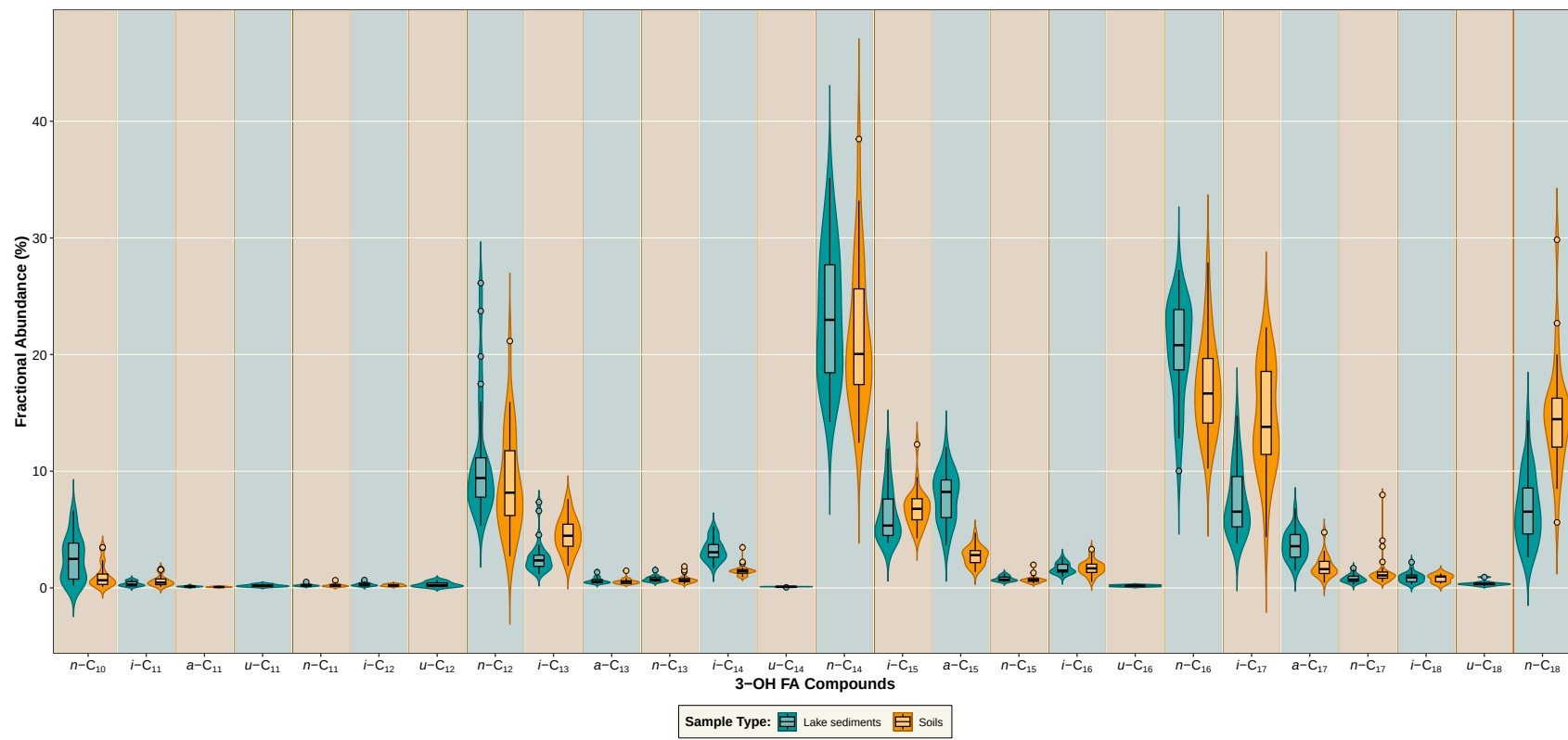


Figure 4.3: 3-OH FA distributions in SW U.S. lake sediments and soils reported as % fractional abundance for individual 3-OH FA compounds

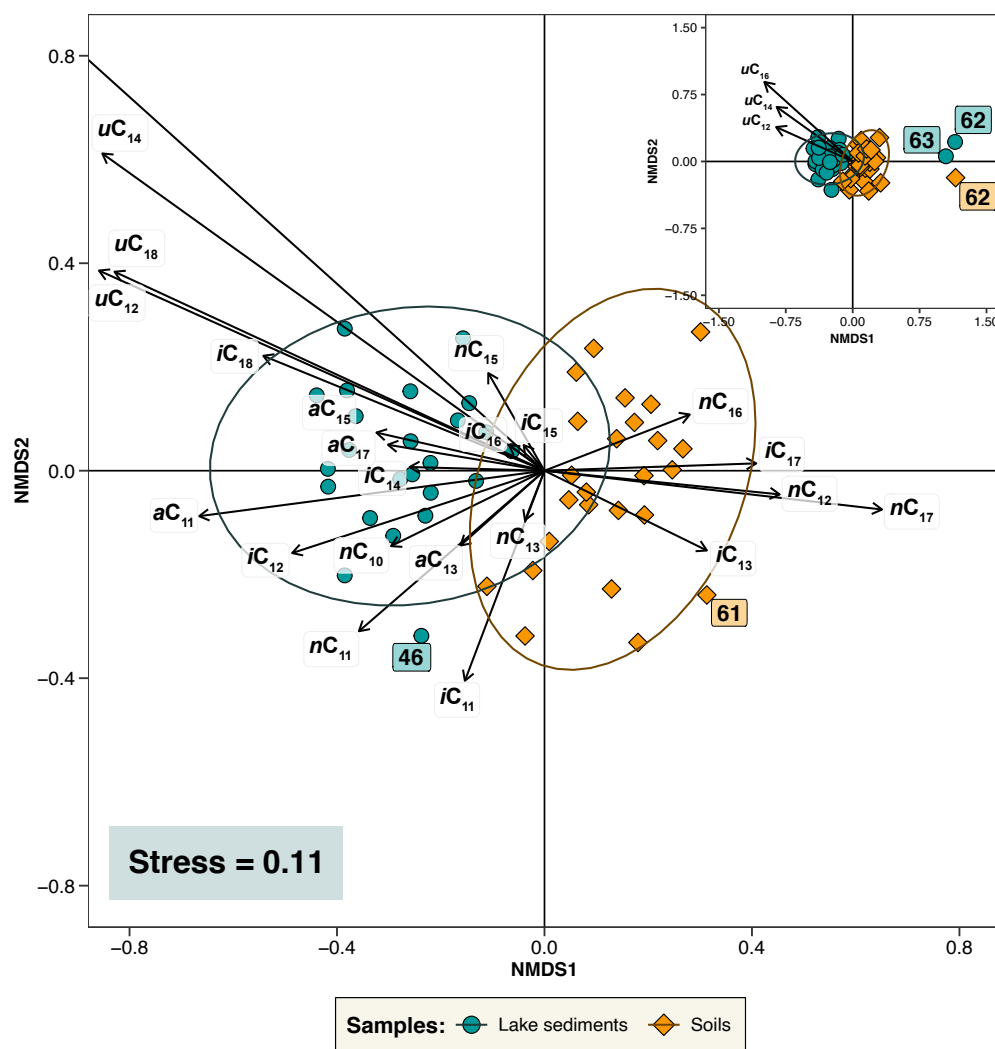


Figure 4.4: NMDS plot of a Bray-Curtis dissimilarity matrix of SW U.S. lake sediment and soil 3-OH FA fractional abundance. 3-OH FA compounds that have significant variability amongst the samples ($p < 0.05$) are visualised following permutation analysis

Sediment from Lake Caballo (USA046) also has a unique 3-OH FA distribution that falls outside the scope of general SW lacustrine 3-OH FA variability, as does soil from Lake Nasworthy (USA061) for SW soil 3-OH FA variability (Fig. 4.4). Pooled distributions of 3-OH FA fractional abundance significantly differ in mean ranking ($w=101298$, $p < 0.001$) and variance ($f=5.77$, $p < 0.05$) between lake sediments and soils (Fig. 4.5), which consolidate the significant ANOSIM

result, and indicate that the intrinsic distribution of 3-OH FA fractional abundance differs between lake sediments (mean=4.77 %, SD=6.75 % composite fractional abundance) and soils (mean=5.99 %, SD=7.27 % composite fractional abundance).

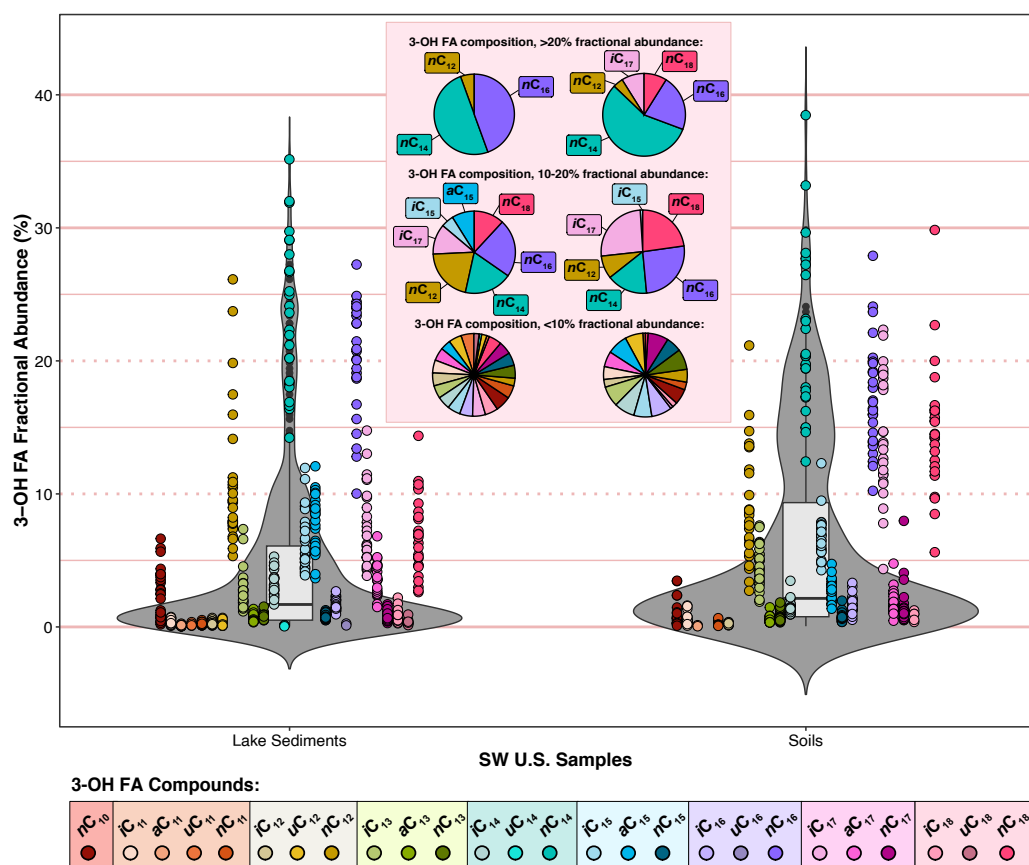


Figure 4.5: Pooled 3-OH FA 3-OH FA fractional abundance in SW U.S. lake sediments and soils. The fractional distributions of individual 3-OH FA compounds are visualised as a jitter and percentage 3-OH FA representation summarised as pie charts for three fractional abundance thresholds: (1) <10 %, (2) 10-20 % abundance, (3) >20 % abundance

For both sediment types, most data points for 3-OH FA fractional abundance fall below the 10 % abundance threshold, where a broad range of 3-OH FA compounds are distributed (Fig. 4.5); representation of 3-OH FA compounds here is very similar between the sediment types, only nC_{14}

and $n\text{-C}_{16}$ are absent within this threshold for both, and $n\text{-C}_{10}$ ($w=552$, $p<0.001$), $i\text{-C}_{14}$ ($w=604$, $p<0.001$), $a\text{-C}_{15}$ ($w=621$, $p<0.001$), and $a\text{-C}_{17}$ ($w=572$, $p<0.001$) are enriched in lake sediments, whilst $i\text{-C}_{13}$ ($w=105$, $p<0.001$) is enriched in soils (Fig. 4.3). Furthermore, there are more 3-OH FA compounds below this threshold for lake sediments due to the occurrence of *unsaturated*-3-OH FAs which are not present in soils, and are identified as a distinguishing feature between lake sediment and soil 3-OH FA distributions by NMDS (Fig. 4.4). Above 10 % fractional abundance, the distributions of 3-OH FA compounds become more distinguished between soils and lake sediments. $i\text{-C}_{17}$ ($w=73$, $p<0.001$) and $n\text{-C}_{18}$ ($w=66$, $p<0.001$) have higher representation between 10 and 20 % abundance in soils because these compounds are enriched in soils relative to lake sediments. The upper fractional abundance range of $a\text{-C}_{15}$ enrichment in lake sediments and higher representation of $i\text{-C}_{15}$ and $n\text{-C}_{12}$ between 10 and 20 % abundance in lake sediments occurs at the expense of $i\text{-C}_{17}$ and $n\text{-C}_{18}$.

A smaller number of 3-OH FA compounds occupy high abundance data points (Fig. 4.5). 3-OH FA representation is dominated by $n\text{-C}_{14}$ above 20 % fractional abundance in both lake sediments and soils, a compound which has no statistical difference in abundance between the two sedimentary groups in either mean ranking ($w=407$, $p=0.21$) or variance ($f=0.06$, $p=0.80$). In addition to $n\text{-C}_{14}$, $n\text{-C}_{16}$ is also produced in dominant proportions (over 20 % total 3-OH FAs) in lake sediments, in which $n\text{-C}_{16}$ is enriched ($w=407$, $p=0.21$). For soils however, $i\text{-C}_{17}$ and $n\text{-C}_{18}$ have representation amongst the high abundance data points at the expense of $n\text{-C}_{16}$ (Fig. 4.5).

Overall, $n\text{-C}_{14}$ is produced as a dominant 3-OH FA compound at high fractional abundance in both lake sediments and soils, and $n\text{-C}_{12}$ is also capable of being produced at relatively high fractional abundance at a comparable capacity in both sediment types. Although $n\text{-C}_{16}$ is a dominant 3-OH FA compound in lake sediments that is consistently produced at high fractional abundance alongside $n\text{-C}_{14}$, more 3-OH FA compounds are consistently produced at a higher fractional abundance (notably $i\text{-C}_{17}$ and $n\text{-C}_{18}$) in soils relative to lake sediments, whilst more

3-OH FAs are produced at lower fractional abundances (<10 % total 3-OH FAs) in lake sediments. *α*-C₁₅ has the most distinguished difference in mean fractional abundance between lake sediments and soils (it is highly enriched in lake sediments), and predominantly accounts for cumulatively higher proportion of *anteiso*-3-OH FAs in lake sediments relative to soils along the SW transect (Appendix B1). *iso*-3-OH FAs are cumulatively enriched in soils relative to lake sediments (Appendix B1), which is predominantly attributed to *i*-C₁₇. However the overall mean of branched and normal 3-OH FAs is comparable between the sediment types (Appendix B1). The composition of 3-OH FA distributions therefore differs substantially between lake sediments and soils across our SW U.S. transect due to the differing distributions of a small number of 3-OH FA compounds, resulting in a significant difference in mean ranking and variance of pooled fractional abundance, in addition to distinguished clustering, between the two sedimentary groups.

4.3.2.3 GDGT distributions

As with the SW soils, a range of brGDGTs and isoGDGTs were detected in our SW lake sediments. The proportion of brGDGTs ranged from 41.5 to 94.9 % total GDGTs (mean=73.9 %, SD=14.8 %), and isoGDGTs from 5.1 to 58.5 % (mean=26.0 %, SD=14.8 %), brGDGTs are therefore generally more abundant in the lake sediments than isoGDGTs. Of the brGDGT distributions, brGDGT-Ia is dominant (24.1 % total mean brGDGT composition) in addition to brGDGT-IIa (18.5 % total brGDGTs) and its 6-methyl counterpart brGDGT-IIa' (18.1 % total brGDGTs) (Fig. 4.6), which accounts for the dominance of brGDGT compounds with no cyclopentyl moieties (denoted with 'a'; 71.4 % total brGDGTs). brGDGTs with one ('b') or two ('c') cyclopentyl moieties are far less abundant, particularly 'c' compounds (4.19 % total brGDGT composition). Compounds with one methyl branch (denoted as 'II') are most abundant when brGDGT distributions are classed by degree of methylation (49.4 % total brGDGTs), and compounds with two or more methyl branches (denoted as 'III') are least abundant (14.7 % total brGDGTs). brGDGTs in the lake sediments are predominantly composed of 5-methyl compounds (71.2 %

total brGDGTs), in comparison to 6-methyl compounds (denoted with ' ; 28.8 % total brGDGTs), which consistently fall below detection limit in addition to brGDGT-IIIc which was detected in one sample (Conchas Lake, USA055). brGDGT-IIIc' was not detected in any SW lake sediment.

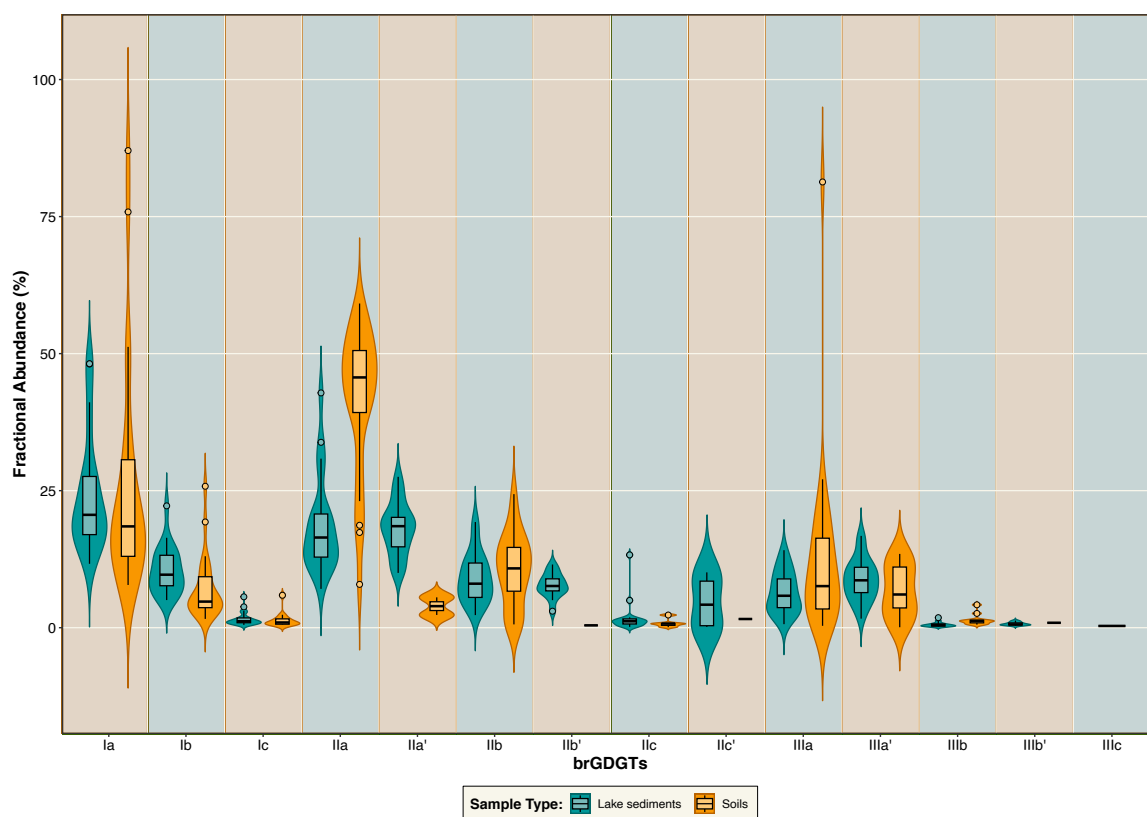


Figure 4.6: brGDGT distributions in SW U.S. lake sediments and soils reported as % fractional abundance for individual brGDGT compounds

4.3.2.4 Comparison of GDGT distributions between SW U.S. lake sediments and soils

Overall, GDGT distributions between SW lake sediments and soils are fairly consistent; brGDGTs are generally predominant over isoGDGTs, and 5-methyl brGDGTs (compounds with no methyl groups ('a')) and those with one cyclopentane moiety ('II') are dominant amongst both sediment

types (Appendix B2). However, beyond these groupings brGDGT distributions between the sediment types become more variable. Soil brGDGTs have a significantly higher mean fractional abundance and variability (mean=15.6 %,SD=17.8 %) than those from lake sediments (mean=9.9 %,SD=9.0 %) as revealed by a Man-Whitney U Test ($w=15586, p<0.05$) and Fligner-Kileen Test ($f=21.56, p<0.001$) performed upon brGDGT fractional abundance grouped by sediment type. Higher brGDGT fractional abundance in soils is attributed to an enrichment of brGDGT-IIa ($w=53, p<0.001$) relative to lake sediments (Fig. 4.6). Furthermore, brGDGT-IIa is the dominant compound in soils, whereas brGDGTs-Ia, -IIa, and -IIa' are almost mutually dominant in lake sediments (Fig. 4.6); although brGDGT groups 'II' and 'a' are dominant in both sediment types, the primary contributing compounds to these groups differ in proportion. Although 5-methyl brGDGTs are dominant in both lake sediments and soils relative to 6-methyl compounds, 5-methyl brGDGTs are enriched in soils relative to lake sediments, and therefore by definition 6-methyl brGDGTs are enriched in lake sediments relative to soils (Appendix B2). 'a' brGDGTs are also enriched in soils relative to lake sediments despite this being the dominant methylation grouping for both sediment types, whilst 'b' compounds are enriched in lake sediments relative to soils (Appendix B2). There is no significant difference in any other brGDGT grouping between the two sediment types (Appendix B2). Therefore brGDGT distributions between the SW soils and lake sediments have both similar and distinct qualities. The abundance of structural groupings relative to each other (intrinsic structural groupings) are very similar between sediment type. However the fractional abundance of certain brGDGT compounds differs between the sediment types, which has led to an overarching difference in brGDGT composition between SW lake sediments and soils, which was also observed amongst 3-OH FA distributions.

4.3.3 Lipid biomarker-based proxy distributions in SW U.S. lake sediments and comparisons with soils

4.3.3.1 3-OH FAs

The established 3-OH FA-based proxies for temperature (RAN_{15} , RAN_{17} , RAN_{13} , RAN_s , RIN_{17} , RIN) and pH (BR , $RIAN$) that were calculated for the 3-OH FA distributions of SW soils in Chapter 3 are now calculated for the SW lake sediments and visualised in Figure 4.7. RAN_{15} values range between 3.66 and 20.9 ± 0.21 (*SE*) ($n=23/26$), RAN_{17} between 2.11 and 12.4 ± 0.29 (*SE*) ($n=24/26$), RAN_{13} between 0.48 and 1.40 ± 0.02 (*SE*) ($n=24/26$), RAN_s between 0.06 and 0.21 ± 0.001 (*SE*) ($n=24/26$), RIN_{17} between 3.92 and 27.8 ± 0.56 (*SE*) ($n=24/26$), RIN between 0.13 and 0.59 ± 0.003 (*SE*) ($n=26$), BR between 0.13 and 0.78 ± 0.004 (*SE*) ($n=26$), and $RIAN$ between 0.11 and 0.90 ± 0.003 (*SE*) ($n=26$). Each proxy produces a range of values amongst the lake sediments which is indicative of a proxy response to a biological or environmental driving mechanism. Lake sediments from Lake Junction and Medina Lake consistently fail to produce values for non-cumulative proxies such as the RAN proxies and RIN_{17} due to their unique 3-OH FA compositions.

RAN_{15} , RAN_{17} , RAN_s and RIN_{17} have distributions that significantly differ in mean and variance (and only in mean for RIN) between SW lake sediments and soils (Fig. 4.7). The aforementioned RAN proxies are consistently higher and more variable in value in lake sediments relative to soils, which is consistent with *anteiso*- enrichment previously observed in the lake sediment 3-OH FA distributions, and may suggest these proxies are more influenced by environmental or biological forcing in lacustrine rather than pedogenic environments. Conversely, higher RIN and RIN_{17} values in soils relative to lake sediments (Fig. 4.7) is consistent with the *iso*- enrichment observed in soil 3-OH FA distributions. There is no significant difference in the distributions of RAN_{13} , BR , or $RIAN$ which have homogeneity in mean and variance between the lake sediments and soils (Fig. 4.7). Overall, proxies that employ predominantly *anteiso*-3-OH FAs

tend to have higher values and more variability in lake sediments, whilst proxies that employ predominantly *iso*-3-OH FAs tend to have higher values and variability in soils. Cumulative proxies that employ all branched 3-OH FA compounds (BR and RIAN) - that are also employed for pH reconstruction - have homogeneous distributions between sediment type likely related to the homogeneous proportion of branched 3-OH FA fractional abundance between SW lake sediments and soils previously observed (Appendix B1).

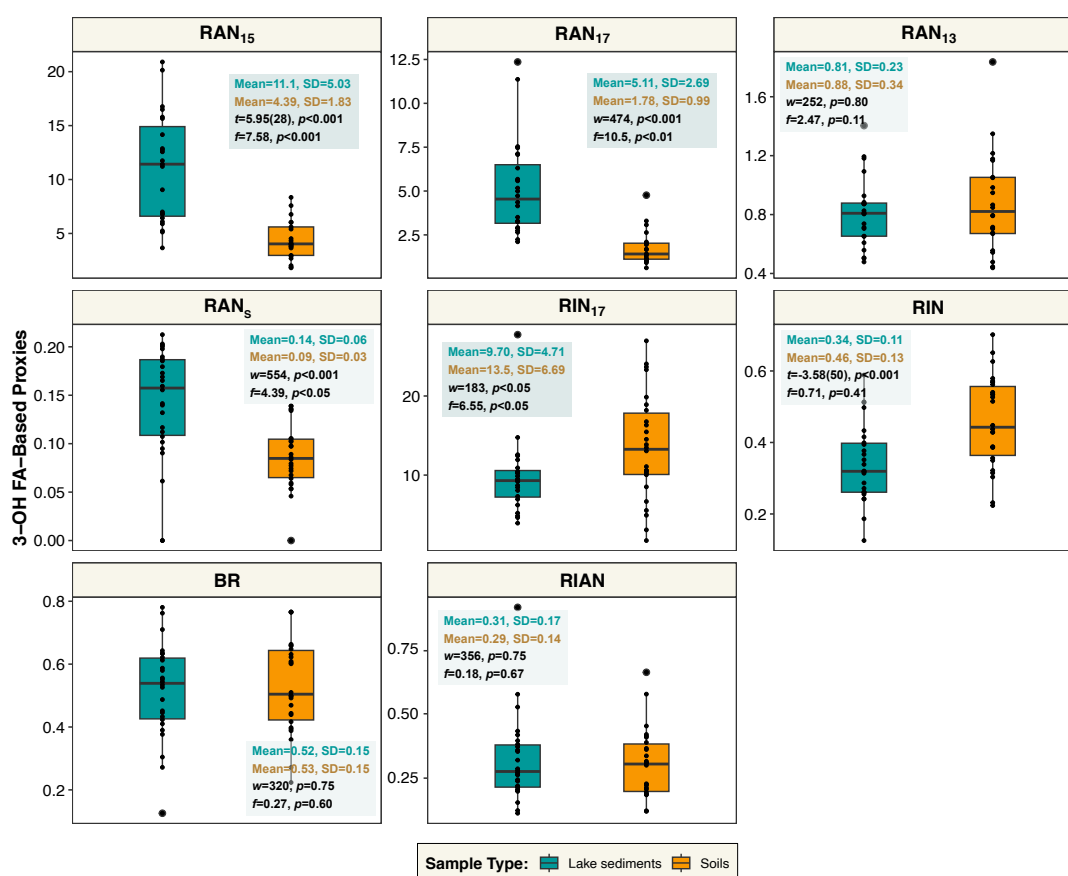


Figure 4.7: Boxplots visualising 3-OH FA-based proxy distributions between SW lake sediments and soils. Descriptive statistics are visualised.

4.3.3.2 brGDGTs

Unlike the brGDGT distributions of EC soils reported in Chapter 3 after [Dirghangi et al. \(2013\)](#), brGDGTs from our SW soils and lake sediments were extracted using the methods of [De Jonge et al. \(2014a\)](#) and [Hopmans et al. \(2016\)](#) that allow separation of 5-methyl and 6-methyl isomers. Therefore a more recent suite of brGDGT-based environmental proxies can be calculated for the SW soils and lake sediments than the MBT and CBT proxies reported by [Dirghangi et al. \(2013\)](#) and from our SW soils in Chapter 3. The more recent suite of GDGT proxies that are subsequently reported comprise MBT'_{5ME} for temperature reconstruction that omits pH-dependent 6-methyl brGDGT isomers, and CBT' which incorporates pH-dependent 6-methyl brGDGT isomers in the numerator ([De Jonge et al. 2014b](#)). MBT'_{6ME} describes the pH-dependent component of the MBT-CBT temperature proxy that is omitted from MBT'_{5ME}, whilst DC' and IR_{6ME} are descriptive indices respective for degree of cyclisation (relative proportion of brGDGT compounds with cyclopentane moieties) and the relative proportion of hexamethylated brGDGT isomers. BIT and R_{i/b} that were described and reported for SW soils in Chapter 3 are reported here for SW lake sediments.

In the SW lake sediments brGDGTs are more enriched relative to isoGDGTs, leading to R_{i/b} values that range between 0.05 and 1.41 (mean=0.42, SD=0.35) and BIT values between 0.55 and 1.0 (mean=0.94, SD=0.99). As R_{i/b} values are generally <0.5, brGDGTs consistently predominate over isoGDGTs in the lake sediments. Sediments from Santa Rosa Lake (NM), Crescent Lake (AZ), and Horseshoe Reservoir (AZ), have a mixed brGDGT and isoGDGT composition more abundant in isoGDGTs (R_{i/b}>1) which is unique amongst the SW lake sediments. The lake sediments have significantly higher and less variable BIT values compared to SW soils (Fig. 4.8). BIT values are known to be highly variable in lake sediments due to the influence of detrital input from the surrounding catchment ([Hopmans et al. 2004](#)) which is also variable. However this lake-soil comparison is unique because the SW lake and soil transect is situated in an arid to

semi-arid environment. BIT is commonly reported to be higher in value in soils relative to lake sediments, and higher BIT values in lake sediments are the result of soil detritus into the lake basin (Dang et al. 2016) and higher abundance of crenarchaeol in lacustrine environments (Hopmans et al. 2004). There is no significant difference in distribution of $R_{i/b}$ between the lake sediments and soils, indicating that crenarchaeol abundance is changing between the sediment types. However limited soil moisture in the arid soils promotes archaeal GDGT production lowering $R_{i/b}$ values, and low BIT values are commonly reported from arid soils (Dirghangi et al. 2013). For the lake sediments to have higher BIT values relative to the soils, an in-situ lacustrine bacterial community is likely to be producing GDGTs. Overall, BIT and $R_{i/b}$ indicate bacterial GDGT production is more prominent in lake sediments relative to soils.

The recent suite of brGDGT-based environmental proxies produced a range of values amongst the SW soils and lake sediments (Fig. 3.6): for soils, MBT'_{5ME} (0.11-0.91, n=24/26), MBT'_{6ME} (0.50-1.00, n=24/26), CBT' (0.01-0.21, n=16/26), DC' (0.03-0.51, n=24/26), IR (0.01-0.21, n=13/26); for lake sediments, MBT'_{5ME} (0.29-0.74, n=22/26), MBT'_{6ME} (0.35-0.93, n=22/26), CBT' (0.09-1.29, n=22/26), DC' (0.15-0.49, n=22/26), IR (0.09-0.65, n=22/26). These distributions indicate that all proxies are responsive to either environmental or biological forcing across each transect. Each brGDGT-based proxy (excluding MBT'_{6ME}) produces consistently higher values in SW lake sediments relative to soils (Fig. 3.6). Higher MBT'_{5ME} and IR in lake sediments and lower MBT'_{6ME} is consistent with enrichment of 6-methyl brGDGT compounds in lake sediments relative to soils. Higher CBT' values in lake sediments is indicative that brGDGTs containing one or more cyclopentane moieties are more abundant in lake sediments particularly amongst 6-methyl isomers. Indeed, CBT' and IR did not produce values for almost half of the soils because 6-methyl brGDGTs were not present. Absence of 6-methyl brGDGTs was reported by (De Jonge et al. 2014b) in 17 out of 239 globally distributed soils, however the SW U.S. represents an extreme arid environment where unique brGDGT distributions are likely to be observed. All proxies excluding $R_{i/b}$ have different distributions between lake sediments and soils,

and differing intrinsic response between the sediment types as variance significantly differs between the sediment types for all proxies except for MBT'_{5ME} and MBT'_{6ME}, even though the samples are subject to the same range of air temperature. This suggests the environmental or biological drivers (or both) of 3-OH FA variability differ between the sediment types, which is plausible because lacustrine and pedogenic environments have distinguished biogenic processes.

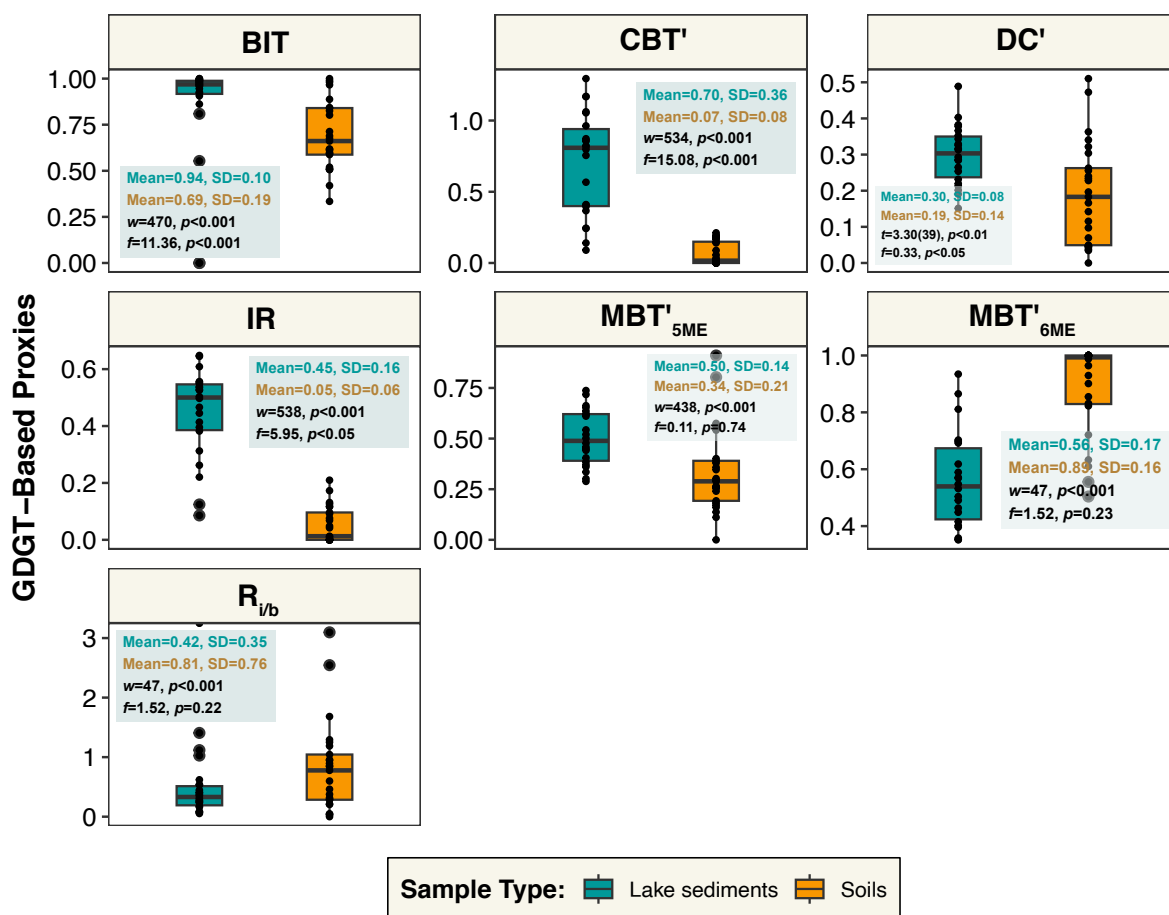


Figure 4.8: Boxplots visualising brGDGT-based proxy distributions between SW lake sediments and soils. Descriptive statistics are visualised.

4.3.4 Application of linear and multivariate analyses to identify environmental drivers of 3-OH FA variability in SW U.S. lake sediments and soils

4.3.4.1 Environmental drivers of 3-OH FA variability in SW U.S. lake sediments

There is surmounting evidence that lake sediment and soil 3-OH FA distributions are distinguished from each other, and so we isolate and perform PCA analysis upon 3-OH FA distributions of SW lake sediments alone. 3-OH FA compounds that consistently fall below detection limit are removed from the lake sediment 3-OH FA dataset following the data wrangling method of [De Jonge et al. \(2024\)](#) in preparation for PCA analysis, and comprise all *unsaturated*-3-OH FAs (*u*-C₁₁, *u*-C₁₂, *u*-C₁₄, *u*-C₁₆, *u*-C₁₈). The first two principal components describe 49.3 % variance within the lake sediment 3-OH FA distributions. 33.0 % of this variance is described by PC1, to which the primary contributing 3-OH FAs are *i*-C₁₂, *n*-C₁₃, *a*-C₁₃, *n*-C₁₆, *i*-C₁₃, *n*-C₁₁, *i*-C₁₁, *n*-C₁₈, and *n*-C₁₇. PC2 describes 16.3 % of 3-OH FA variance amongst the lake sediments, for which the primary contributing 3-OH FA compounds are *n*-C₁₄, *i*-C₁₆, *i*-C₁₅, *i*-C₁₇, *a*-C₁₇, *n*-C₁₀, *i*-C₁₁, and *n*-C₁₅, in which all compounds excluding *n*-C₁₄ positively load onto PC2, and have a negative relationship with *n*-C₁₄ (Fig. 4.9). Environmental variables associated with lacustrine environments (displayed in Figure 4.9) were added to the PCA following permutation analysis. Lake sediment pH is associated with PC1, however does not significantly explain any lacustrine 3-OH FA variance. [MAAT](#) ($r^2=0.64$, $p<0.05$), [T_{DJF}](#) ($r^2=0.25$, $p<0.05$), and [MAP](#) ($r^2=0.32$, $p<0.05$) are the only environmental variables to explain lacustrine 3-OH FA variability across the SW U.S. and negatively load onto PC2.

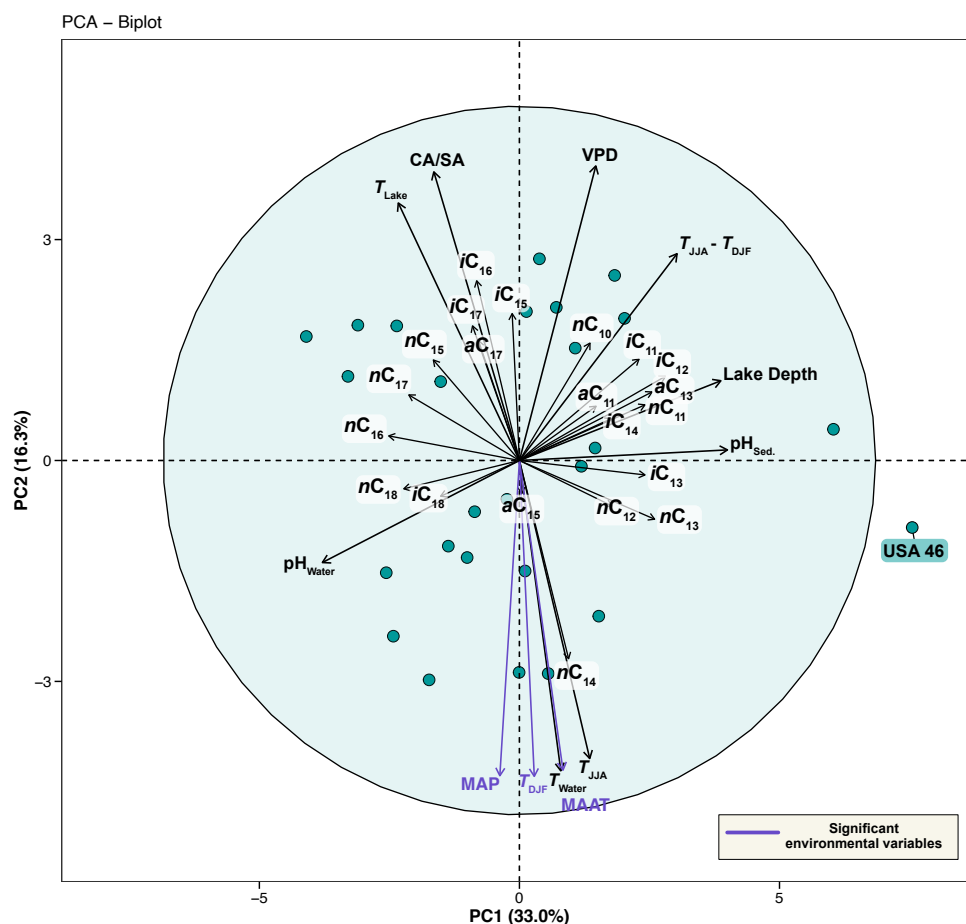


Figure 4.9: PCA of relationships between lake sediment 3-OH FA compounds and environmental variables. Environmental variables were added *a posteriori* following permutation analysis. Significant environmental variables ($p < 0.05$) are indicated.

The SW U.S. environmental transect comprises two cold mountain lake catchments (lakes Mary's (USA043) and Crescent (USA053) (Fig. 4.1)) that form cold outliers amongst the general temperature variability of the SW U.S. lake catchments, whereby 24 lakes are normally distributed in **MAAT** between 14.3 and 21.9 °C (Appendix B3). Although the 3-OH FA variability of sediments from these lakes falls within the general variability of all SW lakes (Fig. 4.9), considering two outlier cold sites in the development of a temperature calibration is not robust without the inclusion of more cold lakes to fully represent 3-OH FA variability under cold **MAAT**,

and form a normally distributed temperature dataset that extends to 6.1 °C (MAAT at Crescent Lake). Furthermore, established 3-OH FA-based proxy values in soils from the two cold lake catchments were observed to have a more comparable environmental response to EC soils than to SW soils, and SW soils exhibited proxy variability in response to MAAT in most soils aside from the two cold catchments, particularly for RAN_S (Appendix A8), which supports a case to omit the cold outliers from both SW lake sediment and soil transects and explore 3-OH FA variability and established proxy performance without confounding influence from these sites.

4.3.4.2 Environmental drivers of 3-OH FA variability in SW U.S. lake sediments omitting unique cold lake catchments

Omitting lakes Mary's and Crescent (cold, high-altitude montane lake catchments) samples from the SW lake sediment and soil transects has a substantial influence on the proportion of 3-OH FA variance explained by environmental parameters in PCA analysis for both sediment types. For lake sediments, the first two principal components describe 48.4 % 3-OH FA variance (Fig. 4.10A). PC1 describes the largest proportion of variance at 30.9 % and is loaded by the 3-OH FA compounds *i*-C₁₂, *a*-C₁₃, *n*-C₁₃, *i*-C₁₃, *n*-C₁₆, *n*-C₁₁, *i*-C₁₁, *n*-C₁₈, and *n*-C₁₇ (Appendix B4C). PC2 describes 17.5 % 3-OH FA variance and is associated with the 3-OH FA compounds *n*-C₁₄, *a*-C₁₇, *n*-C₁₀, *i*-C₁₆, and *i*-C₁₁ (Appendix B4D). The 3-OH FA compounds responsible for the largest degree of 3-OH FA variance (i.e. those contributing to PC1) are the same regardless of whether the two cold sites are omitted from the lake dataset or not (Appendix B4A&C). However, omitting cold sites does affect the 3-OH FA compounds responsible for variance along PC2; the same subset of compounds remain associated with PC2 regardless of whether cold sites are omitted or not, but compound contribution to PC2 differs (Appendix B4B&D). *n*-C₁₄ is predominantly associated with 3-OH FA variance along PC2 for both datasets and maintains negative loading on PC2 for both, however *a*-C₁₇ is more influential on PC2 variance and a smaller number of compounds load onto PC2 when cold sites are omitted (Appendix B4B&D). Permutation analysis

on environmental variables associated with lacustrine environments reveals that **MAP** ($r^2=0.31$, $p<0.05$), **T_{DJF}** ($r^2=0.28$, $p<0.05$), **MAAT** ($r^2=0.27$, $p<0.05$), and **$T_{JJA}-T_{DJF}$** ($r^2=0.25$, $p<0.05$) significantly describe 3-OH FA variance amongst the lake sediments when cold sites are omitted (Fig. 4.10A). Most compounds that significantly correlate with **MAAT** (excluding *a*-C₁₇) are associated with PC1, consolidating that **MAAT** is the primary environmental driver of 3-OH FA variability in SW lake sediments. RDA confirms that **MAAT** is the only environmental variable that significantly explains any constrained 3-OH FA variance amongst the SW lake sediments ($f=2.37$, $p<0.05$) (Appendix B5A).

Pearson's correlation analysis visualised as a correlation matrix, in addition to the RDA, reveals *a*-C₁₃, *i*-C₁₂, *a*-C₁₇, and *n*-C₁₈ have significant linear correlations with **MAAT** (Fig. 4.11). Additionally, a 3-OH FA temperature response in lake sediments is potentially seasonally influenced; *a*-C₁₃ has consistent variability in response to seasonal air temperatures, whereas *i*-C₁₂ and *a*-C₁₇ have negative linear correlations with most seasonal air temperatures excluding **T_{JJA}** (Fig. 4.11). *i*-C₁₄ only has variability in response to **T_{DJF}** and **T_{MAM}** and not **MAAT**, whilst *n*-C₁₈ has most significant variability with **T_{DJF}** and **T_{MAM}** (Fig. 4.11). Therefore, it is possible that the production of different 3-OH FA compounds or subsets of 3-OH FA compounds could be seasonally biased. No 3-OH FA compounds have any significant linear correlation with pH in SW lake sediments.

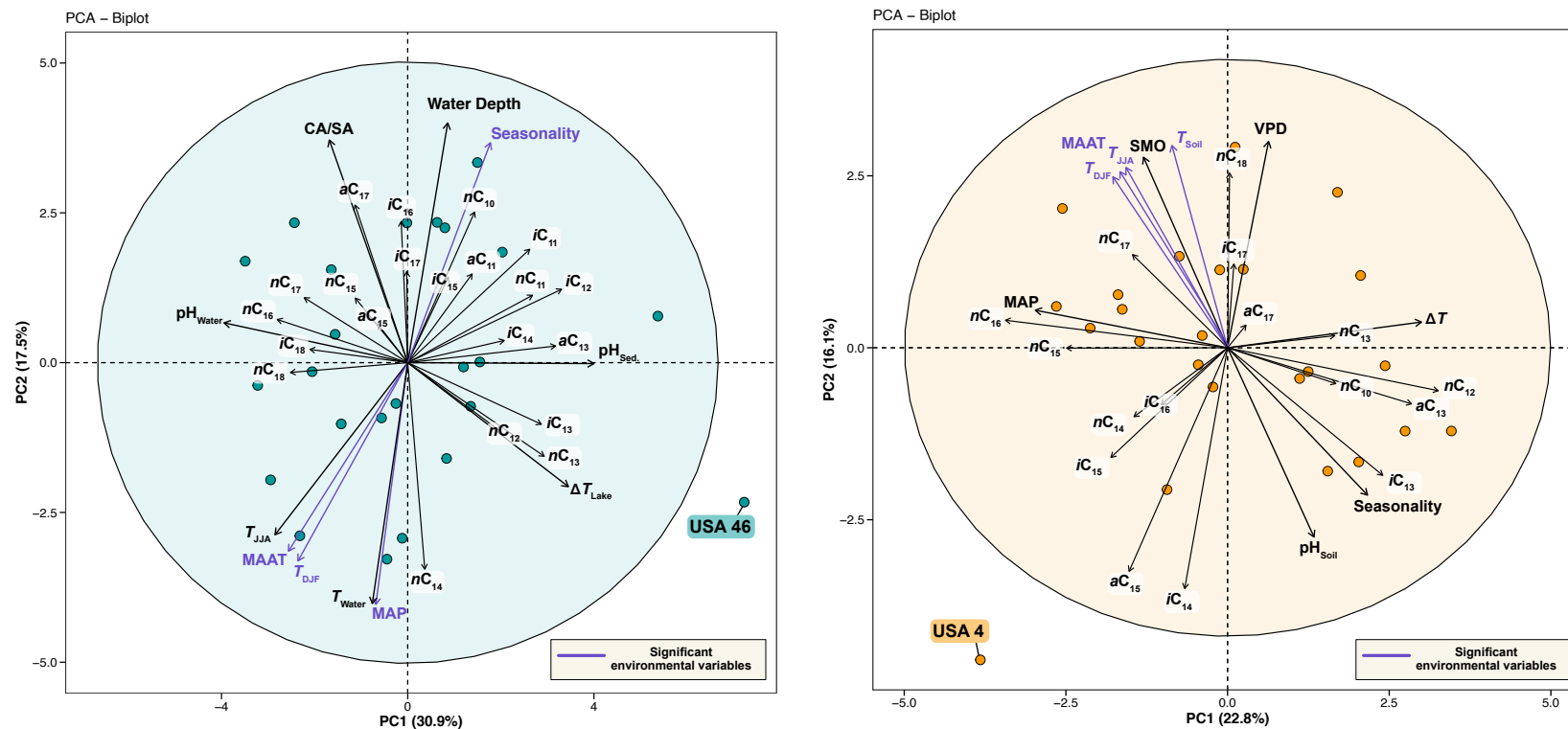


Figure 4.10: PCAs of 3-OH FA fractional distributions from (Left) SW lake sediments and (Right) soils, with cold sites omitted. Environmental variables were added *a posteriori* following permutation analysis. Significant environmental variables ($p < 0.05$) are indicated.

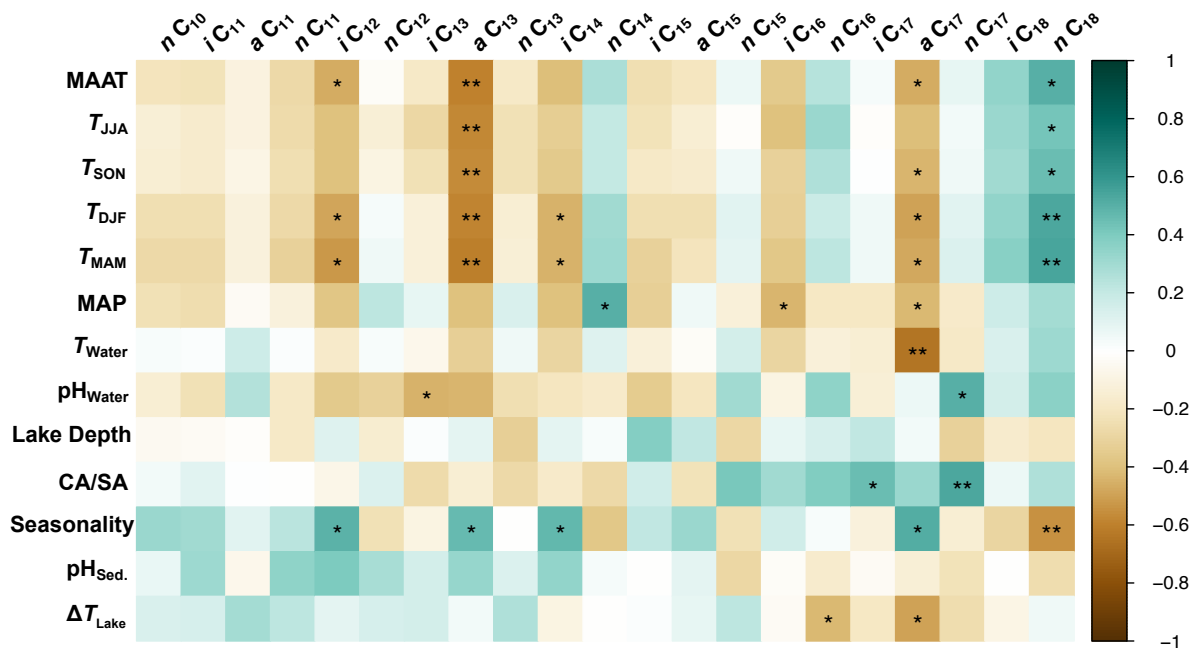


Figure 4.11: Pearson's correlation matrix between 3-OH FA fractional abundance of SW lake sediments and environmental variables, omitting cold lakes. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$).

For SW soils, the first two principal components of PCA analysis performed on 3-OH FA distributions excluding cold soils describe 38.9 % of 3-OH FA variance (Fig. 4.10B). PC1 describes the largest proportion of variance at 22.8 % and is associated with the 3-OH FA compounds $n\text{-C}_{16}$, $n\text{-C}_{12}$, $a\text{-C}_{13}$, $n\text{-C}_{15}$, and $i\text{-C}_{13}$ (Appendix B4E), whilst PC2 describes 16.1 % variance and is associated with the 3-OH FA compounds $i\text{-C}_{14}$, $a\text{-C}_{15}$, $n\text{-C}_{18}$, and $i\text{-C}_{13}$ (Appendix B4F). Environmental variables associated with pedogenic environments were applied to the PCA biplot following permutation analysis (Fig. 4.10B), which reveals that T_{JJA} ($r^2=0.40$, $p<0.001$), MAAT ($r^2=0.36$, $p<0.05$), T_{DJF} ($r^2=0.32$, $p<0.05$), and T_{Soil} ($r^2=0.27$, $p<0.05$) - all temperature variables - are significantly associated with 3-OH FA variance in SW soils. RDA confirms that MAAT ($f=2.80$, $p<0.01$) and T_{Soil} ($f=2.30$, $p=0.01$) are the primary environmental drivers of significant and constrained 3-OH FA variance amongst SW soils (Appendix B5B). The correlations between 3-OH FA compounds and environmental variables amongst the SW soils are

visualised using a Pearson's correlation matrix (Fig. 4.12), which reveals *i*-C₁₃ as the only 3-OH FA compound to have a significant negative linear correlation with **MAAT** ($r^2=0.36$, $p<0.01$). ***T*_{Soil}** additionally has significant negative linear correlations with both *i*-C₁₃ and *a*-C₁₅, however a direct response between the 3-OH FA distributions of SW soils and temperature appears to be fairly limited.

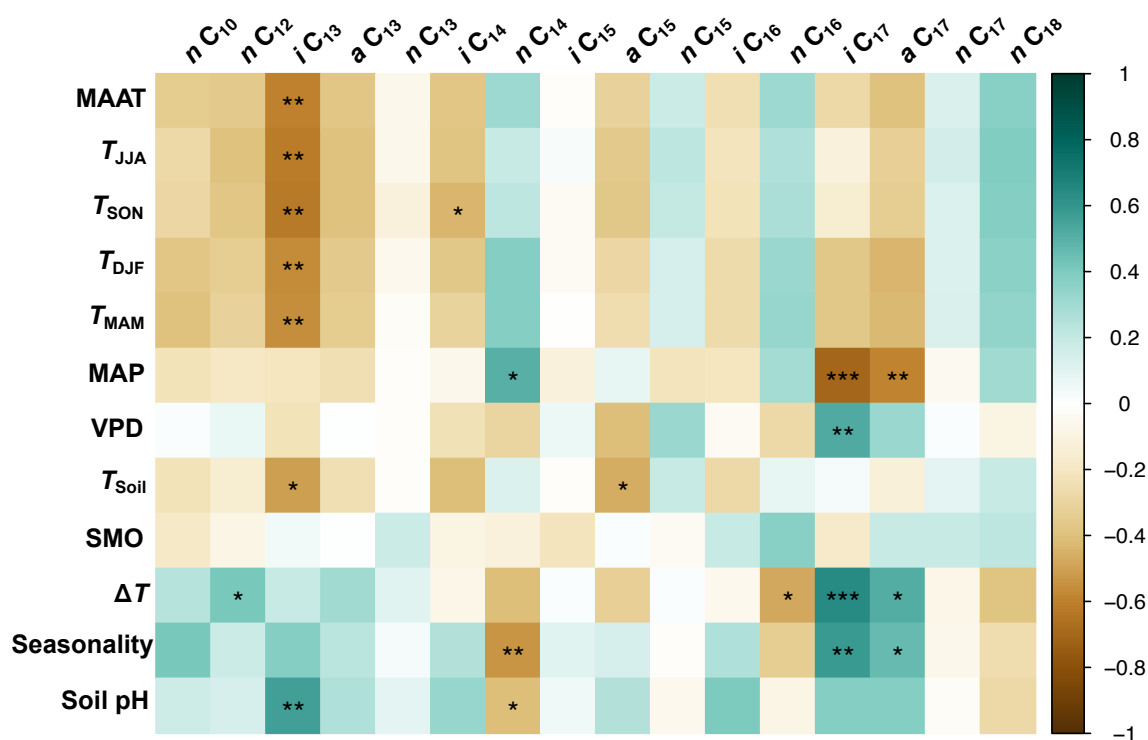


Figure 4.12: Pearson's correlation matrix between 3-OH FA fractional abundance of SW soils and environmental variables, omitting cold lakes. Significance is indicated by asterisks ($p: \leq 0.05 = *$, $\leq 0.01 = **$, $\leq 0.001 = ***$).

4.3.4.3 Quantifying 3-OH FA temperature variability revealed by multivariate and linear correlation analysis

A wider range of 3-OH FA compounds exhibit variability in response to temperature in SW lake sediments relative to soils, as indicated by PCA analysis, RDA analysis, and Pearson's correlation analysis, which suggests that 3-OH FA distributions in lake sediments have more variability in response to temperature than those of soils. Therefore identification of a suite of temperature-responsive 3-OH FAs and development of an index to reflect this 3-OH FA variability is likely to be easier to deconvolve for lake sediments rather than soils. The 3-OH FA compounds that correlate with **MAAT** in lake sediments with cold sites omitted have been reported and are visualised in Figure 4.11. Regressing the summation of *i*-C₁₂, *a*-C₁₃, and *a*-C₁₇ (which negatively correlate with **MAAT**) against *n*-C₁₈ (the only compound to positively correlate with **MAAT**) - expressed as Equation 4.1 - produces a significant negative linear correlation with **MAAT** ($r^2=0.46$, $p<0.001$), indicating that the branched 3-OH FA compounds employed as the numerator decrease in abundance relative to *n*-C₁₈ at higher **MAAT**. Lake sediment from Elephant Butte Reservoir (USA045) produces a high index value outlier outside the constraints of the 95 % confidence bounds for the linear regression, and when omitted a stronger correlation is produced ($r^2=0.55$, $p<0.001$) with a RMSE of 1.63 °C (Fig. 4.13A). Proxy values do have a weak but significant negative correlation with **MAP** ($r^2=0.22$, $p<0.05$), indicating that the sum of *i*-C₁₂, *a*-C₁₃, and *a*-C₁₇ has a tendency to decrease at sites that experience higher **MAP**, however there is no direct variability of this index in response to **MAP**. Residuals of this regression do not significantly correlate with any pH variable (lake surface water pH, lake sediment pH) or **MAP** ($p>0.05$), indicating $(i\text{-}C_{12} + a\text{-}C_{13} + a\text{-}C_{17}) / n\text{-}C_{18}$ is representative of a direct temperature response.

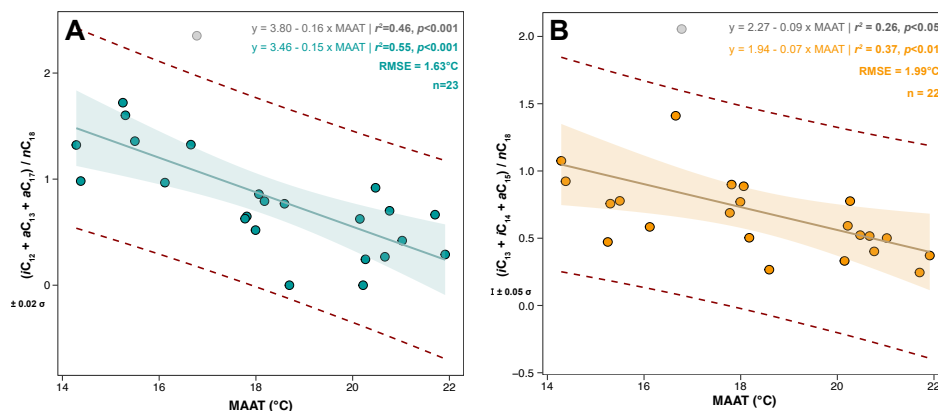


Figure 4.13: Scatterplots between **MAAT** and **A** $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ index in SW U.S. lake sediments, and **B** $[(i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18}]$ index in SW U.S. soils, all omitting two cold lakes. Shaded graph area represents 95 % functional bounds. Dashed lines represent 95 % observational bounds.

$$\text{Lake 3-OHFA temperature response} = (i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18} \quad (4.1)$$

$$\text{MAAT } (^{\circ}\text{C}) = 21.1 - 3.78 \times [(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}] \quad (4.2)$$

$$\text{Soil 3-OHFA temperature response} = (i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18} \quad (4.3)$$

For SW soils, PC2 scores are the only between PC1 to 4 to have a significant but weak positive correlation with **MAAT** ($r^2=0.22$, $p<0.05$) (Appendix B9), however of the four 3-OH FA compounds associated with PC2 for the soils, only $i-C_{13}$ has a significant negative correlation with **MAAT** (Fig. 4.12). Due to the significant PC2 variability in response to **MAAT**, 3-OH FA compounds that are associated with PC2 and have significant or broadly negative correlations with **MAAT** ($i-C_{13}$, $a-C_{15}$, $i-C_{14}$) were summed together and employed as a numerator in a regression against $n-C_{18}$ as the only compound associated with PC2 to positively correlate with **MAAT**. This

regression is expressed in Equation 4.3, and summarises variability of PC2-related 3-OH FA compounds in response to soil pH, which produces a weak significant negative linear correlation with MAAT ($r^2=0.26$, $p<0.05$) (Fig. 4.13B). Soil from the catchment of Lake Junction failed to produce an index value, and soil from that of Elephant Butte Reservoir produced an extremely high index value that falls outside the 95 % confidence bounds of the linear regression, due to their unique 3-OH FA compositions. Omitting these soils improves the correlation to $r^2=0.37$, $p<0.01$, and produces an RMSE of 1.99 °C based on 22 SW soils (Fig. 4.13B). This index for temperature variability does however produce a significant exponential correlation with soil pH ($R=0.49$, <0.05) (Appendix B10), and therefore is dependent on pH which is expected because soil pH and MAAT covary amongst SW soils when $\text{pH}>7$.

$n\text{-C}_{18}$ is employed in both lake sediment and soil temperature indices, and undergoes an increase in fractional abundance with increasing MAAT in both. However, this is the only mutual 3-OH FA temperature response between SW lake sediments and soils. $i\text{-C}_{12}$ is negatively responsive to increasing temperature in lake sediments, however is not physically present in soils to produce any temperature variability. $a\text{-C}_{15}$ and $a\text{-C}_{17}$ are respectively negatively responsive to temperature in lake sediments and soils, and have been employed in the respective RAN_{15} and RAN_{17} temperature proxies developed from soils (Wang et al. 2016). Thus, the 3-OH FA temperature responses observed in lake sediments and soils have both comparable and distinct components. The particular inclusion of $i\text{-C}_{12}$ in the lacustrine temperature index provides further evidence of in-situ 3-OH FA production, as $i\text{-C}_{12}$ is not produced consistently in soils. Overall, a number of multivariate and linear analytical methods have been used to identify specific trends of 3-OH FA variability in response to temperature in both SW U.S. lake sediments and soils, and a specific temperature index has been reported for each.

4.3.4.4 Response of established 3-OH FA-based proxies to environmental variables in SW U.S. lake sediments and soils

RDA and Pearson's correlation analysis were applied to assess the performance of established 3-OH FA-based environmental proxies in our SW lake sediments and soils. RDA was performed on the lake sediments and soils as separate datasets, with the same cold outliers omitted for each sediment group. Environmental variables respectively applicable for lacustrine and pedogenic environments were applied in each RDA as explanatory variables, and the 3-OH FA based proxies RAN₁₅, RAN₁₇, RAN₁₃, RAN_S, RIN₁₇, RIN, BR, and RIAN applied as response variables, and are visualised in Figure 4.14. Environmental variables applied in RDA for SW lake sediments collectively explain 72 % constrained variance amongst the 3-OH FA proxies. The RDA model is globally significant to a high degree ($f=3.32$, $p<0.001$), attributed to significant proxy variability predominantly in response to MAAT ($f=8.67$, $p<0.001$) as well as T_{JJA} ($f=3.80$, $p<0.05$), MAP ($f=4.38$, $p<0.05$), and ΔT_{Lake} ($f=4.47$, $p<0.05$). Temperature variables are strongly associated with RDA1 amongst the lake sediments as is RAN₁₃, and their opposed ordination in respect to each other indicates that RAN₁₃ has a negative relationship with temperature whereby *a*-C₁₃ decreases in abundance relative to *n*-C₁₃ with increasing temperature. The RAN₁₃ temperature response is confirmed by a Pearson's correlation matrix between lake sediment proxy distributions and environmental variables, revealing a weak negative linear correlation between RAN₁₃ and MAAT that is highly significant ($r^2=0.37$, $p<0.01$) (Fig. 4.15A, Appendix B11, Appendix B12). RAN₁₃ additionally has significant variability with lake depth ($r^2=0.22$, $p<0.05$) (Appendix B11), however this is the result of high values in three lakes between 20 to 25 m depth, and not consistent variance amongst the lakes. A number of 3-OH FA proxies are additionally revealed to have significant negative linear correlations with MAAT, namely RAN₁₇, RAN_S, RIN, and BR (Appendix B11), whilst RAN₁₅ and RIN₁₇ have no significant variability in response to any environmental variable amongst the SW lake sediments, despite RAN₁₅ being a primary temperature proxy in soils (Wang et al. 2016, 2021) and RIN₁₇ in lake sediments (Yang et al.

2021). Correlations between 3-OH FA proxies and MAAT for SW U.S. lake sediments and soils including the two cold lakes are shown in Appendix B13 for context.

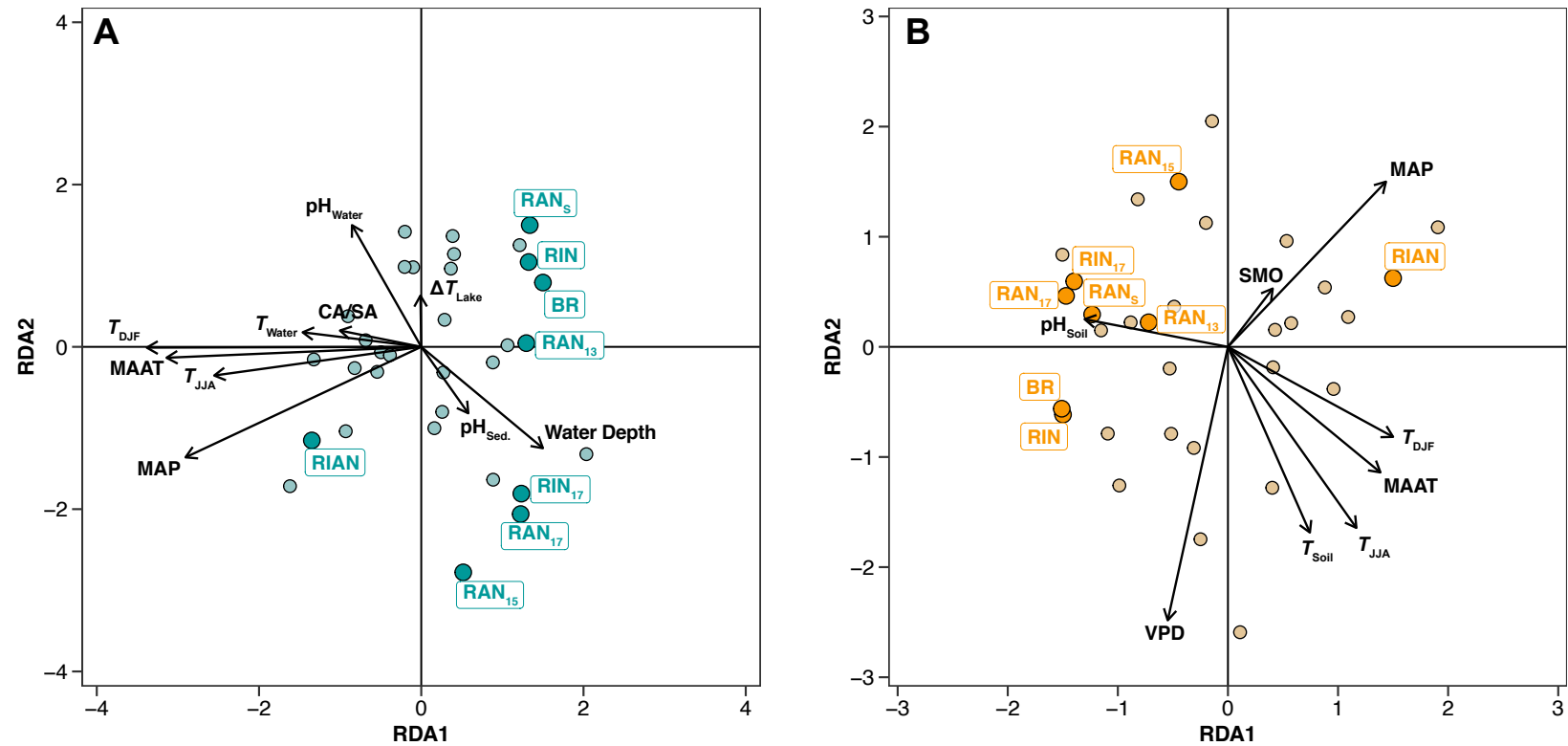


Figure 4.14: RDA using fractional abundance of established 3-OH FA proxies as response variables and environmental conditions as explanatory variables for **A** SW lake sediments and **B** SW soils

A notable observation from the Pearson's correlation matrix for environmental proxy response in lake sediments is the 3-OH FA proxies that significantly respond to **MAAT** have a seasonally biased response to T_{DJF} and T_{MAM} and are often more responsive to these seasonal temperatures compared to **MAAT**, particularly for T_{MAM} with which RAN_{13} portrays most significant variability ($r^2=0.45$, $p<0.001$) (Fig. 4.15B). Additionally, $RIAN$ has significant positive linear correlations with T_{DJF} and T_{MAM} , but not **MAAT** (Appendix B11). The influence of seasonal temperature on 3-OH FA proxies in the SW lake sediments is consolidated by the significant responses between RAN_{17} , RAN_{13} , RAN_5 , BR , $RIAN$, and $T_{JJA}-T_{DJF}$, which are positive for all proxies excluding $RIAN$ (Appendix B11). Generally, these results indicate that *anteiso*-3-OH FAs have a tendency to increase in value at lake catchments that experience larger offsets between T_{JJA} and T_{DJF} . As *anteiso*- and *normal*-3-OH FAs have opposed but significant responses to T_{DJF} , T_{MAM} , and seasonality, that is not restricted to one RAN rendition, and Hellequin et al. (2023) report that representatives of Bacteroidetes are capable of adjusting the relative abundance of all *anteiso*-3-OH FAs in response to growth temperature, we also investigate the response of RAN (Eq. 4.4) to temperature variables. RAN describes the summed fractional abundance of *anteiso*-3-OH FAs regressed against that of *normal*-3-OH FAs, and produces a significant negative linear correlation with **MAAT** ($r^2=0.29$, $p=0.01$) and T_{MAM} ($r^2=0.32$, $p<0.01$), but does not supersede the correlation coefficient of RAN_{13} when regressed with either. Therefore in SW lake sediments, RAN_{13} represents the most significant variability in response to temperature of the established 3-OH FA proxies, and is independent of pH. RAN_{13} covers a **MAAT** range of 14.3-21.9 °C (mean=18.2 °C, SD=2.4 °C) producing an RMSE of 1.98 °C when regressed with **MAAT** (Fig. 4.15A), and a T_{MAM} range of 9.7-17.2 °C (mean=13.6 °C, SD=2.3 °C) producing an RMSE of 1.88 °C when regressed with T_{MAM} (Fig. 4.15B). RAN_{13} did not produce values in lake sediments from Lake Junction (USA062) or Medina Lake (USA063) because *anteiso*-3-OH FAs are not present in their unique 3-OH FA distributions.

$$RAN = \sum [anteiso-] / \sum [normal-] \quad (4.4)$$

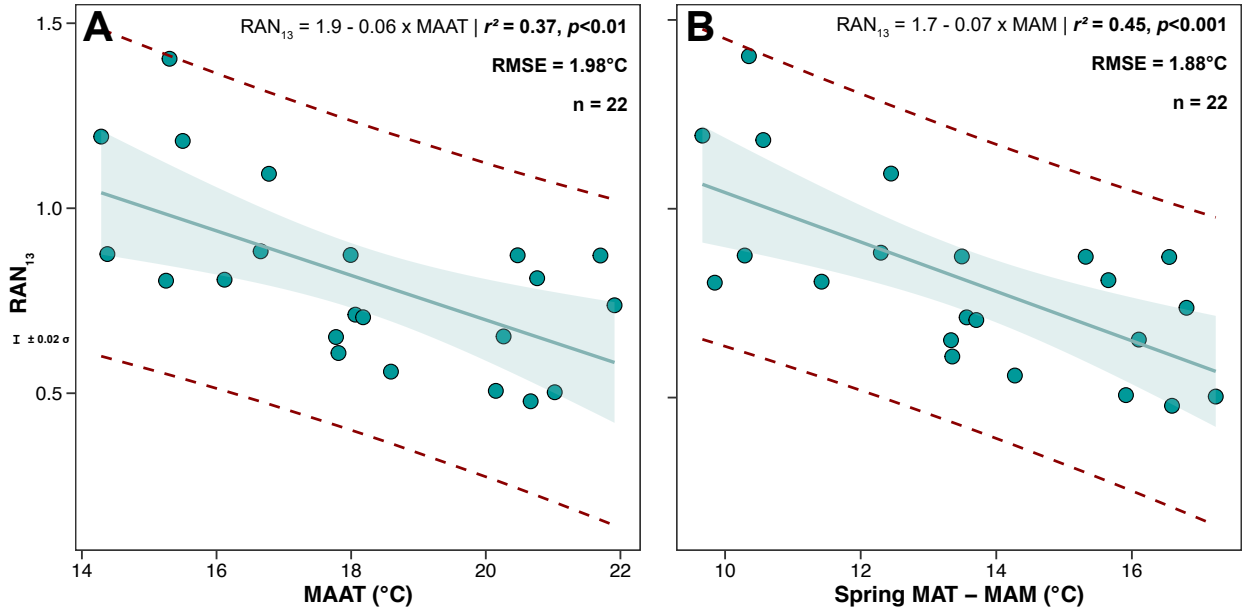


Figure 4.15: Scatterplots between RAN_{13} and **A** $MAAT$ and **B** T_{MAM} in SW lake sediments

$$MAAT (^{\circ}C) = 23.1 - 6.11 \times RAN_{13} \quad (4.5)$$

$$T_{MAM} (^{\circ}C) = 19.2 - 6.92 \times RAN_{13} \quad (4.6)$$

Environmental variables applied in RDA analysis for SW soils explain 53 % constrained proxy variance (Fig. 4.14B). Although less constrained variance is explained by pedogenic environmental variables for SW soils, than by lacustrine environmental variables for SW lake

sediments, the RDA model for soils remains globally significant ($f=2.12$, $p<0.01$). **MAAT** is the predominant environmental driver of 3-OH FA proxy variance in soils ($f=7.78$, $p<0.001$), however T_{DJF} ($f=2.83$, $p<0.05$) and T_{JJA} ($f=2.85$, $p<0.05$) are also revealed as significant environmental descriptors of proxy variance. Pearson's correlation analysis indicates significant correlations between all 3-OH FA proxies and **MAAT** excluding RAN_{15} and RAN_{13} , the latter of which produced the most significant variability in response to temperature in lake sediments (Appendix B14). Instead, RAN_S conveys the most significant temperature response of all proxies in soils, having a moderate and highly significant negative linear correlation with **MAAT** ($r^2=0.49$, $p<0.001$). 3-OH FA-based proxies in SW soils also appear to have a complex relationship with seasonal temperature regime that is potentially distinguished between RAN proxies (i.e. those that employ single 3-OH FA compounds as a numerator and denominator) and the cumulative proxies RIN, BR, and RIAN. RAN_{17} , RAN_S , and RIN_{17} have variability in response to **MAAT** and the seasonal renditions (T_{JJA} , T_{SON} , T_{DJF} , T_{MAM}) from which it is derived, whereas RIN, BR, and RIAN have a seasonally biased response to T_{DJF} and T_{MAM} (Appendix B14). This seasonal bias occurs in the cumulative proxies because they employ *iso*-3-OH FAs which alongside *normal*-3-OH FAs convey seasonal bias, which is not observed for *anteiso*-3-OH FAs (Appendix B14). However, almost all proxies that respond to **MAAT** also have significant variability in response to several other environmental variables, which are commonly **MAP** and soil pH (Appendix B14). For example, RAN_S also has a weak but highly significant positive correlation with soil pH ($r^2=0.36$, $p<0.01$), and therefore a legitimate 3-OH FA temperature response is difficult to deconvolve for the SW soils.

In short, the SW soils arise from environmental variability that is so heterogeneous in climate regime that a legitimate 3-OH FA-based proxy response is likely not identifiable with the current subset of soils despite the distinct variability of RAN_S with **MAAT**; a large proportion of SW soils are associated with an arid climate regime with low **MAP** and alkaline pH, and a small number of SW soils are associated with a more humid climate regime and neutral to acidic pH, for example.

In Chapter 3, the latter minority of SW soils - characterised by a threshold of $\text{pH} < 7.2$ - were found to have a comparable 3-OH FA proxy response to humid and acidic EC soils. However a temperature response in the omitted SW soils with $\text{pH} > 7.2$ was not investigated, which we now proceed to do.

Omitting acidic SW soils with $\text{pH} < 7.2$ from the SW soil dataset condenses 3-OH FA proxy response to environmental variables considerably. RAN_S is the only proxy to have a weak but significant negative correlation with MAAT ($r^2 = 0.27$, $p < 0.05$) (Appendix B15), and this regression is visualised in Figure 4.16. Furthermore, RAN_S does not significantly vary in response to $T_{\text{JJA}} - T_{\text{DJF}}$, which is consistent with no seasonal bias to either T_{JJA} or T_{DJF} (Appendix B15). RAN_S exhibits no significant variability in response to soil pH ($r^2 = 0.11$, $p = 0.17$) despite moderate covariance between MAAT and pH ($r^2 = 0.44$, $p < 0.01$), suggesting RAN_S likely represents a legitimate temperature response in this arid-influenced subset of SW soils. RIN , BR , RIAN , and RAN_{17} only retain their significant responses to MAP and $T_{\text{JJA}} - T_{\text{DJF}}$ (excluding RAN_{17}) in this soil subset, which is consistent with the cumulative variability of *iso*- and *anteiso*-3-OH FAs to MAP (Appendix B15) and may suggest that SW soils are influenced more so by MAP compared to lake sediments. No proxies have a significant response to pH in this soil subset likely due to the restricted pH range.

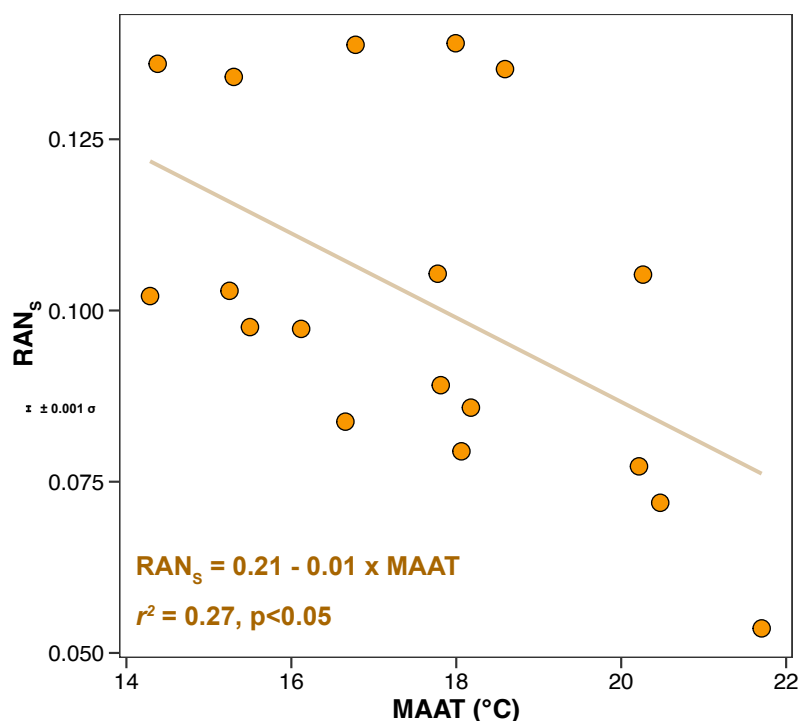


Figure 4.16: Scatterplot between RAN_s in SW soils with pH>7.2.

Overall, established 3-OH FA proxies have a different temperature response between SW lake sediments and soils, as was observed with the non-established temperature indices. RAN₁₃ describes the most significant 3-OH FA temperature variability amongst established proxies in lake sediments, whilst RAN_s was identified as the most viable temperature response amongst established proxies in SW soils with pH>7.2, the alkaline and low MAP subset of soils that did not conform to the RAN₁₇ temperature response of humid acidic U.S. soils in Chapter 3. RAN_s does also have a significant relationship with temperature in lake sediments, however is superseded in correlation coefficient and significance by RAN₁₃. Conversely, RAN₁₃ does not produce any significant variability in response to temperature in any subset of SW soils; the primary 3-OH FA temperature response in soils is observed in lake sediments, but not vice versa, which is suggestive of an in-situ aquatic bacterial community with a different 3-OH FA temperature response to

bacterial communities that reside in soils.

4.3.5 Performance of GDGT-based environmental proxies in SW lake sediments and soils

MBT'_{5ME} conveys significant positive linear relationships with MAAT in both SW lake sediments and soils (Fig. 4.17A) which has been observed in both sediment groups across other transects (De Jonge et al. 2014b, Martínez-Sosa et al. 2021, Russell et al. 2018, Bauersachs et al. 2024).

The MBT'_{5ME}-MAAT regression produces a stronger correlation coefficient in SW soils than lake sediments (Fig. 4.17A). MBT also produces a significant correlations with MAAT for both sediment types, however the slightly improved regressions of MBT'_{5ME} suggest that omitting 6-methyl brGDGTs from MBT (and MBT' by association) does remove a small amount of confounding temperature bias. However MBT'_{5ME} has a significant negative linear relationship with pH in SW soils (Fig. 4.17B), which suggests that removing 6-methyl brGDGTs from MBT does not remove the pH bias of the proxy. As 6-methyl brGDGTs were not detected in a large number of SW soils, a possible cause of the pH bias is that separation of 6-methyl brGDGTs was not achieved during analysis. However this is unlikely because 6-methyl brGDGTs were detected in half of the SW soils and most lake sediments, and all were subjected to the same method. Furthermore, MBT'_{6ME} does not generate a significant response to pH (or lake surface water pH for lake sediments) in either soils or lake sediments despite successful separation. No significant variability was observed between any other GDGT-based proxy and either MAAT or pH.

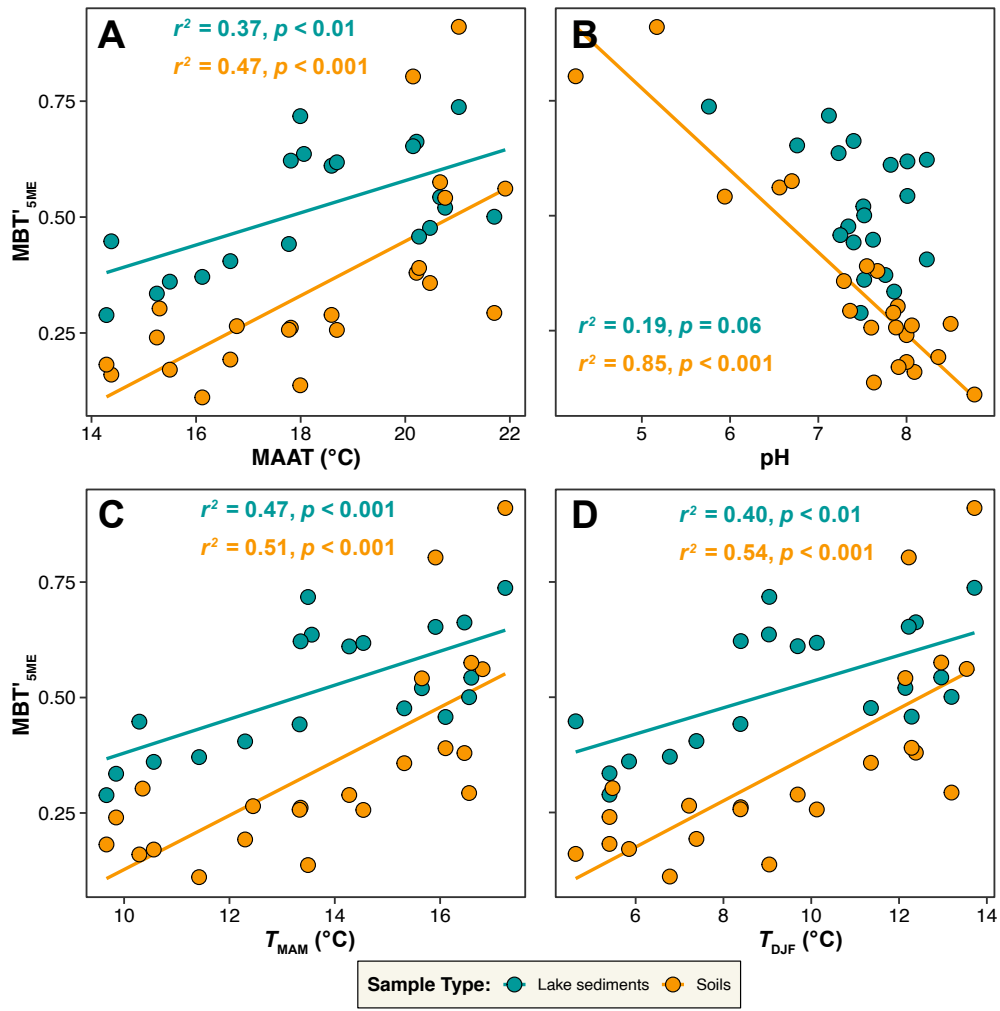


Figure 4.17: Scatterplot matrix between MBT'_{5ME} and **A** MAAT, **B** respective lake sediment and soil pH, **C** T_{MAM} , and **D** T_{DJF} in SW lake sediments and soils.

$$MAAT \text{ (}^{\circ}\text{C)} = 12.7 + 10.8 \times MBT'_{5ME} \quad (4.7)$$

$$T_{MAM} \text{ (}^{\circ}\text{C)} = 7.23 + 12.7 \times MBT'_{5ME} \quad (4.8)$$

A particularly interesting observation is that in lake sediments, the most substantial MBT'_{5ME}

temperature response occurs with T_{MAM} (Fig. 4.17C), and in soils with T_{DJF} (Fig. 4.17D). The seasonal T_{MAM} bias in lake sediments is consistent with that observed for RAN_{13} , and so the bias is consistent amongst different biomarker groups. However no seasonal T_{DJF} bias was observed for RAN_5 , although the cumulative proportion of *iso*- and *normal*-3-OH FAs were significantly responsive to MAP , and winter-spring rainfall is an important environmental characteristic of the SW U.S. climate.

4.4 Discussion

4.4.1 3-OH FA distributions across lacustrine catchments in the context of global variability

4.4.1.1 3-OH FA distributions in lake sediments

3-OH FA distributions are relatively unexplored in lake sediments at the global scale, and have only been reported from one regional lake transect comprising 20 lakes from central China (Yang et al. 2021) aside from the present study. As the Chinese lake transect comprises multiple sediments from the same lake (Lake Daihai $n=3$, Lake Zhenzhu $n=2$, Lake Hongu $n=2$), 3-OH FA data were firstly averaged between the samples originating from each of these three lakes following the data wrangling method of Martínez-Sosa et al. (2021) who only retained distinguished GDGT compositions from multiple sediments for Lake Victoria (Africa) due to its nature as a global Great Lake, in preparation for development of a global lacustrine MBT'_{5ME} temperature calibration using a Bayesian statistical model. Standardising 3-OH FA compositions across multiple sediments from the same lake in this way reduces the confounding influence of intra-lake biomarker variability on inter-lake 3-OH FA proxy-based calibrations, which is particularly important as intra-lake 3-OH FA variability has not been investigated previously for 3-OH FAs, however is not investigated in this study as intra-lake variability is investigated using a detailed sediment transect sampled across a lake which is not the purpose of this study as a regional-scale calibration study.

Amongst the SW U.S. lakes, lake sediment pH was uncoupled from lake surface water pH,

which is a common phenomena as pH (particularly when highly acidic or alkaline) can be neutralised by sedimentary biogeochemical processes (Herlihy & Mills 1986) particularly with increasing depth from the sediment-water interface where microbial-mitigated reduction of cations such as Fe^{3+} and SO_4^{2-} occurs, which has been observed amongst a number of lakes influenced by acid mine drainage (Bachmann et al. 2001). Furthermore, lake water pH is much more ambient to local bedrock chemistry and seasonal lake thermodynamics (Zhao et al. 2022, Bachmann et al. 2001), whilst lake sediments are more likely to record average lake water chemistry (Geladi et al. 1999).

Total 3-OH FA concentrations are much higher in SW U.S. lake sediments (<0.001 to 3.70 ± 0.89 mg/g LOI; mean= 0.64 ± 0.87 mg/g LOI (*SD*)) compared to those reported from Chinese lake sediments (2.28 to 85.4 $\mu\text{g/g}$ TOC, mean= 31.5 ± 24.1 $\mu\text{g/g}$ TOC (*SD*)). Despite the vast difference in concentration, 3-OH FA distributions in lake sediments are fairly similar between the two transects; there is no significant difference in mean composition of 3-OH FA fractional abundance ($w=105714$, $p=0.67$), although $n\text{-C}_{16}$ is predominant in Chinese lake sediments (mean= 22.3 %, $\text{SD}=3.56$ % total 3-OH FAs) unlike U.S. lake sediments in which $n\text{-C}_{16}$ is dominant in a number of samples but cumulatively $n\text{-C}_{14}$ yields the highest 3-OH FA fractional abundance on average (mean= 23.4 %, $\text{SD}=5.63$ % total 3-OH FAs). Variance of composite 3-OH FA fractional abundance is also homogeneous between the two lake transects ($f=<0.001$, $p=0.98$). In an NMDS of lake sediment 3-OH FA distributions compiling both transects, the variability of Chinese lake sediments (indicated by a 95 % confidence ellipsis) falls within that of SW U.S. lakes (Fig. 4.18). Sediments from lakes Junction (USA062) and Medina (USA063) of the SW U.S. transect have such unique 3-OH FA distributions amongst all lake sediments that the initial NMDS computation failed, and could only be computed upon second attempt by omitting these samples from the lake sediment dataset. An ANOSIM test additionally confirms homogeneity in 3-OH FA distribution between Chinese and SW U.S. lake sediments as the low ANOSIM $R = 0.02$ statistic ($p=0.25$) indicates grouping of sediments based on transect has a negligible capability of describing

dissimilarity in 3-OH FA composition amongst the lakes. Overall, the two lake sediment transects do not produce distinguished clusters of 3-OH FA distributions, and although 3-OH FA distributions are more variable amongst SW U.S. lakes (larger 95 % confidence ellipsis in NMDS), this variability is not significantly different from that of Chinese lakes.

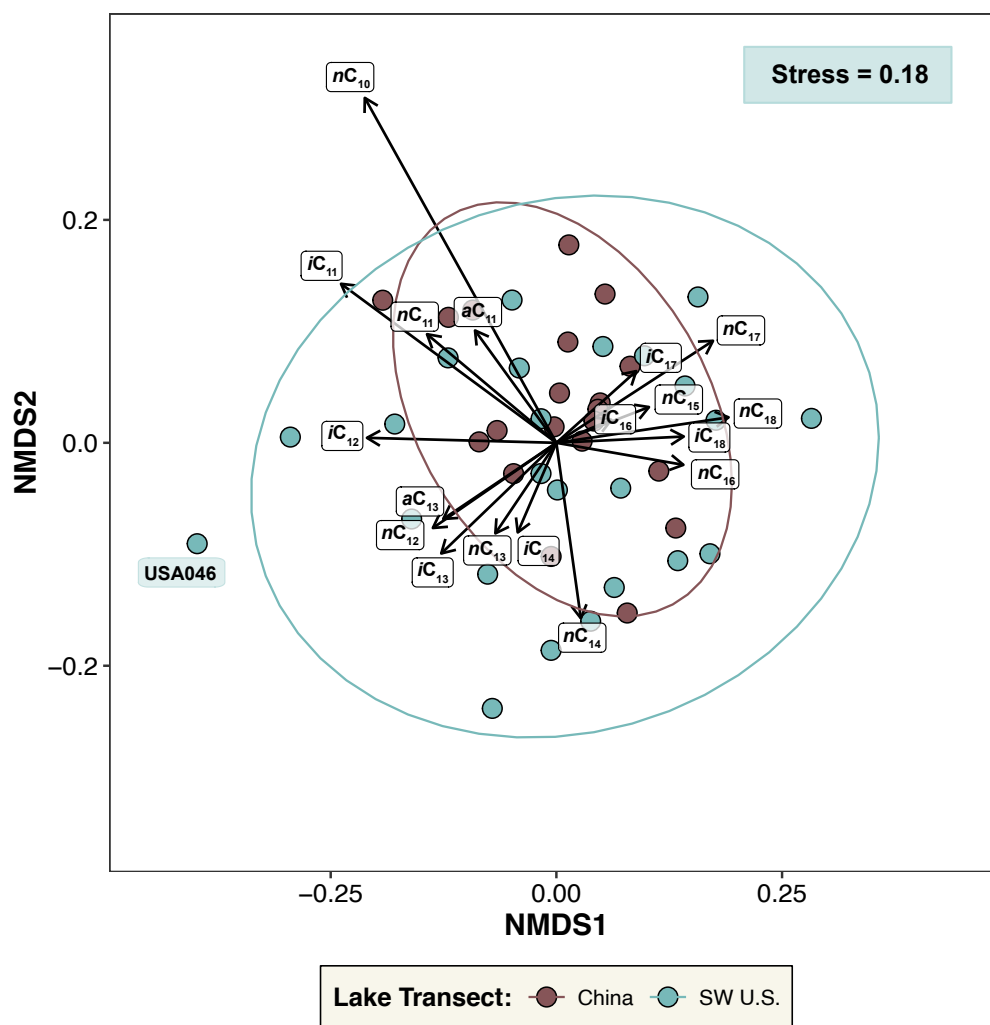


Figure 4.18: NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of SW lake sediments and the Chinese lake sediments of [Yang et al. \(2021\)](#). 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised.

Higher 3-OH FA concentrations and more variable (albeit not significant) 3-OH FA distributions across SW U.S. lake sediments suggests the environmental and biological conditions of the SW lakes promote more 3-OH FA production with more opportunity for variability than those of the Chinese lakes. To address the former, the environmental conditions that each transect encompass differ considerably. All lakes across both transects were alkaline in nature at the time of sampling (Yang et al. 2021), however U.S. lakes are exposed to lower MAP, higher MAAT, and are deeper compared to the Chinese lakes, and each variable has potential to influence bacterial productivity in lacustrine environments. Deeper lakes paired with increased surface area will have a larger photic zone for microbial production, more volume in which turbidity can occur, and more complex stratification regimes (van Bree et al. 2020), overall influencing the productive capacity of lakes (Van Bree et al. 2018). However, MAP in combination with VPD (and relative humidity by association) can particularly influence lake water salinity which is a primary driver of bacterial productivity in lakes (Pedrós-Alió et al. 2000, Yang et al. 2016); in an arid region such as the SW U.S., periods of drought that are becoming increasingly pronounced and uncertain (Apurv & Cai 2021) provoke lake water evaporation which is known to cause extremely variable active storage capacity of the lakes, and has been reported in detail for reservoirs of the Texan Canadian River watershed within which lakes Ute, Meredith, and Conchas occur (Brauer et al. 2015).

Variability of the active storage capacity of the SW lakes can in turn influence water salinity, and we speculate that whilst the Chinese lakes were freshwater in nature at the time of sampling (Yang et al. 2021), many of the SW lakes had the potential to be saline. Water chemistry data such as dissolved oxygen, redox potential, and conductivity are not available from the time of sampling for U.S. lakes as they were for the Chinese lakes, however historical water chemistry reports for a select number of Texan reservoirs (lakes Houston, Colorado City, Medina, Buffalo Springs, E.V. Spence, Somerville) (Burley et al. 2011) reveal that lakes Buffalo Springs, Colorado City, Meredith, and E.V. Spence have a history of high specific conductance (SC) (2030 ± 16 - $4649 \pm 285 \mu\text{S cm}^{-1}$ (SE)) and chloride (Cl^-) concentration (332 ± 14.6 - $1274 \pm 112 \text{ mg L}^{-1}$ (SE)) which

are both indicative measures for salinity. These reported values have not been standardised in consideration of outliers caused by unique phenomena (e.g. fertiliser spill events) or measurement depth, however [Gelca et al. \(2016\)](#) did standardise the same water quality data across a plethora of Texan reservoirs including those mentioned above, and report that although water quality is not driven by a specific environmental predictor, increasing SC and lake water salt content are consistent with increasing aridity in the region.

Biologically, many water bodies across the SW U.S. are either known or suspected to have algal blooms ([Ketelle & Uttormark 1971](#), [Backer et al. 2013](#), [Patiño et al. 2023](#)). Water quality data from SW lakes, particularly in Arizona and New Mexico, is limited, however there are frequent reports of canine illness related to algal blooms in SW U.S. water bodies and many cases which are likely unreported ([Backer et al. 2013](#)). Algal blooms have been consistently and unofficially reported amongst SW water bodies, and are supported by the presence of cyanobacteria in the bacterial communities of most SW U.S. lakes ([Thorpe 2023](#)), which are not obviously abundant in bacterial community compositions from Lake Liangzi, the only Chinese lake from which lake sediments have been subject to 16S RNA sequencing ([Yang et al. 2021](#)). Persistent annual-scale algal blooms have been reported from Lake Colorado City and E.V. Spence Reservoir, and a sub-decadal scale from Buffalo Springs Lake ([Keith et al. 2018](#)). All suggest the SW U.S. lakes can be biologically productive; as algal blooms are commonly induced by an abundance of cyanobacteria in lakes, a productive algal community is indicative of a productive bacterial community. It is therefore plausible that highly productive bacterial communities could produce the higher 3-OH FA concentrations observed in SW U.S. lakes. The lake transects are suspected to have substantial differences in salinity related to climate regime, however despite the environmental heterogeneity between the transects, 3-OH FA distributions are fairly homogeneous. We additionally note that a large proportion of environmental variability in lacustrine environments is not described by those reported for SW U.S. lakes, which include but are not limited to salinity, conductivity and dissolved oxygen. Nutrient availability may also play a

role in biological productivity of the lakes, however have not been measured in our study and has not been reported from any previous monitoring of the SW U.S. lakes. Lakes Junction (USA062) and Medina (USA063) produced the most unique 3-OH FA distributions amongst all lakes, however the available environmental conditions for these sites reveal no unique variability.

4.4.1.2 3-OH FA distributions between lake sediments and soils

3-OH FA distributions are different between SW U.S. lake sediments and soils, which also occurs globally between lake sediments (Yang et al. 2021), soils (Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021a,b, Wang et al. 2021), and marine sediments (Yang et al. 2020, Dong et al. 2023) as revealed by an NMDS plot in which 3-OH FA distributions form distinguished clusters based on sediment type (Fig. 4.19). This is supported by ANOSIM analysis which indicates that grouping global 3-OH FA distributions based on sample type significantly describes clustering within the 3-OH FA dataset (ANOSIM $R=0.68$, $p<0.01$) (Fig. 4.19).

Samples with 3-OH FA distributions plotting outside the 95 % confidence ellipses of their respective sediment type were omitted from the 3-OH FA dataset and the fractional abundances for each 3-OH FA compound was averaged by sample type and visualised in Figure 4.20. Visualising 3-OH FA distributions between soils, lake, and marine sediments in this way highlights fundamental similarities and differences in composition between the sediment types. The high abundance of $n\text{-C}_{14}$ and *normal*-3-OH FAs is consistent between all sediment types, for example. The $i\text{-C}_{17}$ and $n\text{-C}_{18}$ enrichment observed in SW soils also prevails across 324 globally distributed soils (Fig. 4.20), albeit $i\text{-C}_{17}$ is enriched to a smaller degree in global soils and from our Chapter 3 discussion, $i\text{-C}_{17}$ enrichment is possibly associated with arid soils (Wang et al. 2021). The $n\text{-C}_{16}$ and $a\text{-C}_{17}$ enrichment observed in SW lake sediments also remains when the SW lake sediments are combined with those from China (Yang et al. 2021) to produce a dataset of 43 lake sediments (Fig. 4.20). 81 marine sediments from the marginal seas of China also produce a distinct 3-OH FA distribution from lake sediments and soils with enrichment of $i\text{-C}_{13}$, $i\text{-C}_{14}$, and $a\text{-C}_{15}$ compared to

lake sediments and soils ((Yang et al. 2020, Dong et al. 2023); Fig. 4.20). However marine sediment 3-OH FA distributions have a high degree of variability compared to lake and soil 3-OH FA distributions (Fig. 4.19). Overall, marine and lake sediments and soils each have different 3-OH FA distributions, which are cumulatively summarised in Figure 4.20.

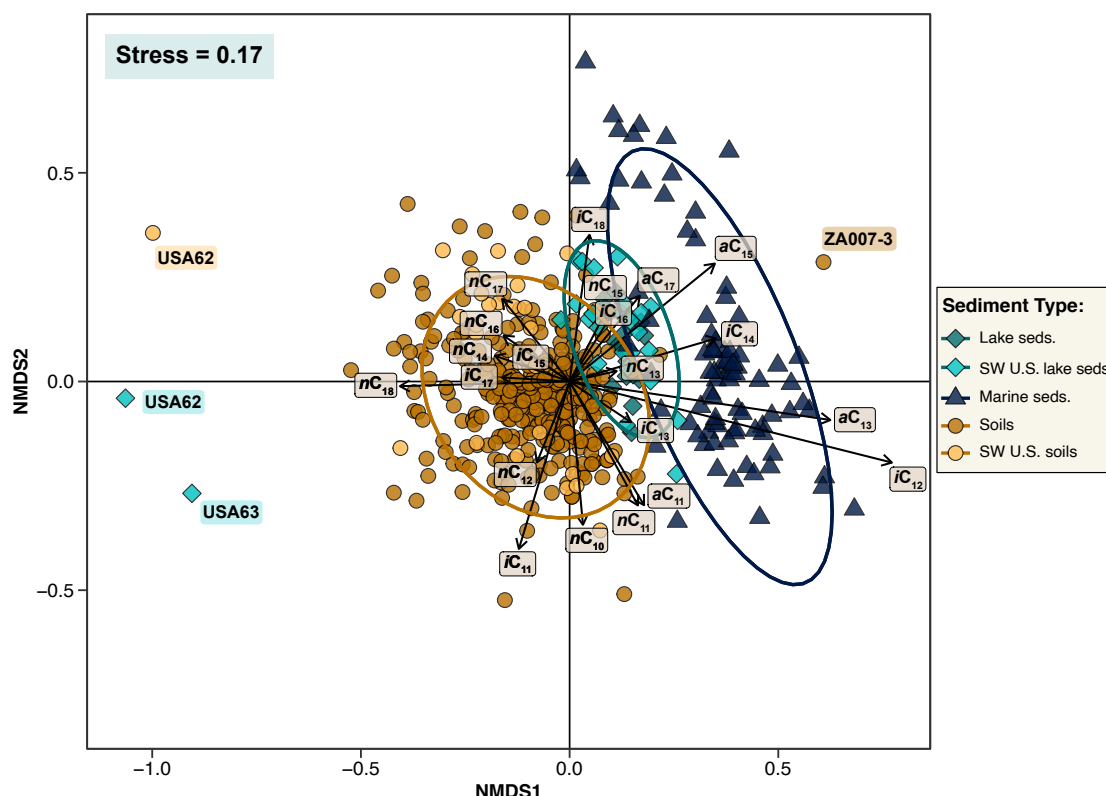


Figure 4.19: NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of global sediments investigated for 3-OH FAs. 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised. Soils are derived from Wang et al. (2021), Véquaud et al. (2021b), Véquaud et al. (2021a), Wang et al. (2016) and Huguet et al. (2019), and SW soils. Lakes sediments comprise SW lakes in this study and those of Yang et al. (2021). Marine sediments comprise those from the marginal seas of China reported by Yang et al. (2020) and Dong et al. (2023).

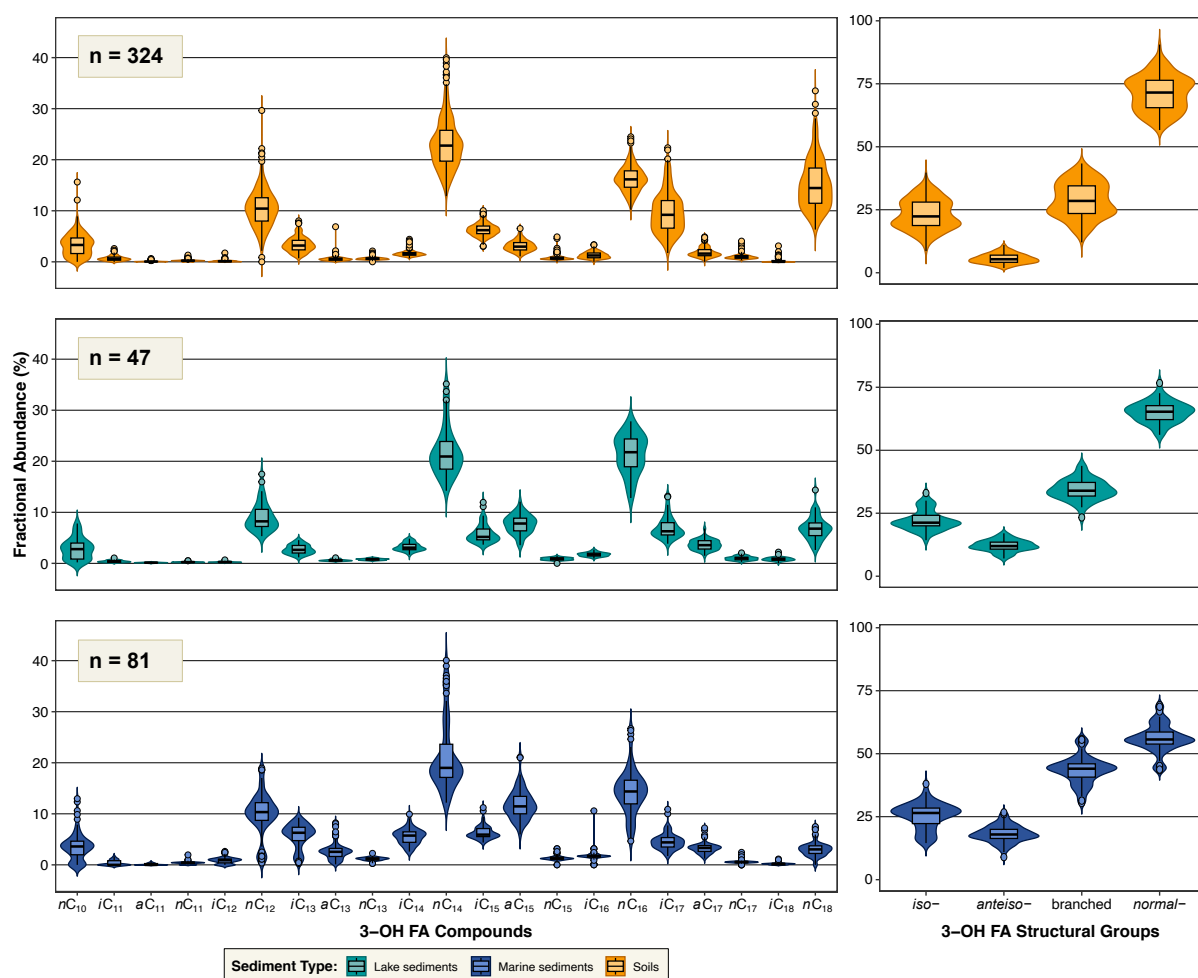


Figure 4.20: Distributions of 3-OH FA fractional abundance amongst global sediments for (Upper) soils, (Middle) lake sediments, and (Lower) marine sediments. Plots to the left visualise individual 3-OH FA compounds, and plots to the right visualise 3-OH FA structural groups.

The 3-OH FA compositional profiles based on sediment type have an important role as a global 3-OH FA database. These profiles could facilitate initial work to differentiate 3-OH FA compositions between sediment type, and those for soils and lake sediments could be used as standard compositions to identify strong regional 3-OH FA variability. The fractional proportion of summed *iso*-3-OH FAs is homogeneous between soils and lake sediments ($p=0.40$). However

that of *anteiso*- ($H(2)=266, P<0.001$) and *normal*-3-OH FAs ($H(2)=180, P<0.001$) significantly differ between all sediment types as confirmed by Kruskal-Wallis rank sum tests (non-parametric ANOVA) and subsequent paired samples Wilcoxon test to identify significance of pairwise relationships (Fig. 4.20); *anteiso*-3-OH FAs become progressively more enriched from soils through to marine sediments, and the reverse is observed for *normal*-3-OH FAs. This 3-OH FA variability suggests the ratio of *anteiso*- to *normal*-3-OH FAs (RAN) should produce variability based on sediment type (Fig. 4.21), and indeed, the distribution of RAN does significantly differs between each of the three sediment groups in question ($H(2)=266, P<0.001$), RAN confidence intervals for each sediment type are presented in Figure 4.21. Overall, there is a wealth of evidence indicating lake sediments produce 3-OH FA distributions that are intermediary between those of soils and marine sediments, and are more comparable to those of soils than marine sediments possibly because lakes and soils are terrestrial environments and lakes directly receive input from catchment soil erosion (Yang et al. 2021). 3-OH FA distributions between soils and marine sediments are completely be distinguished.

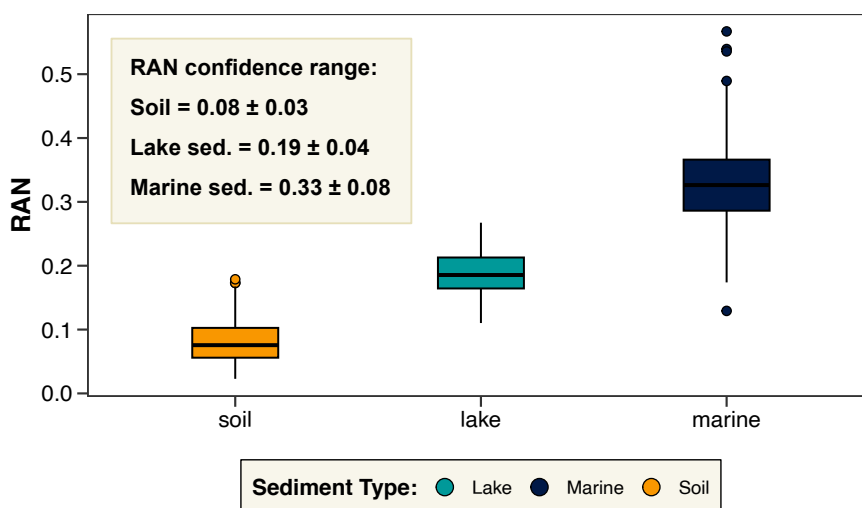


Figure 4.21: RAN distributions amongst globally distributed samples as boxplots. The RAN confidence ranges are also shown, calculated as mean and standard deviation.

3-OH FA compositional clustering based on sediment type suggests that the sources of 3-OH FAs differ between soils, lake, and marine sediments, which may be the result of in-situ biological production unique to each sediment type. Evidence for in-situ 3-OH FA production in lakes arises from significantly higher total 3-OH FA concentrations ($w=567$, $p<0.001$) in SW lake sediments (mean= 0.64 ± 0.89 mg g⁻¹ LOI, SD=0.87) relative to soils from their surrounding catchments (mean= 0.06 ± 0.37 mg g⁻¹ LOI, SD=0.07). However, the opposite was reported by (Yang et al. 2021) from lakes Yuqiao, Chagan, and Daihai whereby catchment soils had higher total 3-OH FA concentrations (mean= 0.64 mg g⁻¹ TOC, SD=0.37) than lake sediments (mean= 0.04 mg g⁻¹ TOC, SD=0.02), albeit concentrations from soils samples within the same catchment were extremely variable (Yang et al. 2021). High 3-OH FA concentrations in SW lake sediments relative to counterpart catchment soils may result from the arid to semi-arid climate that most SW soils are subject to; if soil moisture and MAP are limiting factors for bacterial respiration in many soils, lower 3-OH FA concentrations relative to lake sediments would be expected. However higher brGDGT concentrations have been observed in lake sediments relative to adjacent catchment soils amongst tropical lake catchments where aridity is not an environmental limitation for microbial growth (Tierney & Russell 2009, Loomis et al. 2011). Additionally, *i*-C₁₂ and *i*-C₁₈ are abundant in lake sediments, but not soils from surrounding lake catchments and have a limited presence in soils globally (Fig. 4.19) suggesting the bacterial producers of these compounds are more abundant in aquatic environments.

Gram-negative bacteria are the predominant biological producers of 3-OH FAs (Wollenweber & Rietschel 1990, Alexander & Zähringer 2002), therefore investigating the composition of bacterial communities that reside in environmental samples can also deconvolve how bacterial 3-OH FA sources differ between sediment types and whether in-situ 3-OH FA production occurs in lacustrine environments. Bacterial community structures that have been reported from a small number of soils, lake and marine sediments have been highly distinguished between sediment type at the phylum level, with marine sediments from the South China Sea conveying an almost

complete dominance of Proteobacteria (>90 %) (Yang et al. 2020). A more diverse range of bacterial phyla have been reported to compose community structures in soils and lake sediments (Yang et al. 2020, 2021). The bacterial communities of soils and lake sediments from the catchment of Lake Liangzi (China) retain dominance of Proteobacteria (ca. 55 %), which is abundant at almost equal proportion between the two sedimentary groups, however Chloroflexi (ca.10 %) and Bacteroidetes (ca.14 %) are additional primary contributing phyla in the lake sediments, whilst Acidobacteria is more abundant in soils than lake sediments (Yang et al. 2021). This bacterial composition data derived from 16S RNA sequencing is reported from a limited number of soils (3 each for lake and marine sediments, and soils), however such data is highly detailed for SW U.S. lake sediments and soils, all for which bacterial community structures have been reported (Thorpe 2023). Bipartite cross-correlation network analysis reveals distinct bacterial ASVs (excluding one) between lake sediments and soils; although similar phyla can compose bacterial structures for different sediments, the phyla may be composed of completely different groups of taxa (i.e. ASVs). The phyla Desulfobacteria, Chloroflexi, and Cyanobacteria are highly associated with SW lake sediments and not soils (Thorpe 2023). Instead, SW soils have more homogeneous bacterial community structures that are dominated by Proteobacteria and Acidobacteria (Thorpe 2023), which was also observed in Lake Liangzi soils, and are dominant phyla in soils globally (Wei et al. 2018). However, Cyanobacteria and Desulfobacteria, which are characteristic phyla for SW U.S. lake sediments, are not characteristic for Lake Liangzi (Yang et al. 2021). The composite differences in Gram-negative bacterial community structure between sediment types suggest why clustering of 3-OH FA distributions might occur; marine sediment bacterial community structures have a completely different composition to the more heterogeneous compositions of lake sediments and soils, whilst those of lake sediments and soils were more comparable in heterogeneity with differing abundance of more minor phyla.

Although bacterial community structures are distinguished at the phylum level between SW lake sediments and soils, and appear so in a limited number of external samples to the U.S., it is a

complex endeavour to directly associate composite 3-OH FA distributions within each sediment type to a bacterial community. Lipodomic assessments of pure bacterial cultures are often used to investigate the endogenous lipid groups produced by specific bacterial strains, and commonly report fatty acid composition ([Kaneda 1991](#)). However a small number of studies have specifically assessed 3-OH FA composition, often as a growth experiment in response to specific growth conditions such as temperature; a fundamental endeavour to understand the biological mechanism responsible for microbial lipid-based temperature proxies. Endogenous C₁₅ to C₁₈ 3-OH FAs have been identified in strains of soil-derived Bacteroidetes ([Hellequin et al. 2023](#)). Strains of Bacteroidetes have also been associated with *i*-C₁₅ and *i*-C₁₇ ([Nordström & Laakso 1992](#), [Oyaizu & Komagata 1981](#), [Yamanaka et al. 1988](#)) in addition to *i*-C₁₆ and *n*-C₁₆ ([Mayberry 1980](#)). However lipodomic culture assessments are highly specific, and culture isolates are not representative of the ecological niches and biotic interactions that occur with a natural bacterial community ([Al-Awadhi et al. 2013](#), [Stubbendieck et al. 2016](#)). Thousands of bacterial strains can be associated with a phylum and strains under the same phylum can have different lipodomic compositions. Therefore, although SW lake sediments are associated with a high abundance of Chloroflexi, Desulfobacteria, and Cyanobacteria relative to soils, it is currently impossible to conclude that bacterial strains within these phyla are directly responsible for the uniquely clustered 3-OH FA composition from lake sediments. Nevertheless, evidence for in-situ 3-OH FA production in lakes is surmounting.

Although lacustrine 3-OH FAs are produced by Gram-negative bacteria, and likely by an in-situ lacustrine community as indicated by 16S RNA sequencing ([Thorpe 2023](#)), it remains unknown whether 3-OH FAs are produced by bacteria residing in the water column, sediment, or both. For GDGTs, sources of in-situ lacustrine production have been investigated in detail at Lake Chala ([Baxter et al. 2021](#)). A detailed investigation of aquatic brGMGT production in Lake Chala revealed high brGMGT production in anoxic lake sediments in comparison to suspended particulate matter and both had substantially different brGMGT distributions ([Baxter et al. 2021](#)).

Lake Chala is also permanently stratified which would promote continuous preservation of organic matter below the oxycline in permanently anoxic lake sediments, and the unique conditions of this lake are not applicable to all lakes. Overall, in-situ production of 3-OH FAs in the SW lakes is likely to be occurring. In-situ biomarker production in lakes has been reported previously ([Loomis et al. 2011](#)), however the most detailed studies are reported from lakes that have been studied consistently for a long period of time.

4.4.2 Consistency of temperature response amongst lipid biomarker distributions derived from lake sediments

4.4.2.1 3-OH FA-based temperature calibrations

Although 3-OH FA distributions are similar between lake sediment transects and the composite variability of Chinese lake sediment 3-OH FA distributions falls within that of SW U.S. lake sediments, deconvolving 3-OH FA variability in response to temperature between the transects is more complex. [Yang et al. \(2021\)](#) proposed two 3-OH FA-based temperature calibrations from the Chinese lake sediments, one for RAN_{13} and one for RIN_{17} . As the RAN_{13} regression with $MAAT$ comprises a high degree of scatter particularly at low temperatures as a result of intra-lake 3-OH FA variability, we revised RAN_{13} and RIN_{17} calibrations from the Chinese lakes using averaged values across lake sediments from the same lake that were calculated in this study. We note that RAN_{13} produces significant negative correlations when regressed against $MAAT$ amongst both Chinese and U.S. lakes, which occurs across a temperature range of 14.3-21.9 °C for U.S. lakes and 4.9-17.0 °C for Chinese lakes ([Yang et al. 2021](#)). The RAN_{13} regression differs in slope and intercept between the transects, although both are significant to the same moderation ($r^2=0.37$, $p<0.01$). However, the most viable proxy amongst Chinese lake sediments remains RIN_{17} , with revised RIN_{17} values comprising: Lake Daihai = 9.3 ± 0.16 , Lake Zhenzhu = 8.3 ± 0.05 , Lake Honghu = 4.2 ± 0.22 (Eq. 4.9).

$$MAAT (^{\circ}C) = 23.3 - 1.70 \times RIN_{17} (r^2 = 0.52, p < 0.001; RMSE = 2.78^{\circ}C) \quad (4.9)$$

We investigate three 3-OH FA-based temperature calibrations calculated in the present study: (i) a RAN_{13} - $MAAT$ linear regression from SW U.S. lakes with cold sites (USA043, USA053) omitted (Fig. 4.22A), (ii) a RIN_{17} - $MAAT$ linear regression from Chinese lakes (Fig. 4.22B), (iii) a $(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}$ - $MAAT$ linear regression revealed following PCA analysis on SW U.S. lakes with cold sites omitted (Fig. 4.22C). RAN_{13} was initially developed as a proxy for SST reconstruction in the marine realm, and has a substantially different temperature response in marine and lacustrine environments, having an exponential relationship with SST and $RMSE=2.55^{\circ}C$ (Yang et al. 2020), compared to the linear relationship and lower $RMSE$ in lake sediments. Of the three calibrations, the latter - the PCA-derived U.S. calibration - produces the lowest $RMSE$ ($1.63^{\circ}C$) compared to that of RAN_{13} ($1.98^{\circ}C$) and RIN_{17} ($2.78^{\circ}C$). The two calibrations based on SW U.S. lakes severely overestimate $MAAT$, which occurs $<13^{\circ}C$ for RAN_{13} , and $<ca.14.5^{\circ}C$ for the PCA-derived calibration (Fig. 4.22A&C).

Mean ΔT (offset between reconstructed temperature and $MAAT$) is lower for the RAN_{13} calibration (-0.01 , $SD=1.93$) compared to the PCA-derived calibration (-0.22 , $SD=1.80$) for the SW lakes. Although RAN_{13} predictions are more centred around 0, both of these calibrations supersede the performance of RIN_{17} in SW lakes which severely under-predicts temperature in SW lakes (mean $\Delta T=-11.2$, $SD=8.11$). However, a number of Chinese lake sediments fall within the ΔT confidence variability of lakes from which the calibration is derived for both calibrations. RIN_{17} consistently reconstructs $MAAT$ across the $4.9-17.0^{\circ}C$ temperature range of the calibration, whereby ΔT is consistently distributed around 0, and is capable of reconstructing $MAAT$ for 7 U.S. lakes (including cold sites) (Fig. 4.22B). $MAAT$ in U.S. lakes is otherwise consistently underestimated, particularly for Elephant Butte Reservoir (USA045) for which $\Delta T=-40.7^{\circ}C$.

Overall, the RIN₁₇-based calibration is robust at reconstructing MAAT in the Chinese and cold lakes, whilst the PCA-derived and RAN₁₃ calibrations had a comparable performance in SW lakes; although the PCA-derived calibration has a lower RMSE, the RAN₁₃ calibration produces a mean ΔT more centered around 0. All calibrations were able to reconstruct temperature in a number of sediments from the opposed lake transect to which the calibration was derived.

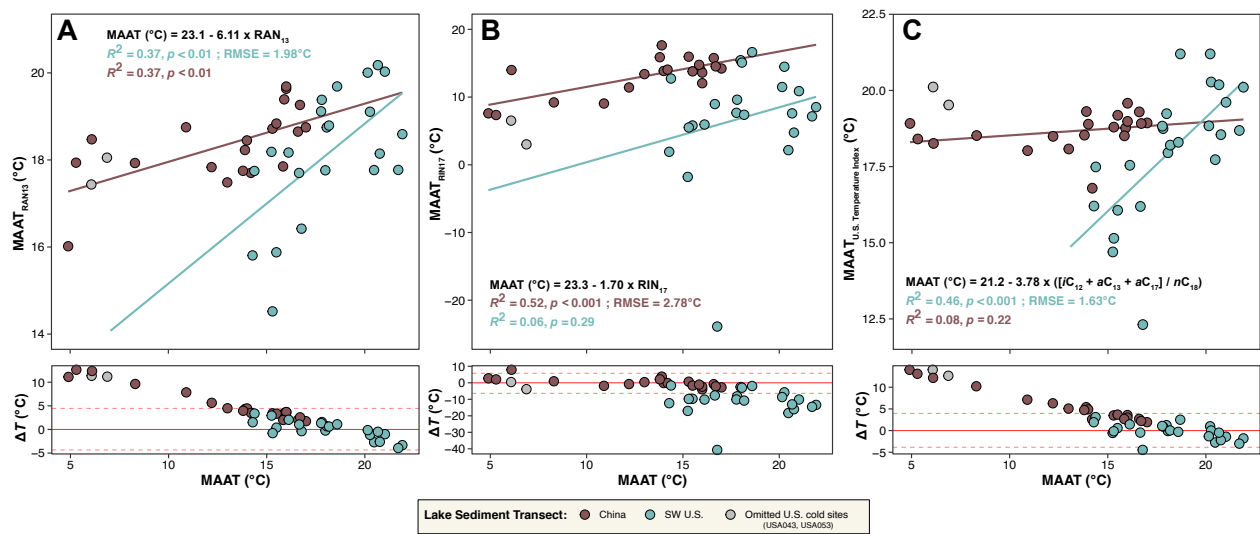


Figure 4.22: Scatterplots showing performance of three 3-OH FA-based calibrations in reconstructing MAAT in U.S. and Chinese lake sediments. Plots comprise reconstructed temperature vs MAAT of: **A** RAN₁₃ calibrated directly from SW lakes, **B** RIN₁₇ calibrated from Chinese lakes of Yang et al. (2021) with modifications outlined in main text, and **C** $[(i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18}]$ index calibrated directly from SW lakes. Below are plots of ΔT (reconstructed MAAT - MAAT vs MAAT).

Linear regression models are not the only calibration method to generate proxy-based temperature reconstructions; multiple linear regression (MLR) models have been applied in a number of studies following Tierney et al. (2010) and (Pearson et al. 2011) to develop transfer functions for temperature and pH reconstruction in lipid biomarker-based proxy calibrations. Stepwise forward selection (SFS) is a form of multiple linear regression invoking the sequential addition of variables (i.e. 3-OH FA or brGDGT compounds) to the regression until the addition of

compounds does not improve the model significance further (Chowdhury & Turin 2020, Bauersachs et al. 2024). MLR models including SFS have been applied in a number of calibration studies related to brGDGTs (Bauersachs et al. 2024, De Jonge et al. 2014a, Tierney et al. 2010, Russell et al. 2018) and 3-OH FAs (Véquaud et al. 2021b). We initially applied MLR across the entire lacustrine 3-OH FA dataset to identify any mutual temperature variability between the transects which did not yield any coherent results, however the best calibration was obtained from U.S. samples alone and is expressed in Equation 4.10:

$$MAAT (^{\circ}C) = 36.8 - 5.65 \times aC_{13} - 1.11 \times aC_{15} - 6 \times nC_{17} - 1.37 \times iC_{13} + 0.26 \times nC_{18}$$

$$(r^2 = 0.84, p < 0.001; RMSE = 1.12^{\circ}C) \quad (4.10)$$

This MLR-based calibration employs a -C₁₃, a -C₁₅, i -C₁₃, n -C₁₇, and n -C₁₈ (Fig. 4.23A). Both a -C₁₃ and n -C₁₈ are directly influenced by temperature across SW U.S. lakes, and 3-OH FA compounds from each of RAN₁₃, RIN₁₇, and the PCA-derived U.S. lakes calibration are incorporated into the MLR calibration. This calibration has a lower RMSE (1.12°C) than those presented in Figure 4.22, which is consistent with previous development of MLR-based calibrations where lower RMSE and higher correlation coefficients were observed relative to standard linear regression (Tierney et al. 2010, Pearson et al. 2011). ΔT amongst U.S. lakes ranges between -2.96 and 2.44°C (mean=-0.14, SD=1.14°C) the average of which falls between that for RAN₁₃ and the PCA-derived temperature estimates for SW soils. However the improved performance of the MLR regression is observed in the smaller degree of ΔT variability compared to other calibrations. The reconstructed MAATs generated from the MLR calibration are therefore fairly consistent with MAAT and closely distributed around $\Delta T=0^{\circ}C$ (Fig. 4.23B). However, like

the previous two U.S.-based calibrations, reconstructed **MAAT** is over-predicted for a number of Chinese lakes (n=8), which becomes more severe amongst colder temperatures (**MAAT**<11°C) whereby reconstructed **MAAT** for cold U.S. lakes is also severely over-predicted (Fig. 4.23B).

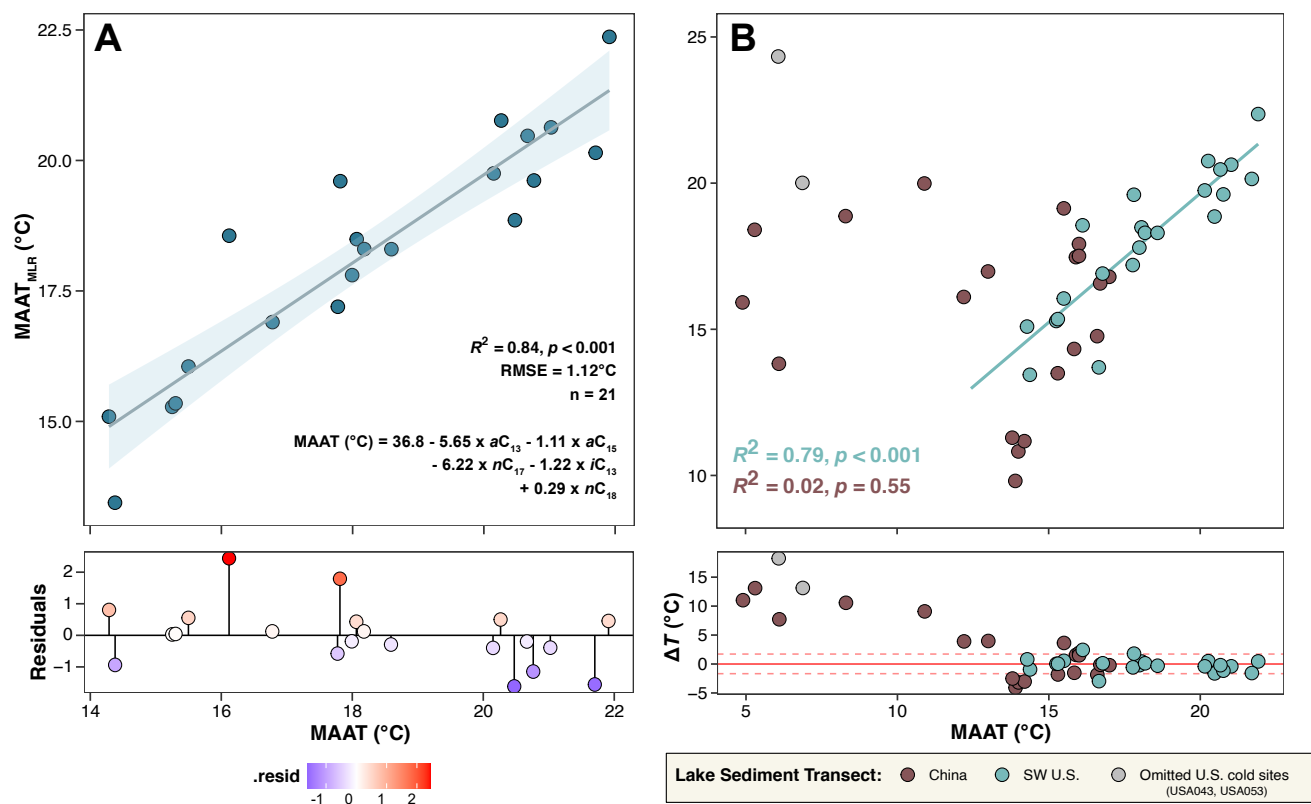


Figure 4.23: Calibration and results of MLR calibration developed from SW lake sediments. **A** MLR calibration visualised as MLR reconstructed **MAAT** for calibration samples vs **MAAT**, with residuals below. **B** MLR reconstructed **MAAT** for SW U.S. (present study) and Chinese (Yang et al. 2021) lake sediments vs **MAAT**. Below is a plot of ΔT (reconstructed **MAAT** - **MAAT**) vs **MAAT**.

Although MLR has been observed to improve RMSE and correlation coefficients of temperature calibrations in this study and others (Russell et al. 2018, Pearson et al. 2011, Tierney et al. 2010) compared to standard linear regression, this method can incorporate lipid compounds into a temperature calibration that do not have a direct response to temperature (i.e brGDGT-IIb in

a SFS calibration for East African lakes (Russell et al. 2018), and *a*-C₁₅, *i*-C₁₃, and *n*-C₁₇ in this study), so the calibration may deviate from a genuine biological process (Chowdhury & Turin 2020). However, the MLR calibration does have a greater statistical integrity compared to the $[(i\text{-}C_{12} + a\text{-}C_{13} + a\text{-}C_{17}) / n\text{-}C_{18}]$ linear regression which incorporates all compounds that significantly respond to MAAT.

Over-prediction of MAAT amongst cold lakes from 'warm' U.S.-based temperature calibrations developed from sediments with a general MAAT range of 14.3-21.9 °C is expected due to the steep slope of calibration regressions ($r^2=0.37\text{-}0.84$) and cold MAAT over-prediction is expected to become more pronounced as correlation coefficient increases amongst U.S. 3-OH FA-based temperature calibrations; the MLR-based calibration over-predicts cold MAAT to a greater degree than the RAN₁₃-based calibration, for example. Overall, the best performing 3-OH FA calibration for U.S. lakes is that based on MLR due to the higher correlation coefficient and lower RMSE compared to other calibrations considered in this discussion. This calibration is able to generate sensible MAAT estimates for a number of Chinese lake sediments where the MAAT range of the Chinese and U.S. lakes overlaps, however MAAT is severely overestimated for cold lakes.

4.4.2.2 brGDGT-based temperature calibrations

3-OH FA-based temperature proxies are not the only organic biomarker to have been investigated amongst the SW U.S. and Chinese lake sediments; those derived from brGDGTs also have a variable temperature response between the lake transects. MBT'_{5ME} is the most widely applied brGDGT-based temperature proxy in both pedogenic and lacustrine environments, having demonstrated strong positive variability in response to MAAT in many lakes across regional (Bauersachs et al. 2024, Loomis et al. 2011) and global (Martínez-Sosa et al. 2021, Raberg et al. 2021) transects. Amongst SW U.S. lakes, MBT'_{5ME} also demonstrates significant regional variability in response to MAAT ($r^2=0.37$, $p<0.01$) that was not observed amongst Chinese lakes,

in which MBT'_{6ME} - the 6-methyl rendition of MBT'_{5ME} - had a significant temperature response particularly when regressed with temperature for months above freezing T_{MAF} (Dang et al. 2018).

The regional MBT'_{5ME} calibration developed from U.S. lakes is based on a linear regression model with a weak but highly significant correlation coefficient and a RMSE of 1.90°C, which is most comparable to the RAN₁₃-based U.S. lakes temperature calibration which shares a correlation coefficient of the same calibre ($r^2=0.37$). MBT'_{5ME} produces sensible temperature estimates amongst U.S. lakes that are reasonably distributed around $\Delta T=0^\circ\text{C}$ (Fig. 4.24A). However, ΔT decreases with MAAT indicating a systematic under-prediction of MAAT >20°C (Fig. 4.24A). As observed with the U.S.-based 3-OH FA calibrations, the warm bias of the lakes from which the calibration was developed means that MAAT estimates for cold lakes are over-estimated. Regional MBT'_{5ME} calibrations are highly variable in correlation coefficient and RMSE. The MBT'_{5ME} calibration from SW U.S. lakes produces a lower coefficient compared to other regional calibrations, however the RMSE falls within their variability ($r^2=0.64$, RMSE=0.92°C (European Lakes; Bauersachs et al. (2024)), $r^2=0.92$, RMSE=2.43°C (East African Lakes; Russell et al. (2018))).

Regressions between 3-OH FA and brGDGT estimated MAAT reveal that the three U.S.-based 3-OH FA calibrations previously discussed have varying degrees of consistency with MBT'_{5ME} temperatures (Fig. 4.24B). The most consistent biomarker-based temperature estimates are those generated from RAN₁₃ and MBT'_{5ME} ($r^2=0.44$, $p=0.001$) which is expected due to the comparable correlation coefficients of the calibrations. Temperature estimates between the PCA-derived 3-OH FA-based calibration and MBT'_{5ME} have a weak consistency because the former has a lower RMSE and slightly higher correlation coefficient than the latter, whilst there is no relationship between MLR-based temperature estimates and those generated by MBT'_{5ME} because the MLR-based 3-OH FA calibration far exceeds the MBT'_{5ME} calibration in predictive potential, having a substantially higher correlation coefficient and lower RMSE (Fig. 4.24B). As a result, ΔT

of $MAAT_{RAN13} - MAAT_{MBT'5ME}$ are consistently distributed around 0°C across the temperature range of calibration (14.3-21.9 $^{\circ}\text{C}$) whereas ΔT of $MAAT_{MLR} - MAAT_{MBT'5ME}$ reveals offsets between the 3-OH FA-based and brGDGT-based calibrations amongst low and high end member temperatures of the calibration (Fig. 4.24B). The greatest offset between reconstructive potential of 3-OH FA-based calibrations relative to the MBT'_{5ME} -based calibration occurs amongst the two cold U.S. lakes (Fig. 4.24B), which are not included in the calibration dataset due to their nature as temperature outliers amongst the lakes. Of all biomarker-based calibrations considered, the MLR-based 3-OH FA calibration is the most functional at reconstructing temperature amongst SW U.S. soils, and RAN_{13} and MBT'_{5ME} are the most comparable calibrations between biomarker groups.

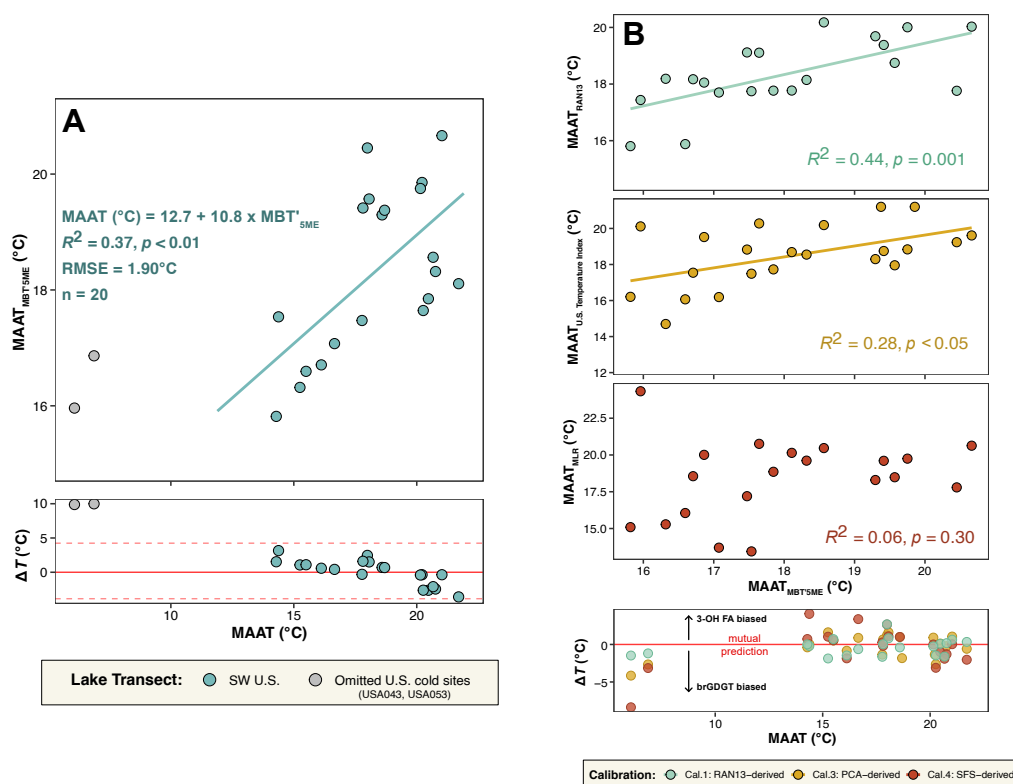


Figure 4.24: Scatterplots of **A** $\text{MBT}'_{5\text{ME}}$ MAAT reconstructions and MAAT for SW soils. Below is a plot of ΔT (reconstructed MAAT - MAAT vs MAAT. **B** comparison of estimated MAAT between the three previously considered 3-OH FA-based calibrations and the SW U.S. $\text{MBT}'_{5\text{ME}}$ calibration, amongst SW lake sediments. Below is a plot of ΔT (3-OH FA reconstructed MAAT - brGDGT reconstructed MAAT) vs MAAT

4.4.2.3 Seasonality of lipid biomarker-based temperature calibrations amongst SW U.S. lakes

A number of studies have reported warm season bias in mid to high latitude lacustrine temperature reconstructions from $\text{MBT}'_{5\text{ME}}$ calibrations (Shanahan et al. 2013, Bauersachs et al. 2024, Sun et al. 2011), which is likely related to the restricted productivity of microbial GDGT producers to warmer months where temperature is above freezing on average in these environments. MBT' and $\text{MBT}'_{6\text{ME}}$ produced correlations that differ in slope and intersect between cold ($\text{MAAT} < 5^\circ\text{C}$) and warm ($\text{MAAT} > 5^\circ\text{C}$) Chinese lakes when regressed with MAAT, and a mutual temperature

regression between the two lake subsets was only produced upon regression with the average temperature of months above freezing (T_{MAF}) (Dang et al. 2018), a temperature iteration that has been applied in recent lacustrine calibrations to improve MBT'_{5ME} temperature estimates in polar regions (Dearing Crampton-Flood et al. 2020, Martínez-Sosa et al. 2021, Raberg et al. 2021, Shanahan et al. 2013, Sun et al. 2011, Foster et al. 2016).

There is no indication that RIN₁₇ is incapable of producing reliable MAAT estimates in cold Chinese lakes, even though there is warm bias amongst GDGTs for these lakes. However only a subset of the Chinese lakes investigated for GDGTs by Dang et al. (2018) were investigated for 3-OH FAs by Wang et al. (2021), which comprises only two of the cold lakes. For the SW U.S. lakes, a reliable temperature calibration could only be developed by omitting two outlier cold sites in this transect, however 3-OH FA proxy values from the U.S. cold lakes do produce consistent temperature estimates to the Chinese cold lakes, they just do not fall within the trajectory of the warmer U.S. lakes. A fundamental issue with addressing if warm season bias influences 3-OH FA calibrations as is the case for GDGT calibrations is that considerably more warm lakes have been investigated for 3-OH FA composition compared to cold lakes (n=4) which are a minority. Therefore a comprehensive assessment of how 3-OH FA temperature proxy functionality differs between cold and warm lakes - as has occurred in detail for brGDGT-based temperature proxies (Raberg et al. 2021, Dang et al. 2018) - is not possible at this time.

Although there is no comprehensive evidence for warm season bias in the current database of lakes investigated for 3-OH FAs, both RAN₁₃ and MBT'_{5ME} produced higher and more significant correlation coefficients when regressed with T_{MAM} rather than MAAT, indicating a potential systematic spring seasonal temperature bias for 3-OH FA production in SW U.S. lakes. We have noted the use of T_{MAF} in recent MBT'_{5ME} calibrations, which is calculated by averaging temperature of months above freezing (>0 °C) (Dearing Crampton-Flood et al. 2020, Martínez-Sosa et al. 2021, Raberg et al. 2021), however, of the U.S. lakes, the cold montane lakes

Crescent and Mary's are the only that consistently freeze in winter months (T_{DIF} -0.76 and -0.97, respectively).

Additionally, T_{MAM} does not describe the warmest temperatures encountered by SW U.S. lakes. Calculation of $\Delta T_{T_{MAM}-MAAT}$ reveals that the offset between T_{MAM} and $MAAT$ can substantially differ amongst lake catchments that have similar $MAAT$ particularly when $MAAT > 20^{\circ}\text{C}$. Lakes with $MAAT$ s above this threshold originate from semi-arid to humid central to eastern Texas (lakes Houston, Medina, Bastrop, Fayette, Somerville) and arid Arizona (lakes Theodore Roosevelt, Horseshoe, Saguaro, Bartlett). The Texan group of lakes has an average $\Delta T_{T_{MAM}-MAAT}$ of $-4.0 \pm 0.2^{\circ}\text{C}$, which is $-5.1 \pm 0.2^{\circ}\text{C}$ for the Arizonan lakes. Temperature seasonality therefore substantially differs amongst the U.S. lakes due to variability of climate regime amongst the SW U.S. lakes and severe $\Delta T_{T_{MAM}-MAAT}$ differences also occur around the threshold $MAAT = 15^{\circ}\text{C}$. Ultimately the U.S. lakes comprise lakes from semi-arid and humid climate regime where seasonality is more pronounced, whereas lakes from arid regions have less seasonality and are more likely to have exposure to a wet and dry season. As a result, the variability of $MAAT$ in lakes differs from that of T_{MAM} .

The pronounced biomarker proxy response to T_{MAM} in SW U.S. lakes may be related to the preferential growth conditions of bacterial 3-OH FA producers, and may be related to the unique environmental conditions these lakes are subject to. Tree ring growth across the American Southwest is strongly associated with winter-spring precipitation (Díaz et al. 2002, George 2014, Sheppard et al. 2002) which has a tendency to infiltrate soils, providing a higher degree of wetting compared to the flash floods associated with Summer monsoonal rainfall which quickly runs off dry ground (Sheppard et al. 2002). Essentially, vegetation growth across the SW U.S. is responsive to the moisture regime of this arid to semi-arid region, not the temperature regime, and it is possible that bacterial growth in SW lake sediments is also associated with this. spring represents a seasonal period in which active reservoir capacity is likely to be higher due to winter runoff,

compared to higher evaporation in summer. Furthermore, higher reservoir inflows in Winter and spring can cause microbes such as cyanobacteria and diatoms to occupy dominant niches in lacustrine phytomicrobial ecosystems (Blinn et al. 1998). Overall, a T_{MAM} seasonal temperature proxy bias is likely related to the arid climate of the SW U.S. and winter-spring precipitation, and not the warmest temperatures in Summer.

4.4.3 Different bacterial strategies for temperature response between and within lake sediments and soils

4.4.3.1 Lacustrine environments

Previous studies have indicated that different 3-OH FA temperature proxies, and different calibrations of the same proxy, can be required for application between lake sediments and soils (Wang et al. 2016, Yang et al. 2021), and between different soil transects (Véquaud et al. 2021b, Wang et al. 2021). Globally, two regional lake transects have been investigated for 3-OH FAs thus far: the U.S. lakes of the present study, and the Chinese lakes of Yang et al. (2021). RAN_{13} , $[i\text{-C}_{12} + a\text{-C}_{13} + a\text{-C}_{17}) / n\text{-C}_{18}]$, and a MLR multiple linear regression involving $a\text{-C}_{13}$, $a\text{-C}_{15}$, $i\text{-C}_{13}$, $n\text{-C}_{17}$, and $n\text{-C}_{18}$ produced significant relationships with MAAT in U.S. lakes, whereas RIN_{17} was capable of reconstructing MAAT in Chinese and two cold U.S. lakes, but underestimated MAAT in the majority of U.S. lakes. This suggests the bacterial communities producing 3-OH FAs amongst each lake transect adjust their 3-OH FA membrane composition in response to temperature using different methods.

The MLR calibration employs selected 3-OH FA compounds that are also employed in RAN_{13} and $[i\text{-C}_{12} + a\text{-C}_{13} + a\text{-C}_{17}) / n\text{-C}_{18}]$, and describes the summation of the branched 3-OH FA compounds $a\text{-C}_{13}$, $a\text{-C}_{15}$ and $i\text{-C}_{13}$ regressed against that of the *normal*-3-OH FAs $n\text{-C}_{17}$ and $n\text{-C}_{18}$. RAN_{13} was initially developed as a temperature proxy for SST in marine sediments (Yang et al. 2020), and the $a\text{-C}_{13}$ component of all SW lacustrine temperature proxies may be due to the

aquatic nature of these environments. MBT'_{5ME} produces significant variability in response to temperature in SW lakes, but MBT'_{6ME} produced significant variability in response to [MAAT](#) in Chinese lakes; an opposed proxy response was also observed amongst GDGT proxies.

The MLR-derived proxy thus describes a degree of 3-OH FA methyl branching in response to temperature, which is also the case for RIN_{17} . Both proxies have significant negative linear relationships with [MAAT](#), however preference for RIN_{17} in Chinese lakes suggests bacterial communities residing in these lakes physiologically adapt their cell membranes to growth temperature by reducing the proportion of *i*-C₁₇ relative to *n*-C₁₇ at higher temperatures. Instead, the bacterial communities residing in U.S. lakes have an apparent preference to decrease the proportion of *a*-C₁₃, *a*-C₁₅ and *i*-C₁₃ relative to *n*-C₁₇ and *n*-C₁₈ at higher growth temperatures. These diverging responses to temperature indicate that although 3-OH FA compositions are homogeneous between the two lacustrine transects, environmentally-induced variability can differ.

4.4.3.2 Pedogenic environments

In soils, RAN_{15} and RAN_{17} have been proposed as proxies for temperature reconstruction ([Wang et al. 2016](#)), however the performance of these proxies is highly variable amongst globally distributed soils ([Véquaud et al. 2021b](#)) to the extent that neither have any temperature response in some soils (i.e. French Alps ([Véquaud et al. 2021a](#))). Furthermore, a limited number of arid soils have been investigated for 3-OH FAs, and so the unique environmental conditions from which the SW U.S. soils are derived are largely underrepresented in the global soil 3-OH FA dataset.

Soils across the SW U.S. have a highly variable 3-OH FA response to temperature; in Chapter 3, we determined that RAN_{17} is the best descriptor of 3-OH FA variability in response to temperature in U.S. soils that form under acidic and humid conditions. These conditions encompassed all U.S. EC soils, and a minority of soils from Eastern Texas and two soils from the montane regions of Arizona and New Mexico. However the remaining subset of SW U.S. arid

soils appear to convey a 3-OH FA temperature response best described by RAN_S, which was initially developed as a temperature proxy for SST reconstruction from marine sediments across the marginal seas of China (Yang et al. 2020). However RAN_S is not fully representative of temperature variability in arid SW soils as this proxy employs *i*-C₁₂ which is below detection limit in most SW soils. PCA and Pearson's correlation analysis yielded $[(i\text{-}C_{13} + i\text{-}C_{14} + a\text{-}C_{15})/n\text{-}C_{18}]$ which supersedes RAN_S in correlation coefficient and significance ($r^2=0.37$, $p<0.01$), and retains the negative variability of *i*-C₁₄ in response to increasing MAAT, incorporating only 3-OH FA compounds that have a direct response to temperature in the soils. *a*-C₁₅ is additionally employed RAN₁₅ which has been calibrated as a temperature proxy amongst global soils, so the temperature response of SW soils is not completely distinct.

$[(i\text{-}C_{13} + i\text{-}C_{14} + a\text{-}C_{15})/n\text{-}C_{18}]$ describes the decrease in abundance of selected methyl-branched 3-OH FA compounds relative to *n*-C₁₈ with increasing MAAT, which RAN_S, too, conveys relative to the sum of *n*-C₁₃ and *n*-C₁₄ at the expense of *n*-C₁₈. According to these results, Gram-negative bacteria residing in acidic and humid soils potentially instigate a different 3-OH FA-based physiological response to growth temperature compared to bacterial communities residing in alkaline and arid soils.

4.4.3.3 Lacustrine vs pedogenic environments

3-OH FA distributions differ between lake sediments and soils, as does 3-OH FA temperature variability within their distributions. Temperature proxies such as RAN₁₃ and RIN₁₇ are most functional in U.S. and Chinese lakes, respectively, and the best 3-OH FA temperature calibration from U.S. lakes was identified using MLR which predominantly exploits the significant temperature variability of *a*-C₁₃ and *n*-C₁₈. 3-OH FA temperature proxy response in soils - at least from the U.S. - appears distinguished between pH and moisture; RAN₁₇ - a global temperature proxy in soils (Wang et al. 2021) is functional in acidic humid soils, whereas RAN_S and $[(i\text{-}C_{13} + i\text{-}C_{14} + a\text{-}C_{15})/n\text{-}C_{18}]$ are more responsive to temperature in Arid alkaline soils. The stronger

3-OH FA proxy responses to [MAAT](#) in the SW U.S. lakes compared to soils is possibly related to a stronger response of lacustrine Gram-negative bacterial communities to temperature, compared to soil bacterial communities which are subject to moisture stress in the arid environment of the SW U.S. MBT'_{5ME} is functional as a temperature proxy in both SW U.S. lake sediments and soils, but difference in slope and intercept of MBT'_{5ME}-[MAAT](#) regressions between the sediment types calls for lacustrine and soil-specific calibrations. [T_{MAM}](#) seasonal temperature bias is more pronounced in lake sediment 3-OH FA and brGDGT proxy responses to temperature than for soils where no consistent seasonal bias was identified. However MBT'_{5ME} is biased to [T_{DJF}](#) in soils, suggesting that the microbial producers of brGDGTs in soils may be influenced more-so by seasonal climate regime than those of 3-OH FAs. Overall, the difference in 3-OH FA proxies producing variability in response to temperature between lake sediments and soils suggests that temperature proxies developed from soils do not fully describe the 3-OH FA temperature response in lacustrine environments, likely due to the substantial difference in Gram-negative bacterial producers between the two sediment types.

Regardless of different bacterial production, all 3-OH FA proxies considered for application between and within lake sediments and soils describe a decrease in abundance of methyl-branched 3-OH FA compounds, relative to selected *normal*-3-OH FA compounds that do not comprise a methyl branch, with increasing temperature. It is well established that single-celled organisms such as bacteria undergo lipodomic adjustment of their cell membranes to promote optimum membrane fluidity in response to ambient conditions of their external environment ([Berry & Foegeding 1997](#), [Kaneda 1991](#)). Such adaptation is required to maintain osmotic gradients between the external and internal environment of the microorganism, and most biologically relevant processes involving cell membranes occur when the phospholipid bilayer is in a liquid-crystalline ($L\alpha$) phase. In turn, mechanical processes responsible for maintaining $L\alpha$ such as membrane thickness and permeability are governed by temperature ([Garcia-Manyes et al. 2005](#)). Lipidomic adjustments in response to temperature often involve incorporation of lipids with lower melting points into the

phospholipid bilayer of a cell membrane at lower growth temperatures ([Annous et al. 1999](#), [Kaneda 1991](#)). This response includes an increased abundance of methyl-branched lipids relative to *normal*-lipids at lower growth temperature because branched fatty acids have lower melting points than *normal* fatty acids, particularly branching in the *anteiso* position ([Kaneda 1991](#)). Therefore all 3-OH FA-based proxy responses observed could be representative of a legitimate biological process - homeoviscous adaptation through branched:straight-chain 3-OH FA adjustment in the lipid bilayer of Gram-negative bacteria - and convey regional and sediment-specific renditions of homeoviscous adaptation, likely due to the diverse heterogeneity of Gram-negative bacterial communities between and within lake sediments and soils.

4.4.4 Limitations of the regional stature of 3-OH FA-based temperature calibrations developed from SW U.S. lacustrine environments

We have investigated 3-OH FA variability in response to temperature in both lacustrine and pedogenic environments, and the most applicable temperature calibrations are derived from regional transects; the RIN₁₇ calibration developed from Chinese lakes under-predicts temperatures in 3-OH FA distributions from U.S. lakes, and the RAN₁₃, PCA-derived, and MLR-derived calibrations considered for U.S. lakes have a tendency to over-predict temperature in a number of Chinese lakes, particularly those influenced by colder [MAAT](#). No mutual proxy could be developed between the two lake transects despite fairly homogeneous 3-OH FA distributions between all lakes, therefore different 3-OH FA-based proxies were calibrated to contemporary [MAAT](#) between the transects.

As a result, the 3-OH FA-based temperature calibrations proposed in this study are regional, having potential to reconstruct temperature in lacustrine environments with environmental variability in which the calibrations were developed from. Geographically-restricted regional temperature proxy calibrations are useful to capture unique biomarker variability resulting from endemic microbial communities that are under stress due to environmental conditions (often

extreme [MAAT](#), pH, or [MAP](#) in the case of Arid SW soils) that diverge from global trends ([Bauersachs et al. 2024](#), [Yang et al. 2015](#), [McClymont et al. 2023](#)), and are recommended in certain marine ([Bendle et al. 2005](#)) (unique salinity in water masses), lacustrine ([Loomis et al. 2014](#)) (seasonal climate-induced mixing regimes), and pedogenic ([Peterse et al. 2012](#), [Dirghangi et al. 2013](#)) environments. The regional RAN_{15} linear regression model developed from Mt Shennongjia (China) soils ([Wang et al. 2021](#)) has been most successful at reconstructing [MAAT](#) in a 9.0 Ka stalagmite record from Heshang Cave (China) ([Wang et al. 2018](#)), in comparison to global RAN_{15} calibrations based on random forest and multiple linear regression models ([Véquaud et al. 2021b](#)). However regional calibrations often have limited applicability outside the environmental variability from which they were developed ([Zhao et al. 2023](#)). Recent efforts to reduce confounding environmental variability in global temperature proxy calibrations have involved the development of Bayesian ([Dearing Crampton-Flood et al. 2020](#), [Martínez-Sosa et al. 2021](#), [Raberg et al. 2021](#)), and machine learning ([Eley et al. 2019](#), [Wang et al. 2021](#), [Véquaud et al. 2022](#), [Martínez-Sosa et al. 2023](#))-based statistical models. The use of T_{MAF} in recent global MBT'_{5ME} calibrations is an effort to reduce confounding proxy variability from seasonal bias at the global scale ([Raberg et al. 2021](#), [Martínez-Sosa et al. 2021](#)).

3-OH FAs have not been investigated in lacustrine environments globally, and as highly heterogeneous environments even with the same region, capacity to understand how 3-OH FA distributions and temperature proxies perform in lacustrine environments globally is currently limited. Regional calibrations are currently required for 3-OH FA-based temperature reconstruction in lacustrine environments, as the lake transects that have been investigated for 3-OH FA were found to have different temperature variability despite homogeneous 3-OH FA distributions. However it remains uncertain if this is also the case for lacustrine environments globally, and is likely to remain so without further investigation of 3-OH FA distributions in lacustrine environments in a global assessment. The SW U.S. is a unique environment; bacterial communities residing in soils here are subject to moisture stress, and lakes in this region likely

have highly variable and unpredictable water chemistry regimes that have not been investigated. However despite this heterogeneity we have shown that 3-OH FA-based temperature proxies can be functional in this unique environment, and ultimately recommend a 3-OH FA-based MLR-derived calibration for temperature reconstruction in SW U.S. lakes and similar environs.

There is also little understanding of how 3-OH FA diagenesis might affect lacustrine and pedogenic 3-OH FA-based proxy calibrations, and their performance in temporal sediments. Unbound short-chain, unsaturated, and branched fatty acids and hydroxy fatty acids are more unstable than their saturated normal counterparts (Ali & Mudge 2008), and contribute to the labile organic matter fraction in sediments, and so they can undergo diagenesis. However 3-OH FAs predominantly occur bound within LPS, making them more resistant to diagenesis, and less accessible to microbial degradation (Ali & Mudge 2008). 3-OH FAs are therefore expected to be fairly robust to diagenesis, and the potential influence of diagenesis should be investigated in future studies of 3-OH FAs in temporal lacustrine sediments.

4.5 Conclusions

Using 26 co-sampled lake sediments and catchment soils from the SW U.S., we investigated how 3-OH FA distributions differ between and within lacustrine and pedogenic environments. 3-OH FAs substantially differed between the SW lake sediments and soils in both yield and distribution. Concentrations were substantially higher in lake sediments, and lacustrine 3-OH FA distributions were particularly enriched in *a*-C₁₅ and *n*-C₁₆, and *i*-C₁₇ and *n*-C₁₈ for soils. *i*-C₁₂ and *i*-C₁₈ were abundant in lake sediments, but consistently fell below detection limit in soils. These trends in 3-OH FA distribution were observed between the two sediment types globally.

There is substantial evidence for in-situ lacustrine 3-OH FA production in SW lakes; 3-OH FA distributions and proxy response differed between lake sediments and soils, *i*-C₁₂ and *i*-C₁₈ were present in lake sediments and not catchment soils, RAN₁₃ and MBT'_{5ME} convey a seasonal T_{MAM}

bias in lake sediments but not in soils, and the bacterial communities are completely distinguished between the two sediment types. Spring seasonal bias indicates that 3-OH FA production in SW lakes is not responsive to warm season temperatures, as has been the case for recent MBT'_{5ME} calibrated to T_{MAF} . Different 3-OH FA indices were responsive to temperature in U.S. (this study) and Chinese (Yang et al. 2021) lakes, $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ in U.S. lakes and RIN₁₇ in Chinese lakes potentially indicating regionally variable bacterial responses to climate in lacustrine environments. Despite seasonal bias, the performance of a number of temperature proxies calibrated to MAAT was assessed. The best temperature response was generated from a regional MLR model from U.S. lakes.

Overall, different 3-OH FA temperature proxies were functional between and within lake sediments and soils, however all proxies describe a relative decrease in branched 3-OH FA compounds relative to *normal*- compounds with increasing MAAT, which could represent a homeoviscous adaptation response by Gram-negative bacterial. All temperature calibrations reported are regional to either the SW U.S. or China, and applicability outside of these regions are unknown due to the limited global coverage of 3-OH FA lacustrine studies. Overall we have revealed that in-situ 3-OH FA production is likely occurring in lacustrine environments which require lacustrine-specific temperature calibrations, and we recommend the need to investigate lacustrine 3-OH FA variability further in a global 3-OH FA calibration, as has occurred for soils.

**CHAPTER 5: INITIAL DEVELOPMENT OF A LOW N' SLOW ACID HYDROLYSIS
METHOD FOR 3-OH FA EXTRACTION**

5.1 Introduction

Acid hydrolysis is a fundamental procedure for the extraction of 3-OH FAs, a lipid group of increasing interest in the field of palaeoclimatology ([Wang et al. 2016, 2021](#), [Véquaud et al. 2021b](#), [Yang et al. 2020, 2021](#)). Bound by ester and amide linkages within the Lipid A component of Lipopolysaccharide (LPS), 3-OH FAs are constituents of Gram-negative bacterial outer cell membranes ([Rietschel 1976](#), [Wollenweber et al. 1982](#), [Alexander & Zähringer 2002](#), [Galanos et al. 1985](#)). There is surmounting evidence that Gram-negative bacteria physiologically respond to growth conditions such as temperature and pH through the biosynthesis of different 3-OH FA structural isomers ([Hellequin et al. 2023](#)), which has been reported for other lipid groups (particularly bacterial FAs) for decades ([Kaneda 1991](#)). Such evidence arises from the development of palaeoclimate proxies based on empirical relationships between 3-OH FAs and environmental parameters (temperature and pH), which has advanced in recent years; 3-OH FA distributions have been explored in a range of sediments, both terrestrial ([Wang et al. 2016](#), [Huguet et al. 2019](#), [Véquaud et al. 2021a](#), [Wang et al. 2021](#), [Véquaud et al. 2021b](#), [Yang et al. 2021](#)) and marine ([Yang et al. 2020](#), [Dong et al. 2023](#)) in origin, at a range of scales. As a result, there is increasing interest in the detection of 3-OH FAs.

Optimum recovery of 3-OH FA TMSi-FAME derivatives, detected by GC-MS, is currently achieved by performing an acid hydrolysis pre-treatment upon aliquots of homogenised sediment as the initial step of 3-OH FA extraction. During the hydrolysis procedure, the ester and amide bonds connecting 3-OH FAs to glucosamine disaccharide moieties within Lipid A ([Keinanen et al. 2003](#)) are broken upon reaction with H₂O in the presence of HCl catalyst. The hydrolysis reaction breaks 3-OH FAs out of the Lipid A complex, producing 3-OH FA monomers that are amenable to organic solvent extraction ([Joslyn 1970](#)). As 3-OH FAs naturally occur bound in Lipid A, and palaeoclimatic interpretations based on 3-OH FAs rely on investigation of the relative distribution of isolated 3-OH FA monomers, acid hydrolysis is an essential process in 3-OH FA extraction to

release these fatty acids from their natural bound state and improve 3-OH FA recovery.

In practice, a wide range of hydrolysis conditions are employed as temperature, duration, and reagent type and concentration influence the performance of hydrolysis for different sample media (Milinsk et al. 2008). Acid hydrolysis has been optimised for 3-OH FA recovery from environmental samples (stalagmites (Wang et al. 2012) and soils (Yang et al. 2016c)), the conditions for which are 130 °C temperature, 3 hr duration, and 3M HCl reagent (Wang et al. 2012), which combines the parameters of concentrated acid catalyst, high reaction temperature, and short reaction duration. Acid hydrolysis is preferable to saponification (alkaline hydrolysis using KOH) due to its improved recovery of amide-bound 3-OH FAs which are more resistant to hydrolysis than ester-bound 3-OH FAs (Yang et al. 2016c, Kawamura & Ishiwatari 1982). However methanolysis using HCl-MeOH reagent for direct methylation of Lipid A-bound 3-OH FAs using the conditions of 100 °C for 18 hrs was also used to extract 3-OH FAs from atmospheric dust samples (Lee et al. 2004, Mielniczuk et al. 1993). The optimised hydrolysis conditions of Wang et al. (2012) have been used in all recent studies investigating the palaeoclimatic potential of 3-OH FAs (Wang et al. 2016, 2018, Huguet et al. 2019, Véquaud et al. 2021b,a, Wang et al. 2021, Yang et al. 2020, 2021, Dong et al. 2023), however 3-OH FAs have not been investigated across a broad range of organic geochemistry laboratories since proposal of the first 3-OH FA-based palaeoclimate proxies in 2016 (Wang et al. 2016). This is potentially because application of the optimised hydrolysis method for 3-OH FA extraction solicits use of heating mantle and Allihn condenser apparatus, required to heat the acid-sample mixture above 100 °C and cool and condense the vapour produced from boiling; this equipment is expensive and inefficient for batch sample processing, which potentially limits the wide application of this method, particularly as other biomarker groups such as *n*-alkanes and GDGTs do not require acid hydrolysis for extraction.

The harsh conditions of acid hydrolysis at 130 °C are capable of promoting thermal alteration of short-chain, branched, and unsaturated FAs as they have lower melting points relative to their

long chain, un-branched, and saturated counterparts ([Kaneda 1991](#), [Français et al. 2021](#)), which might also be applicable to 3-OH FAs. Thermal alteration of 3-OH FA distributions during extraction could be problematic to 3-OH FA palaeoclimate proxies which are dependent on ratios between 3-OH FA structural groups that differ in thermal properties, however the extent to which 3-OH FA proxies might be influenced by hydrolysis temperature remains to be investigated.

We therefore aimed to develop an acid hydrolysis method that does not require specialised equipment, and has an improved technical efficiency that is more suited for processing large numbers of samples often required for high resolution sample analysis and could be more applicable in a wider range of laboratories. To eliminate use of a reflux system (Allihn condensers), a lower hydrolysis temperature whereby the acid does not boil and produce extensive vapour is required. Although higher hydrolysis temperature has been demonstrated to increase yield of 3-OH FAs ([Wang et al. 2012](#), [Montone et al. 2021](#)), there are also numerous reports of low temperature hydrolysis (<100 °C) performed for longer durations (14 hrs +) ([Montone et al. 2021](#), [Ichihara & Fukubayashi 2010](#)); high temperatures and short durations are not the only conditions under which hydrolysis is viable, and the kinetics of low temperature hydrolysis is compensated with longer duration ([Youssef et al. 2005](#)). 80 % recovery of 3-OH FAs from *Chlorella vulgaris* microalgal biomass was achieved using 80 °C hydrolysis overnight ([Montone et al. 2021](#)) for example, and both high and low temperature hydrolysis methods are capable of producing similar geochemical results ([Ichihara & Fukubayashi 2010](#)). Overall, there is no universal method for hydroxy fatty acid extraction. Extraction methods depend on the lab-specific equipment available and the samples being analysed, which determine the requirements of a cost-effective method that can produce reliable results. However, many method development studies for FAME and OH-FAME extraction and quantification seek to extract lipid monomers from small concentrated samples such as tissues ([Haan et al. 1979](#)), microbial biomass ([Edlund et al. 1985](#), [Youssef et al. 2005](#), [Montone et al. 2021](#)), vegetable oils ([Kapoor et al. 1979](#), [Jham et al. 1982](#)), biofuels ([Laurens et al. 2012](#)), atmospheric dust particulate ([Mielniczuk et al. 1993](#)), and combinations of

all ([Ichihara & Fukubayashi 2010](#)) - typically small (μL scale) samples composed purely of FA monomers, and low temperature long duration acid hydrolysis has not been tested on environmental samples with complex organic matrices (such as soils and sediments). Although [Wang et al. \(2012\)](#) compared 3-OH FA yields from stalagmites extracted under 80 °C and 130 °C hydrolysis and reported an increase in total 3-OH FA yield from 0.5 to 0.9 $\mu\text{g g}^{-1}$ with increasing hydrolysis temperature, the longest hydrolysis duration considered was 4 hrs and used for both experiments at 80 °C and 130 °C; long hydrolysis durations for 80 °C were not considered, and only total 3-OH FA yield was assessed, not the relative abundance of different 3-OH FA compounds and structural groups composing the total yield which most 3-OH FA proxies are dependent on.

We therefore compare the performance of the optimised hydrolysis method for 3-OH FA extraction developed by [Wang et al. \(2012\)](#) - used in previous chapters to extract 3-OH FAs from EC soils and SW soils and lake sediments - with a low temperature (70 °C) long duration (14 hrs) acid hydrolysis method (termed 'Low n' Slow' hydrolysis) by re-extracting all SW soils and lake sediments using the latter method (26 samples each) and comparing total 3-OH FA recovery and intrinsic standardised fractional abundances between methods and sediment types. The hydrolysis conditions were selected after considering an array of literature focusing on the development of fatty acid extraction methods ([Ichihara & Fukubayashi 2010](#), [Haan et al. 1979](#), [Edlund et al. 1985](#), [Youssef et al. 2005](#), [Montone et al. 2021](#), [Kapoor et al. 1979](#), [Jham et al. 1982](#)).

5.2 Experimental

This chapter compares two 3-OH FA datasets from our SW U.S. samples that are extracted by two different acid hydrolysis methods: SW 3-OH FA dataset 1 comprises 3-OH FAs from SW soils and lake sediments extracted by the optimised hydrolysis method of [Wang et al. \(2012\)](#), and SW 3-OH FA dataset 2 comprises 3-OH FAs from SW soils and lake sediments extracted by Low n' Slow

acid hydrolysis. 3-OH FA results from dataset 1 were reported and discussed in Chapter 4, whereas 3-OH FA results from Low n' Slow hydrolysis are reported and discussed exclusively in the present chapter.

5.2.1 SW samples dataset 1: original optimised acid hydrolysis method for 3-OH FA extraction

Refer to Chapter 2 for a description of the optimised hydrolysis method of [Wang et al. \(2012\)](#) employing the conditions of 3M HCl, 130 °C, and 3 hrs, that was used as a standard 3-OH FA extraction method for EC soils and SW soils and lake sediments reported in Chapters 3 and 4.

5.2.2 SW samples dataset 2: 3-OH FA extraction method employing Low n' Slow acid hydrolysis

The conditions for Low n' Slow acid hydrolysis were selected based on reported conditions for low temperature hydrolysis (aq.HCl) and methanolysis (HCl-MeOH) which generally range from 45-80 °C ([Ichihara & Fukubayashi 2010](#)). A hydrolysis temperature of 70 °C was selected from this range. A duration of 14 hrs was selected based on long overnight durations for hydrolysis reported in literature ([Ichihara & Fukubayashi 2010](#)). The 3M HCl reagent was kept constant to the hydrolysis method of [Wang et al. \(2012\)](#). The conditions for Low n' Slow hydrolysis are therefore 70 °C, 14 hrs, 3M HCl.

5 g of homogenised SW U.S. sample was inundated in 15 mL 3M HCl in centrifuge tubes and heated on a heating block at 70 °C for 14 hours overnight. HCl was added slowly to the lake sediments; where sediments with high carbonate content effervesced upon addition of HCl, glass rods were used to mix the sediment and acid. This procedure exposed more sedimentary surface area to the acid, promoting controlled effervescence and limited sample loss. Upon cooling, the supernatant was treated following the 3-OH FA extraction method outlined in Chapter 2, to yield 3-OH FA TMSi-FAME derivatives for [GC-MS](#) analysis.

5.3 Results

5.3.1 Low n' Slow 3-OH FA extraction of SW U.S. lake sediments and soils

5.3.1.1 3-OH FA distributions

Total 3-OH FA concentrations are $<0.001-2.03 \pm 0.18 \text{ mg g}^{-1} \text{ LOI}$ (mean=0.41,SD=0.48 $\text{mg g}^{-1} \text{ LOI}$) in SW lake sediments, and $0.02-0.88 \pm 0.16 \text{ mg g}^{-1} \text{ LOI}$ (mean=0.26,SD=0.22 $\text{mg g}^{-1} \text{ LOI}$) in SW soils. The lowest 3-OH FA concentrations in lakes are from lakes Medina and Junction, which also produce amongst the lowest concentrations in soils, in addition to Crescent Lake. The differences in 3-OH FA concentration between counterpart lake sediments and soils ($\Delta_{\text{Conc.}} = \text{lake sediment 3-OH FA conc.} - \text{soil 3-OH FA conc.}$) are highly variable; $\Delta_{\text{Conc.}}$ ranges between -0.71 and 1.96 ± 0.34 , however a mean of 0.15 (SD=0.50) indicates 3-OH FA concentrations have a tendency to be higher in lake sediments relative to soils. Overall, 3-OH FA concentrations between Low n' Slow soils and lake sediments are homogeneous ($w=343, p=0.56$).

SW U.S. 3-OH FA distributions extracted by Low n' Slow acid hydrolysis (visualised in Appendix C1) follow a similar general compositional structure observed in both lake sediments and soils distributed globally (Wang et al. 2016, Yang et al. 2021, 2020, Wang et al. 2021); *normal*-3-OH FAs with even C numbers are the most abundant structural group in both lake sediments (mean=54.2 %, SD=5.09 % total 3-OH FAs) and soils (mean=59.0 %,SD=6.79 % total 3-OH FAs), of which *n*-C₁₄ is dominant in soils (mean=20.3 %, SD=4.71 % total 3-OH FAs), and *n*-C₁₄ (mean=22.2 %, SD=8.04 % total 3-OH FAs) and *n*-C₁₆ (mean=20.2 %, SD=7.10 % total 3-OH FAs) are almost mutually dominant in lake sediments (Appendix C1). Mono-unsaturated 3-OH FAs are detected in both lake sediments (mean=0.87 %, SD=1.01 % total 3-OH FAs) and soils (mean=0.70 %, SD=0.46 % total 3-OH FAs) - *u*-C₁₈, *u*-C₁₁, and *u*-C₁₂ are the most abundant *unsaturated* 3-OH FA compounds in lake sediments, whilst *u*-C₁₂ and *u*-C₁₆ are the dominant *unsaturated* compounds in soils.

NMDS analysis of a Hellinger-transformed Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances reveals that 3-OH FA distributions form distinguished clusters between SW lake sediments and soils when extracted by Low n' Slow hydrolysis which is consistent with originally extracted SW U.S. lake sediments and soils; lake sediments negatively load on NMDS1, and the majority of soils positively load on NMDS1 (Fig. 5.1). ANOSIM analysis confirms this observation as the significant ANOSIM R statistic of 0.71 ($p < 0.001$) indicates a high degree of dissimilarity between 3-OH FA distributions when grouped by sediment type. ANOSIM analysis also reveals a higher degree of 3-OH FA dissimilarity is described by sediment type grouping under Low n' Slow hydrolysis (ANOSIM $R = 0.72$) compared to standard hydrolysis (ANOSIM $R = 0.68$), however both methods produce significantly ($p < 0.001$) and highly (ANOSIM $R > 0.5$) distinguished 3-OH FA compositions between the sediment types (Fig. 5.1). The 95 % confidence ellipses for each sediment type indicate that 3-OH FA distributions are much more variable amongst soils than for lake sediments (larger ellipsis), which is supported by the positive ordination of more 3-OH FA compounds with soils than with lake sediments (Fig. 5.1). Instead, a smaller number of methyl-branched 3-OH FA compounds such as α -C₁₅, i -C₁₈, α -C₁₇, and i -C₁₄ negatively load onto NMDS1 and are therefore positively associated with lake sediments, and not soils (Fig. 5.1). Parametric and non-parametric t-tests confirm the enrichment of these compounds in lake sediments relative to soils ($p < 0.05$) (Appendix C2).

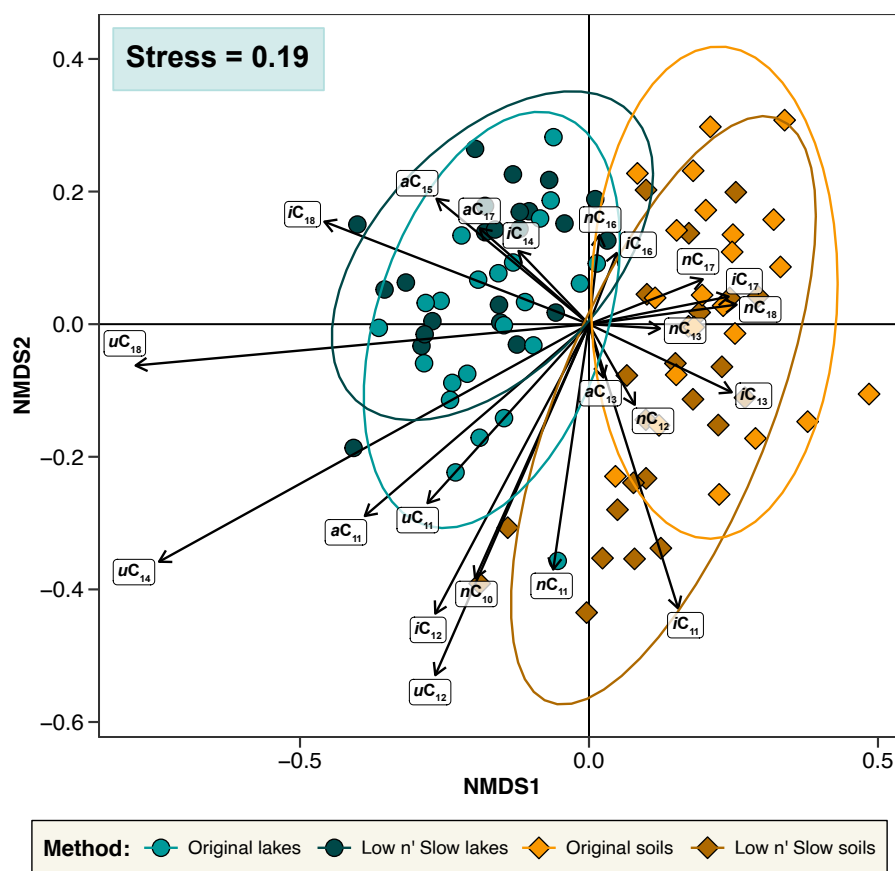


Figure 5.1: NMDS plot of a Bray-Curtis dissimilarity matrix of SW U.S. lake sediment and soil 3-OH FA fractional abundance for samples extracted by Low n' Slow hydrolysis and standard acid hydrolysis. 3-OH FA compounds that have significant variability amongst the samples ($p < 0.05$) are visualised following permutation analysis.

The higher ANOSIM R for lake sediments indicates that 3-OH FA distributions differ more between hydrolysis method than for soils, and indeed, a number of differences in 3-OH FA composition are observed amongst lake sediments and soils extracted by different hydrolysis methods (Appendix C4). $a\text{-C}_{15}$ ($t=4.13(44), p < 0.001$), $a\text{-C}_{17}$ ($t=3.80(44), p < 0.001$), and $i\text{-C}_{16}$ ($t=2.13(44), p < 0.05$) are enriched at the expense of $n\text{-C}_{10}$, $n\text{-C}_{12}$, and $n\text{-C}_{16}$ in lake sediments extracted by Low n' Slow hydrolysis relative to those that were not (Appendix C4). Instead, $n\text{-C}_{13}$ ($w=373, p < 0.05$), $n\text{-C}_{15}$ ($w=334, p < 0.01$), and $i\text{-C}_{13}$ ($t=2.13(44), p < 0.05$) are enriched at the expense

of $n\text{-C}_{18}$ in soils extracted by Low n' Slow hydrolysis compared to those that were not (Appendix C4). As a result, *anteiso*-3-OH FAs are enriched in lake sediments and shorter chain 3-OH FAs in soils extracted by Low n' Slow hydrolysis. *anteiso*-3-OH FA enrichment in SW U.S. lakes is best visualised as a ternary plot. The ternary plot visualisation employing $\sum(\textit{anteiso}\text{-}3\text{-OH FAs})$ along the x-axis, $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{10}\text{-}\text{C}_{14})$ along the y-axis, and $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{15}\text{-}\text{C}_{18})$ along the z-axis was first used by Yang et al. (2020), and can summarise 3-OH FA compositional differences between the SW U.S. soils and lake sediments and between extraction methods. Lake sediments extracted by Low n' Slow hydrolysis are generally more enriched in *anteiso*-3-OH FAs (mean=17.2 % total 3-OH FAs) compared to lake sediments that were extracted by the optimised acid hydrolysis method of Wang et al. (2012) (mean=12.4 % total 3-OH FAs). Instead, $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{15}\text{-}\text{C}_{18})$ and $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{10}\text{-}\text{C}_{14})$ promote a high degree of variability amongst both soil datasets (Fig. 5.2).

In Chapter 4, RAN values were found to vary based on 3-OH FA distributions between sediment types, and RAN confidence thresholds for each sediment type were 0.05-0.12 for soils, 0.15-0.23 for lake sediments, and 0.25-0.41 for marine sediments. The average RAN of SW soils extracted by Low n' Slow hydrolysis is 0.08 which falls within the RAN threshold of soils extracted by original hydrolysis. However, the average RAN of Low n' Slow lake sediments is 0.31, which exceeds the lacustrine RAN confidence range extracted by the original hydrolysis method, and is more comparable to the *anteiso*- composition of marine sediments. Despite *anteiso*-3-OH FA enrichment, 3-OH FA compositions of Low n' Slow lake sediments differ from that of marine sediments due to higher $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{15}\text{-}\text{C}_{18})$ in the former, and generally higher $\sum(\textit{iso}\text{- and } \textit{normal}\text{-}\text{C}_{10}\text{-}\text{C}_{14})$ in the latter. Overall, Low n' Slow hydrolysis has influenced the 3-OH FA composition of lake sediments to a greater degree than soils. Overall, the differences in intrinsic 3-OH FA distribution between extraction methods for both lake sediments may be caused by adoption of a lower hydrolysis temperature.

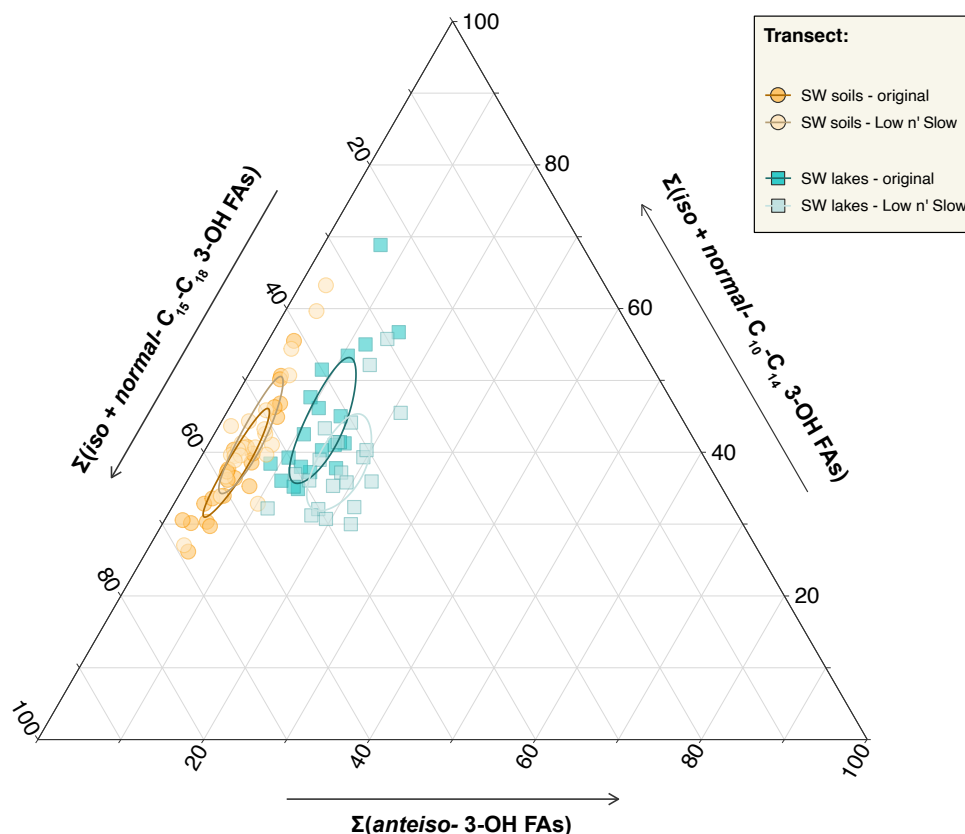


Figure 5.2: Ternary plot of SW U.S. lake sediments and soils 3-OH FA distributions extracted by two different acid hydrolysis methods.

5.3.2 3-OH FA concentrations

To assess the performance of Low n' Slow acid hydrolysis, 3-OH FA recovery was firstly compared within and between extraction methods for SW lake sediments and soils. When extracted by the standard hydrolysis method, a higher 3-OH FA recovery was obtained from lake sediments relative to soils, however when extracted by Low n' Slow hydrolysis, 3-OH FA concentrations were homogeneous between lake sediments and soils (Appendix C3). Between lake sediments extracted by the two methods, 3-OH FA recovery was also homogeneous in mean and variance (Appendix C3), and $\Delta_{\text{Conc.}}$ between identical lake sites were distributed around 0 (mean $\Delta_{\text{Conc.}} = -0.25 \pm 1.07$, $\text{SD} = 0.87 \text{ mg g}^{-1} \text{ LOI}$) despite a bias towards the standard hydrolysis method

due to substantially higher 3-OH FA concentrations for lakes Colorado City (USA058) and Champion Creek (USA059) compared to Low n' Slow hydrolysis (Fig. 5.3). Overall, this indicates that Low n' Slow hydrolysis is just as efficient at extracting 3-OH FAs from lake sediments as the standard acid hydrolysis method, and hydrolysis by the latter allows more opportunity for variability. Amongst soils, 3-OH FA recovery was substantially higher when extracted by Low n' Slow hydrolysis than standard hydrolysis (Appendix C3), and $\Delta_{\text{Conc.}}$ was higher in all soils extracted by Low n' Slow hydrolysis (mean $\Delta_{\text{Conc.}} = 0.20 \pm 0.53$, $\text{SD} = 0.18 \text{ mg g}^{-1} \text{ LOI}$) (Fig. 5.3). This indicates the performance of acid hydrolysis using high temperatures can be highly variable between different sample media, and that Low n' Slow hydrolysis promotes more consistent 3-OH FA recovery between lake sediments and soils. Overall, it seems that there are 3-OH FAs present in soils that acid hydrolysis performed at 130 °C was unable to extract.

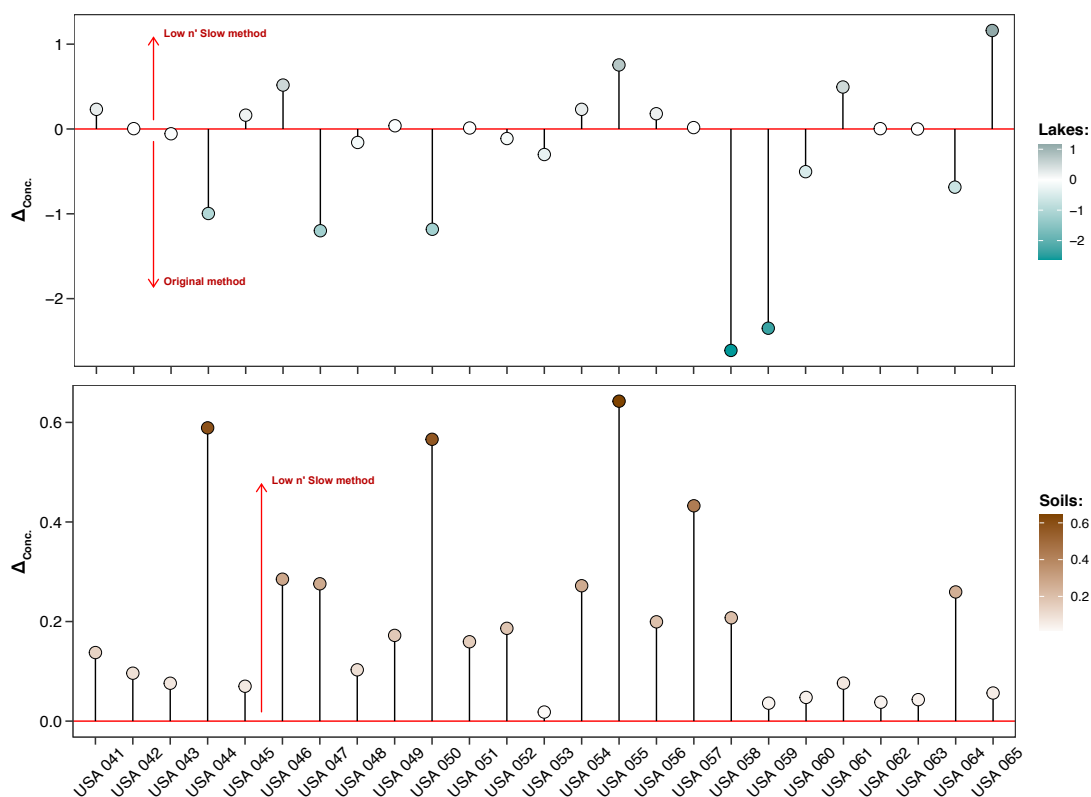


Figure 5.3: $\Delta_{\text{Conc.}}$ (L. sediment 3-OH FA conc. - soil 3-OH FA conc.) of total 3-OH FAs for SW (Upper) lake sediments and (Lower) soils between hydrolysis methods for counterpart samples.

5.3.2.1 3-OH FA-based environmental proxy distributions

The established 3-OH FA-based proxies for temperature (RAN_{15} , RAN_{17} , RAN_{13} , RAN_S , RIN_{17} , RIN) and pH (BR , RIAN) that were calculated previously for the 3-OH FA distributions of SW lake sediments and soils extracted by the standard 3-OH FA hydrolysis method of (Wang et al. 2016) are calculated for those extracted by Low n' Slow hydrolysis and visualised in Figure 5.4. Proxy values from Low n' Slow hydrolysis are reported in Appendix E3. Sediments from lakes Junction and Medina consistently fail to produce proxy values due to their unique 3-OH FA compositions.

Almost all proxy distributions significantly differ in mean, however despite this proxy variance is fairly homogeneous between the sediment types (Fig. 5.4). RAN_{13} and RIN_{17} are the only

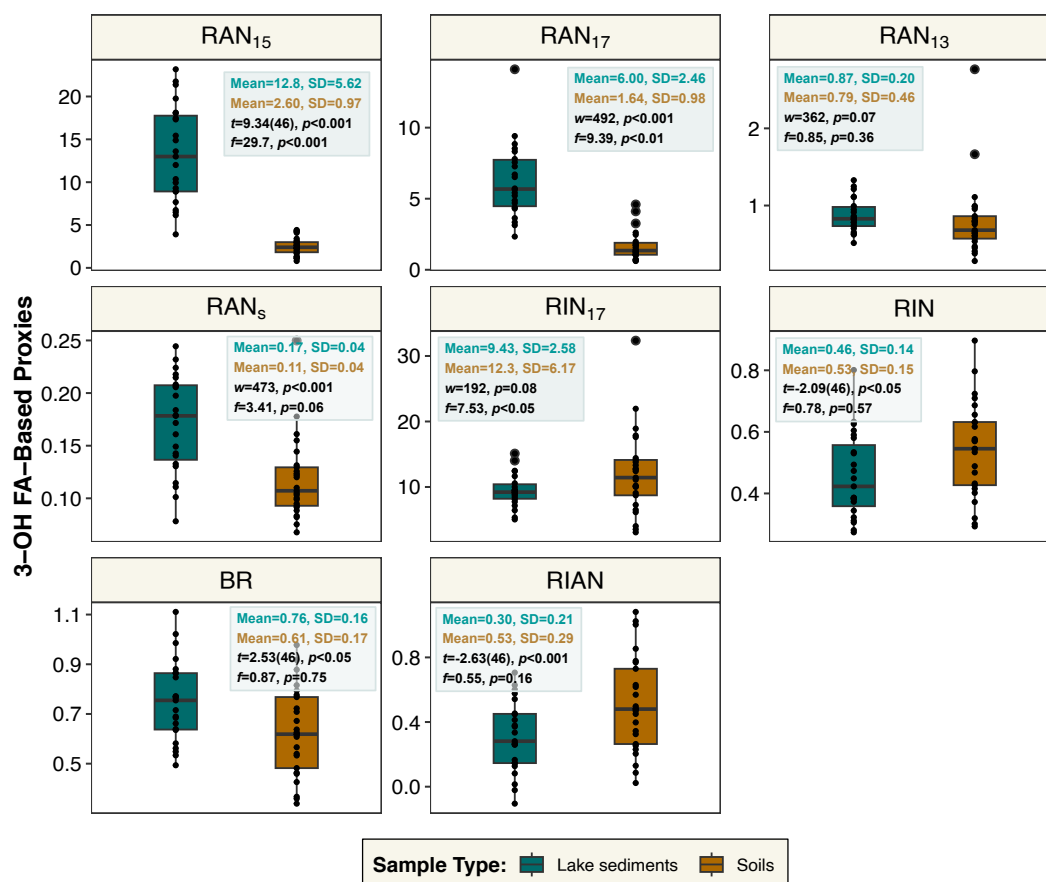


Figure 5.4: Boxplots visualising 3-OH FA-based proxy distributions in SW U.S. lake sediments and soils extracted by Low n' Slow acid hydrolysis.

proxies to have homogeneity in mean between lake sediments and soils, however RIN₁₇ values are more variable in soils than in lake sediments. RAN₁₅ and RAN₁₇ distributions are the only that are significantly distinguished between lake sediments and soils in both mean and variance (Fig. 5.4); distributions for both are higher on average and more variable in lake sediments, which is consistent with *anteiso*-3-OH FA enrichment in lake sediments relative to soils. The more cumulative proxies BR, RIAN, and RIN which employ a wide range of 3-OH FA compounds in both numerator and denominator do not significantly differ in variability between the two sediment

types, however significantly higher average values in for RIN in soils is consistent with *iso*-3-OH FA enrichment in soils relative to lake sediments. Higher average BR values in lake sediments relative to soils indicates that *anteiso*- enrichment in lake sediments supersedes *iso*- enrichment in soils. RAN₁₃ is the only proxy to produce homogeneous distributions between lake sediments and soils that do not statistically differ in either mean or variance (Fig. 5.4). Overall, RAN proxies have a tendency to produce higher values in lake sediments relative to soils, which is consistent with the 3-OH FA distributions reported, and was also observed between SW lake sediments and soils extracted using standard acid hydrolysis. Cumulative proxies that employ more 3-OH FA compounds (RIN, BR, and RIAN) have a tendency to produce higher values in soils.

During our investigation of temperature-induced 3-OH FA variability in the SW samples, RAN₁₇ was found to have a significant response to MAAT in soils that are generally acidic and humid in nature, whilst RAN_S was the best descriptor of 3-OH FA variability in response to MAAT in generally alkaline and arid soils. In lake sediments, RAN₁₃ conveyed the most prominent temperature response of the established 3-OH FA proxies from previous 3-OH FA studies, however the performance of RAN₁₃ as a temperature proxy was superseded by calibrations directly developed from SW lake sediment 3-OH FA distributions themselves, and comprised $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ and a multiple linear regression based on $a-C_{13}/n-C_{18}$.

The aforementioned proxies were functional in SW samples extracted by standard acid hydrolysis at 130 °C, but do not all convey the same temperature response in counterpart samples extracted by Low n' Slow hydrolysis. We first examine the established 3-OH FA-based temperature (and pH) proxies from previous studies which comprise RAN₁₅, RAN₁₇, BR, and RIAN (Wang et al. 2016), RAN₁₃ and RIN (Yang et al. 2020), RAN_S (Dong et al. 2023), and RIN₁₇ (Yang et al. 2021). In lake sediments extracted by Low n' Slow hydrolysis, every established 3-OH FA-based environmental proxy produced significant variability in response to MAAT (Appendix C5), and the most prominent variability was not described by RAN₁₃ as for the

original batch of sediments, but by BR which describes the ratio of summed *iso*- and *anteiso*-3-OH FAs relative to *normal*-3-OH FAs ($r^2=0.49, p<0.001$) (Appendix C5) and is applied as a pH proxy in soils (Wang et al. 2016, 2021). BR is more responsive to MAAT in Low n' Slow sediments as the cumulative proportion of *normal*-3-OH FAs have a higher, moderate, and more significant correlation coefficient ($r^2=0.49, p<0.001$) compared to the weak coefficient of original sediments ($r^2=0.20, p<0.05$) (Appendix C5). *anteiso*-3-OH FAs also have a stronger temperature response in Low n' Slow sediments. The BR-MAAT correlation also differ predominantly in intercept but also slope, the former because *anteiso*-3-OH FAs are enriched in Low n' Slow sediments and the cumulative proportion of *normal*-3-OH FAs is consistently lower relative to original sediments, producing higher BR values in Low n' Slow sediments. The latter (steeper slope) is associated with the enhanced response of *anteiso*-3-OH FAs to MAAT compared to original sediments. BR is slightly more responsive to T_{MAM} ($r^2=0.50, p<0.001$) than to MAAT, however a T_{MAM} response rather negligible amongst most established 3-OH FA proxies and highly variable as RAN_{13} and RAN_{17} are more responsive to T_{MAM} but the cumulative proportion of *anteiso*-3-OH FAs is not (Appendix C6).

$[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ which was deconvolved as a temperature response through PCA, RDA, and pairwise regression analysis using the 3-OH FA distributions of the original SW sediments has a comparable temperature response in Low n' Slow sediments ($r^2=0.48, p<0.001$), despite producing uniquely high values in sediments from lakes Meredith, Elephant Butte, Theodore Roosevelt, and Houston (Appendix C5). The $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ -MAAT regression has a comparable slope for sediments between methods, however the intercept is higher for Low n' Slow sediments because proxy values are consistently higher. This proxy is also more responsive to T_{MAM} in both lake sediment groups. However the MLR calibration consistently underestimates temperature in Low n' Slow sediments which becomes more severe with increasing MAAT. Therefore, 3-OH FA response to temperature does differ for most proxies between lake sediments extracted by the different hydrolysis methods, however the variability of $[(i-C_{12} + a-C_{13}$

+ $a-C_{17}$) / $n-C_{18}$] in response to temperature was consistent between methods.

For Low n' Slow soils, no established 3-OH FA-based temperature proxy produced any significant variability in response to temperature (Appendix C7). RAN_S is highly responsive in the original sediments, having a significant negative linear correlation with $MAAT$, however no such response is observed under Low n' Slow hydrolysis ($r^2 < 0.01, p < 0.001$). $[(i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18}]$ was identified as a temperature proxy in the original SW soils using PCA, RDA and pairwise correlation analysis, however this index also does not convey any significant temperature response in Low n' Slow soils.

Overall, employing a lower hydrolysis temperature during 3-OH FA extraction does influence 3-OH FA variability in response to temperature. BR was most responsive to $MAAT$ in Low n' Slow lake sediments, superseding the temperature response of RAN_{13} in the original lakes. However $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ had a consistent temperature response in lake sediments regardless of hydrolysis method. BR had a negligible difference in variability between $MAAT$ and T_{MAM} , however $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ was substantially more responsive to T_{MAM} which was also observed amongst proxies from original sediments. No proxies were responsive to temperature in Low n' Slow soils despite the significant response of RAN_S in the original soils. Therefore, employing a lower hydrolysis temperature for 3-OH FA extraction is capable of influencing 3-OH FA variability and proxy functionality in response to $MAAT$, which is related to the intrinsic differences in 3-OH FA distribution between the same sediment groups extracted by different methods.

5.4 Discussion

5.4.1 Effect of lower hydrolysis temperature on 3-OH FA recovery and distributions

A number of 3-OH FA compounds significantly differed in relative abundance between hydrolysis methods for both lake sediments and soils, suggesting intrinsic 3-OH FA compositions

(standardised fractional abundances) can be influenced by hydrolysis temperature during extraction. Enrichment of branched (*iso*- and *anteiso*-3-OH FAs was a notable feature of SW lake sediment 3-OH FA distributions extracted by Low n' Slow hydrolysis. Whereas *iso*- and *unsaturated*-3-OH FAs were enriched in SW soils extracted by Low n' Slow hydrolysis; the *unsaturated*-3-OH FA enrichment is of particular interest because none were detected in any SW soil sample extracted by 130 °C hydrolysis.

Acid hydrolysis at low temperature (70 °C, but <100 °C generally) is capable of promoting an increase in *unsaturated*-3-OH FA compounds, and enrichment of branched relative to *normal*-3-OH FA compounds particularly through increased relative abundance of *anteiso*-3-OH FAs in the case of SW U.S. lake sediments, and the selective promotion of these structural groups may be associated with an underlying chemical mechanism. It is well established that *unsaturated*- and branched FAs and 3-OH FAs have lower melting points relative to their *normal*- (straight-chain) structural counterparts (Yao & Hammond 2006) - unsaturated FAs particularly are less stable than saturated FAs (Jacqueline 2013) - and as a result, the more stable straight-chain saturated FAs and 3-OH FAs can be more resistant to pre-analytical treatments such as acid hydrolysis which can promote thermal and oxidative lipid degradation (Ulmer et al. 2021). *anteiso*- fatty acids in particular are known to have a large difference in melting point and phase transition temperature relative to normal 3-OH FAs, for example, melting point and phase transition temperature of *a*-C₁₅ fatty acid is 23.0 and -16.5 °C, respectively, those of *i*-C₁₅ are 51.7 and 6.5 °C, and those of *n*-C₁₅ are 52.5 and 34.2 °C (Kaneda 1991). There branched 3-OH FAs are more susceptible to thermal degradation than straight-chain fatty acids, and *anteiso*-fatty acids more so than *iso*-fatty acids. Under lower hydrolysis temperatures, *anteiso*-3-OH FAs are less susceptible to thermal degradation, which may account for their higher fractional abundance under Low n' Slow acid hydrolysis, particularly in lake sediments. Studies optimising hydrolysis parameters often report recovery of the lipid group of interest alone as a measure of method quality (Wang et al. 2012), and incentive to optimise extraction commonly involves analysis of composite FAME profiles in

foodstuffs ([Jham et al. 1982](#)) or tissue samples from biopsies ([Haan et al. 1979](#)), not the calculation of environmental proxies based on highly specific compounds of a particular lipid group.

Under cold growth temperatures and as a cold shock mechanism to rapid temperature decline, bacteria incorporate lipids with lower melting points into their cell membranes. This lipodomic-based physiological response to ambient growth temperature forms the foundation of lipid-based palaeoclimate proxies which are designed to describe the relative increase of lower melting point lipids (i.e. methyl-branched 3-OH FAs) compared to more thermally stable lipids at lower temperature. In short, there is a thermal relationship between branched and unsaturated 3-OH FAs and their normal counterparts. Therefore, an increase in relative abundance of branched and unsaturated 3-OH FAs at lower hydrolysis temperature is not surprising, as hydrolysis can become more selective at higher temperatures, and lower hydrolysis temperatures would be expected to promote a more holistic assortment of low and high melting point fatty acids. Temperature sensitivity of amino acid racemization has been observed during acid hydrolysis ([Weiss et al. 1998](#)), and the relative proportion of C₁₆ and C₁₈ dicarboxylic acids has also been reported to differ as a result of extraction efficiency between hydrolysis methods and sample media ([Angst et al. 2017](#)). Additionally, above a temperature of 100 °C, the yield of *n*-alkanes and sterols decreased during optimisation of ASE-mediated lipid extraction from soils ([Quenea et al. 2012](#)), further highlighting the susceptibility of lipids to thermal alteration at high temperature. Overall, differences in 3-OH FA fractional abundance due to different hydrolysis temperature can be expected.

Furthermore, 3-OH FAs are bound within Lipid A by ester and amide bonds, however some studies have reported that amide-bound 3-OH FAs predominate over ester-bound 3-OH FAs ([Kawamura & Ishiwatari 1982](#), [Yang et al. 2016c](#)). As amide bonds are stronger than ester bonds, harsher temperature and reagent concentrations have been applied to break amide relative to ester bonds ([Kawamura & Ishiwatari 1982](#)), and amide-bound 3-OH FAs are likely to be more resistant

to hydrolysis (Yang et al. 2016c). Kawamura & Ishiwatari (1982) report saponification temperatures >150 °C are required for the hydrolysis of amide bonds, however acid hydrolysis at 130 °C promoted optimum recovery of suspected amide-bound 3-OH FAs relative to ester-bound 2-OH FAs that are favoured by the milder saponification conditions (Yang et al. 2016c). Acid hydrolysis at mild temperatures (65 °C) with the presence of HCl catalyst was also capable of breaking amide bonds in pure wool and egg albumin samples (proteins), although HCl was not the most favourable catalysing reagent (Steinhardt & Fugitt 1942). Therefore, we cannot discount the possibility that lower hydrolysis temperatures preferentially extract ester-bound 3-OH FAs which is responsible for differing fractional 3-OH FA distributions between hydrolysis methods, however amide bond hydrolysis is possible at mild temperatures, and optimised recovery at low temperatures potentially depends on catalyst used.

However, the effects of lower hydrolysis temperature on 3-OH FA yield and fractional composition were not consistent between lake sediments and soils as sample media. Low n' Slow acid hydrolysis produced lower 3-OH FA yields (not significantly) in SW lake sediments, but substantially higher yields in SW soils comprising *unsaturated*-3-OH FAs relative to those extracted by 130 °C hydrolysis. Additionally, *iso*-3-OH FAs had a higher fractional abundance in Low n' Slow soils, whereas all branched 3-OH FAs had a higher fractional abundance in lake sediments. This suggests that acid hydrolysis is capable of having a different performance between sample media.

Lake sediments and soils do differ in composition; lake sediments comprise an assortment of autochthonous organic material from eutrophic biological processes (e.g. biological faeces and decay debris), erosional debris (lithic and organic) from the surrounding watershed, and suspended material from fluvial inlets all of which is more complex for reservoirs (Bengtsson et al. 2012). Therefore, the composition of organic carbon, and particulates (sand, clay, and silt) can differ between lake sediments and soils, however without soil compositional data this is difficult to

investigate for the U.S. samples. % LOI is comparable between SW lake sediments and soils. However, C_{Org} can be resistant to acid hydrolysis due to protection from microaggregates in the silt and clay fractions of soil, and the performance of hydrolysis can differ between these fractions (Plante et al. 2006). Additionally, waterlogged micropores in clay can trap solvent of methanolic hydrolysis reagents, affecting the performance of hydrolysis, which is particularly problematic when pressure is applied during the hydrolysis procedure (e.g. when performed in pressure chambers such as microwave extractor or accelerated solvent extractor) (Angst et al. 2017, Jansen et al. 2006). Nevertheless, the comparable recovery of 3-OH FAs between lake sediments and soils from Low n' Slow hydrolysis is more consistent with LOI (an estimation of organic C content) which is comparable between the two sediment types (5 % in soils, 7 % in lakes).

5.4.2 Implications of differing acid hydrolysis temperatures on palaeoclimatic interpretation of 3-OH FAs

Using the 130 °C optimised acid hydrolysis method of Wang et al. (2012), 3-OH FA recovery was substantially higher in SW lake sediments relative to catchment soils, which is substantial evidence for in-situ 3-OH FA production in lacustrine environments across the SW U.S. However when extracted by Low n' Slow hydrolysis at 70 °C, 3-OH FA recovery was more consistent between the sediment types, and although $\Delta_{Conc.}$ between counterpart lake sediments and catchment soils was slightly higher on average for lake sediments ($\Delta_{Conc.}=0.5$), there was no consistent significant enrichment of 3-OH FAs in lake sediments relative to soils under Low n' Slow hydrolysis as was the case for 130 °C hydrolysis. Therefore using a different hydrolysis temperature has the potential to influence environmental interpretations of 3-OH FAs.

Variable interpretation of 3-OH FA datasets extracted by different hydrolysis temperatures is further supported by differing proxy performance between hydrolysis methods. BR and RAN_5 conveyed substantial variability in response to MAAT in Low n' Slow lake sediments compared to RAN_{13} in originally extracted sediments, for example, and RAN_5 did not have any significant

response to [MAAT](#) in Low n' Slow SW soils despite having a highly significant negative correlation with [MAAT](#) in soils extracted by 130 °C hydrolysis. Low n' Slow hydrolysis therefore promotes BR as a 3-OH FA-based temperature proxy in lake sediments, which has been reported as a pH proxy in soils ([Wang et al. 2016, 2021](#)). The RAN proxies which have been proposed as 3-OH FA-based temperature proxies in soils (RAN₁₅ and RAN₁₇; ([Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021b, Wang et al. 2021](#))) and marine sediments (RAN₁₃ and RAN_S; ([Yang et al. 2020, Dong et al. 2023](#))) describe a relative increase in *anteiso*- to *normal*-3-OH FAs at lower bacterial growth temperature, and is hypothesised to represent a physiological homeoviscous adaptation response to temperature in the outer cell membranes of Gram-negative bacteria ([Wang et al. 2016](#)).

Although BR describes an increase in the proportion of all branched 3-OH FAs (*iso*- and *anteiso*-) relative to normal 3-OH FAs at lower temperature, homeoviscous adaptation is a general descriptor for any relative increase in lower-melting point lipids at lower temperatures, which can include *iso*-3-OH FAs ([Kaneda 1991](#)). Furthermore a recent experiment culturing selected Bacteroidetes taxa at different growth temperatures revealed that taxa are capable of increasing the relative proportion of all *anteiso*-3-OH FAs in response to low temperature ([Hellequin et al. 2023](#)), not just one as each RAN index describes. The temperature-dependent components of BR are $\Sigma(\textit{anteiso-})$ and $\Sigma(\textit{normal-})$ 3-OH FAs - $\Sigma(\textit{iso-})$ 3-OH FAs does not have a significant response to [MAAT](#) outside of BR - and therefore BR could be a viable temperature proxy by incorporating the temperature response of cumulative *anteiso*-3-OH FAs observed in the culture experiment. The significant negative correlation of BR with [MAAT](#) in Low n' Slow SW lake sediments could therefore be regarded as a viable bacterial temperature response. Not all temperature proxies were affected by hydrolysis describes a 3-OH FA temperature response directly from SW lake sediments temperature; $[(i\text{-}C_{12} + a\text{-}C_{13} + a\text{-}C_{17}) / n\text{-}C_{18}]$ describes a 3-OH FA temperature response directly from SW lake sediments is the only proxy to retain a temperature response between hydrolysis methods. Additionally, RAN₁₃ remains most responsive to [T_{MAM}](#), therefore some fundamental

3-OH FA-based climate interpretations remain the same between extraction methods.

Ultimately, palaeoclimatic interpretations of 3-OH FAs are dependent on the calculation of proxies based on the fractional abundance of specific compounds. Our results show that hydrolysis temperature can affect the recovery and fractional proportion of 3-OH FAs from environmental samples, and interpretations regarding in-situ lacustrine 3-OH FA production and proxy functionality differ depending on hydrolysis method used, which raises some complex questions. Differing relative proportions of 3-OH FA compounds based on hydrolysis temperature is not surprising because 3-OH FA-based temperature proxies describe the preferential production of low-temperature 3-OH FAs at lower temperatures, and 3-OH FAs are thermally treated during acid hydrolysis. The confounding influence of different acid hydrolysis temperatures is a problem specific to 3-OH FAs as the acid hydrolysis method is not strictly required for extraction of other biomarker groups such as *n*-alkanes and brGDGTs ([Hou et al. 2008](#), [Tipple & Pagani 2013](#), [De Jonge et al. 2014a](#)).

Development of a widely applicable standard extraction method for 3-OH FAs intent for palaeoclimatic interpretation could mitigate against variable 3-OH FA-based proxy results between hydrolysis method. Consistency of extraction methods has been a problem for other research areas in which geochemical results are dependent on complex pre-treatment methods which highly vary between study, an example of which is the study of conjugated linoleic acid (CLA) isomers in animal research trails, which are highly susceptible to degradation by different hydrolysis methods using methanolic HCl ([Jenkins 2010](#)). To date, 3-OH FA investigations in light of palaeoclimate research have utilised a consistent hydrolysis temperature of 130 °C ([Wang et al. 2016](#), [Huguet et al. 2019](#), [Véquaud et al. 2021a](#), [Wang et al. 2018](#), [Véquaud et al. 2021b](#), [Wang et al. 2021](#), [Yang et al. 2020, 2021](#), [Dong et al. 2023](#)), however since the proposal of the first 3-OH FA temperature proxies in 2016 ([Wang et al. 2016](#)) 3-OH FAs have not been widely investigated across many laboratories, potentially because acid hydrolysis is a complex procure unique to 3-OH

FAs compared to common palaeoclimate biomarker groups. A revised hydrolysis procedure employing lower temperature such as our Low n' Slow method which improves on technical efficiency and use of non-specialised equipment is potentially a more viable option for wide uptake of 3-OH FA extraction amongst geochemistry laboratories. Further method development for 3-OH FAs is required in the future, and we proceed to describe some potential directions for future work to further streamline 3-OH FA extraction.

5.4.3 Streamlining 3-OH FA extraction: potential to further optimise 3-OH FA extraction from environmental media

A number of parameters affect the performance of acid hydrolysis: temperature, duration, and reagent concentration ([Jham et al. 1982](#)). Although these parameters have been optimised for 3-OH FA extraction before using stalagmite samples ([Wang et al. 2012](#)) and resulted in the optimised conditions of 130 °C, 3 hrs, 3M HCl, the impact of long hydrolysis durations such as 14 hrs on 3-OH FA yield and composition were not considered (longest duration considered was 4 hrs) or paired alongside lower temperatures (80 °C) to compare the effect of long-duration low temperature hydrolysis with short-duration high temperature hydrolysis. Substantially increasing duration of thermal treatment is usually necessary for low temperature hydrolysis to have a comparable effect to high temperature hydrolysis. Although longer duration for acid hydrolysis was considered in the present study, hydrolysis conditions were selected, but have not been optimised by testing a range of long durations and low temperatures. In practice, hydrolysis conditions are extremely diverse and are often adapted to optimise analyte yields from the specific samples in question ([Haan et al. 1979](#), [Jham et al. 1982](#)). Therefore future method development should rigorously investigate the optimal conditions of Low n' Slow hydrolysis for 3-OH FA recovery and fractional distributions.

Decreasing temperature and increasing duration to improve technical efficiency of hydrolysis pre-treatment is not the only way to streamline 3-OH FA extraction. 3-OH FA extraction even

using Low n' Slow hydrolysis involves two procedures which have the potential to be combined: hydrolysis using 3M HCl, and methylation of FA and 3-OH FA monomers to produce FAMES and OH-FAMES using HCl-MeOH reagent. Hydrolysis is breaking of ester and amide bonds using a H₂O molecule, and methylation is substitution of COOH to COCH₃. Heating sample with methanolic HCl (HCl-MeOH) instead of aq. HCl during hydrolysis combines these mechanisms. The methanolysis procedure is commonly performed in a pressure chamber such as a microwave extractor to keep the solvent in a liquid state. However, a problem associated with methanolysis is that HCl-MeOH is not stable and can itself undergo halogenation, producing chloromethane and water (Jham et al. 1982, Hardouin & Lemaitre 2020). Therefore, the concentration of HCl-MeOH can change over long storage times. Additionally the efficiency of methanolysis is affected by the presence of water which can depend on preparation method; anhydrous HCl-MeOH prepared by bubbling hydrogen chloride gas through methanol is a more favourable and efficient methanolysis reagent than HCl-MeOH prepared by addition of MeOH to aq. HCl due to the absence of water.

Overall, there is a vast scope to further streamline 3-OH FA extraction, which should be investigated in future work in addition to the palaeoclimatic interpretations of 3-OH FAs.

5.5 Conclusions

Our SW U.S. soils and lake sediments were re-extracted using a Low n' Slow acid hydrolysis method, to assess how 3-OH FA recovery and composition is influenced by hydrolysis temperature. Differences in 3-OH FA composition were observed between the hydrolysis methods and were not consistent between lake sediments and soils. 3-OH FA recovery was substantially higher in Low n' Slow soils relative to conventional acid hydrolysis at 130 °C, and was homogeneous for the lake sediments between methods. Additionally, the relative proportion of branched 3-OH FAs were enriched in lake sediments extracted by Low n' Slow hydrolysis relative to those that were not, particularly apparent for *anteiso*-3-OH FAs. Instead, *iso*-3-OH FAs were

cumulatively enriched in Low n' Slow soils.

The differences in 3-OH FA recovery and composition between hydrolysis methods did affect the palaeoclimate interpretations that are possible from 3-OH FAs. BR was the most responsive proxy to temperature in Low n' Slow lake sediments, compared to RAN₁₃ in the original sediments, and no proxy conveyed a significant response to temperature in Low n' Slow soils, despite the strong temperature response of RAN₅ in the original soils. Additionally, a strong case for in-situ 3-OH FA production could be made from the substantially higher 3-OH FA recovery in lake sediments relative to soils from the original SW U.S. dataset, however homogeneous 3-OH FA recovery between lake sediments and soils under Low n' Slow hydrolysis meant the same conclusion could not be made. Despite these differences, RAN₁₃ was still most responsive to T_{MAM} and $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ had a comparable response to temperature in SW lake sediments regardless of hydrolysis method.

Differences in 3-OH FA composition between hydrolysis methods can be expected as thermal treatment of lipid monomers can become selective at higher temperatures due to thermal alteration of low melting point lipids. The difference in performance for Low n' Slow hydrolysis between lake sediments and soils could be related to differences in particle composition between the two media. We could not rule out potential preferential hydrolysis of ester-bound relative to amide-bound 3-OH FAs at lower hydrolysis temperature. Each of these matters are ultimately complex to deconvolve.

There are complex limitations that could arise from the use of different 3-OH FA hydrolysis methods in terms of palaeoclimate interpretation. Currently all palaeoclimate-related 3-OH FA investigations have utilised the optimised 130 °C optimised hydrolysis conditions proposed by (Wang et al. 2012). However the Low n' Slow hydrolysis method has a number of benefits; 3-OH FA recovery appears more consistent between different sample media, and complex reflux equipment is not required for lower hydrolysis temperature which has the potential to make the

method more accessible to a wider range of organic geochemistry laboratories.

Ultimately, this study has revealed that Low n' Slow hydrolysis could be a viable 3-OH FA extraction method. However there is potential for future method development work; the optimal temperature, reagent concentration, and duration conditions for Low n' Slow hydrolysis should be investigated, and 3-OH FA extraction could be further streamlined by the use of methanolysis using HCl-MeOH instead of acid hydrolysis. We hope this study provides insights for further method development in the future.

**CHAPTER 6: APPLICATION OF LOW N' SLOW ACID HYDROLYSIS FOR 3-OH FA
EXTRACTION IN A LAKE SEDIMENT CORE FROM ESTHWAITE WA-
TER**

6.1 Introduction

Esthwaite Water (UK) is one of the most studied lakes in the world, having been continuously monitored since 1945 ([Maberly et al. 2017](#)). Paired with a well-documented history of eutrophication and land use management ([Dong et al. 2012](#), [Thorpe et al. 2022](#)), Esthwaite Water offers a unique opportunity to investigate temporal changes in 3-OH FA composition alongside instrumental measurements of water quality. Water quality parameters measured at Esthwaite comprise surface water temperature, pH, DO, alkalinity, ammonium (NH₄-N), nitrate (NO₃-N), soluble reactive phosphate (SRP; PO₄-P), total phosphorous (TOT P), dissolved reactive silicon (SiO₂), and phytoplankton chlorophyll ([Maberly et al. 2017](#)) (Appendix D1).

The environmental conditions of Esthwaite Water over the ca.60-year monitoring period are summarised in detail by [Dong et al. \(2012\)](#) and [Thorpe et al. \(2022\)](#). Of particular interest was an increase in Winter SRP and NO₃-N from 1975 due to installment of a waste-water treatment plant and fish farm. Winter SRP was calculated from SRP (PO₄-P) in both studies as a measure of nutrient availability to microbes associated with Spring blooms ([Talling & Heaney 1988](#), [Thackeray et al. 2008](#)). Declining concentrations of SRP and NO₃-N since the late 1990s is attributed to water quality restoration efforts, and by 2015 nutrient concentrations were comparable to pre-1970 levels ([Thorpe et al. 2022](#)). Refer to [Thorpe et al. \(2024\)](#) Supplementary Figure S2 for an overview of mean annual environmental conditions at Esthwaite Water between 1945 and 2015. Environmental records for Esthwaite Water are summarised in Appendix D1.

Three sediment cores were collected from the deepest point of Esthwaite Water in 2016 where long-term monitoring is also conducted, and subject to 16S RNA sequencing by [Thorpe et al. \(2022\)](#). Changes in bacterial community structure in Esthwaite Water are documented by [Thorpe et al. \(2022\)](#) and provide a detailed context of 3-OH FA Gram-negative bacterial producers, which alongside temporal environmental records provide a unique opportunity to investigate how lacustrine 3-OH FA distributions change on a multi-decadal temporal scale in response to

environmental and biotic conditions. Therefore in 2021, a sediment core was sampled from the deepest point of Esthwaite Water to provide sediments for joint 3-OH FA extraction and 16S RNA sequencing. The sequencing data is reported by [Thorpe et al. \(2024\)](#), and 3-OH FA data over the 105 year record is reported in this Chapter as a short data report to investigate (i) the resolution of temporal variability in 3-OH FA distribution, (ii) whether established 3-OH FA-based temperature proxies can track temperature across the temporal record at Esthwaite, (iii) how changes in 3-OH FA distribution might relate to changes in bacterial community structure at the temporal scale, as a more detailed study in the future.

6.2 Summary of 3-OH FA data

A full spectrum of *normal*-, *iso*-, and *anteiso*-3-OH FAs between C₁₀ and C₁₈ were detected in Esthwaite Water lake sediments (Fig. 6.1); no compound was undetected in any sample.

Mono-unsaturated 3-OH FAs were additionally detected, representing a minor proportion of total 3-OH FA fractional abundance (1.58 ± 0.24 % (*SD*)), and were associated with even C numbers (*u*-C₁₂, *u*-C₁₄, *u*-C₁₆, *u*-C₁₈). 3-OH FAs with even C numbers and *normal*- structures were most abundant, comprising 56.7 ± 3.77 % (*SD*) total 3-OH FAs on average, and *n*-C₁₆ was the dominant 3-OH FA compound on average (17.7 ± 1.41 % (*SD*)), although *n*-C₁₂ was uniquely co-dominant with *n*-C₁₆ in ESW28 (sedimentation in 1965) and *n*-C₁₄ was co-dominant with *n*-C₁₆ in ESW29 and ESW33 (sedimentation in 1963 and 1935, respectively).

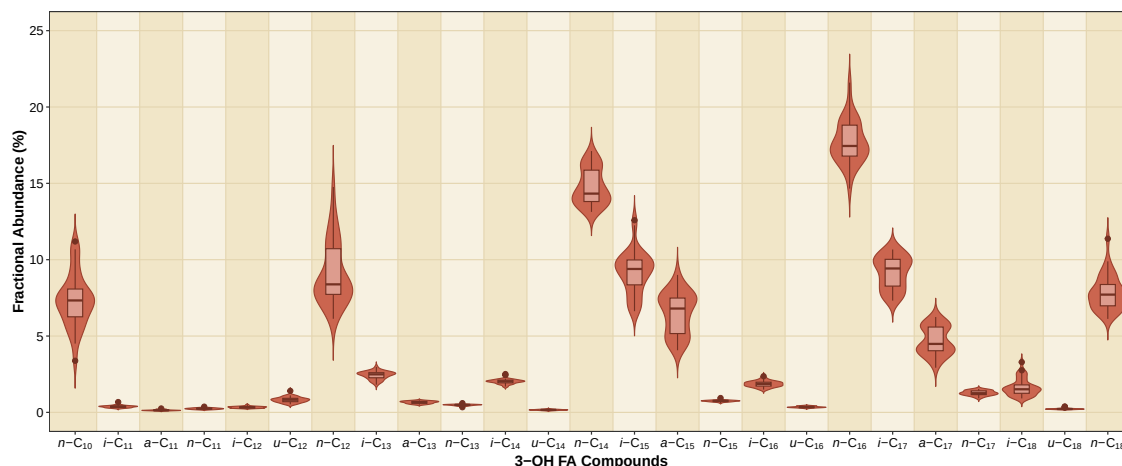


Figure 6.1: 3-OH FA distributions Esthwaite Water lake sediments reported as % fractional abundance of individual 3-OH FA compounds

The fractional abundance of most 3-OH FA compounds is highly variable throughout the 105-year Esthwaite Water sediment core record, which is indicated by the significant ($p < 0.05$) result of permutation tests performed between the NMDS matrix and fractional abundance of individual 3-OH FA compounds; all 3-OH FA compounds that have significant variability amongst the sediments are visualised in Figure 6.2. With a standard deviation in fractional abundance of 2.05 %, *n*-C₁₂ is the most variable 3-OH FA compound amongst the Esthwaite sediments. The variability of most 3-OH FAs in the sediment core suggests that 3-OH FA compositions are responding to either environmental or biological forcing or both during the 105-year period of sedimentation represented by the core.

Composite 3-OH FA distributions throughout the Esthwaite core are visualised in Figure 6.2 as an NMDS plot of a Bray-Curtis distance matrix of the distribution of 3-OH FA fractional abundance in each sediment section. 3-OH FA variability along NMDS1 is attributed to short-chain 3-OH FAs such as *n*-C₁₀, *i*-C₁₁, *a*-C₁₁, and *n*-C₁₂ (Fig. 6.2). Younger sediments deposited from 1995 to recent have a tendency to negatively load onto NMDS2, and those representing sedimentation prior to 1995 generally load positively onto NMDS2 (Fig. 6.2).

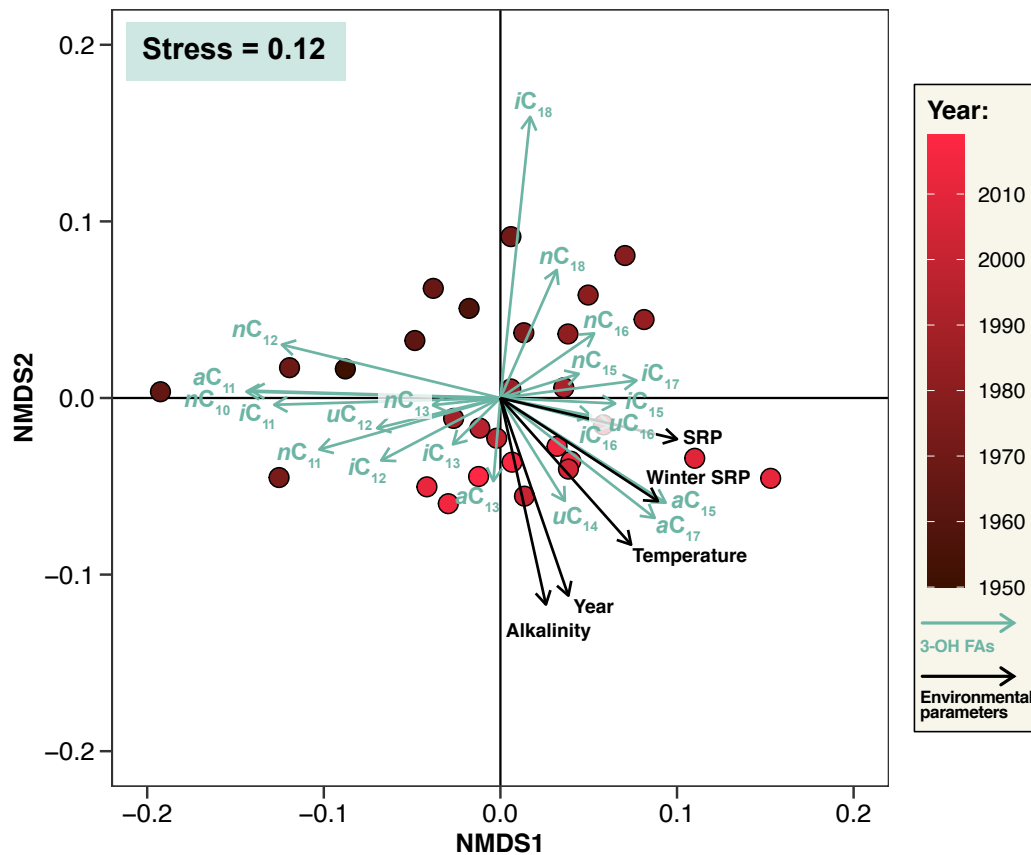


Figure 6.2: NMDS plot of a Bray-Curtis dissimilarity matrix of 3-OH FA fractional abundances of Esthwaite Water lake sediments. 3-OH FA compounds that have significant ($p < 0.05$) variability amongst the samples following permutation analysis are visualised. This is also applicable for environmental variables.

Permutation tests between environmental variables and the dissimilarity matrix reveal that 3-OH FA compositions are best structured by sample age (year of sedimentation) ($r^2 = 0.79, p = 0.001$), however alkalinity ($p < 0.01$), lake surface water temperature ($p < 0.01$), SRP ($p < 0.05$), and Winter SRP ($p < 0.05$) all significantly describe matrix dissimilarity, and therefore 3-OH FA variability between samples throughout the sediment core. The comparable ordination of all significant environmental variables, particularly between alkalinity and year of sedimentation indicates that nutrient concentration has generally increased throughout the sediment record, and highest

alkalinity, temperature, and SRP measurements are observed during more recent periods of sedimentation.

6.3 3-OH FA distributions in context of environmental variability at Esthwaite Water

The established 3-OH FA-based environmental proxies RAN_{15} , RAN_{17} , RAN_{13} , RAN_S , RIN_{17} , RIN , and BR were calculated throughout the Esthwaite Water sediment record based on 3-OH FA distributions from 35 core sections. RAN ($\sum[anteiso-3-OH\ FAs] / \sum[normal-3-OH\ FAs]$) was additionally calculated to summarise the combined variability of the individual RAN proxies. Proxy variability is as follows: RAN_{15} 5.49-12.48 (8.67 ± 1.88 (SD)), RAN_{17} 2.40-6.39 (3.79 ± 1.03 (SD)), RAN_{13} 1.12-1.67 (1.37 ± 0.16 (SD)), RAN_S 0.14-0.20 (0.17 ± 0.01 (SD)), RIN_{17} 5.44-10.9 (7.35 ± 1.36 (SD)), RIN 0.33-0.60 (0.46 ± 0.07 (SD)), RAN 0.12-0.27 (0.20 ± 0.05 (SD)), BR 0.50-0.86 (0.66 ± 0.11 (SD)). Of these proxies, RAN_{15} and RIN_{17} convey the highest degree of variability due to their high standard deviations.

3-OH FA proxy variability was investigated across the time-series of the lake record. Values were generally stable for most proxies, and RAN_{15} , RAN_{17} , and RAN_{13} steadily declined in value until roughly 1970 when all proxies underwent an increase in value. The onset of increasing proxy values around the 1970s and magnitude of increase differs between proxies. RAN_{15} values rapidly increase from 5.49 to 12.0 between 1970 and 1988 and stabilise around 11.2 until re-stabilising at 9.41 from 2009 onwards (Fig. 6.3). Rapidly changing values are not observed for RAN_{17} which fluctuate throughout the 1970s but without any particular trend. Instead, RAN_{17} values rapidly increase from 3.44 to 5.80 between 1985 and 2005 when RAN_{15} is in a period of stability. From approximately 2009, RAN_{17} values become fairly stable around 4.57. The trend of RAN_{13} values is more comparable to those of RAN_{17} , however exhibits a steady increase from 1973 up until approximately 2015 where values stabilise at pre-1945 levels. RAN_S values are more dynamic through the sequence than any other RAN proxy. RAN_S values start increasing from 1965, earlier

than other RAN proxies, and sharply decline between ca.1975 and 1982, when other RAN proxies are stable (RAN₁₇) or increasing in value. From 1982, RAN_S enters a second phase of increasing value until ca.2001, and from 2012 RAN_S values stabilise. These trends in RAN proxies are summarised by RAN where values are fairly stable until rapidly increasing from 1973 through to 2001 after which values stabilise. RAN values do not return back to pre-1973 levels. RIN₁₇ values highly fluctuate throughout the sequence, and RIN values steadily increase from as early as 1945. BR portrays the combined trend of RIN and RAN.

Overall, RAN₁₅ exhibits the most drastic changes in value throughout the Esthwaite sediment record, and a number of proxies (particularly RAN₁₅) broadly track historical events at Esthwaite Water. Proxy values are observed to increase around the early 1970s when a wastewater treatment plant was installed (1973) - RAN, RIN, and RAN₁₅ start increasing in value from their pre-1970s baseline almost exactly at this time - which was associated with nutrient enrichment at Esthwaite during the 1970s ([Thorpe et al. 2022](#), [Dong et al. 2012](#)). Similarly, most proxies stabilise in value from approximately 2009 onwards coinciding with efforts to reduce nutrient enrichment, such as fish farm closure (2009) and upgrades to wastewater treatment (2010). Of notable importance, however, is the high-resolution sub-decadal variability of the 3-OH FA proxies (particularly RAN₁₅) throughout this lacustrine sequence.

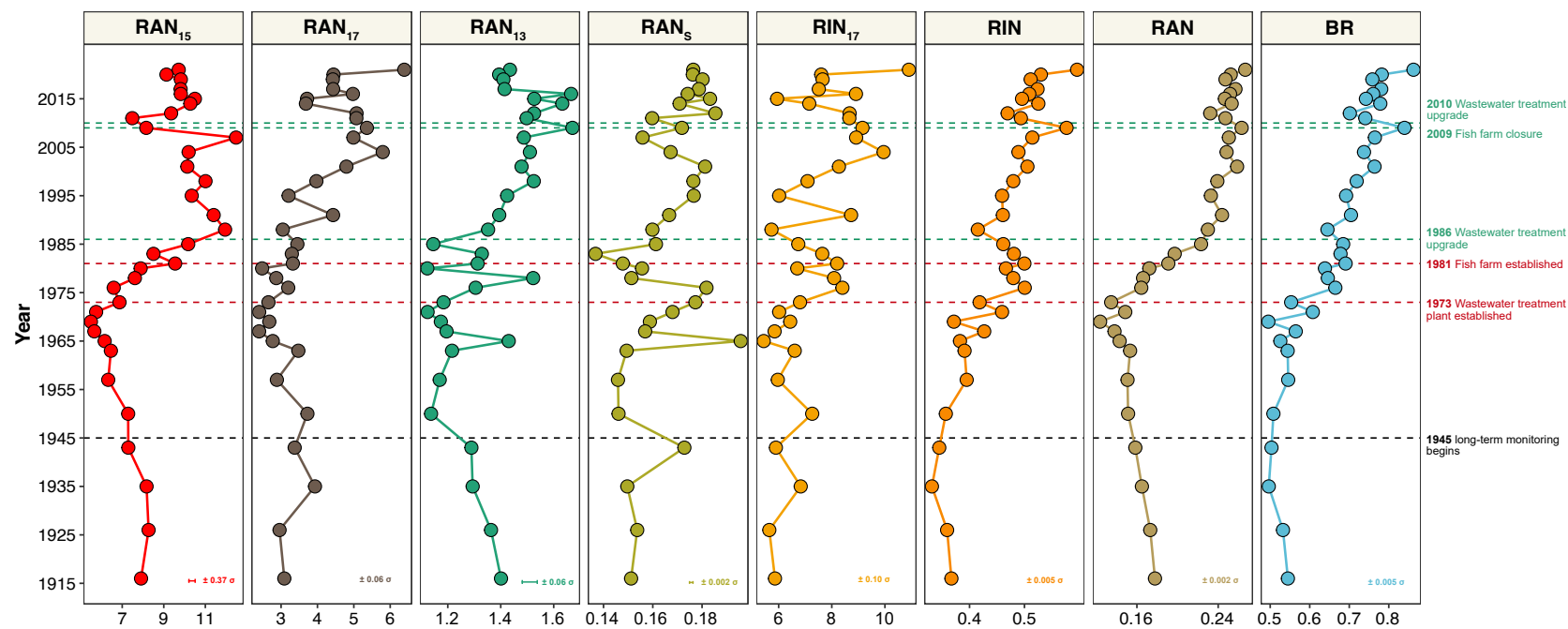


Figure 6.3: Variability of 3-OH FA proxies throughout the Esthwaite Water lacustrine sequence. Chronology of important historic events related to water quality are indicated. In **RED** are events promoting nutrient enrichment and in **GREEN** are efforts to mitigate nutrient enrichment.

The significant permutations previously observed between sample age, alkalinity, SRP, Winter SRP, and lake surface water temperature are consolidated by a Pearson's correlation heatmap in which these environmental variables have significant positive correlations with all or most 3-OH FA proxies (Fig. 6.4). Interestingly, all proxies excluding RAN_S have a positive relationship with temperature. RAN₁₅, RAN₁₇, RAN₁₃, RIN₁₇, and RIN have previously conveyed either negative or no correlations with temperature in soils (Wang et al. 2016, Huguet et al. 2019, Véquaud et al. 2021a,b, Wang et al. 2021), lake (Yang et al. 2021) and marine sediments (Yang et al. 2020, Dong et al. 2023), and this Esthwaite lacustrine record is the first occurrence of any positive proxy-temperature relationship particularly for the RAN proxies. This result cannot be explained by the principles of homeoviscous adaptation of bacterial cell membranes because low melting point *anteiso*- and *iso*- compounds do not increase relative to *normal*- compounds at lower temperatures - the opposite is occurring.

The positive proxy response to temperature potentially occurs because temperature change across the sequence is fairly small (9.97-12.6 °C, $\Delta T=2.63$ °C) compared to surface transects which are designed to describe 3-OH FA variability over wide temperature gradients (e.g. $\Delta T=12.1$ °C for Chinese lakes (Yang et al. 2021), $\Delta T=7.62$ °C for SW U.S lakes), and proxies such as RIAN have optimum functionality over wide environmental gradients (pH for RIAN) (Huguet et al. 2019). Esthwaite Water also has a long history of substantial nutrient enrichment which could supersede temperature as an environmental driver of bacterial community structure and growth (Newton & McMahon 2011, Thorpe et al. 2022); RAN₁₅ is highly responsive to Winter SRP in addition to temperature, for example (Fig. 6.4). Overall, the response of 3-OH FA-based proxies to environmental conditions at Esthwaite Water is complex, and requires further investigation to deconvolve.



Figure 6.4: Pearson's correlation heatmap between 3-OH FA proxies and environmental variables. Significance is indicated by asterisks ($p \leq 0.05 = *$, $\leq 0.01 = **$, $\leq 0.001 = ***$)

6.4 Brief summary

Initial assessment of 3-OH FA distributions thorough at 105-year sedimentary record from Esthwaite Water that represents sedimentation across a multi-decadal time-series between 1916 and 2021 reveals 3-OH FA-proxies can track environmental changes at a high-resolution sub-decadal scale. Low n' Slow acid hydrolysis was successful in extracting 3-OH FAs consistently from the lacustrine sediment core. Almost all proxies produced variability in response

to nutrient concentrations and lake surface water temperature, but all proxy-environmental correlations were positive even for temperature which contradicts the variability of 3-OH FA temperature proxies in sediments globally, and is not representative of homeoviscous adaptation of Gram-negative bacterial cell membranes at Esthwaite. A recent study has shown sedDNA has potential to track nutrient conditions of Esthwaite Water (Thorpe et al. 2024, 2022), and bacterial communities are responsive to environmental conditions at Esthwaite (Thorpe et al. 2022). By 2015, the abundance of cyanobacteria declined back to pre-1970s levels due to efforts to reduce nutrient enrichment at Esthwaite, but not all bacterial phyla recovered their pre-1970s baseline, such as Bacteroidetes (Thorpe et al. 2022). Most 3-OH FA-based proxies (RAN₁₅, RAN₁₇, RIN, RAN) also did not recover to their pre-1970s baseline, which may suggest bacterial phyla that did not recover to their pre-1970s baseline abundance are potentially the predominant bacterial producers of 3-OH FAs at Esthwaite Water. Overall, we have briefly summarised 3-OH FA distributions along a lake core from Esthwaite Water, and further investigation is required to fully interpret 3-OH FA response to environmental variability throughout the record.

CHAPTER 7: CONCLUSIONS

7.1 Main findings of this thesis

This thesis has presented novel geochemical data from soils and lake sediments subject to contrasting environmental conditions in the contiguous U.S., in addition to a data report from a 105 year sediment sequence from Esthwaite Water (UK), and has revealed some important insights into the environmental drivers of 3-OH FA variability in pedogenic and lacustrine environments.

The first aim of this thesis was to *assess the potential for temperature proxy development and calibration from 3-OH FAs in pedogenic environments at the continental scale*. In Chapter 3, 3-OH FA distributions from 62 soils across the U.S. EC and SW U.S. were investigated for variability in response to temperature and pH. pH was the primary environmental driver of 3-OH FA variability in the soils. $-\log(\text{RIN}_{\text{USA}}) [(i\text{-C}_{13} + i\text{-C}_{16} + i\text{-C}_{17}) / (n\text{-C}_{14} + n\text{-C}_{18})]$ conveyed the most robust 3-OH FA pH response, and had a higher correlation coefficient when regressed with pH compared to RIAN in globally compiled soils from the studies of Wang et al. (2021) and Véquaud et al. (2021b). RAN_{17} was not the first-order 3-OH FA temperature response in all U.S. soils; environmental investigation revealed a subset of SW soils associated predominantly with an arid climate regime and alkaline soil pH (>7.2) in which RAN_5 was the primary 3-OH FA response to MAAT (Chapter 4). The pH 7.2 threshold may be associated with Acidobacteria which have a preference for acidic and anaerobic soils. RAN_{17} and $-\log(\text{RIN}_{\text{USA}})$ can be described as homeoviscous adaptation responses to growth temperature and pH, and are likely the result of both membrane lipodome phenotypic plasticity and bacterial community structure.

Comparison of RAN_{17} -based and brGDGT-based MBT temperature estimates revealed both proxy calibrations were able to track MAAT in EC soils, but RAN_{17} estimates were more consistent with MAAT in SW soils compared to MBT which severely underestimated MAAT in these soils, indicating that environmental drivers promoting a shift in bacterial to archaeal

brGDGT production may be more confounding to the functionality of MBT than a shift in bacterial community for the functionality of RAN₁₇.

Re-classing the U.S. soils by a 7.2 pH threshold instead of their geographic classifications was successful in generating an optimised RAN₁₇ calibration from EC soils and a subset of SW soils, integrating both transects. Although the pH threshold improved the correlation coefficient between RAN₁₇ and MAAT for global soils, RMSE did not improve indicating much scatter still remains which should be investigated in the future.

The second aim of this thesis was to **investigate 3-OH FA distributions and responses to environmental variability within and between lacustrine and pedogenic environments**. In Chapter 4, 3-OH FA distributions from 26 lake sediments (co-sampled with the SW catchment soils reported in Chapter 3) were investigated for environmentally-driven 3-OH FA variability. $[(i-C_{12} + a-C_{13} + a-C_{17}) / n-C_{18}]$ best described 3-OH FA temperature variability in SW lake sediments, not RAN₁₇ as for the Chinese lakes of Yang et al. (2021). A MLR-based 3-OH FA calibration employing $a-C_{13}$, $a-C_{15}$, $i-C_{13}$, $n-C_{17}$, and $n-C_{18}$ superseded the performance of any linear calibration considered. A number of biomarker proxies including RAN₁₃ and MBT'_{5ME} were most responsive to T_{MAM} in SW lakes. We additionally unveiled evidence for in-situ lacustrine 3-OH FA production in SW lakes, which is consistent with differing 3-OH FA responses to temperature between SW lake sediments and catchment soils.

3-OH FAs appear to have regionally distinct responses to environmental conditions in lacustrine environments, which is not surprising due to their heterogeneous nature. However this raises implications for the development of a consistent 3-OH FA temperature calibration based on one proxy for application in lacustrine environments, for which efforts are being made for the brGDGT-based MBT'_{5ME}. Two regional lake transects have been investigated for 3-OH FAs, and we recommend the investigation of lacustrine 3-OH FAs across globally distributed transects in future studies to fully investigate the complexity of 3-OH FA temperature variability in lacustrine

environments.

A limitation to the U.S. soil and lake studies is that pedogenic and lacustrine environments are complex environments influenced by a system of many variables, and not all environmental variables that could influence 3-OH FA production could be included in this thesis. Physical conditions of soils such as particle size and cation chemistry are important soil properties that are known to influence GDGT distributions (De Jonge et al. 2024) however could not be measured for this study. Additionally, lake water chemistry (DO, salinity, cation and anion concentrations), turbidity regimes, and LSWT and BWT are important conditions to fully understand the occurrence of biomarker distributions in lacustrine environments, however were not recorded for the SW U.S. lakes. Another limitation is that MAAT is used for SW U.S. lacustrine calibrations due to absence of long-term LSWT records, however this is an acknowledged limitation for many biomarker-based lacustrine temperature calibrations (Raberg et al. 2021, Bauersachs et al. 2024, Yang et al. 2021).

The third aim of this thesis was to **assess the performance of a low temperature acid hydrolysis method in 3-OH FA extraction**. In Chapter 5, the SW lake sediments and soils were re-extracted using a Low n' Slow hydrolysis method (70° C, 14 hrs, 3M HCl) and compared with counterpart samples presented in Chapter 4. The lower hydrolysis temperature successfully streamlined 3-OH FA extraction, yielded 3-OH FAs, and reduced risk of thermal modification of low melting point lipid monomers. However, a number of implications were revealed; differing hydrolysis temperature for 3-OH FA extraction can alter fractional proportion of branched and *normal*- structural groups, which can affect palaeoclimatic interpretations from a dataset if a consistent hydrolysis temperature is not used for 3-OH FA extraction.

Chapter 6 exhibits a short study investigating 3-OH FA distributions in a 105 yr lacustrine sequence from Esthwaite Water. This initial study reveals 3-OH FA proxies are able to track environmental conditions in this sequence at a sub-decadal resolution. However, the RAN proxies

uniquely have significant positive correlations with surface water temperature, which have been negative amongst all other global sediments. This is potentially due to a Gram-negative bacterial response to nutrient enrichment (i.e. Winter SRP) rather than temperature.

The overarching aim of this thesis was to produce novel regional temperature and pH calibrations based on 3-OH FAs in pedogenic and lacustrine environments, and particularly to further investigate the environmental controls on 3-OH FA distributions in lakes. Although all calibrations proposed are regional, our $-\log(\text{RIN}_{\text{USA}})$ pH proxy from U.S. soils is a better descriptor of pH-induced 3-OH FA variability amongst globally compiled soils compared to RIAN. Additionally, our study of 3-OH FA distributions in SW U.S. lakes is a novel comparison of 3-OH FA proxies and calibrations between lake sediments and soils and indicates that development and application of lacustrine 3-OH FA proxies are much more localised compared to the soil-based RAN_{17} proxy which has a more widespread applicability in humid acidic soils. This thesis has yielded novel insights into 3-OH FA distributions in pedogenic and lacustrine environments which have the potential to improve 3-OH FA temperature and pH proxy calibrations in the future, and has produced regional 3-OH FA temperature and pH calibrations applicable to U.S. soils (T and pH) and lakes (T).

7.2 Recommendations for future work

There is still much to learn about 3-OH FA distributions in pedogenic and lacustrine environments following on from the findings of this thesis. 3-OH FA distributions have now been investigated across two regional lake transects in the SW U.S. (this study) and China (([Yang et al. 2021](#))), where different 3-OH FA-based proxies were found to convey a potential Gram-negative bacterial response to temperature. There is now potential for a global-scale investigation of 3-OH FAs in lakes to determine the extent to which regional 3-OH FA-based temperature proxies would be required for applicability in lacustrine environments. Furthermore, the sources of 3-OH FAs in

lakes remain uncertain; the extent to which 3-OH FAs are produced in the water column or in lake sediment is unknown, as is the influence of 3-OH FAs from detrital input from catchment soils on lake 3-OH FA distributions. We recommend the high-resolution sampling of sediments within lakes and their catchment soils to investigate the variability of 3-OH FAs within single lake catchments.

Our calibrations are regionally applicable to the U.S. however their applicability outside of the U.S. should still be considered and tested. The Esthwaite Water investigation revealed 3-OH FA proxies are variable at a sub-decadal timescale, and based on other lacustrine organic biomarker studies, 3-OH FAs should be applicable to at least Quaternary timescales (hundreds of thousands of years) however 3-OH FAs have not yet been explored in temporal lake cores, and we recommend that a range of regional to global 3-OH FA-based temperature calibrations are considered when investigating 3-OH FA distributions in a lake core in future studies. We additionally highlight the importance of measuring water chemistry conditions of lakes such as DO, salinity, ionic concentrations, LSWT and [BWT](#), and knowledge of turbidity and temperature regimes of lakes used in calibrations, as well as soil chemistry (cations, particle size) in soils which can provide data to quantify soil type.

Finally, there is also much scope for further streamlining the 3-OH FA extraction method: there is potential to combine the processes of acid hydrolysis and methylation using methanolysis, and optimisation of the molarity, temperature, and duration conditions for Low n' Slow hydrolysis. Overall, this thesis has produced novel 3-OH FA temperature and pH calibrations for application in pedogenic and lacustrine environments, and there is much potential for further 3-OH FA-based work in these environments in the future, as well as to streamline 3-OH FA extraction.

REFERENCES

- Al-Awadhi, H., Dashti, N., Khanafer, M., Al-Mailem, D., Ali, N. & Radwan, S. (2013), 'Bias problems in culture-independent analysis of environmental bacterial communities: a representative study on hydrocarbonoclastic bacteria', *SpringerPlus* **2**, 1–11.
- Alexander, C. & Zähringer, U. (2002), 'Chemical structure of Lipid A-the primary immunomodulatory center of bacterial lipopolysaccharides', *Trends in Glycoscience and Glycotechnology* **14**(76), 69–86.
- Ali, M. M. & Mudge, S. M. (2008), 'Diagenesis of Fatty Acids in Conwy Estuary, United Kingdom', *Sains Malaysiana* **37**(4), 299–305.
- Angst, G., Cajthaml, T., Angst, Š., Mueller, K. E., Kögel-Knabner, I., Beggel, S., Krieger, S. & Mueller, C. W. (2017), 'Performance of base hydrolysis methods in extracting bound lipids from plant material, soils, and sediments', *Organic geochemistry* **113**, 97–104.
- Annous, B. A., Kozempel, M. F. & Kurantz, M. J. (1999), 'Changes in membrane fatty acid composition of *Pediococcus* sp. strain NRRL B-2354 in response to growth conditions and its effect on thermal resistance', *Applied and Environmental Microbiology* **65**(7), 2857–2862.
- Apurv, T. & Cai, X. (2021), 'Regional drought risk in the contiguous United States', *Geophysical Research Letters* **48**(5), e2020GL092200.
- Armitage, D. W. (2023), 'Global maps of lake surface water temperatures reveal pitfalls of air-for-water substitutions in ecological prediction', *Ecography* **2023**(2), e06595.
- Bachmann, T., Friese, K. & Zachmann, D. (2001), 'Redox and pH conditions in the water column and in the sediments of an acidic mining lake', *Journal of Geochemical Exploration* **73**(2), 75–86.
- Backer, L. C., Landsberg, J. H., Miller, M., Keel, K. & Taylor, T. K. (2013), 'Canine cyanotoxin

- poisonings in the United States (1920s-2012): Review of suspected and confirmed cases from three data sources', *Toxins* **5**(9), 1597–1628.
- Bartram, A. K., Jiang, X., Lynch, M. D., Masella, A. P., Nicol, G. W., Dushoff, J. & Neufeld, J. D. (2014), 'Exploring links between pH and bacterial community composition in soils from the Craibstone Experimental Farm', *FEMS Microbiology Ecology* **87**(2), 403–415.
- Bauersachs, T., Rochelmeier, J. & Schwark, L. (2015), 'Seasonal lake surface water temperature trends reflected by heterocyst glycolipid-based molecular thermometers', *Biogeosciences* **12**(12), 3741–3751.
- Bauersachs, T., Russell, J. M., Evans, T. W., Schwalb, A. & Schwark, L. (2021), 'A heterocyste glycolipid-based calibration to reconstruct past continental climate change', *Nature Communications* **12**(1), 2406.
- Bauersachs, T., Schubert, C. J., Mayr, C., Gilli, A. & Schwark, L. (2024), 'Branched GDGT-based temperature calibrations from Central European lakes', *Science of The Total Environment* **906**, 167724.
- Baxter, A., Peterse, F., Verschuren, D. & Damsté, J. S. (2021), 'Anoxic in situ production of bacterial GMGTs in the water column and surficial bottom sediments of a meromictic tropical crater lake: Implications for lake paleothermometry', *Geochimica et Cosmochimica Acta* **306**, 171–188.
- Beck, H. E., McVicar, T. R., Vergopolan, N., Berg, A., Lutsko, N. J., Dufour, A., Zeng, Z., Jiang, X., van Dijk, A. I. & Miralles, D. G. (2023), 'High-resolution (1 km) Köppen-Geiger maps for 1901-2099 based on constrained CMIP6 projections', *Scientific data* **10**(1), 724.
- Bendle, J., Rosell-Melé, A. & Ziveri, P. (2005), 'Variability of unusual distributions of alkenones in the surface waters of the Nordic seas', *Paleoceanography* **20**(2).

- Bengtsson, L., Herschy, R. W. & Fairbridge, R. W. (2012), 'Encyclopedia of lakes and reservoirs', *Monographiae Biologicae* **53**, 10–26.
- Berry, E. D. & Foegeding, P. M. (1997), 'Cold temperature adaptation and growth of microorganisms', *Journal of Food Protection* **60**(12), 1583–1594.
- Blinn, D. W., Shannon, J. P., Benenati, P. L. & Wilson, K. P. (1998), 'Algal ecology in tailwater stream communities: the Colorado River below Glen Canyon Dam, Arizona', *Journal of Phycology* **34**(5), 734–740.
- Brauer, D., Baumhardt, R. L., Gitz, D., Gowda, P. & Mahan, J. (2015), 'Characterization of trends in reservoir storage, streamflow, and precipitation in the Canadian River watershed in New Mexico and Texas', *Lake and Reservoir Management* **31**(1), 64–79.
- Burley, T. E., Asquith, W. H. & Brooks, D. L. (2011), 'Spatially pooled depth-dependent reservoir storage, elevation, and water-quality data for selected reservoirs in Texas, January 1965-January 2010', *Reston (Virginia): US Geological Survey. Data Series* **594**, 13.
- Bush, R. T. & McInerney, F. A. (2015), 'Influence of temperature and C₄ abundance on n-alkane chain length distributions across the central USA', *Organic Geochemistry* **79**, 65–73.
- Chowdhury, M. Z. I. & Turin, T. C. (2020), 'Variable selection strategies and its importance in clinical prediction modelling', *Family medicine and community health* **8**(1).
- Dang, X., Ding, W., Yang, H., Pancost, R. D., Naafs, B. D. A., Xue, J., Lin, X., Lu, J. & Xie, S. (2018), 'Different temperature dependence of the bacterial brGDGT isomers in 35 Chinese lake sediments compared to that in soils', *Organic Geochemistry* **119**, 72–79.
- Dang, X., Yang, H., Naafs, B. D. A., Pancost, R. D. & Xie, S. (2016), 'Evidence of moisture control on the methylation of branched glycerol dialkyl glycerol tetraethers in semi-arid and arid soils', *Geochimica et Cosmochimica Acta* **189**, 24–36.

- Davidson, A. & McKerrow, A. (2016), 'GAP/LANDFIRE national terrestrial ecosystems 2011', *US Geological Survey: Reston, VA, USA*.
- De Jonge, C., Guo, J., Hällberg, P., Griepentrog, M., Rifai, H., Richter, A., Ramirez, E., Zhang, X., Smittenberg, R. H., Peterse, F. et al. (2024), 'The impact of soil chemistry, moisture and temperature on branched and isoprenoid GDGTs in soils: A study using six globally distributed elevation transects', *Organic Geochemistry* **187**, 104706.
- De Jonge, C., Hopmans, E. C., Zell, C. I., Kim, J.-H., Schouten, S. & Damsté, J. S. S. (2014b), 'Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction', *Geochimica et Cosmochimica Acta* **141**, 97–112.
- De Jonge, C., Kuramae, E. E., Radujković, D., Weedon, J. T., Janssens, I. A. & Peterse, F. (2021), 'The influence of soil chemistry on branched tetraether lipids in mid-and high latitude soils: Implications for brGDGT-based paleothermometry', *Geochimica et Cosmochimica Acta* **310**, 95–112.
- De Jonge, C., Radujković, D., Sigurdsson, B. D., Weedon, J. T., Janssens, I. & Peterse, F. (2019), 'Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils', *Organic Geochemistry* **137**, 103897.
- De Jonge, C., Stadnitskaia, A., Hopmans, E. C., Cherkashov, G., Fedotov, A. & Sinninghe Damsté, J. S. (2014a), 'In situ produced branched glycerol dialkyl glycerol tetraethers in suspended particulate matter from the Yenisei River, Eastern Siberia', *Geochimica et Cosmochimica Acta* **125**, 476–491.
- Dearing Crampton-Flood, E., Tierney, J. E., Peterse, F., Kirkels, F. M. & Damsté, J. S. S. (2020), 'BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats', *Geochimica et Cosmochimica Acta* **268**, 142–159.

- DeCelles, P. G., Castañeda, I. S., Carrapa, B., Liu, J., Quade, J., Leary, R. & Zhang, L. (2018), 'Oligocene-Miocene great lakes in the India-Asia collision zone', *Basin Research* **30**, 228–247.
- Denich, T., Beaudette, L., Lee, H. & Trevors, J. (2003), 'Effect of selected environmental and physico-chemical factors on bacterial cytoplasmic membranes', *Journal of microbiological methods* **52**(2), 149–182.
- Díaz, S. C., Therrell, M. D., Stahle, D. W. & Cleaveland, M. K. (2002), 'Chihuahua (Mexico) winter-spring precipitation reconstructed from tree-rings, 1647-1992', *Climate Research* **22**(3), 237–244.
- Ding, S., Xu, Y., Wang, Y., He, Y., Hou, J., Chen, L. & He, J.-S. (2015), 'Distributions of glycerol dialkyl glycerol tetraethers in surface soils of Qinghai-Tibetan Plateau: implications of GDGT-based proxies in cold and dry regions.', *Biogeosciences Discussions* **12**(1).
- Dirghangi, S. S., Pagani, M., Hren, M. T. & Tipple, B. J. (2013), 'Distribution of glycerol dialkyl glycerol tetraethers in soils from two environmental transects in the USA', *Organic Geochemistry* **59**, 49–60.
- Dong, X., Bennion, H., Maberly, S. C., Sayer, C. D., Simpson, G. L. & Battarbee, R. W. (2012), 'Nutrients exert a stronger control than climate on recent diatom communities in Esthwaite Water: evidence from monitoring and palaeolimnological records', *Freshwater Biology* **57**(10), 2044–2056.
- Dong, Z., Yang, Y., Wang, C., Bendle, J. A., Ruan, X., Lü, X. & Xie, S. (2023), 'Development of a novel sea surface temperature proxy based on bacterial 3-hydroxy fatty acids', *Frontiers in Marine Science* **9**, 1050269.
- Dubey, N. & Ghosh, S. (2023), 'The relative role of soil moisture and vapor pressure deficit in affecting the Indian vegetation productivity', *Environmental Research Letters* **18**(6), 064012.
- Edlund, A., Nichols, P. D., Roffey, R. & White, D. C. (1985), 'Extractable and lipopolysaccharide

- fatty acid and hydroxy acid profiles from desulfovibrio species.’, *Journal of Lipid Research* **26**(8), 982–988.
- Eglinton, G. & Calvin, M. (1967), ‘Chemical fossils’, *Scientific American* **216**(1), 32–43.
- Eglinton, T. I. & Eglinton, G. (2008), ‘Molecular proxies for paleoclimatology’, *Earth and Planetary Science Letters* **275**(1-2), 1–16.
- Eley, Y. L., Thompson, W., Greene, S. E., Mandel, I., Edgar, K., Bendle, J. A. & Dunkley Jones, T. (2019), ‘OPTiMAL: a new machine learning approach for GDGT-based palaeothermometry’.
- Emmett, B., Frogbrook, Z., Chamberlain, P., Griffiths, R., Pickup, R., Poskitt, J., Reynolds, B., Rowe, E., Rowland, P., Wilson, J. et al. (2008), ‘Countryside Survey. Soils Manual’.
- Ernst, R., Ejsing, C. S. & Antonny, B. (2016), ‘Homeoviscous adaptation and the regulation of membrane lipids’, *Journal of molecular biology* **428**(24), 4776–4791.
- Fenn, M. E., Huntington, T., McLaughlin, S., Eagar, C., Gomez, A., Cook, R. et al. (2006), ‘Status of soil acidification in North America’, *Journal of Forest Science* **52**(Special Issue), 3–13.
- Feuchtmayr, H., Clarke, M., De Ville, M., Dodd, B., Fletcher, J., James, J., Mackay, E., Pereira, M., Rhodes, G., Thackeray, S. & Maberly, S. (2021), ‘Surface temperature, surface oxygen, water clarity, water chemistry and phytoplankton Chlorophyll a data from Esthwaite Water, 2014 to 2018.’, *NERC Environmental Information Data Centre* .
- Fick, S. E. & Hijmans, R. J. (2017), ‘Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas’, *International journal of climatology* **37**(12), 4302–4315.
- Foster, L. C., Pearson, E. J., Juggins, S., Hodgson, D. A., Saunders, K. M., Verleyen, E. & Roberts, S. J. (2016), ‘Development of a regional glycerol dialkyl glycerol tetraether (GDGT)-temperature calibration for Antarctic and sub-Antarctic lakes’, *Earth and Planetary Science Letters* **433**, 370–379.

- Français, M., Bott, R., Dargaignaratz, C., Giniès, C., Carlin, F., Broussolle, V. & Nguyen-Thé, C. (2021), 'Short-chain and unsaturated fatty acids increase sequentially from the lag phase during cold growth of *Bacillus cereus*', *Frontiers in Microbiology* **12**, 694757.
- Galanos, C., Lüderitz, O., Rietschel, E. T., Westphal, O., Brade, H., Brade, L., Freudenberg, M., Schade, U., Imoto, M., Yoshimura, H. et al. (1985), 'Synthetic and natural *Escherichia coli* free Lipid A express identical endotoxic activities', *European journal of biochemistry* **148**(1), 1–5.
- Gallagher, T. M., Hren, M. & Sheldon, N. D. (2019), 'The effect of soil temperature seasonality on climate reconstructions from paleosols', *American Journal of Science* **319**(7), 549–581.
- Garcia-Manyes, S., Oncins, G. & Sanz, F. (2005), 'Effect of temperature on the nanomechanics of lipid bilayers studied by force spectroscopy', *Biophysical journal* **89**(6), 4261–4274.
- Geladi, P., Bårring, H., Dåbakk, E., Trygg, J., Antti, H., Wold, S. & Karlberg, B. (1999), 'Calibration transfer for predicting lake-water pH from near infrared spectra of lake sediments', *Journal of Near Infrared Spectroscopy* **7**(4), 251–264.
- Gelca, R., Hayhoe, K., Scott-Fleming, I., Crow, C., Dawson, D. & Patiño, R. (2016), 'Climate-water quality relationships in Texas reservoirs', *Hydrological Processes* **30**(1), 12–29.
- George, S. S. (2014), 'An overview of tree-ring width records across the Northern Hemisphere', *Quaternary Science Reviews* **95**, 132–150.
- Guerzoni, M. E., Lanciotti, R. & Cocconcelli, P. S. (2001), 'Alteration in cellular fatty acid composition as a response to salt, acid, oxidative and thermal stresses in *Lactobacillus helveticus*', *Microbiology* **147**(8), 2255–2264.
- Haan, G., Van der Heide, S. & Wolthers, B. (1979), 'Analysis of fatty acids from human lipids by gas chromatography', *Journal of Chromatography B: Biomedical Sciences and Applications* **162**(3), 261–271.

- Halfman, R., Lembrechts, J., Radujković, D., De Gruyter, J., Nijs, I. & De Jonge, C. (2022), 'Soil chemistry, temperature and bacterial community composition drive brGDGT distributions along a subarctic elevation gradient', *Organic Geochemistry* **163**, 104346.
- Hardman, A. L. (2019), The Development of 3-Hydroxy Fatty Acid-Based Palaeoclimate Proxies, MSci thesis, University of Birmingham, Birmingham, UK.
- Hardouin, C. & Lemaitre, S. (2020), 'Safety case study. intrinsic instability of concentrated solutions of alcoholic hydrogen chloride: potential hazards associated with methanol', *Organic Process Research & Development* **24**(5), 867–871.
- Hellequin, E., Collin, S., Seder-Colomina, M., Véquaud, P., Anquetil, C., Kish, A. & Huguet, A. (2023), 'Membrane lipid adaptation of soil Bacteroidetes isolates to temperature and pH', *Frontiers in Microbiology* **14**, 1032032.
- Herlihy, A. & Mills, A. (1986), 'The pH regime of sediments underlying acidified waters', *Biogeochemistry* **2**, 95–99.
- Hopmans, E. C., Schouten, S. & Sinninghe Damsté, J. S. (2016), 'The effect of improved chromatography on GDGT-based palaeoproxies', *Organic Geochemistry* **93**, 1–6.
- Hopmans, E. C., Weijers, J. W., Schefuß, E., Herfort, L., Sinninghe Damsté, J. S. & Schouten, S. (2004), 'A novel proxy for terrestrial organic matter in sediments based on branched and isoprenoid tetraether lipids', *Earth and Planetary Science Letters* **224**(1-2), 107–116.
- Horton, J. D. (2017), 'The state geologic map compilation (SGMC) geodatabase of the conterminous United States'.
- Hou, J., D'Andrea, W. J. & Huang, Y. (2008), 'Can sedimentary leaf waxes record D/H ratios of continental precipitation? Field, model, and experimental assessments', *Geochimica et Cosmochimica Acta* **72**(14), 3503–3517.

- Huang, X., Meyers, P. A., Jia, C., Zheng, M., Xue, J., Wang, X. & Xie, S. (2013), 'Paleotemperature variability in central china during the last 13 ka recorded by a novel microbial lipid proxy in the Dajiuhu peat deposit', *The Holocene* **23**(8), 1123–1129.
- Huguet, A., Coffinet, S., Roussel, A., Gayraud, F., Anquetil, C., Bergonzini, L., Bonanomi, G., Williamson, D., Majule, A. & Derenne, S. (2019), 'Evaluation of 3-hydroxy fatty acids as a pH and temperature proxy in soils from temperate and tropical altitudinal gradients', *Organic Geochemistry* **129**, 1–13.
- Ichihara, K. & Fukubayashi, Y. (2010), 'Preparation of fatty acid methyl esters for gas-liquid chromatography', *Journal of lipid research* **51**(3), 635–640.
- Inglis, G. N., Collinson, M. E., Riegel, W., Wilde, V., Farnsworth, A., Lunt, D. J., Valdes, P., Robson, B. E., Scott, A. C., Lenz, O. K. et al. (2017), 'Mid-latitude continental temperatures through the early Eocene in western Europe', *Earth and Planetary Science Letters* **460**, 86–96.
- IPCC (2021), *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, Vol. In Press, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jacqueline, M. B. (2013), 'Chapter 6-Lipids Basics: Fats and Oils in Foods and Health: Healthy Lipid Choices, Roles and Applications in Nutrition', *Food Science and the Culinary Arts, The Science and Practice of Healthy Cooking, Culinary Nutrition* pp. 231–277.
- Jansen, B., Nierop, K. G., Kotte, M. C., de Voogt, P. & Verstraten, J. M. (2006), 'The applicability of accelerated solvent extraction (ASE) to extract lipid biomarkers from soils', *Applied Geochemistry* **21**(6), 1006–1015.
- Jansen, E., Overpeck, J., Briffa, K., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, W., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, D., Solomina, O.,

- Villalba, R. & Zhang, D. (2007), *Palaeoclimate*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
- Jenkins, T. (2010), 'Common analytical errors yielding inaccurate results during analysis of fatty acids in feed and digesta samples', *Journal of Dairy Science* **93**(3), 1170–1174.
- Jenske, R. & Vetter, W. (2008), 'Gas chromatography/electron-capture negative ion mass spectrometry for the quantitative determination of 2-and 3-hydroxy fatty acids in bovine milk fat', *Journal of agricultural and food chemistry* **56**(14), 5500–5505.
- Jham, G., Teles, F. & Campos, L. (1982), 'Use of aqueous HCl/MeOH as esterification reagent for analysis of fatty acids derived from soybean lipids', *Journal of the American Oil Chemists Society* **59**(3), 132–133.
- Jones, R. T., Robeson, M. S., Lauber, C. L., Hamady, M., Knight, R. & Fierer, N. (2009), 'A comprehensive survey of soil acidobacterial diversity using pyrosequencing and clone library analyses', *The ISME journal* **3**(4), 442–453.
- Joslyn, M. A. (1970), 'Methods in food analysis: physical, chemical, and instrumental methods of analysis', (*No Title*) .
- Kaneda, T. (1991), 'Iso-and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance', *Microbiological reviews* **55**(2), 288–302.
- Kapoor, V., Chawla, A. & Jindal, O. (1979), 'Fatty acid composition of yellow-and violet-flowered sesbania sesban seeds', *Journal of the American Oil Chemists' Society* **56**(10), 904–905.
- Kato, H., Haishima, Y., Iida, T., Tanaka, A. & Tanamoto, K.-i. (1998), 'Chemical structure of Lipid A isolated from *Flavobacterium meningosepticum* lipopolysaccharide', *Journal of bacteriology* **180**(15), 3891–3899.

- Kawamura, K. & Ishiwatari, R. (1982), 'Tightly bound β -hydroxy acids in a recent sediment', *Nature* **297**(5862), 144–145.
- Keinanen, M., Korhonen, L., Martikainen, P., Vartiainen, T., Miettinen, I., Lehtola, M., Nenonen, K., Pajunen, H. & Kontro, M. H. (2003), 'Gas chromatographic-mass spectrometric detection of 2-and 3-hydroxy fatty acids as methyl esters from soil, sediment and biofilm', *Journal of Chromatography B* **783**(2), 443–451.
- Keith, D., Rover, J., Green, J., Zalewsky, B., Charpentier, M., Thursby, G. & Bishop, J. (2018), 'Monitoring algal blooms in drinking water reservoirs using the Landsat-8 Operational Land Imager', *International journal of remote sensing* **39**(9), 2818–2846.
- Kendrick, G. D. (1993), 'The historic american engineering record and re-interpreting the history of the west'.
- Ketelle, M. J. & Uttormark, P. D. (1971), *Problem lakes in the United States*, US Environmental Protection Agency.
- Kim, J.-H., Schouten, S., Hopmans, E. C., Donner, B. & Damsté, J. S. S. (2008), 'Global sediment core-top calibration of the TEX₈₆ paleothermometer in the ocean', *Geochimica et Cosmochimica Acta* **72**(4), 1154–1173.
- Klages, J. P., Salzmann, U., Bickert, T., Hillenbrand, C.-D., Gohl, K., Kuhn, G., Bohaty, S. M., Titschack, J., Müller, J., Frederichs, T. et al. (2020), 'Temperate rainforests near the South Pole during peak Cretaceous warmth', *Nature* **580**(7801), 81–86.
- Klok, J., Baas, M., Cox, H., De Leeuw, J., Rijpstra, W. & Schenck, P. (1988), 'The mode of occurrence of lipids in a Namibian Shelf diatomaceous ooze with emphasis on the β -hydroxy fatty acids', *Organic geochemistry* **12**(1), 75–80.
- Küchler, A. (1964), 'Potential natural vegetation of the conterminous United States (manual and map.) Special Publ. 36, 1965 rev', *New York: American Geographical Society* .

- Lane, N. (2017), 'Proton gradients at the origin of life', *BioEssays* **39**(6), 1600217.
- Laurens, L. M., Quinn, M., Van Wychen, S., Templeton, D. W. & Wolfrum, E. J. (2012), 'Accurate and reliable quantification of total microalgal fuel potential as fatty acid methyl esters by in situ transesterification', *Analytical and bioanalytical chemistry* **403**, 167–178.
- Lee, A. K., Chan, C. K., Fang, M. & Lau, A. P. (2004), 'The 3-hydroxy fatty acids as biomarkers for quantification and characterization of endotoxins and Gram-negative bacteria in atmospheric aerosols in Hong Kong', *Atmospheric Environment* **38**(37), 6307–6317.
- Lei, Y., Strong, D. J., Caballero, M., Correa-Metrio, A., Pérez, L., Schwalb, A., Macario-González, L., Cohuo, S., Lozano-García, S., Ortega-Guerrero, B. et al. (2023), 'Regional vs. global temperature calibrations for lacustrine brGDGTs in the North American (sub) tropics: Implications for their application in paleotemperature reconstructions', *Organic Geochemistry* **184**, 104660.
- Li, Y., Zhao, S., Pei, H., Qian, S., Zang, J., Dang, X. & Yang, H. (2018), 'Distribution of glycerol dialkyl glycerol tetraethers in surface soils along an altitudinal transect at cold and humid Mountain Changbai: Implications for the reconstruction of paleoaltimetry and paleoclimate', *Science China Earth Sciences* **61**, 925–939.
- Li, Z., Huang, X., Naafs, B. D. A., Wang, C. & Xie, S. (2022), 'Evaluation of the wider applications of the alkanol index BNA₁₅ as temperature proxy in a broad distribution of peat deposits', *Organic Geochemistry* **173**, 104486.
- Liang, J., Russell, J. M., Xie, H., Lupien, R. L., Si, G., Wang, J., Hou, J. & Zhang, G. (2019), 'Vegetation effects on temperature calibrations of branched glycerol dialkyl glycerol tetraether (brGDGTs) in soils', *Organic Geochemistry* **127**, 1–11.
- Liu, K.-S. (1994), 'Preparation of fatty acid methyl esters for gas-chromatographic analysis of lipids in biological materials', *Journal of the American Oil Chemists' Society* **71**, 1179–1187.

- Loomis, S. E., Russell, J. M., Heureux, A. M., D'Andrea, W. J. & Sinninghe Damsté, J. S. (2014), 'Seasonal variability of branched glycerol dialkyl glycerol tetraethers (brGDGTs) in a temperate lake system', *Geochimica et Cosmochimica Acta* **144**, 173–187.
- Loomis, S. E., Russell, J. M. & Sinninghe Damsté, J. S. (2011), 'Distributions of branched GDGTs in soils and lake sediments from western Uganda: Implications for a lacustrine paleothermometer', *Organic Geochemistry* **42**(7), 739–751.
- Lozano-Parra, J., Pulido, M., Lozano-Fondón, C. & Schnabel, S. (2018), 'How do soil moisture and vegetation covers influence soil temperature in drylands of Mediterranean regions?', *Water* **10**(12), 1747.
- Maberly, S., Brierley, B., Carter, H., Clarke, M., De Ville, M., Fletcher, J., James, J., Keenan, P., Kelly, J., Mackay, E. et al. (2017), 'Surface temperature, surface oxygen, water clarity, water chemistry and phytoplankton Chlorophyll a data from Esthwaite Water, 1945 to 2013', *NERC Environmental Information Data Centre*.
- Martínez-Sosa, P., Tierney, J. E., Pérez-Angel, L. C., Stefanescu, I. C., Guo, J., Kirkels, F., Sepúlveda, J., Peterse, F., Shuman, B. N. & Reyes, A. V. (2023), 'Development and application of the Branched and Isoprenoid GDGT Machine learning Classification algorithm (BIGMaC) for paleoenvironmental reconstruction', *Paleoceanography and Paleoclimatology* **38**(7), e2023PA004611.
- Martínez-Sosa, P., Tierney, J. E., Stefanescu, I. C., Dearing Crampton-Flood, E., Shuman, B. N. & Routson, C. (2021), 'A global Bayesian temperature calibration for lacustrine brGDGTs', *Geochimica et Cosmochimica Acta* **305**, 87–105.
- Mayberry, W. R. (1980), 'Hydroxy fatty acids in bacteroides species: D-(-)-3-hydroxy-15-methylhexadecanoate and its homologs', *Journal of Bacteriology* **143**(2), 582–587.

- McClymont, E. L., Mackay, H., Stevenson, M. A., Damm-Johnsen, T., Honan, E. M., Penny, C. E. & Cole, Y. A. (2023), 'Biomarker proxies for reconstructing Quaternary climate and environmental change', *Journal of Quaternary Science* **38**(7), 991–1024.
- Menges, J., Huguet, C., Alcañiz, J. M., Fietz, S., Sachse, D. & Rosell-Melé, A. (2014), 'Influence of water availability in the distributions of branched glycerol dialkyl glycerol tetraether in soils of the Iberian Peninsula', *Biogeosciences* **11**(10), 2571–2581.
- Mielniczuk, Z., Mielniczuk, E. & Larsson, L. (1993), 'Gas chromatography-mass spectrometry methods for analysis of 2-and 3-hydroxylated fatty acids: Application for endotoxin measurement', *Journal of microbiological methods* **17**(2), 91–102.
- Milinsk, M. C., Matsushita, M., Visentainer, J. V., Oliveira, C. C. d. & Souza, N. E. d. (2008), 'Comparative analysis of eight esterification methods in the quantitative determination of vegetable oil fatty acid methyl esters (FAME)', *Journal of the Brazilian Chemical Society* **19**, 1475–1483.
- Miller, R. O. & Kissel, D. E. (2010), 'Comparison of soil pH methods on soils of North America', *Soil Science Society of America Journal* **74**(1), 310–316.
- Molnar, P. (2022), 'Differences between soil and air temperatures: Implications for geological reconstructions of past climate', *Geosphere* **18**(2), 800–824.
- Montone, C. M., Aita, S. E., Catani, M., Cavaliere, C., Cerrato, A., Piovesana, S., Laganà, A. & Capriotti, A. L. (2021), 'Profiling and quantitative analysis of underivatized fatty acids in *Chlorella vulgaris* microalgae by liquid chromatography-high resolution mass spectrometry', *Journal of Separation Science* **44**(16), 3041–3051.
- Müller, P. J., Kirst, G., Ruhland, G., Von Storch, I. & Rosell-Melé, A. (1998), 'Calibration of the alkenone paleotemperature index U^k_{37} based on core-tops from the eastern South Atlantic and the global ocean (60 N-60 S)', *Geochimica et cosmochimica Acta* **62**(10), 1757–1772.

- Naafs, B. D. A., Inglis, G. N., Zheng, Y., Amesbury, M., Biester, H., Bindler, R., Blewett, J., Burrows, M., Del Castillo Torres, D., Chambers, F. M. et al. (2017b), 'Introducing global peat-specific temperature and pH calibrations based on brGDGT bacterial lipids', *Geochimica et Cosmochimica Acta* **208**, 285–301.
- Naafs, B., Rohrsen, M., Inglis, G. N., Lhteenoja, O., Feakins, S. J., Collinson, M. E., Kennedy, E. M., Singh, P., Singh, M., Lunt, D. J. et al. (2018), 'High temperatures in the terrestrial mid-latitudes during the early Palaeogene', *Nature Geoscience* **11**(10), 766–771.
- Nachtergaele, F., van Velthuisen, H., Verelst, L., Wiberg, D., Henry, M., Chiozza, F., Yigini, Y., Aksoy, E., Batjes, N., Boateng, E. et al. (2023), *Harmonized World Soil Database version 2.0*, Food and Agriculture Organization of the United Nations.
- Newton, R. J. & McMahon, K. D. (2011), 'Seasonal differences in bacterial community composition following nutrient additions in a eutrophic lake', *Environmental Microbiology* **13**(4), 887–899.
- Nordstrm, K. M. & Laakso, S. V. (1992), 'Effect of growth temperature on fatty acid composition of ten *Thermus* strains', *Applied and environmental microbiology* **58**(5), 1656–1660.
- Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M. O., Lahti, L., McGlinn, D., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C. J. & Weedon, J. (2022), *Vegan: Community Ecology Package*. R package version 2.6-4.
URL: <https://CRAN.R-project.org/package=vegan>
- Orata, F. (2012), 'Derivatization reactions and reagents for gas chromatography analysis',

- Oyaizu, H. & Komagata, K. (1981), 'Chemotaxonomic and phenotypic characterization of the strains of species in the Flavobacterium-Cytophaga complex', *The Journal of General and Applied Microbiology* **27**(1), 57–107.
- Pan, F., Yuan, H., Song, J., Yao, Q., Li, X., Duan, L. & Xing, J. (2023), '3-hydroxy fatty acids as proxies for seawater temperature and pH in the eastern China marginal seas', *Chemical Geology* **617**, 121258.
- Patiño, R., Christensen, V. G., Graham, J. L., Rogosch, J. S. & Rosen, B. H. (2023), 'Toxic algae in inland waters of the conterminous United States—A review and synthesis', *Water* **15**(15), 2808.
- Pearson, E. J., Juggins, S., Talbot, H. M., Weckström, J., Rosén, P., Ryves, D. B., Roberts, S. J. & Schmidt, R. (2011), 'A lacustrine GDGT-temperature calibration from the Scandinavian Arctic to Antarctic: Renewed potential for the application of GDGT-paleothermometry in lakes', *Geochimica et Cosmochimica Acta* **75**(20), 6225–6238.
- Pedrós-Alió, C., Calderón-Paz, J. I., MacLean, M. H., Medina, G., Marrasé, C., Gasol, J. M. & Guixa-Boixereu, N. (2000), 'The microbial food web along salinity gradients', *FEMS Microbiology Ecology* **32**(2), 143–155.
- Peterse, F., van der Meer, J., Schouten, S., Weijers, J. W., Fierer, N., Jackson, R. B., Kim, J.-H. & Damsté, J. S. S. (2012), 'Revised calibration of the MBT-CBT paleotemperature proxy based on branched tetraether membrane lipids in surface soils', *Geochimica et Cosmochimica Acta* **96**, 215–229.
- Pettersson, M. & Bååth, E. (2003), 'Temperature-dependent changes in the soil bacterial community in limed and unlimed soil', *FEMS Microbiology ecology* **45**(1), 13–21.

- Picek, T., Šimek, M. & Šantrčková, H. (2000), 'Microbial responses to fluctuation of soil aeration status and redox conditions', *Biology and Fertility of Soils* **31**, 315–322.
- Plante, A. F., Conant, R. T., Paul, E. A., Paustian, K. & Six, J. (2006), 'Acid hydrolysis of easily dispersed and microaggregate-derived silt-and clay-sized fractions to isolate resistant soil organic matter', *European Journal of Soil Science* **57**(4), 456–467.
- Poger, D., Caron, B. & Mark, A. E. (2014), 'Effect of methyl-branched fatty acids on the structure of lipid bilayers', *The Journal of Physical Chemistry B* **118**(48), 13838–13848.
- Prakash, O., Nimonkar, Y., Shaligram, S., Joseph, N. & Shouche, Y. S. (2015), 'Response of cellular fatty acids to environmental stresses in endophytic *Micrococcus* spp.', *Annals of microbiology* **65**, 2209–2218.
- PRISM Climate Group, O. S. U. (2014), 'CiteDrive brings reference management to Overleaf'.
<https://prism.oregonstate.edu> [Accessed: 05/10/2019].
- Quenea, K., Mathieu, J. & Derenne, S. (2012), 'Soil lipids from accelerated solvent extraction: Influence of temperature and solvent on extract composition', *Organic geochemistry* **44**, 45–52.
- R Core Team (2023), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: <https://www.R-project.org/>
- Raberg, J. H., Harning, D. J., Crump, S. E., de Wet, G., Blumm, A., Kopf, S., Geirsdóttir, Á., Miller, G. H. & Sepúlveda, J. (2021), 'Revised fractional abundances and warm-season temperatures substantially improve brGDGT calibrations in lake sediments', *Biogeosciences* **18**(12), 3579–3603.
- Raetz, C. R., Reynolds, C. M., Trent, M. S. & Bishop, R. E. (2007), 'Lipid A modification systems in Gram-negative bacteria', *Annu. Rev. Biochem.* **76**, 295–329.

- Rietschel, E. T. (1976), 'Absolute configuration of 3-hydroxy fatty acids present in lipopolysaccharides from various bacterial groups', *European Journal of Biochemistry* **64**(2), 423–428.
- Russell, J. M., Hopmans, E. C., Loomis, S. E., Liang, J. & Damsté, J. S. 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', *Organic Geochemistry* **117**, 56–69.
- Russell, N. (1995), Psychrotrophy and adaptation to low temperatures: microbial membrane lipids, *in* 'Proceedings of the 19th International Congress on Refrigeration. Workshop Refrigeration and Microbiology: Health, Food, Drinks and Flowers', Vol. 1, pp. 359–365.
- Russell, N. & Fukunaga, N. (1990), 'A comparison of thermal adaptation of membrane lipids in psychrophilic and thermophilic bacteria', *FEMS Microbiology Letters* **75**(2-3), 171–182.
- Sauer, D., Sponagel, H., Sommer, M., Giani, L., Jahn, R. & Stahr, K. (2007), 'Podzol: Soil of the year 2007. A review on its genesis, occurrence, and functions', *Journal of Plant Nutrition and Soil Science* **170**(5), 581–597.
- Schoon, P. L., Heilmann-Clausen, C., Schultz, B. P., Sluijs, A., Sinninghe Damsté, J. S. & Schouten, S. (2013), 'Recognition of Early Eocene global carbon isotope excursions using lipids of marine Thaumarchaeota', *Earth and Planetary Science Letters* **373**, 160–168.
- Schouten, S., Hopmans, E. C. & Sinninghe Damsté, J. S. (2013), 'The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review', *Organic geochemistry* **54**, 19–61.
- Seki, O., Okazaki, Y. & Harada, N. (2021), 'Assessment of long-chain n-alkanes as a paleoclimate proxy in the Bering Sea sediments', *Progress in Oceanography* **198**, 102687.
- Shanahan, T. M., Huguen, K. A. & Van Mooy, B. A. (2013), 'Temperature sensitivity of branched and isoprenoid GDGTs in Arctic lakes', *Organic Geochemistry* **64**, 119–128.

- Sheppard, P. R., Comrie, A. C., Packin, G. D., Angersbach, K. & Hughes, M. K. (2002), 'The climate of the US Southwest', *Climate Research* **21**(3), 219–238.
- Siliakus, M. F., van der Oost, J. & Kengen, S. W. (2017), 'Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure', *Extremophiles* **21**, 651–670.
- Sinninghe Damsté, J. S., Hopmans, E. C., Pancost, R. D., Schouten, S. & Geenevasen, J. A. (2000), 'Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments', *Chemical Communications* (17), 1683–1684.
- Sinninghe Damsté, J. S., Rijpstra, W. I. C., Foesel, B. U., Huber, K. J., Overmann, J., Nakagawa, S., Kim, J. J., Dunfield, P. F., Dedysh, S. N. & Villanueva, L. (2018), 'An overview of the occurrence of ether-and ester-linked iso-diabolic acid membrane lipids in microbial cultures of the Acidobacteria: Implications for brGDGT paleoproxies for temperature and pH', *Organic Geochemistry* **124**, 63–76.
- Sinninghe Damsté, J. S., Rijpstra, W. I. C., Hopmans, E. C., Weijers, J. W., Foesel, B. U., Overmann, J. & Dedysh, S. N. (2011), '13, 16-Dimethyl octacosanedioic acid (iso-diabolic acid), a common membrane-spanning lipid of Acidobacteria subdivisions 1 and 3', *Applied and environmental microbiology* **77**(12), 4147–4154.
- Sjogren, J., Magnusson, J., Broberg, A., Schnurer, J. & Kenne, L. (2003), 'Antifungal 3-hydroxy fatty acids from *Lactobacillus plantarum* MiLAB 14', *Applied and environmental microbiology* **69**(12), 7554–7557.
- Slessarev, E., Lin, Y., Bingham, N., Johnson, J., Dai, Y., Schimel, J. & Chadwick, O. (2016), 'Water balance creates a threshold in soil pH at the global scale', *Nature* **540**(7634), 567–569.
- Steinhardt, J. & Fugitt, C. H. (1942), 'Catalyzed hydrolysis of amide and peptide bonds in proteins', *J. Res. Natl Bur. Stand* **29**, 315–327.

- Stubbendieck, R. M., Vargas-Bautista, C. & Straight, P. D. (2016), 'Bacterial communities: interactions to scale', *Frontiers in microbiology* **7**, 217534.
- Sun, Q., Chu, G., Liu, M., Xie, M., Li, S., Ling, Y., Wang, X., Shi, L., Jia, G. & Lü, H. (2011), 'Distributions and temperature dependence of branched glycerol dialkyl glycerol tetraethers in recent lacustrine sediments from China and Nepal', *Journal of Geophysical Research: Biogeosciences* **116**(G1).
- Super, J. R., Chin, K., Pagani, M., Li, H., Tabor, C., Harwood, D. M. & Hull, P. M. (2018), 'Late Cretaceous climate in the Canadian Arctic: multi-proxy constraints from Devon Island', *Palaeogeography, palaeoclimatology, palaeoecology* **504**, 1–22.
- Talling, J. & Heaney, S. (1988), 'Long-term changes in some English (Cumbrian) lakes subjected to increased nutrient inputs', *Algae and the aquatic environment* pp. 1–29.
- Thackeray, S., Jones, I. & Maberly, S. (2008), 'Long-term change in the phenology of spring phytoplankton: species-specific responses to nutrient enrichment and climatic change', *Journal of Ecology* **96**(3), 523–535.
- Thorpe, A. C. (2023), *Exploring Past Environments with Microbial Biomarkers: A Molecular Approach Using Sedimentary DNA and 3-Hydroxy Fatty Acids*, University of Birmingham.
- Thorpe, A. C., Anderson, A., Goodall, T., Thackeray, S. J., Maberly, S. C., Bendle, J. A., Gweon, H. S. & Read, D. S. (2022), 'Sedimentary DNA records long-term changes in a lake bacterial community in response to varying nutrient availability', *Environmental DNA* **4**(6), 1340–1355.
- Thorpe, A. C., Mackay, E. B., Goodall, T., Bendle, J. A., Thackeray, S. J., Maberly, S. C. & Read, D. S. (2024), 'Evaluating the use of lake sedimentary DNA in palaeolimnology: A comparison with long-term microscopy-based monitoring of the phytoplankton community', *Molecular ecology resources* **24**(2), e13903.

- Tierney, J. E. (2012), 'GDGT thermometry: Lipid tools for reconstructing paleotemperatures', *The Paleontological Society Papers* **18**, 115–132.
- Tierney, J. E., Mayes, M. T., Meyer, N., Johnson, C., Swarzenski, P. W., Cohen, A. S. & Russell, J. M. (2010), 'Late-twentieth-century warming in Lake Tanganyika unprecedented since AD 500', *Nature Geoscience* **3**(6), 422–425.
- Tierney, J. E. & Russell, J. M. (2009), 'Distributions of branched GDGTs in a tropical lake system: implications for lacustrine application of the MBT/CBT paleoproxy', *Organic Geochemistry* **40**(9), 1032–1036.
- Tipple, B. J. & Pagani, M. (2013), 'Environmental control on eastern broadleaf forest species' leaf wax distributions and D/H ratios', *Geochimica et Cosmochimica Acta* **111**, 64–77.
- Toney, J. L., Huang, Y., Fritz, S. C., Baker, P. A., Grimm, E. & Nyren, P. (2010), 'Climatic and environmental controls on the occurrence and distributions of long chain alkenones in lakes of the interior United States', *Geochimica et Cosmochimica Acta* **74**(5), 1563–1578.
- Troudi, A., Pagès, J. M. & Brunel, J. M. (2021), 'Chemical highlights supporting the role of Lipid A in efficient biological adaptation of gram-negative bacteria to external stresses', *Journal of Medicinal Chemistry* **64**(4), 1816–1834.
- Ulmer, C. Z., Koelmel, J. P., Jones, C. M., Garrett, T. J., Aristizabal-Henao, J. J., Vesper, H. W. & Bowden, J. A. (2021), 'A review of efforts to improve lipid stability during sample preparation and standardization efforts to ensure accuracy in the reporting of lipid measurements', *Lipids* **56**(1), 3–16.
- van Bree, L. G., Peterse, F., Baxter, A. J., De Crop, W., Van Grinsven, S., Villanueva, L., Verschuren, D. & Sinninghe Damsté, J. S. (2020), 'Seasonal variability and sources of in situ brGDGT production in a permanently stratified African crater lake', *Biogeosciences Discussions* **2020**, 1–36.

- Van Bree, L., Peterse, F., Van der Meer, M., Middelburg, J., Negash, A., De Crop, W., Cocquyt, C., Wieringa, J., Verschuren, D. & Damsté, J. S. (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', *Quaternary Science Reviews* **192**, 208–224.
- van Gestel, N. C., Reischke, S. & Bååth, E. (2013), 'Temperature sensitivity of bacterial growth in a hot desert soil with large temperature fluctuations', *Soil Biology and biochemistry* **65**, 180–185.
- Véquaud, P., Derenne, S., Anquetil, C., Collin, S., Poulenard, J., Sabatier, P. & Huguet, A. (2021a), 'Influence of environmental parameters on the distribution of bacterial lipids in soils from the French Alps: Implications for paleo-reconstructions', *Organic Geochemistry* **153**, 104194.
- Véquaud, P., Derenne, S., Thibault, A., Anquetil, C., Bonanomi, G., Collin, S., Contreras, S., Nottingham, A. T., Sabatier, P., Salinas, N. et al. (2021b), 'Development of global temperature and pH calibrations based on bacterial 3-hydroxy fatty acids in soils', *Biogeosciences* **18**(12), 3937–3959.
- Véquaud, P., Thibault, A., Derenne, S., Anquetil, C., Collin, S., Contreras, S., Nottingham, A. T., Sabatier, P., Werne, J. P. & Huguet, A. (2022), 'FROG: A global machine-learning temperature calibration for branched GDGTs in soils and peats', *Geochimica et Cosmochimica Acta* **318**, 468–494.
- Wang, C., Bendle, J. A., Yang, H., Yang, Y., Hardman, A., Yamoah, A., Thorpe, A., Mandel, I., Greene, S. E., Huang, J. et al. (2021), 'Global calibration of novel 3-hydroxy fatty acid based temperature and pH proxies', *Geochimica et Cosmochimica Acta* **302**, 101–119.
- Wang, C., Bendle, J. A., Zhang, H., Yang, Y., Liu, D., Huang, J., Cui, J. & Xie, S. (2018), 'Holocene temperature and hydrological changes reconstructed by bacterial 3-hydroxy fatty acids in a stalagmite from central China', *Quaternary Science Reviews* **192**, 97–105.

- Wang, C., Bendle, J., Yang, Y., Yang, H., Sun, H., Huang, J. & Xie, S. (2016), 'Impacts of pH and temperature on soil bacterial 3-hydroxy fatty acids: Development of novel terrestrial proxies', *Organic Geochemistry* **94**, 21–31.
- Wang, C., Zhang, H., Huang, X., Huang, J. & Xie, S. (2012), 'Optimization of acid digestion conditions on the extraction of fatty acids from stalagmites', *Frontiers of Earth Science* **6**, 109–114.
- Wei, H., Peng, C., Yang, B., Song, H., Li, Q., Jiang, L., Wei, G., Wang, K., Wang, H., Liu, S. et al. (2018), 'Contrasting soil bacterial community, diversity, and function in two forests in China', *Frontiers in Microbiology* **9**, 1693.
- Weijers, J. W., Schouten, S., Hopmans, E. C., Geenevasen, J. A., David, O. R., Coleman, J. M., Pancost, R. D. & Sinninghe Damsté, J. S. (2006), 'Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits', *Environmental Microbiology* **8**(4), 648–657.
- Weijers, J. W., Schouten, S., van den Donker, J. C., Hopmans, E. C. & Damsté, J. S. S. (2007), 'Environmental controls on bacterial tetraether membrane lipid distribution in soils', *Geochimica et Cosmochimica Acta* **71**(3), 703–713.
- Weiss, M., Manneberg, M., Juranville, J.-F., Lahm, H.-W. & Fountoulakis, M. (1998), 'Effect of the hydrolysis method on the determination of the amino acid composition of proteins', *Journal of chromatography A* **795**(2), 263–275.
- Wildung, R., Garland, T. & Buschbom, R. (1975), 'The interdependent effects of soil temperature and water content on soil respiration rate and plant root decomposition in arid grassland soils', *Soil Biology and Biochemistry* **7**(6), 373–378.
- Wollenweber, H.-W., Broady, K. W., Luderitz, O. & Rietschel, E. T. (1982), 'The chemical

- structure of Lipid A: demonstration of amide-linked 3-acyloxyacyl residues in *Salmonella minnesota* Re lipopolysaccharide', *European journal of biochemistry* **124**(1), 191–198.
- Wollenweber, H.-W. & Rietschel, E. T. (1990), 'Analysis of lipopolysaccharide (Lipid A) fatty acids.',
- Wu, Y., Zeng, J., Zhu, Q., Zhang, Z. & Lin, X. (2017), 'pH is the primary determinant of the bacterial community structure in agricultural soils impacted by polycyclic aromatic hydrocarbon pollution', *Scientific reports* **7**(1), 1–7.
- Xie, S., Pancost, R. D., Chen, L., Evershed, R. P., Yang, H., Zhang, K., Huang, J. & Xu, Y. (2012), 'Microbial lipid records of highly alkaline deposits and enhanced aridity associated with significant uplift of the Tibetan Plateau in the Late Miocene', *Geology* **40**(4), 291–294.
- Yamanaka, S., Fudo, R., Kawaguchi, A. & Komagata, K. (1988), 'TAXONOMIC SIGNIFICANCE OF HYDROXY FATTY ACIDS IN MYXOBACTERIA WITH SPECIAL REFERENCE TO 2-HYDROXY FATTY ACIDS IN PHOSPHOLIPIDS', *The Journal of General and Applied Microbiology* **34**(1), 57–66.
- Yang, H., Lü, X., Ding, W., Lei, Y., Dang, X. & Xie, S. (2015), 'The 6-methyl branched tetraethers significantly affect the performance of the methylation index (MBT) in soils from an altitudinal transect at Mount Shennongjia', *Organic Geochemistry* **82**, 42–53.
- Yang, H., Pancost, R. D., Dang, X., Zhou, X., Evershed, R. P., Xiao, G., Tang, C., Gao, L., Guo, Z. & Xie, S. (2014), 'Correlations between microbial tetraether lipids and environmental variables in Chinese soils: Optimizing the paleo-reconstructions in semi-arid and arid regions', *Geochimica et Cosmochimica Acta* **126**, 49–69.
- Yang, J., Ma, L., Jiang, H., Wu, G. & Dong, H. (2016), 'Salinity shapes microbial diversity and community structure in surface sediments of the Qinghai-Tibetan Lakes', *Scientific reports* **6**(1), 25078.

- Yang, Y., Wang, C., Bendle, J. A., Yu, X., Gao, C., Lü, X., Ruan, X., Wang, R. & Xie, S. (2020), 'A new sea surface temperature proxy based on bacterial 3-hydroxy fatty acids', *Organic Geochemistry* **141**, 103975.
- Yang, Y., Wang, C., Bendle, J., Luo, Z., Dang, X., Xue, J., Xiang, X. & Xie, S. (2021), 'Appraisal of paleoclimate indices based on bacterial 3-hydroxy fatty acids in 20 Chinese alkaline lakes', *Organic Geochemistry* **160**, 104277.
- Yang, Y., Wang, C., Zhang, H., Huang, J. & Xie, S. (2016c), 'Influence of extraction methods on the distribution pattern and concentration of fatty acids and hydroxy fatty acids in soil samples: Acid digestion versus saponification', *Geochemical Journal* **50**(5), 439–443.
- Yao, L. & Hammond, E. G. (2006), 'Isolation and melting properties of branched-chain esters from lanolin', *Journal of the American Oil Chemists' Society* **83**(6), 547–552.
- Youssef, N. H., Duncan, K. E. & McInerney, M. J. (2005), 'Importance of 3-hydroxy fatty acid composition of lipopeptides for biosurfactant activity', *Applied and Environmental Microbiology* **71**(12), 7690–7695.
- Zhang, W., Hu, W., Zhu, Q., Niu, M., An, N., Feng, Y., Kawamura, K. & Fu, P. (2023), 'Hydroxy fatty acids in the surface Earth system', *Science of The Total Environment* p. 167358.
- Zhao, B., Russell, J. M., Tsai, V. C., Blaus, A., Parish, M. C., Liang, J., Wilk, A., Du, X. & Bush, M. B. (2023), 'Evaluating global temperature calibrations for lacustrine branched GDGTs: Seasonal variability, paleoclimate implications, and future directions', *Quaternary Science Reviews* **310**, 108124.
- Zhao, S., Shi, X., Sun, B., Liu, Y., Tian, Z. & Huotari, J. (2022), 'Effects of pH on phosphorus form transformation in lake sediments', *Water Supply* **22**(2), 1231–1243.
- Zhu, K., Ding, X., Julotok, M. & Wilkinson, B. J. (2005), 'Exogenous isoleucine and fatty acid shortening ensure the high content of anteiso-C₁₅: 0 fatty acid required for low-temperature

growth of *Listeria monocytogenes*', *Applied and environmental microbiology* **71**(12), 8002–8007.

Zhuang, Y., Dong, J., He, X., Wang, J., Li, C., Dong, L., Zhang, Y., Zhou, X., Wang, H., Yi, Y. et al. (2022), 'Impact of heating temperature and fatty acid type on the formation of lipid oxidation products during thermal processing', *Frontiers in Nutrition* **9**, 913297.

APPENDIX A: SUPPLEMENTARY INFORMATION FOR CHAPTER 3

Appendix A1

Environmental Var.	<i>f</i> Statistic	<i>p</i> val. (<i>f</i>)	<i>f</i> Analysis	<i>f</i> Significance	<i>t</i> Statistic	<i>p</i> val. (<i>t</i>)	<i>t</i> Analysis	<i>t</i> Significance
MAAT	4.574	0.032	Fligner Kileen	sig	280	0.007	Wilcoxon	sig
<i>T</i>_{JJA}	3.957	0.047	Fligner Kileen	sig	206	<0.001	Wilcoxon	sig
<i>T</i>_{SON}	1.897	0.168	Fligner Kileen	not sig	249	0.002	Wilcoxon	sig
<i>T</i>_{DJF}	5.784	0.016	Fligner Kileen	sig	327	0.045	Wilcoxon	sig
<i>T</i>_{MAM}	7.248	0.007	Fligner Kileen	sig	273	0.006	Wilcoxon	sig
MAP	20.612	<0.001	Fligner Kileen	sig	900	<0.001	Wilcoxon	sig
VPD	5.008	0.025	Fligner Kileen	sig	46	<0.001	Wilcoxon	sig
Soil pH	0.028	0.867	Fligner Kileen	not sig	47	<0.001	Wilcoxon	sig
<i>T</i>_{Soil}	2.961	0.085	Fligner Kileen	not sig	195	<0.001	Wilcoxon	sig
SMO	4.439	0.035	Fligner Kileen	sig	750	<0.001	Wilcoxon	sig

Table A1: Statistics for comparison of mean and variance of environmental variables between EC and SW soil transects

Appendix A2

3-OH FA Structures	<i>f</i> Statistic	<i>p</i> val. (<i>f</i>)	<i>f</i> Analysis	<i>f</i> Significance	<i>t</i> Statistic	<i>p</i> val. (<i>t</i>)	<i>df</i>	<i>t</i> Analysis	<i>t</i> Significance
<i>anteiso-</i>	0.804	0.544	<i>f</i> test	not sig	-1.398	0.167	60	<i>t</i> test (equal var.)	not sig
<i>branched</i>	0.325	0.002	<i>f</i> test	sig	-7.619	<0.001	37	<i>t</i> test (non equal var.)	sig
<i>iso-</i>	0.254	<0.001	<i>f</i> test	sig	-8.579	<0.001	34	<i>t</i> test (non equal var.)	sig
<i>normal-</i>	0.335	0.003	<i>f</i> test	sig	7.286	<0.001	37	<i>t</i> test (non equal var.)	sig

Table A2: Statistics for comparison of 3-OH FA structural group fractional abundance distributions between EC and SW soil transects

Appendix A3

3-OH FAs	<i>f</i> Statistic	<i>p</i> val. (<i>f</i>)	<i>f</i> Analysis	<i>f</i> Significance	<i>t</i> Statistic	<i>p</i> val. (<i>t</i>)	<i>t</i> Analysis	<i>t</i> Significance
<i>nC</i> ₁₀	12.592	<0.001	Fligner Kileen	sig	751	<0.001	Wilcoxon	sig
<i>iC</i> ₁₁	5.041	0.025	Fligner Kileen	sig	716	<0.001	Wilcoxon	sig
<i>aC</i> ₁₁	24.813	<0.001	Fligner Kileen	sig	722	<0.001	Wilcoxon	sig
<i>nC</i> ₁₁	4.046	0.044	Fligner Kileen	sig	752	<0.001	Wilcoxon	sig
<i>iC</i> ₁₂	0.227	0.634	Fligner Kileen	not sig	441	0.648	Wilcoxon	not sig
<i>nC</i> ₁₂	2.793	0.095	Fligner Kileen	not sig	622	0.029	Wilcoxon	sig
<i>iC</i> ₁₃	2.229	0.135	Fligner Kileen	not sig	158	<0.001	Wilcoxon	sig
<i>aC</i> ₁₃	0.357	0.550	Fligner Kileen	not sig	355	0.108	Wilcoxon	not sig
<i>nC</i> ₁₃	2.652	0.103	Fligner Kileen	not sig	588	0.088	Wilcoxon	not sig
<i>iC</i> ₁₄	0.007	0.931	Fligner Kileen	not sig	464	0.960	Wilcoxon	not sig
<i>nC</i> ₁₄	0.039	0.843	Fligner Kileen	not sig	676	0.003	Wilcoxon	sig
<i>iC</i> ₁₅	0.518	0.472	Fligner Kileen	not sig	484	0.825	Wilcoxon	not sig
<i>aC</i> ₁₅	0.118	0.731	Fligner Kileen	not sig	408	0.396	Wilcoxon	not sig
<i>nC</i> ₁₅	5.770	0.016	Fligner Kileen	sig	415	0.454	Wilcoxon	not sig
<i>iC</i> ₁₆	17.820	<0.001	Fligner Kileen	sig	81	<0.001	Wilcoxon	sig
<i>nC</i> ₁₆	6.811	0.009	Fligner Kileen	sig	337	0.063	Wilcoxon	not sig
<i>iC</i> ₁₇	18.253	<0.001	Fligner Kileen	sig	43	<0.001	Wilcoxon	sig
<i>aC</i> ₁₇	3.979	0.046	Fligner Kileen	sig	370	0.164	Wilcoxon	not sig
<i>nC</i> ₁₇	6.716	0.010	Fligner Kileen	sig	199	<0.001	Wilcoxon	sig
<i>iC</i> ₁₈	0.007	0.933	Fligner Kileen	not sig	502	0.541	Wilcoxon	not sig
<i>nC</i> ₁₈	1.323	0.250	Fligner Kileen	not sig	647	0.011	Wilcoxon	sig

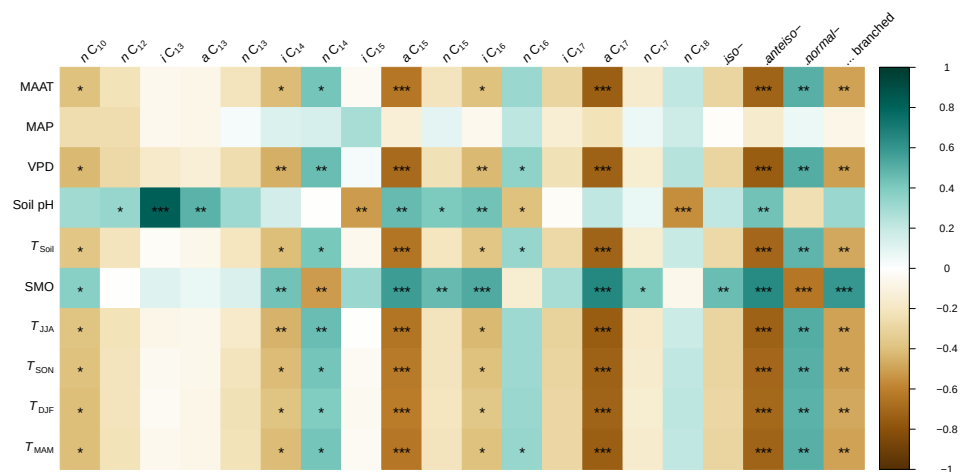
Table A3: Statistics for comparison of 3-OH FA fractional abundance distributions between EC and SW soil transects

Appendix A4

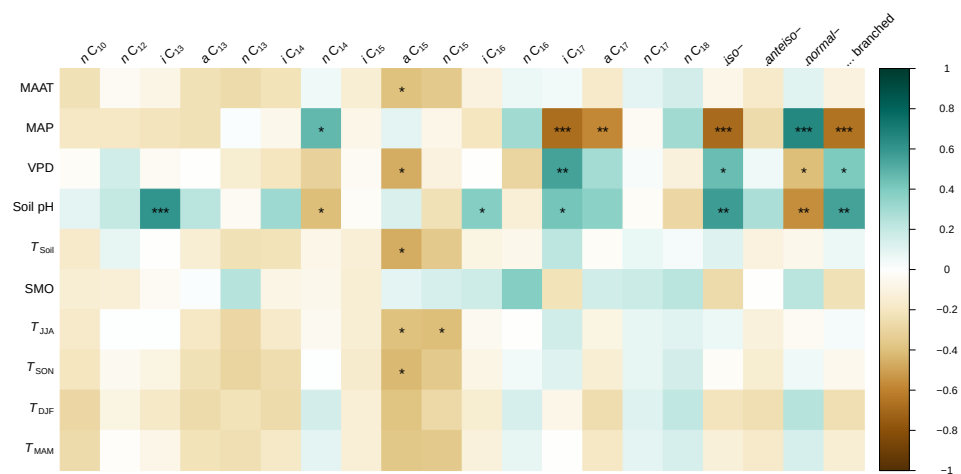
3-OH FAs	r^2	p val.
<i>nC</i> ₁₀	0.30	0.01
<i>iC</i> ₁₁	0.15	0.01
<i>aC</i> ₁₁	0.08	0.09
<i>uC</i> ₁₁	0.04	0.17
<i>nC</i> ₁₁	0.32	<0.001
<i>iC</i> ₁₂	0.49	<0.001
<i>uC</i> ₁₂	0.01	0.41
<i>nC</i> ₁₂	0.32	<0.001
<i>iC</i> ₁₃	0.51	<0.001
<i>aC</i> ₁₃	0.34	<0.001
<i>uC</i> ₁₃	0.04	0.13
<i>nC</i> ₁₃	0.10	0.08
<i>iC</i> ₁₄	0.19	0.02
<i>uC</i> ₁₄	0.05	0.11
<i>nC</i> ₁₄	0.25	<0.001
<i>iC</i> ₁₅	0.32	<0.001
<i>aC</i> ₁₅	0.21	<0.001
<i>nC</i> ₁₅	0.04	0.29
<i>iC</i> ₁₆	0.12	0.04
<i>uC</i> ₁₆	0.17	<0.001
<i>nC</i> ₁₆	0.43	<0.001
<i>iC</i> ₁₇	0.14	0.01
<i>aC</i> ₁₇	0.20	0.03
<i>nC</i> ₁₇	0.73	<0.001
<i>iC</i> ₁₈	<0.001	0.90
<i>uC</i> ₁₈	0.03	0.30
<i>nC</i> ₁₈	0.59	<0.001

Table A4: Permutation test results for NMDS of 3-OH FA distributions amongst U.S. soils

Appendix A5



(a) East Coast soils



(b) SW U.S. soils

Figure A5: Pearson's correlation matrix between the fractional abundance of 3-OH FA compounds and environmental variables grouped by U.S. soil transect

Appendix A6

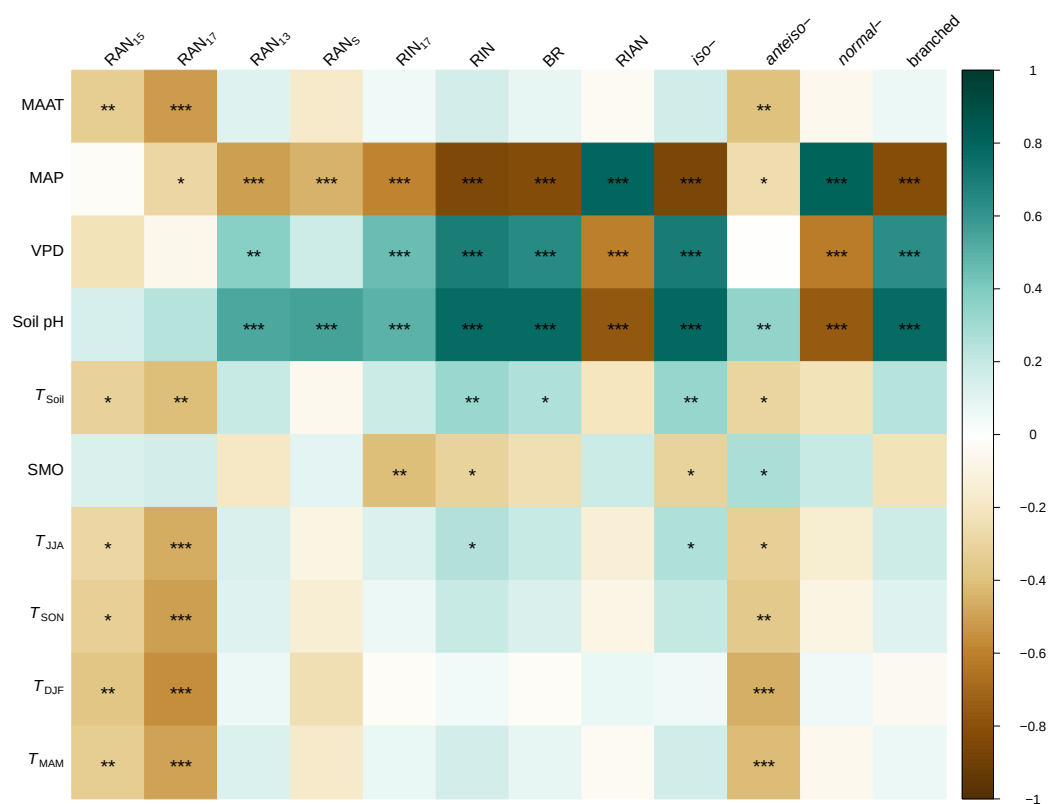
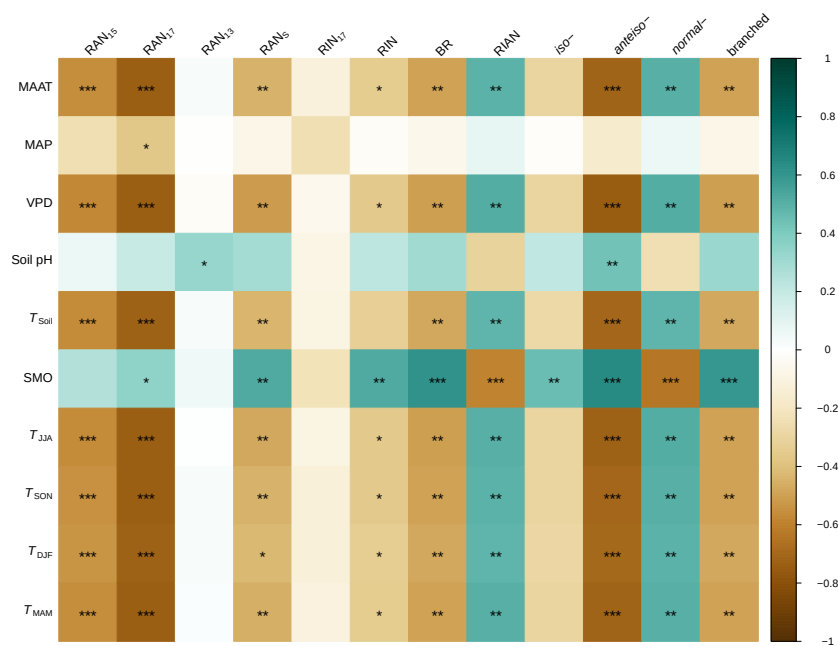
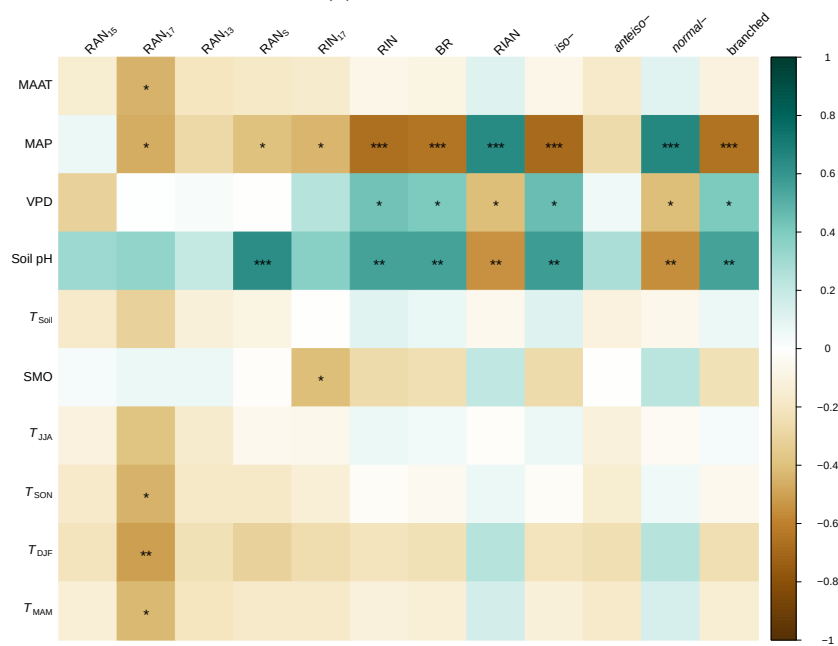


Figure A6: Pearson's correlation matrix between established 3-OH FA-based environmental proxies from U.S. soils and environmental variables.

Appendix A7



(a) East Coast soils



(b) SW U.S. soils

Figure A7: Pearson's correlation matrix between 3-OH FA-based proxies and environmental variables grouped by U.S. soil transect

Appendix A8

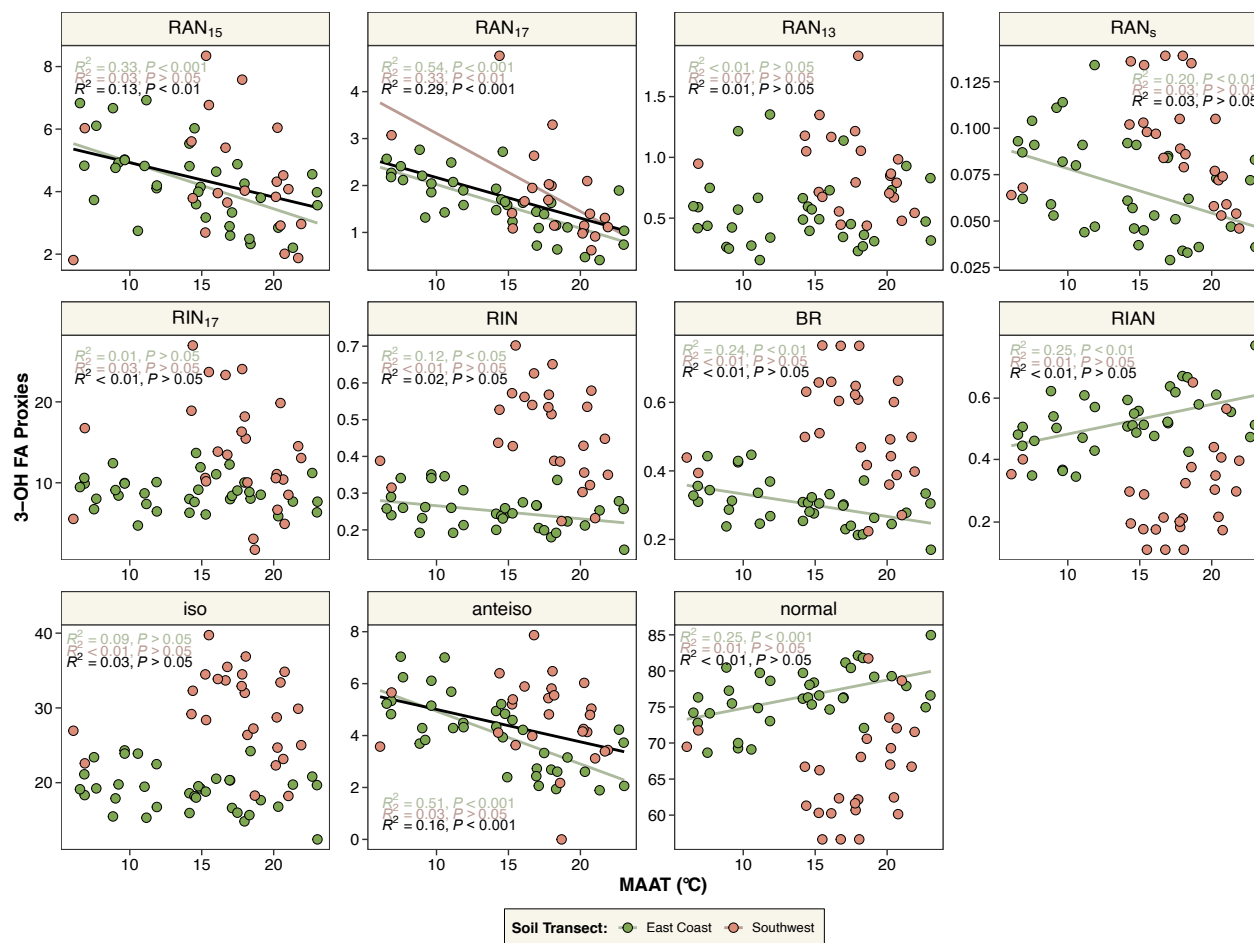


Figure A8: Crossplot matrix of established 3-OH FA-based temperature and pH proxies and **MAAT** in U.S. soils. Black regression lines and r^2 summarise all U.S. soils compiled. Regression lines displayed for significant correlations.

Appendix A9

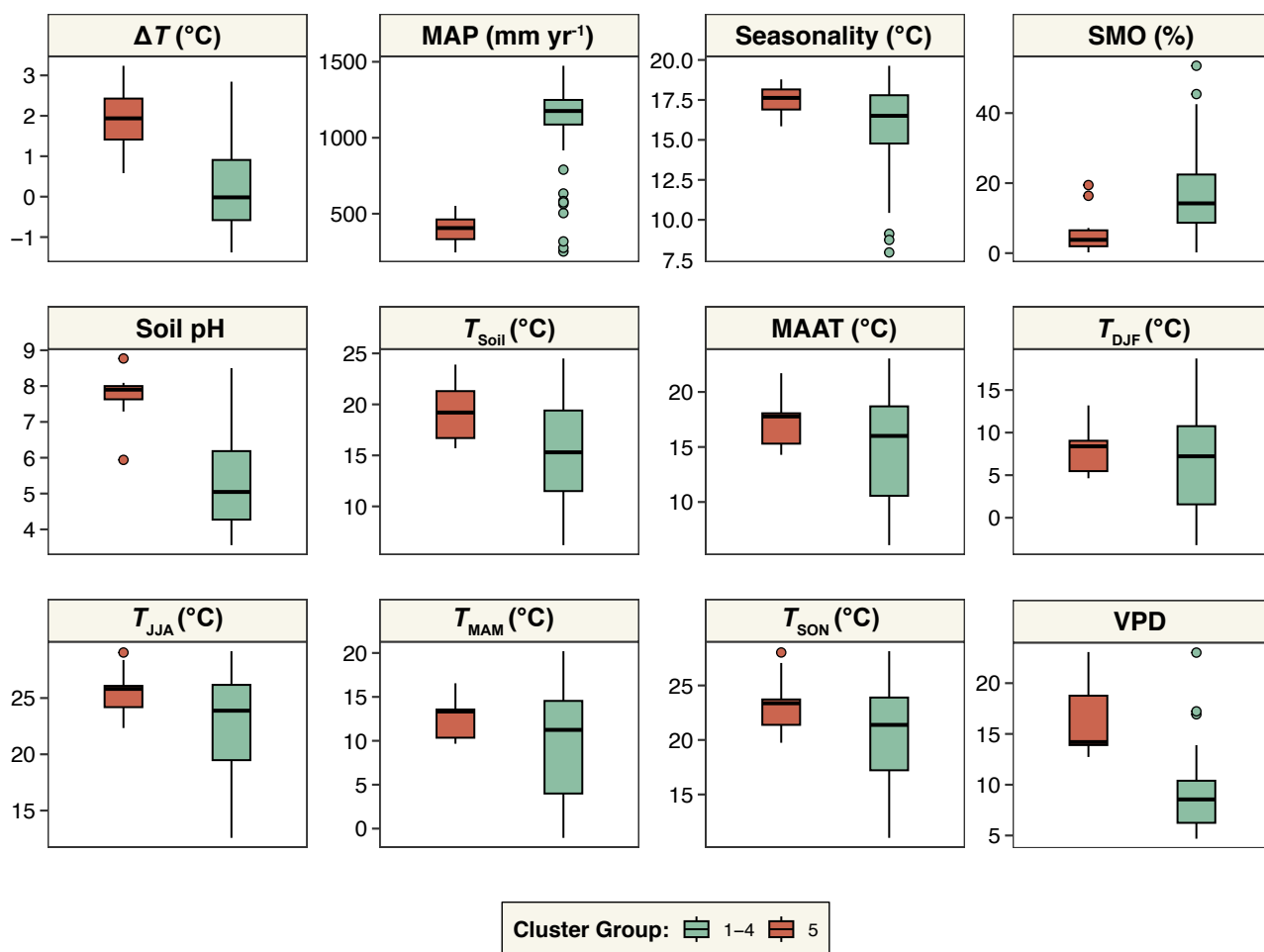


Figure A9: Boxplots visualising distributions in environmental variables between cluster groups 1-4 with cluster group 5 (after Fig.3.17) for U.S. soils

Appendix A10

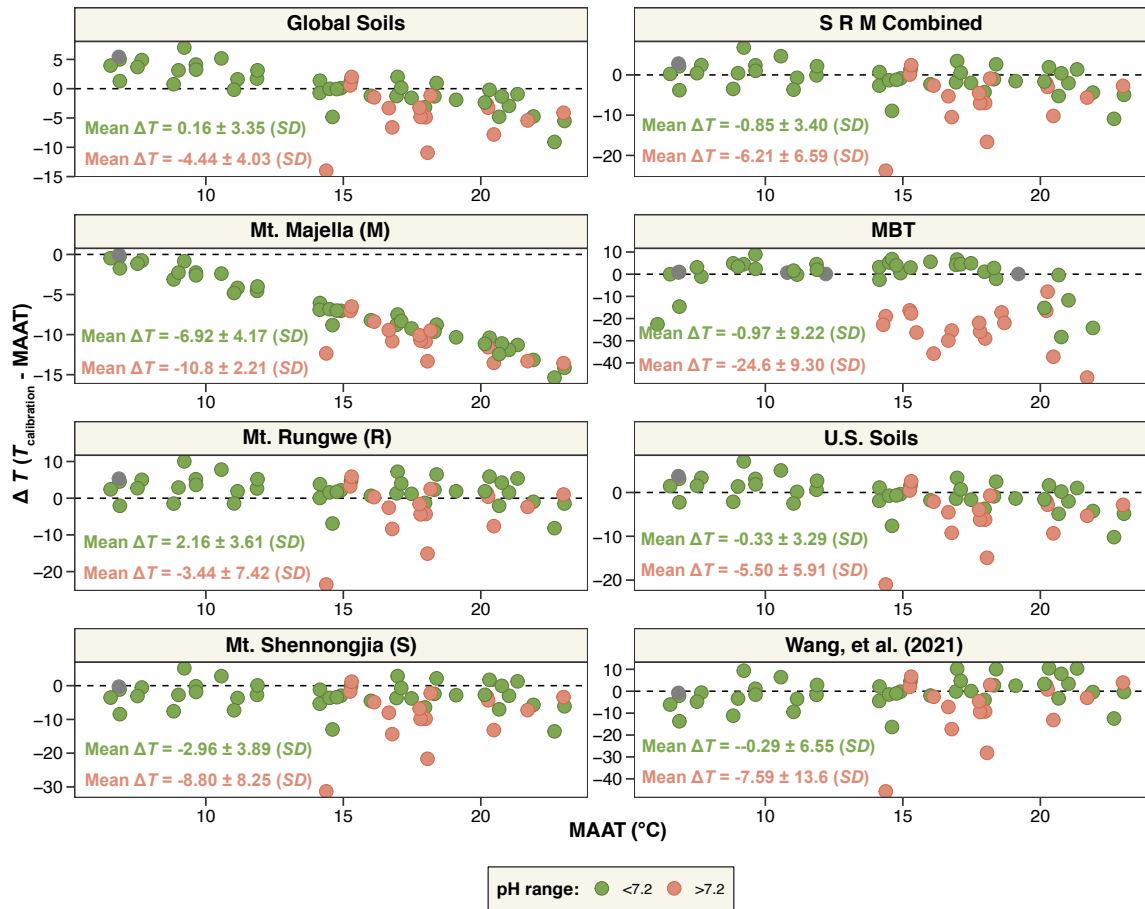


Figure A10: Crossplots between contemporary MAAT and ΔT for various RAN₁₇ (and one MBT) calibrations for U.S. soils grouped by a pH 7.2 threshold. Calibrations refer to the following equations: 'Global Soils' Eq.3.5, 'SRM Combined' Eq.3.9, 'Mt. Majella' Eq.3.7, 'MBT' after Peterse et al. (2012), 'Mt. Rungwe' Eq.3.8, 'U.S. Soils' Eq.3.4, 'Mt. Shennongjia' Eq.3.6, 'Wang, et al. (2021)' Eq.3.10.

Appendix A11

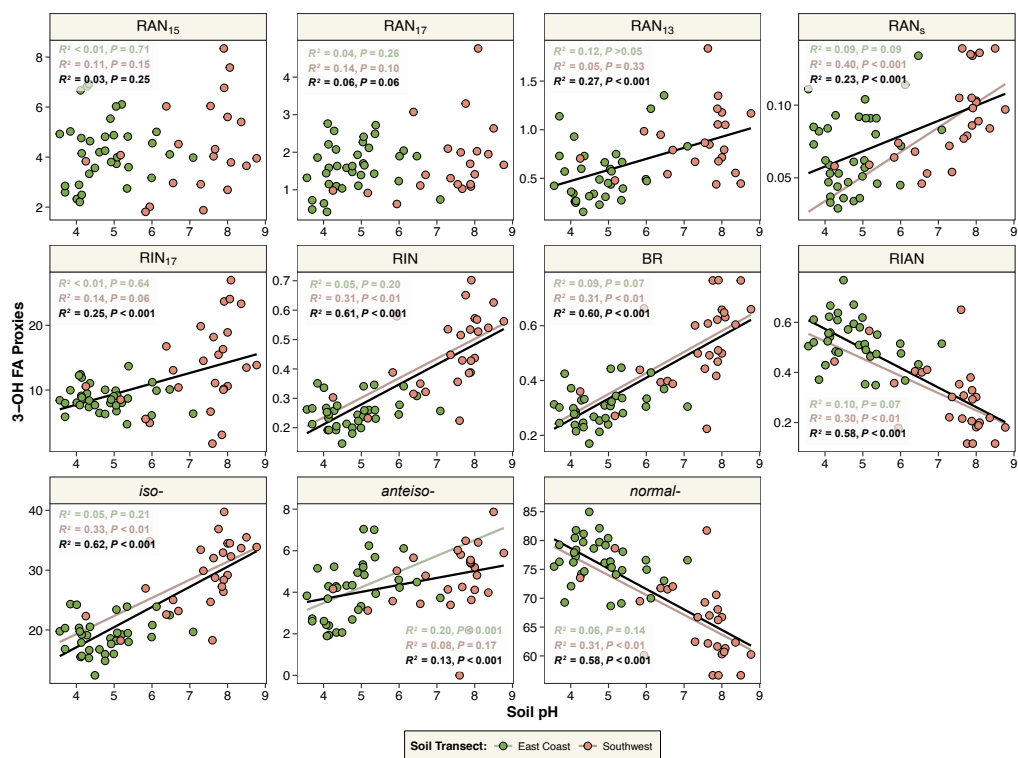


Figure A11: Crossplot matrix of established 3-OH FA-based temperature and pH proxies and soil pH in U.S. soils. Black regression lines and r^2 summarise all U.S. soils compiled. Regression lines displayed for significant correlations.

APPENDIX B: SUPPLIMENTARY INFORMATION FOR CHAPTER 4

Appendix B1

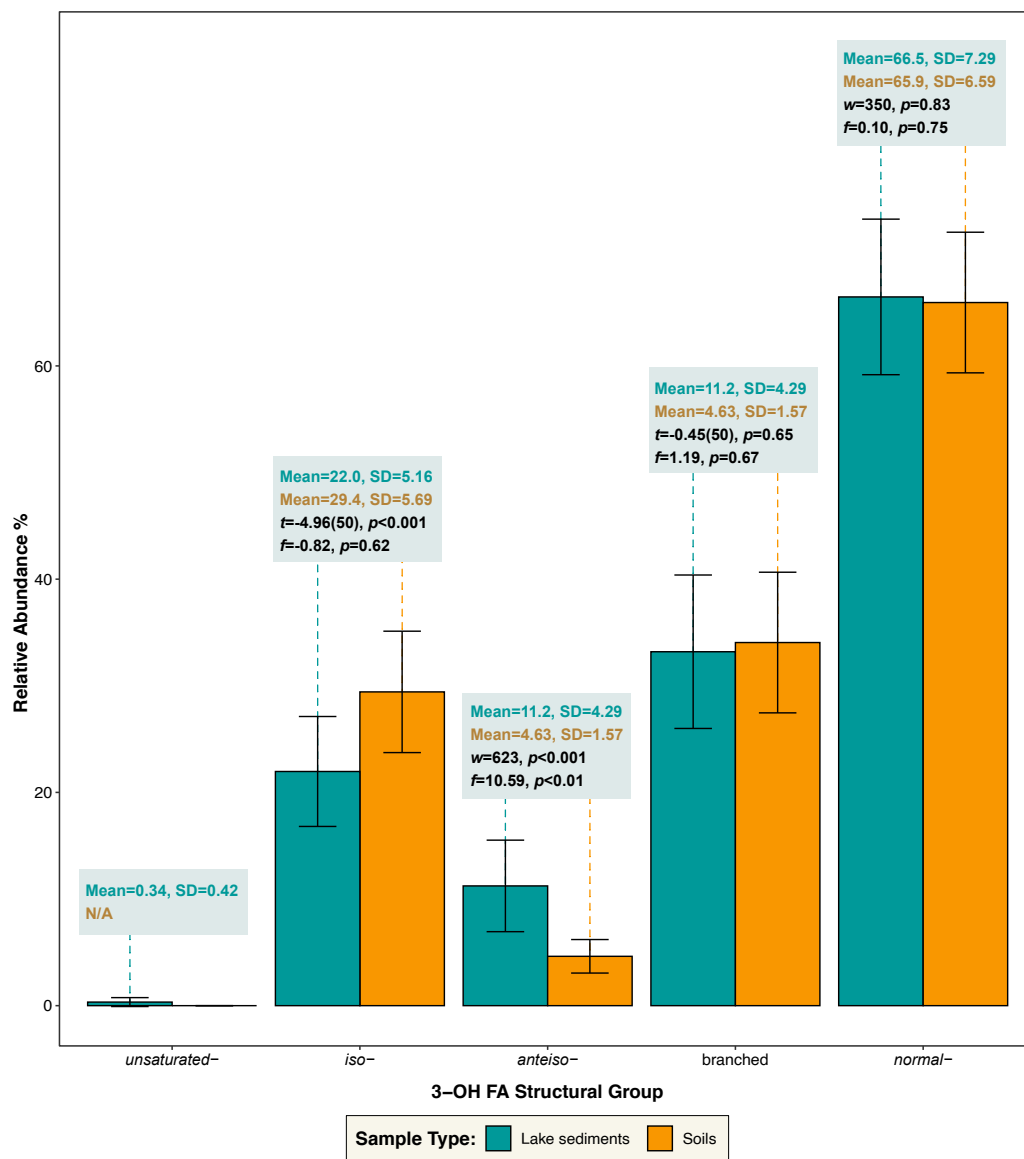


Figure B1: Barplots visualising difference in mean fractional abundance of 3-OH FA structural groups and descriptive statistics between SW U.S. lake sediments and soils.

Appendix B2

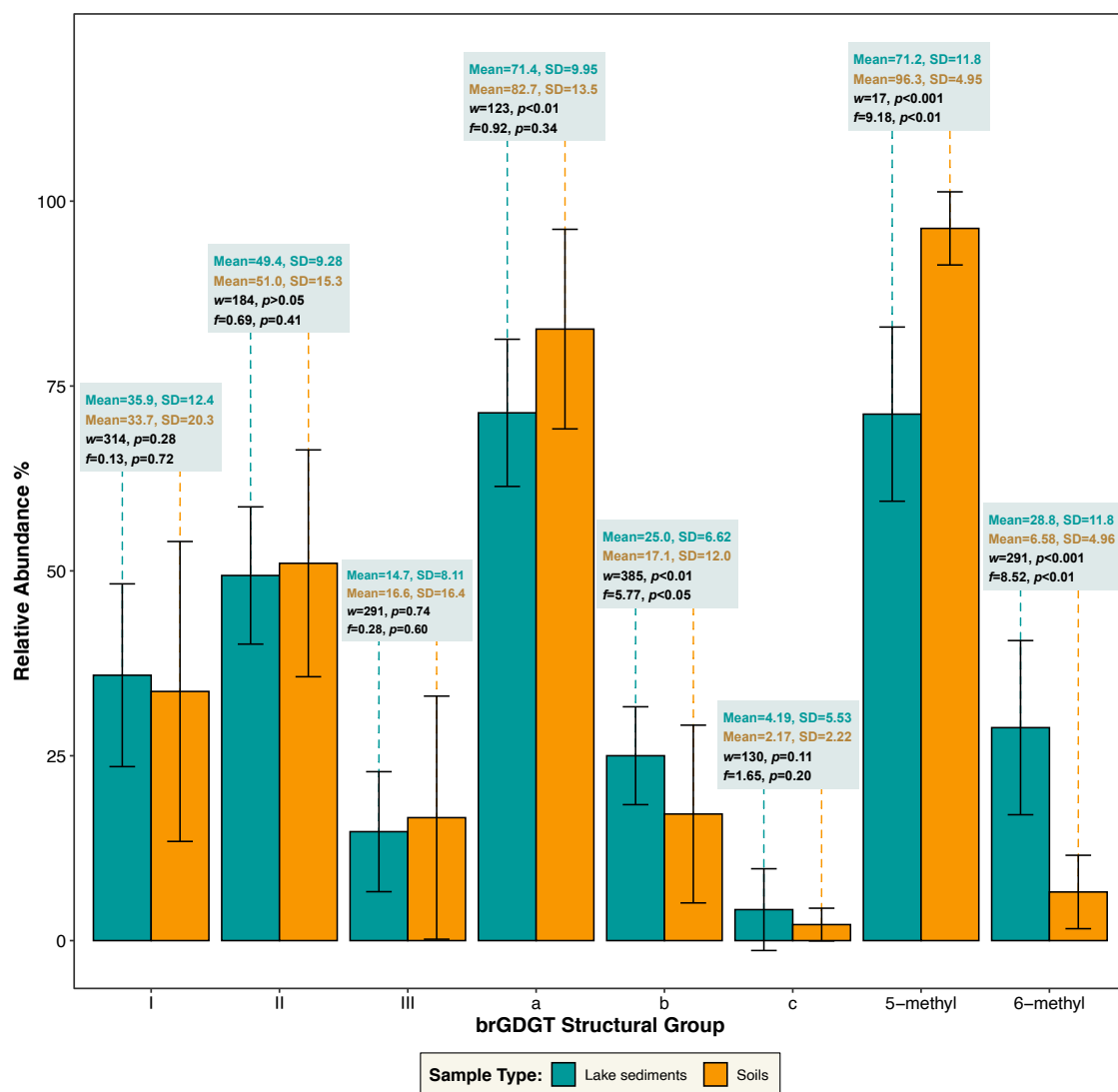


Figure B2: Barplots visualising difference in mean fractional abundance of brGDGT structural groups and descriptive statistics between SW U.S lake sediments and soils.

Appendix B3

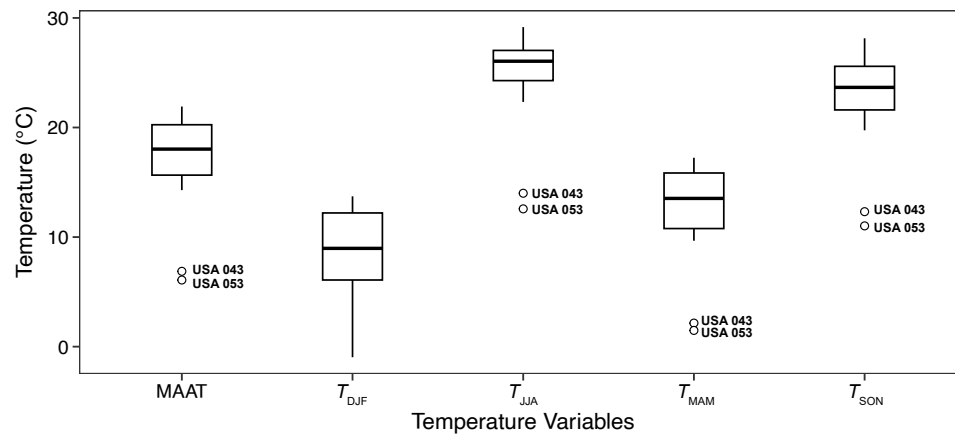
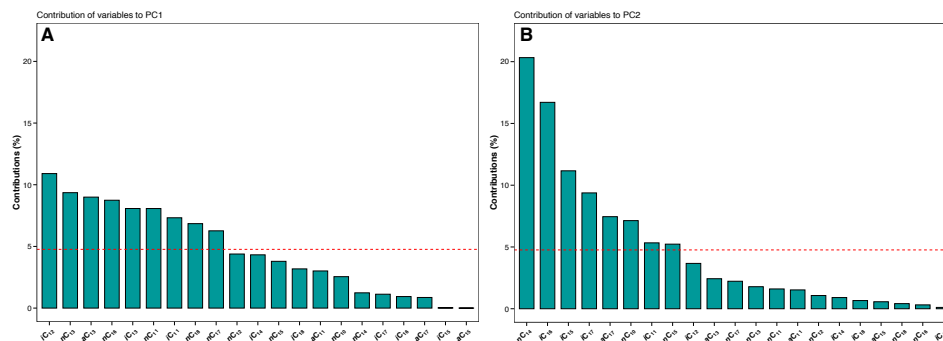


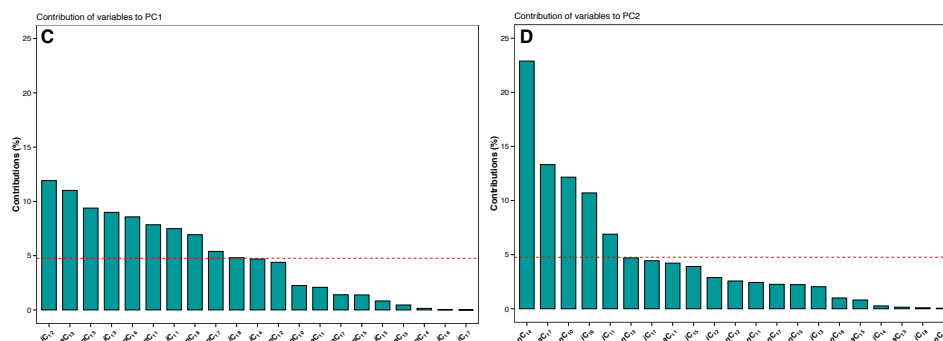
Figure B3: Boxplots showing distribution of air temperature variables along the SW U.S. lake transect.

Appendix B4

Lake sediments:



Lake sediments with cold sites omitted:



Soils with cold sites omitted:

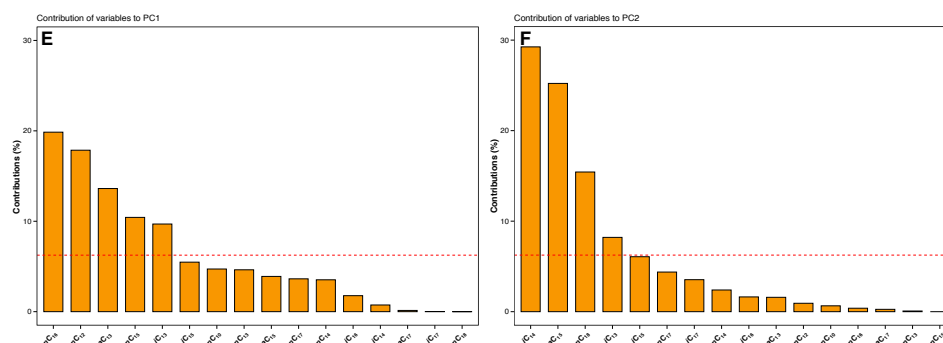
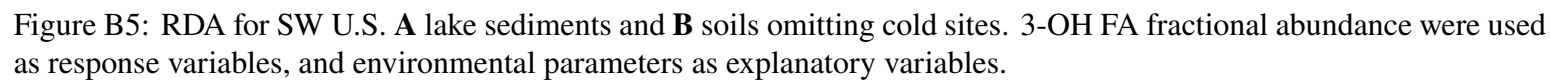


Figure B4: Contribution of 3-OH FA compounds to PC1 and PC2 in **A & B** all U.S. lake sediments (Fig.4.9), **C & D** U.S. lake sediments with cold lakes omitted (Fig.4.10), **E & F** U.S. soils with cold lakes omitted (Fig.4.10)



Appendix B6

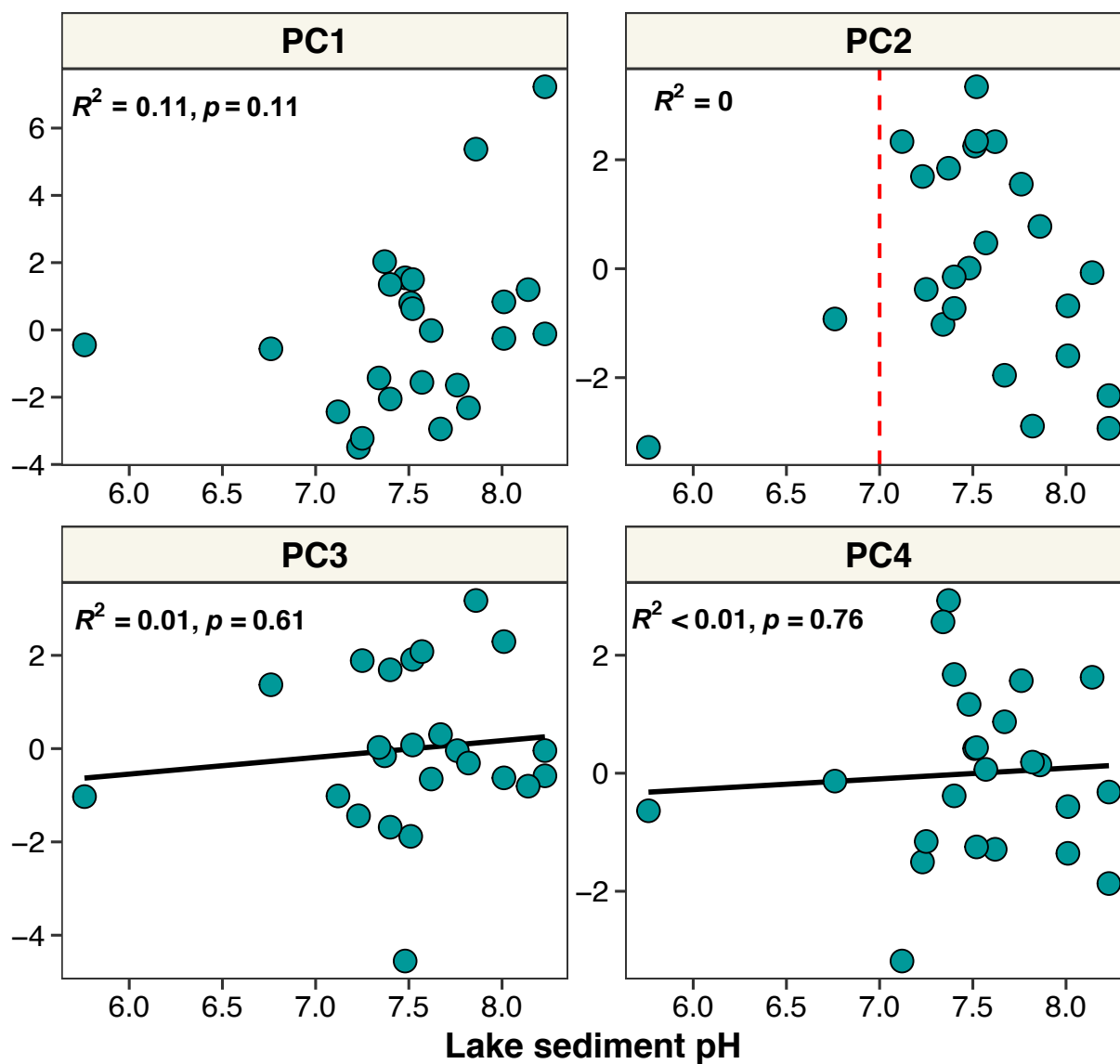


Figure B6: Crossplots between PC1 to 4 from PCA and lake sediment pH from PCA of SW lake sediments with cold sites omitted.

Appendix B7

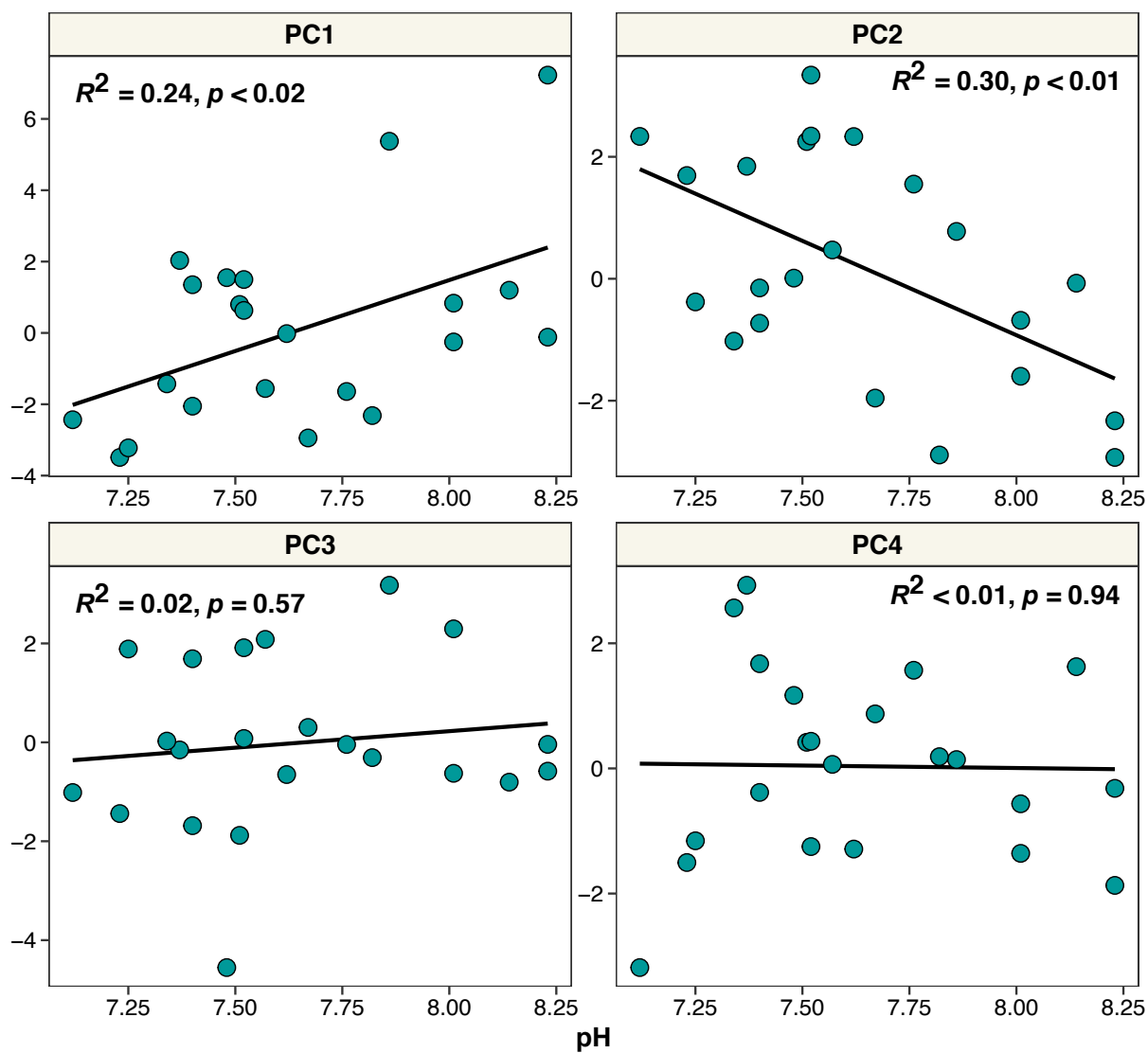


Figure B7: Crossplots between PC1 to 4 from PCA and lake sediment pH omitting soils with $pH < 7$ for SW lake sediments with cold sites omitted.

Appendix B8

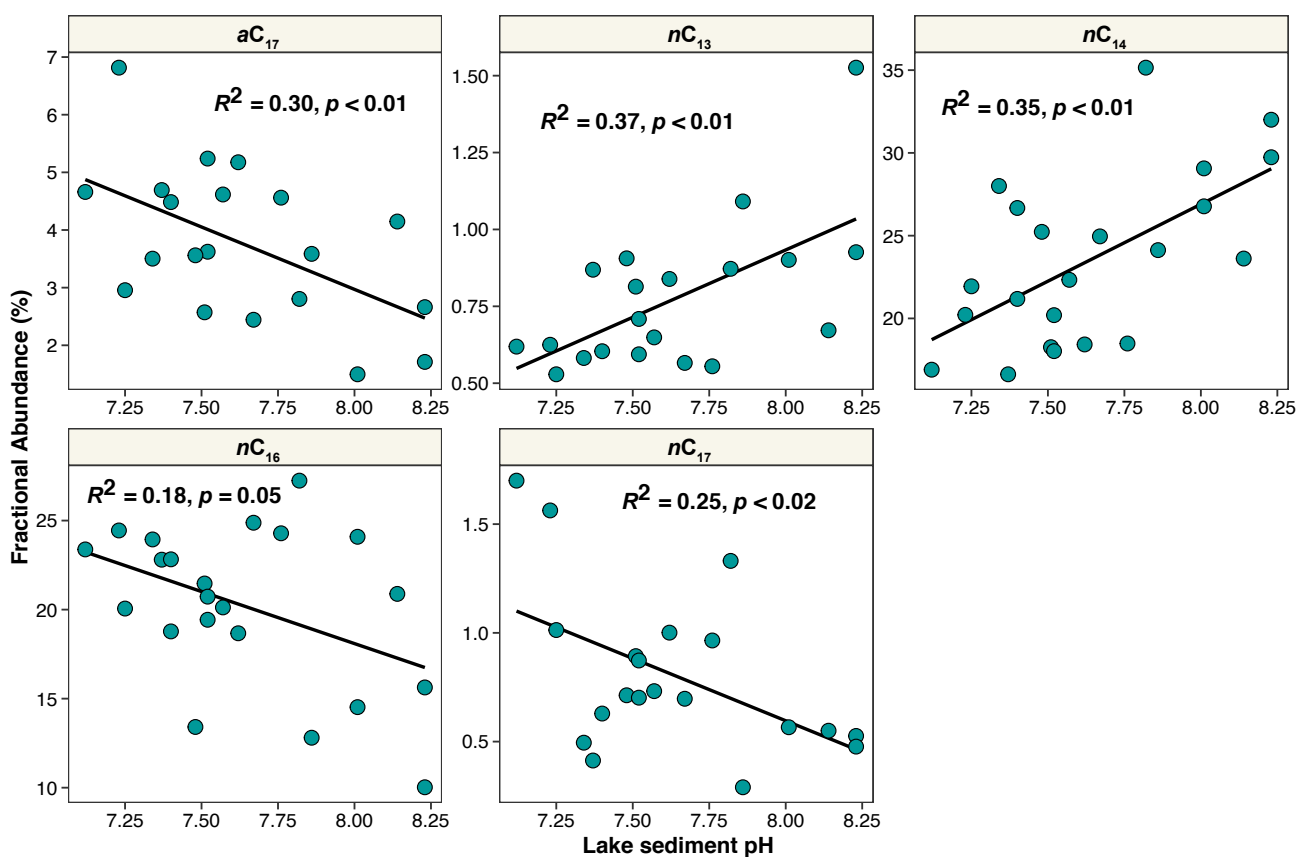


Figure B8: Crossplots between selected 3-OH FA compounds associated with PC1 and PC2 in SW lake sediments and lake sediment pH for SW lake sediments with cold sites omitted.

Appendix B9

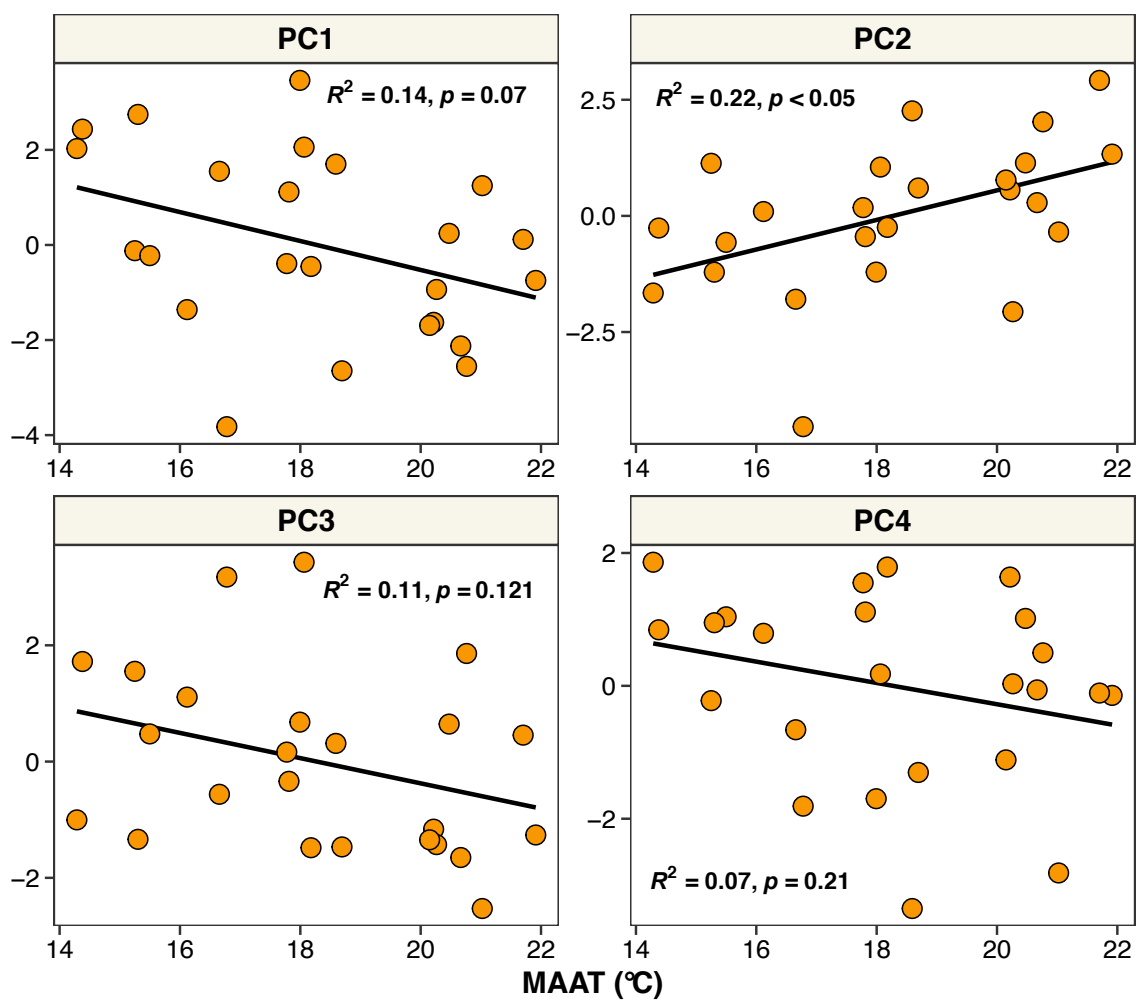


Figure B9: Crossplots between PC1 to 4 from PCA of soils with cold sites omitted vs [MAAT](#).

Appendix B10

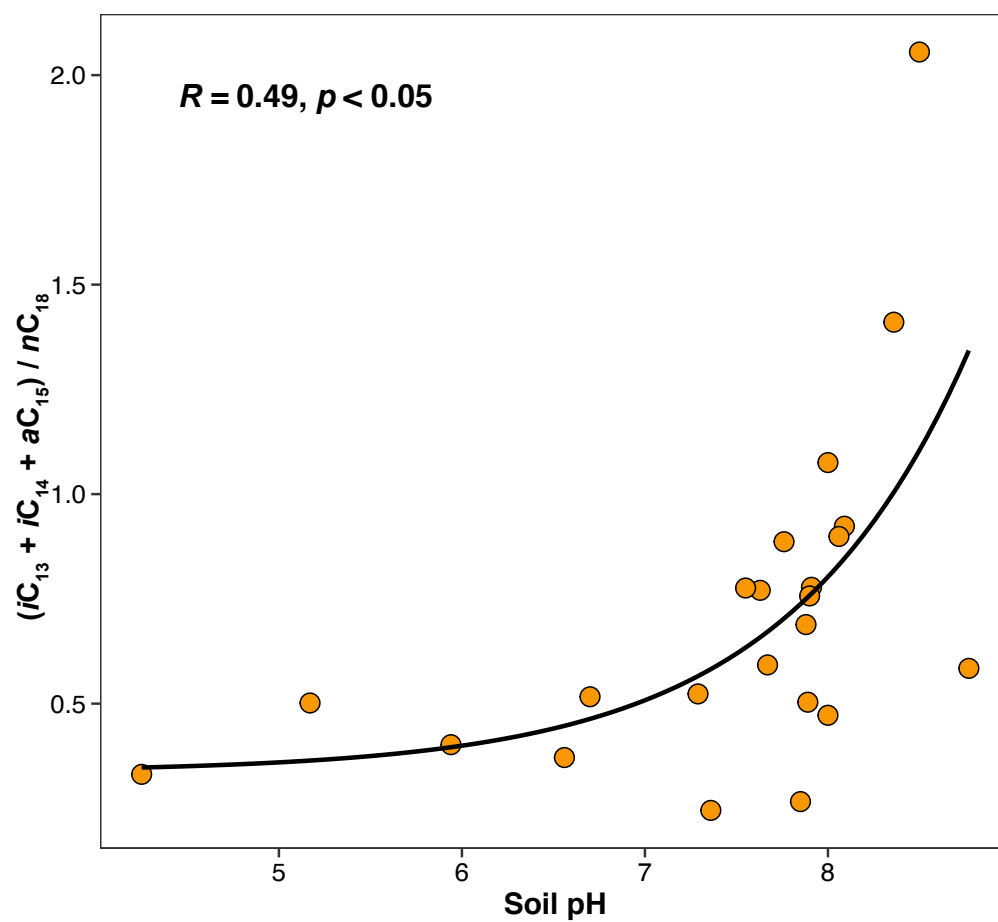


Figure B10: Crossplot visualising relationship between soil pH and $[(i-C_{13} + i-C_{14} + a-C_{15}) / n-C_{18}]$ in SW U.S. soils.

Appendix B11

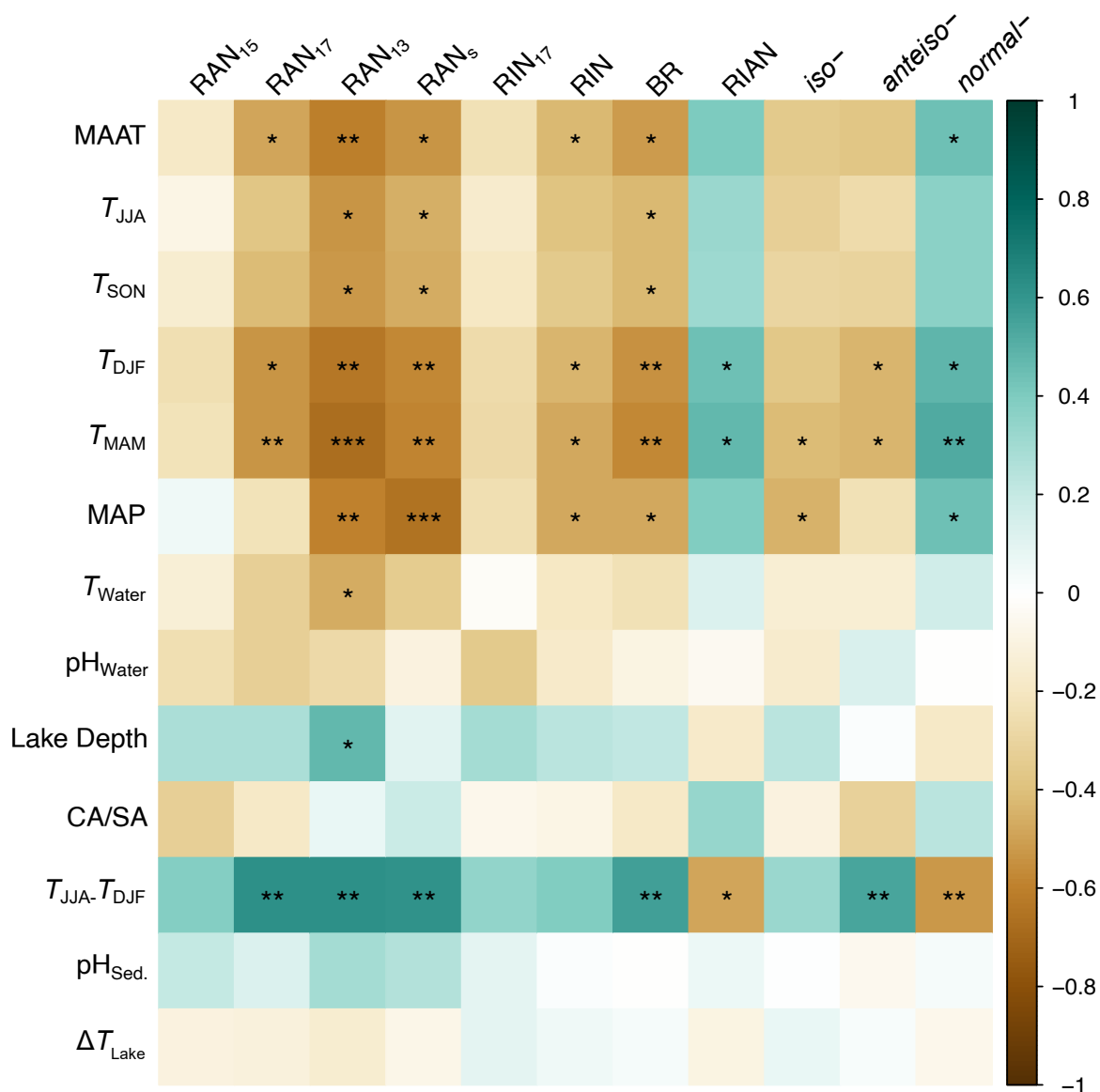


Figure B11: Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW lakes with cold sites omitted. Significance is indicated by asterisks ($p \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)

Appendix B12

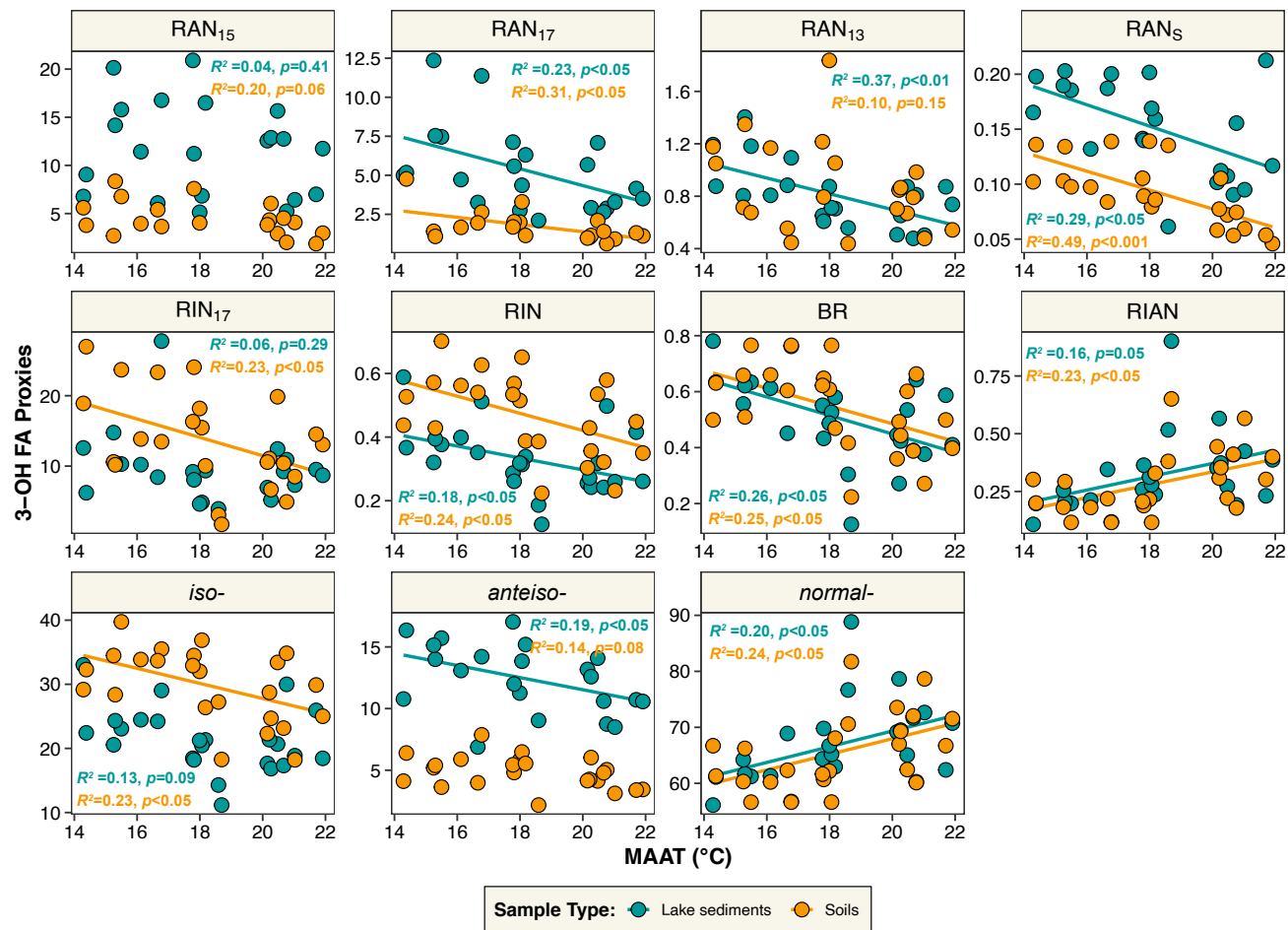


Figure B12: Scatterplots between established 3-OH FA-based proxies and MAAT for SW U.S. lake sediments and soils with two cold lakes omitted.

Appendix B13

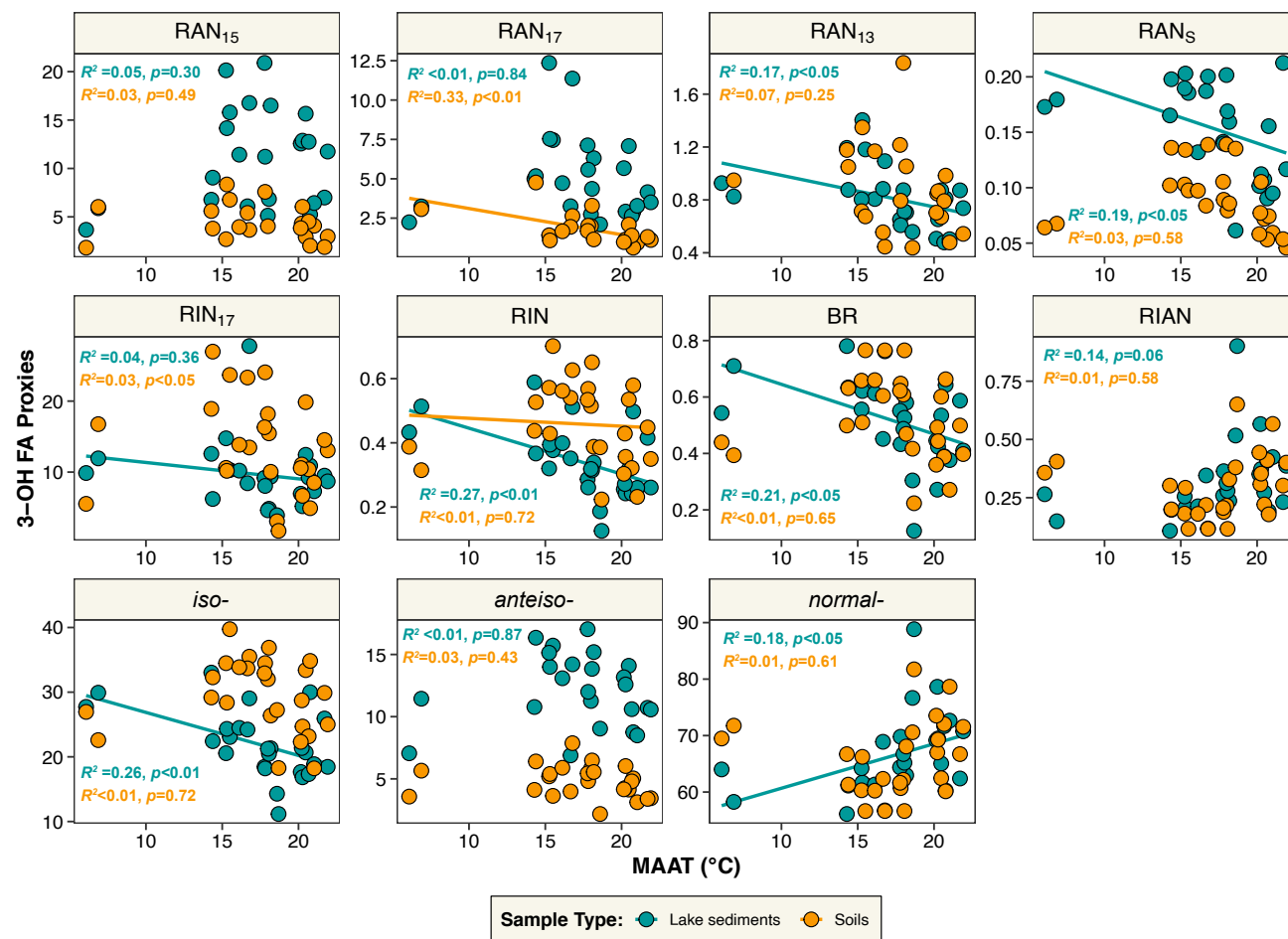


Figure B13: Scatterplots between established 3-OH FA-based proxies and MAAT for SW U.S. lake sediments and soils including two cold lakes.

Appendix B14

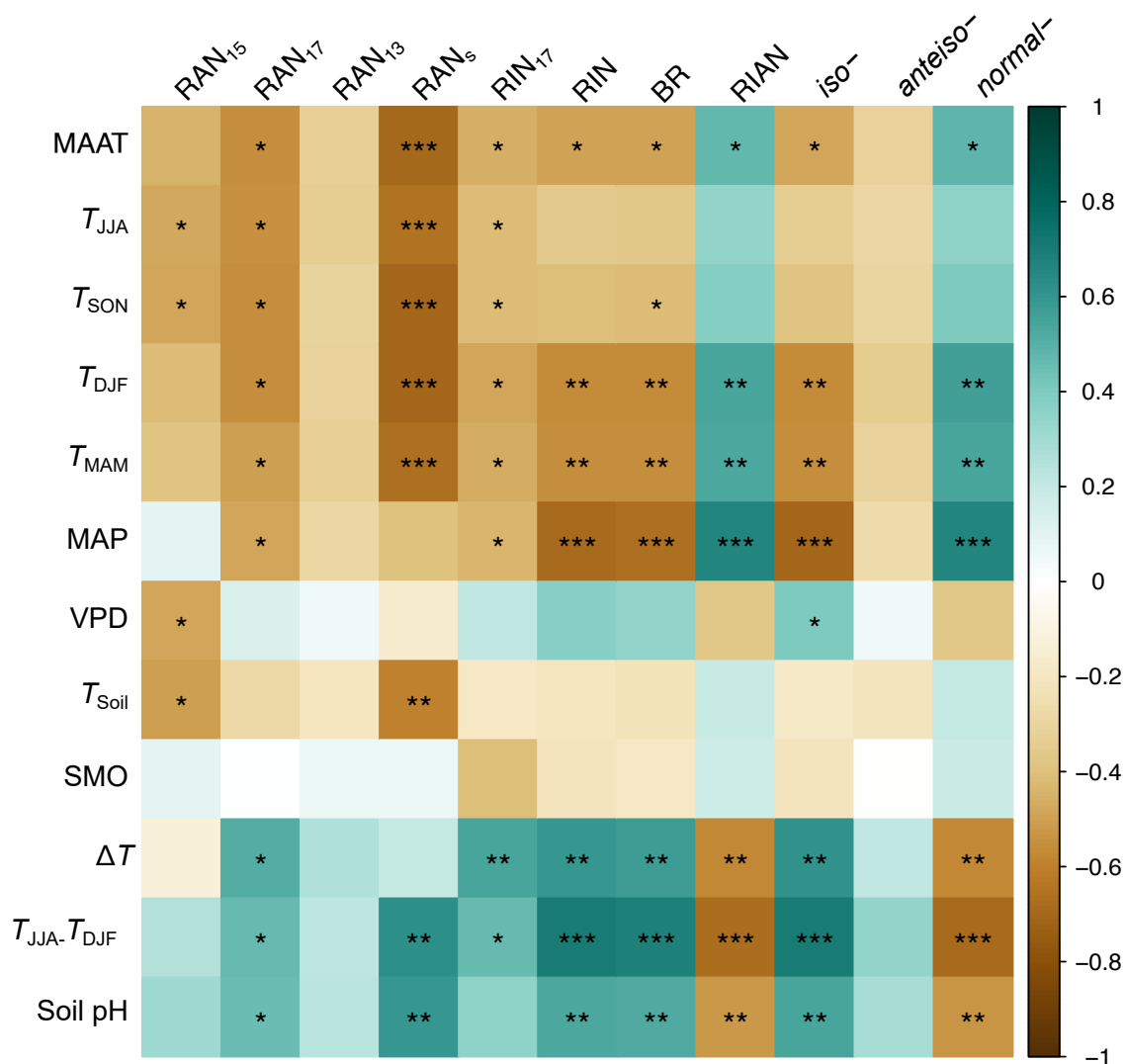


Figure B14: Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW soils with cold sites omitted. Significance is indicated by asterisks ($p: \leq 0.05 = *$, $\leq 0.01 = **$, $\leq 0.001 = ***$)

Appendix B15

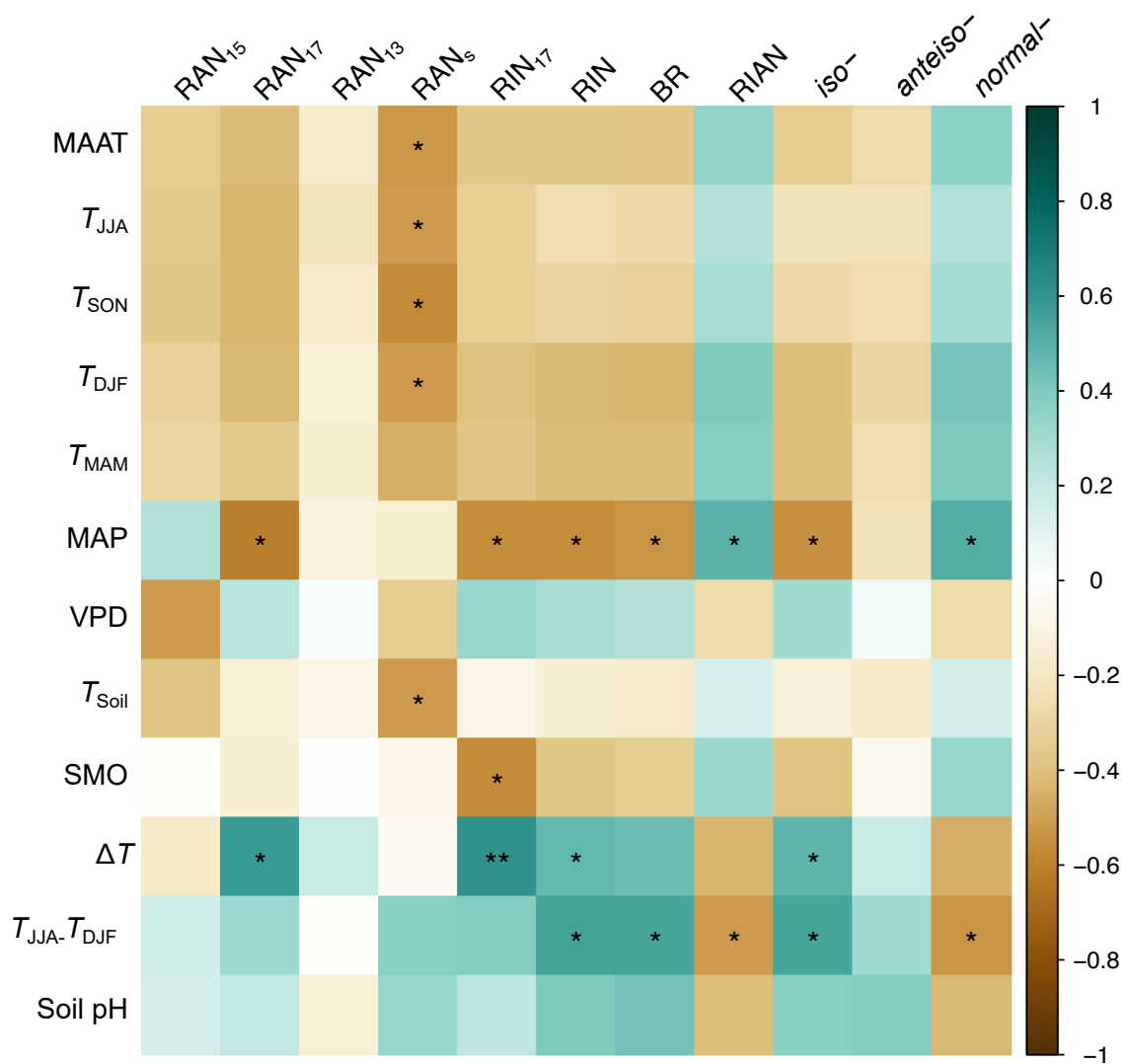


Figure B15: Pearson's correlation heatmap between 3-OH FA proxies and environmental variables in SW lakes with pH>7.2. Significance is indicated by asterisks ($p: \leq 0.5 = *$, $\leq 0.1 = **$, $\leq 0.001 = ***$)

APPENDIX C: SUPPLEMENTARY INFORMATION FOR CHAPTER 5

Appendix C1

280

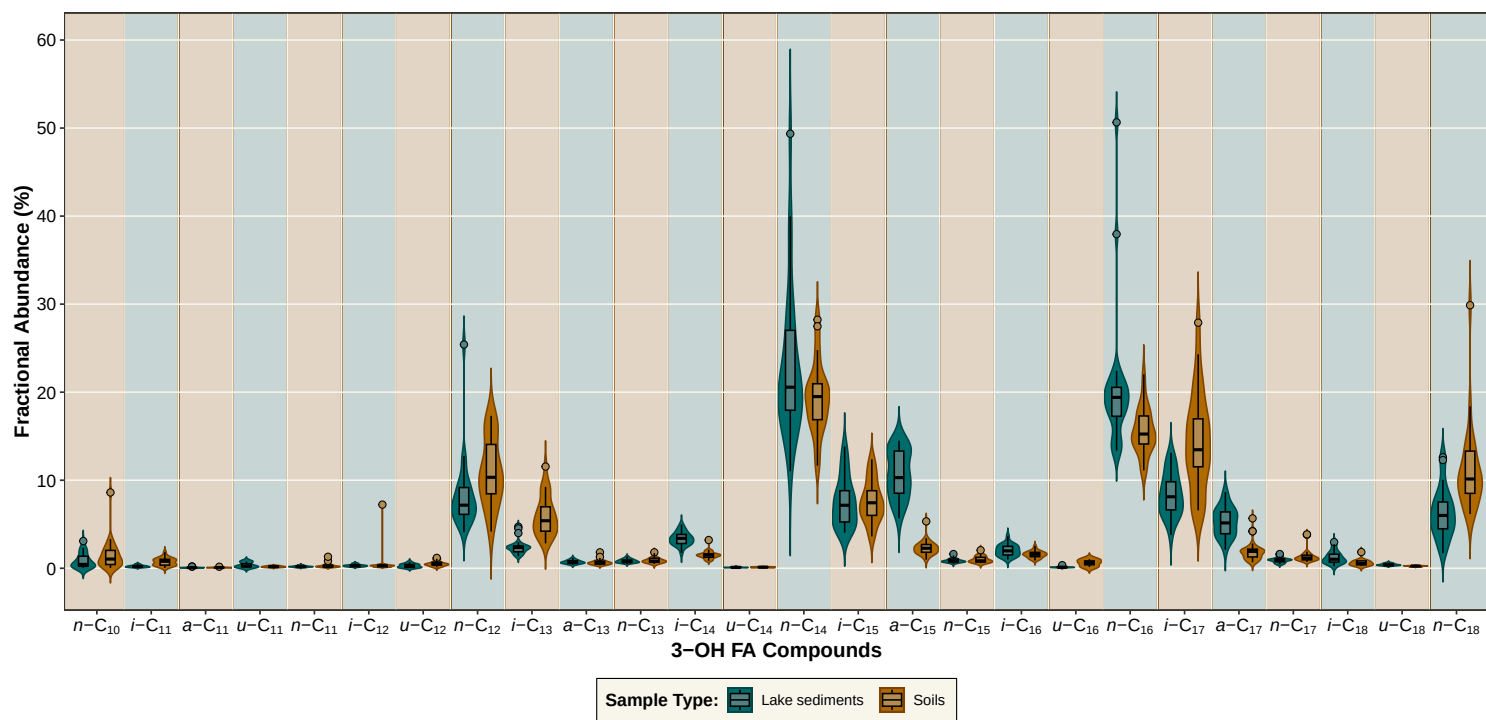


Figure C1: 3-OH FA distributions in SW U.S. lake sediments and soils extracted by Low n' Slow acid hydrolysis reported as % fractional abundance

Appendix C2

3-OH FAs	F stat.	F p value	F analysis	F significance	T stat.	T p value	DF	T analysis	T significance
nC10	1.617	0.203	Fligner Kileen	no	202	0.078	-	Wilcoxon	no
iC11	14.671	0.000	Fligner Kileen	yes	57	0.000	-	Wilcoxon	yes
aC11	0.003	0.954	Fligner Kileen	no	103	0.247	-	Wilcoxon	no
nC11	2.719	0.099	Fligner Kileen	no	175	0.174	-	Wilcoxon	no
iC12	0.195	0.659	Fligner Kileen	no	258	0.716	-	Wilcoxon	no
nC12	2.870	0.090	Fligner Kileen	no	146	0.004	-	Wilcoxon	yes
iC13	12.758	0.000	Fligner Kileen	yes	21	0.000	-	Wilcoxon	yes
aC13	0.660	0.417	Fligner Kileen	no	321	0.337	-	Wilcoxon	no
nC13	1.380	0.240	Fligner Kileen	no	247	0.409	-	Wilcoxon	no
iC14	5.416	0.020	Fligner Kileen	yes	561	0.000	-	Wilcoxon	yes
nC14	3.023	0.082	Fligner Kileen	no	390	0.135	-	Wilcoxon	no
aC15	14.534	0.000	Fligner Kileen	yes	575	0.000	-	Wilcoxon	yes
nC15	1.927	0.165	Fligner Kileen	no	249	0.427	-	Wilcoxon	no
nC16	1.227	0.268	Fligner Kileen	no	491	0.001	-	Wilcoxon	yes
aC17	4.686	0.030	Fligner Kileen	yes	495	0.000	-	Wilcoxon	yes
nC17	0.544	0.461	Fligner Kileen	no	150	0.008	-	Wilcoxon	yes
iC18	4.475	0.034	Fligner Kileen	yes	184	0.017	-	Wilcoxon	yes
nC18	0.028	0.866	Fligner Kileen	no	77	0.000	-	Wilcoxon	yes
iC15	1.415	0.408	f test	no	-0.113	0.911	46	t test, equal var.	no
iC16	2.713	0.019	f test	yes	3.014	0.004	46	t test, non equal var.	yes
iC17	0.193	0.000	f test	yes	-4.836	0.000	46	t test, non equal var.	yes

Table C2: Statistics for comparison of mean and variance of 3-OH FA fractional abundances between Low n' Slow extracted lake sediments and soils

Appendix C3

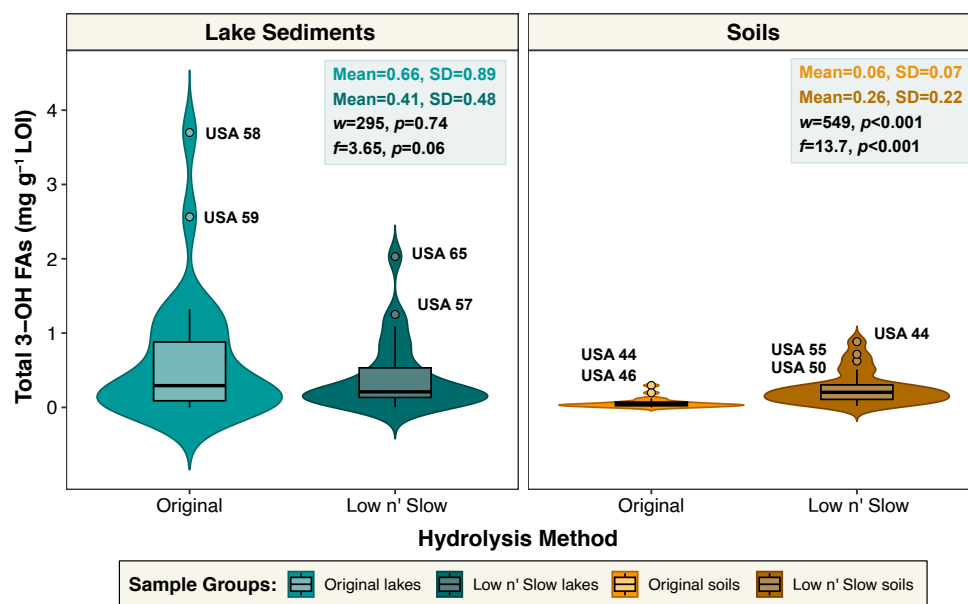


Figure C3: Comparison of 3-OH FA concentrations between SW lake sediments and soils extracted by Low n' Slow and original acid hydrolysis methods.

Appendix C4

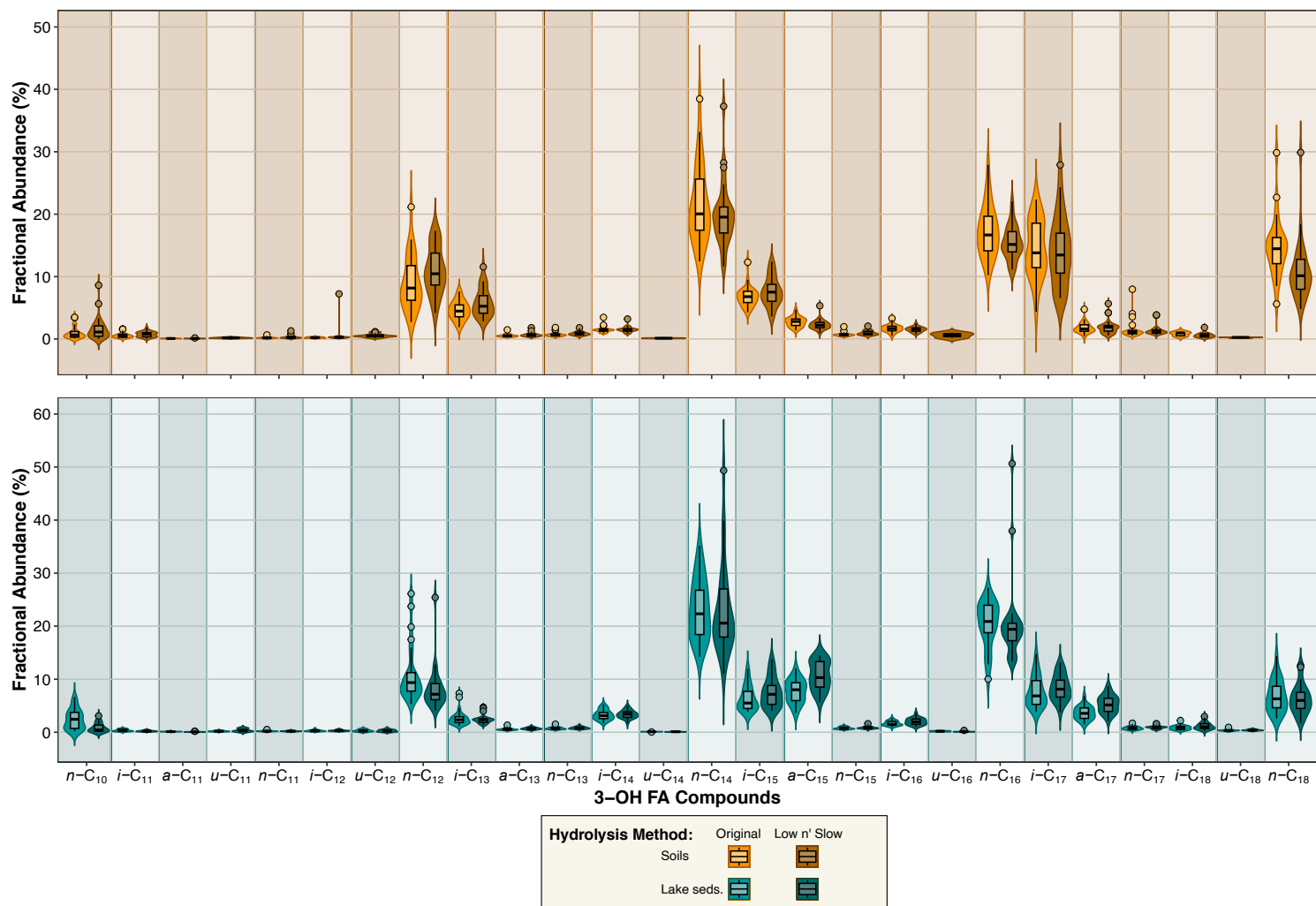


Figure C4: 3-OH FA distributions in SW U.S. soils (Upper) and lake sediments (Lower) extracted by Low n' Slow and original acid hydrolysis reported as % fractional abundance.

Appendix C5

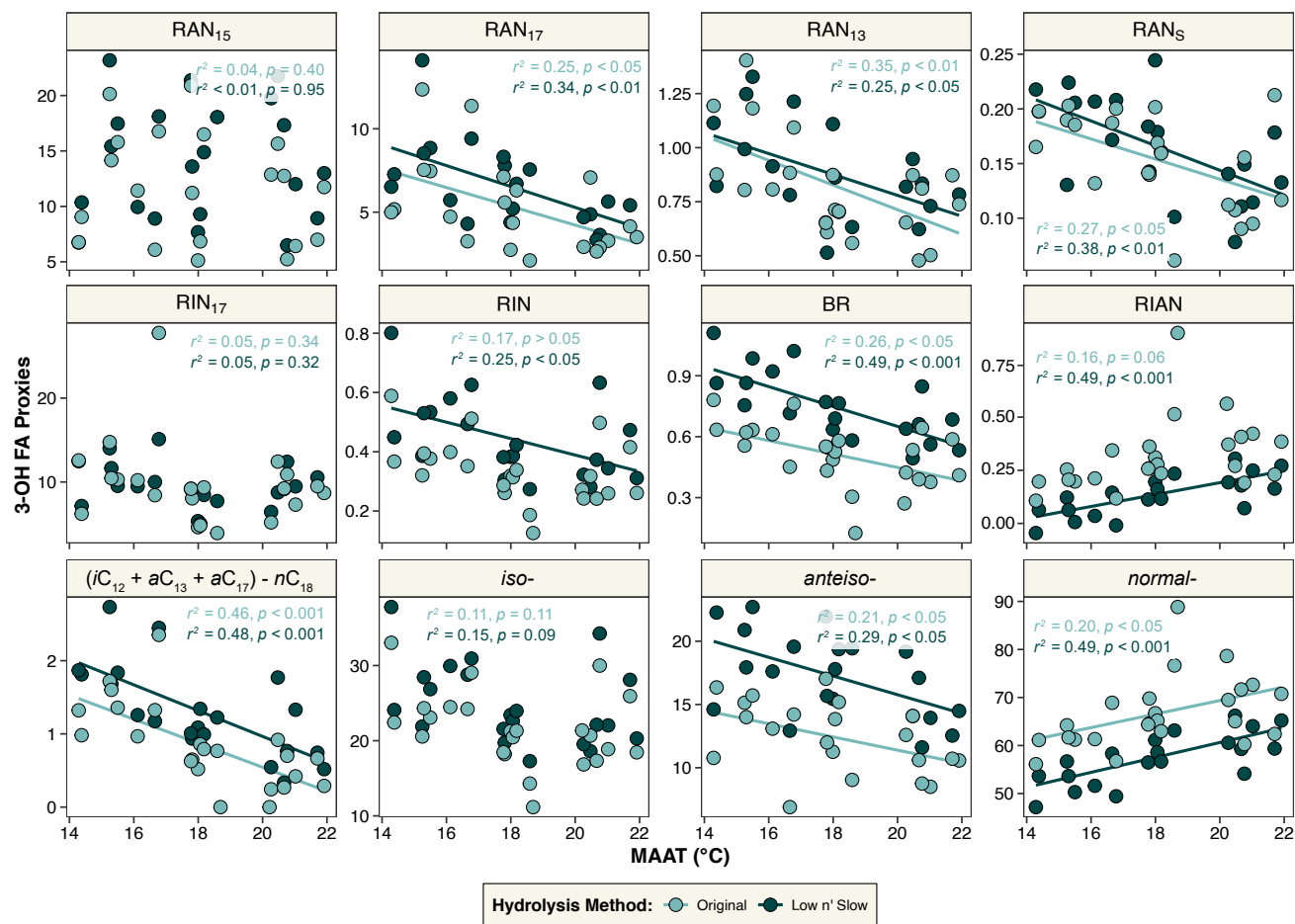


Figure C5: Crossplot matrix of established 3-OH FA-based temperature and pH proxies and MAAT in U.S. lakes extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.

Appendix C6

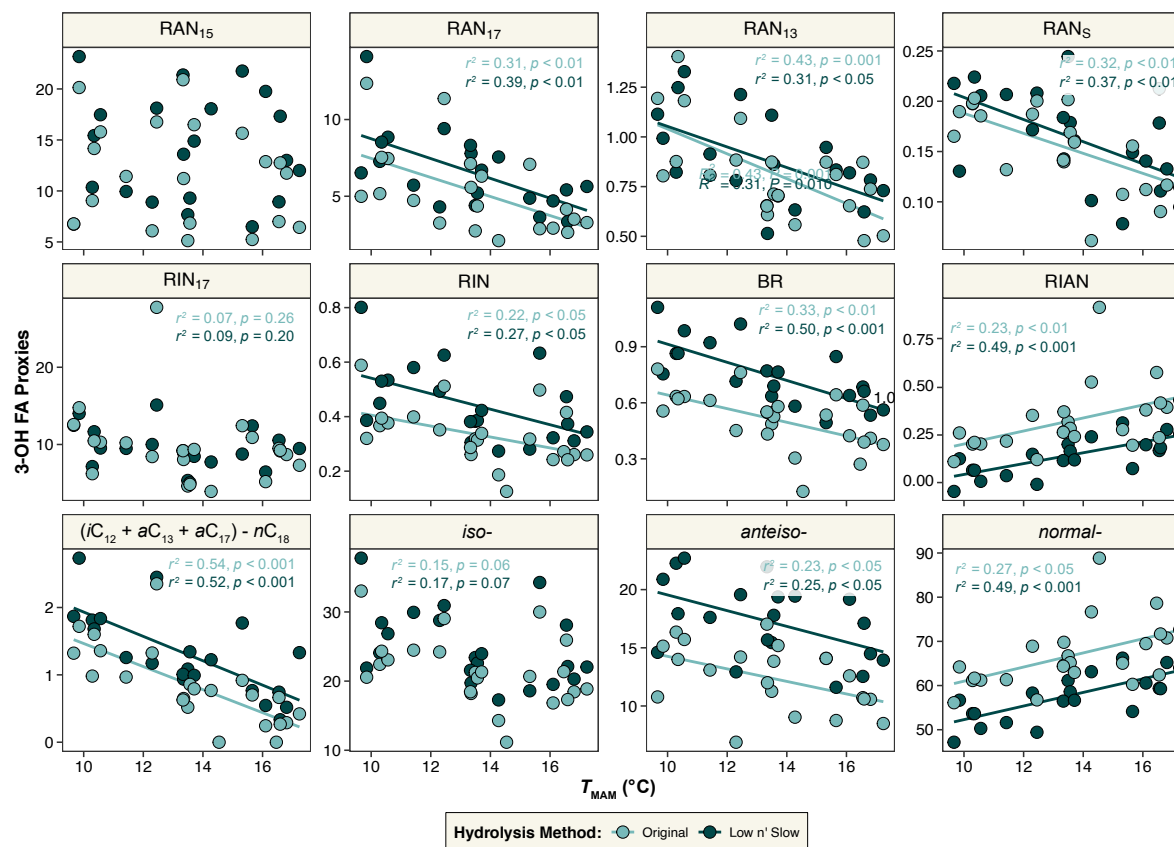


Figure C6: Crossplot matrix of established 3-OH FA-based temperature and pH proxies and T_{MAM} in U.S. lakes extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.

Appendix C7

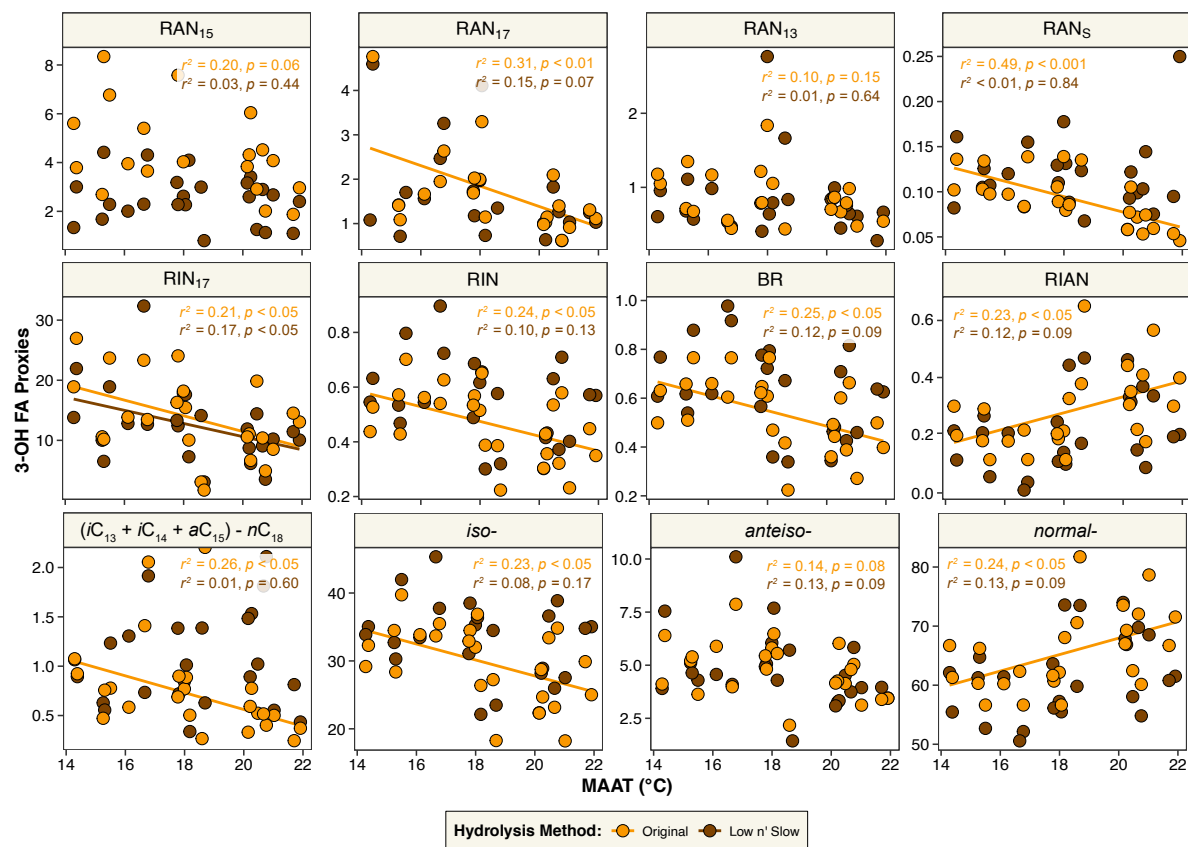


Figure C7: Crossplot matrix of established 3-OH FA-based temperature and pH proxies and MAAT in U.S. soils extracted by Low n' Slow acid hydrolysis. Regression lines displayed for significant correlations.

APPENDIX D: SUPPLEMENTARY INFORMATION FOR CHAPTER 6

Appendix D1

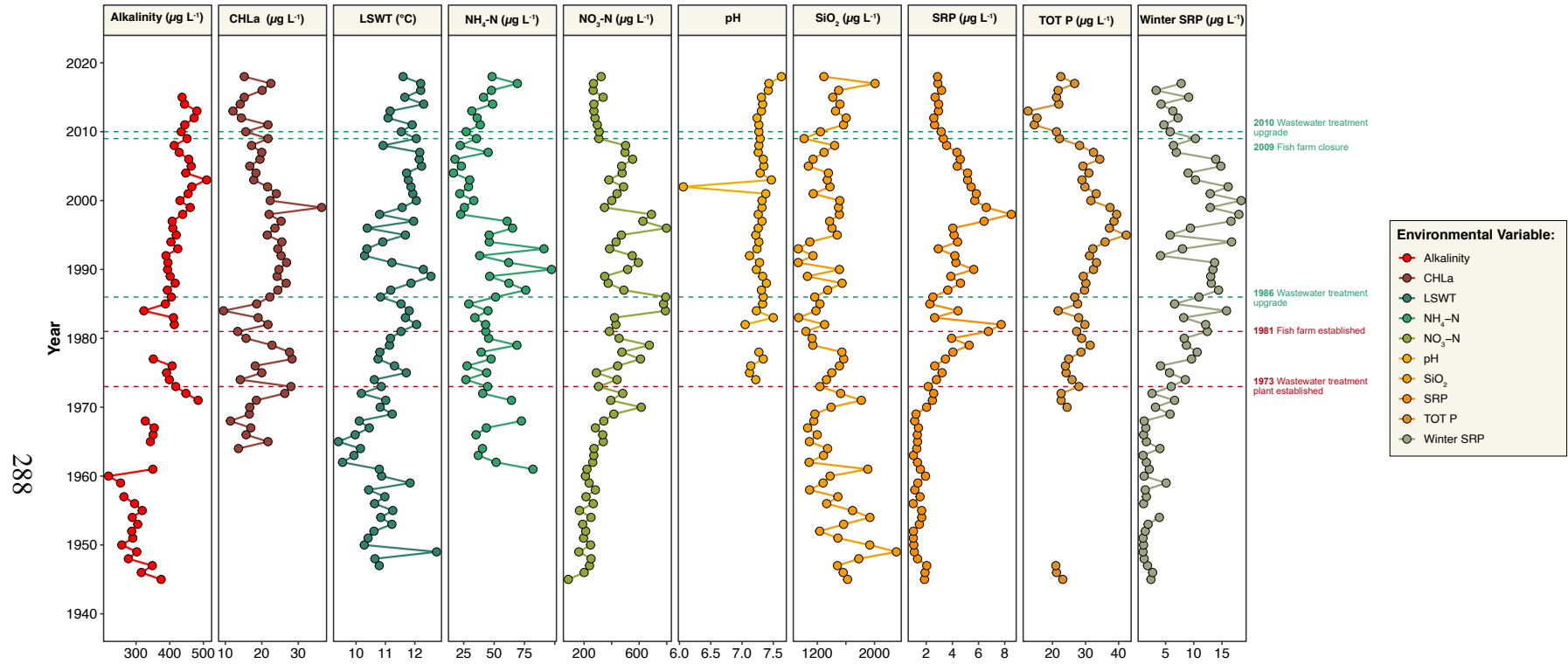


Figure D1: Environmental variables from [UKCEH](#) monitoring records of Esthwaite Water. Alkalinity is measured as $\mu\text{g L}^{-1} \text{CaCO}_3$ and pH. SRP is soluble reactive phosphate ($\text{PO}_4\text{-P}$), and Winter SRP is SRP averaged for winter months which was calculated by [Thorpe et al. \(2022\)](#). Data is integrated between water samples collected from lake surface waters between 0 and 5 m depth in water column. Data after [Maberly et al. \(2017\)](#) and ([Feuchtmayr et al. 2021](#)).

APPENDIX E: DATA TABLES

290

Table E1: U.S. East Coast soils 3-OH FA data (fractional abundance and proxies)

29

Table E2: Southwestern U.S. lake sediments and soils (extracted by original acid hydrolysis method) 3-OH FA data (fractional abundance and proxies)

292

Table E3: Southwestern U.S. lake sediments and soils (extracted by Low n' Slow hydrolysis method) 3-OH FA data (fractional abundance and proxies)

Appendix E5

METADATA			LOI CALCULATION					
Sample	Class	Type	Empty Boat Mass (g)	Sample Mass (g)	Combusted Sample Mass + Boat (g)	Combusted Sample Mass (g)	LOI (g)	LOI (%)
USA 041	Sample	soil	1.24	1.09	2.28	1.04	0.05	4.20
USA 041	Sample	lake sediment	1.24	1.01	2.24	0.99	0.02	2.19
USA 042	Sample	soil	1.25	1.03	2.22	0.98	0.06	5.57
USA 042	Sample	lake sediment	18.78	1.15	19.87	1.10	0.05	4.36
USA 043	Sample	soil	1.24	1.08	2.23	0.99	0.09	8.25
USA 043	Sample	lake sediment	18.27	1.12	19.28	1.01	0.11	9.52
USA 044	Sample	soil	1.24	1.10	2.31	1.07	0.03	2.67
USA 044	Sample	lake sediment	17.83	1.04	18.82	0.99	0.05	4.77
USA 045	Sample	soil	1.25	1.01	2.26	1.01	0.01	0.82
USA 045	Sample	lake sediment	18.11	1.03	19.05	0.94	0.09	8.85
USA 046	Sample	soil	13.72	1.27	14.97	1.25	0.01	1.04
USA 046	Sample	lake sediment	19.21	1.01	20.20	1.00	0.01	0.82
USA 047	Sample	soil	1.24	1.03	2.20	0.96	0.07	7.24
USA 047	Sample	lake sediment	1.26	1.02	2.21	0.95	0.07	7.28
USA 048	Sample	soil	1.24	1.10	2.23	0.99	0.11	9.95
USA 048	Sample	lake sediment	1.24	1.08	2.20	0.96	0.12	11.54
USA 049	Sample	soil	1.25	1.01	2.20	0.96	0.05	5.38
USA 049	Sample	lake sediment	1.23	1.06	2.18	0.95	0.11	10.08
USA 050	Sample	soil	1.24	1.07	2.28	1.04	0.03	2.90
USA 050	Sample	lake sediment	1.24	1.07	2.25	1.01	0.06	5.98
USA 051	Sample	soil	1.24	1.09	2.26	1.02	0.07	6.00
USA 051	Sample	lake sediment	1.26	1.02	2.17	0.91	0.12	11.34
USA 052	Sample	soil	1.24	1.14	2.34	1.10	0.04	3.45
USA 052	Sample	lake sediment	1.24	1.08	2.25	1.00	0.08	7.47
USA 053	Sample	soil	1.24	1.03	2.09	0.85	0.18	17.62
USA 053	Sample	lake sediment	1.26	0.79	1.89	0.63	0.16	19.86
USA 054	Sample	soil	1.24	1.06	2.25	1.01	0.05	4.26
USA 054	Sample	lake sediment	1.24	1.02	2.17	0.94	0.09	8.45
USA 055	Sample	soil	1.25	1.06	2.28	1.03	0.03	2.80
USA 055	Sample	lake sediment	1.26	1.10	2.30	1.03	0.07	6.12
USA 056	Sample	soil	1.24	1.11	2.30	1.06	0.05	4.09
USA 056	Sample	lake sediment	1.26	1.04	2.26	1.00	0.04	3.72
USA 057	Sample	soil	1.24	1.06	2.28	1.04	0.03	2.51
USA 057	Sample	lake sediment	1.26	1.03	2.20	0.95	0.08	7.87
USA 058	Sample	soil	1.24	1.12	2.32	1.08	0.03	3.01
USA 058	Sample	lake sediment	1.25	1.16	2.40	1.15	0.01	0.83
USA 059	Sample	soil	1.25	1.00	2.20	0.95	0.06	5.87
USA 059	Sample	lake sediment	1.26	1.03	2.25	0.98	0.05	4.70
USA 060	Sample	soil	1.24	1.06	2.25	1.01	0.05	4.72
USA 060	Sample	lake sediment	1.25	1.03	2.20	0.95	0.07	7.30
USA 061	Sample	soil	1.26	1.09	2.28	1.03	0.07	6.02
USA 061	Sample	lake sediment	1.25	1.14	2.31	1.06	0.08	6.73
USA 062	Sample	soil	1.25	1.03	2.04	0.79	0.24	23.00
USA 062	Sample	lake sediment	13.60	1.03	14.60	1.00	0.03	3.37
USA 063	Sample	soil	1.24	1.02	2.06	0.82	0.21	20.31
USA 063	Sample	lake sediment	14.36	1.22	15.49	1.13	0.09	7.28
USA 064	Sample	soil	1.24	1.02	2.19	0.95	0.07	6.93
USA 064	Sample	lake sediment	14.09	1.06	15.01	0.92	0.15	13.80
USA 065	Sample	soil	1.24	1.06	2.15	0.91	0.15	14.05
USA 065	Sample	lake sediment	13.77	1.36	15.10	1.34	0.03	2.10
USA 066	Sample	soil	1.25	1.05	2.23	0.99	0.07	6.29
USA 066	Sample	lake sediment	13.98	1.08	14.90	0.93	0.16	14.49
USA 002	Sample	soil	1.25	1.00	2.21	0.96	0.04	4.06
USA 002	Duplicate	soil	13.60	1.06	14.61	1.01	0.05	4.32
USA 002	TriPLICATE	soil	13.98	1.03	14.96	0.98	0.04	4.02
USA 019	Sample	soil	18.78	1.02	19.75	0.97	0.04	4.38
USA 019	Duplicate	soil	13.71	1.00	14.67	0.96	0.04	4.29
USA 019	TriPLICATE	soil	14.09	1.08	15.12	1.03	0.05	4.20
USA 061	Duplicate	soil	1.25	1.09	2.27	1.03	0.07	6.19
USA 061	TriPLICATE	soil	1.24	1.17	2.33	1.09	0.08	6.56
USA 044	Duplicate	soil	1.25	1.23	2.44	1.19	0.03	2.84
USA 044	TriPLICATE	soil	1.26	1.07	2.30	1.04	0.03	3.03
USA 063	Duplicate	soil	1.25	1.06	2.08	0.83	0.23	22.00
USA 063	TriPLICATE	soil	1.24	1.21	2.18	0.94	0.26	21.85
USA 062	Duplicate	soil	1.26	1.10	2.10	0.84	0.26	23.55
USA 062	TriPLICATE	soil	1.25	1.00	2.01	0.76	0.24	24.11
USA 045	Duplicate	lake sediment	1.26	1.20	2.34	1.08	0.12	9.93
USA 045	TriPLICATE	lake sediment	1.26	1.05	2.21	0.95	0.10	9.89
USA 046	Duplicate	lake sediment	1.26	1.11	2.36	1.10	0.01	1.02
USA 046	TriPLICATE	lake sediment	1.26	1.10	2.35	1.09	0.01	1.16
USA 050	Duplicate	lake sediment	1.25	1.03	2.22	0.97	0.06	5.76
USA 050	TriPLICATE	lake sediment	1.26	1.03	2.23	0.97	0.06	5.95
USA 048	Duplicate	lake sediment	1.26	1.10	2.23	0.97	0.13	11.93
USA 048	TriPLICATE	lake sediment	1.26	1.02	2.16	0.90	0.12	11.41
USA 002	SD	soil					0.003	0.163
USA 019	SD	soil					0.001	0.088
USA 061	SD	soil					0.006	0.278
USA 044	SD	soil					0.003	0.180
USA 063	SD	soil					0.028	0.936
USA 062	SD	soil					0.013	0.554
USA 045	SD	lake sediment					0.014	0.612
USA 046	SD	lake sediment					0.002	0.175
USA 050	SD	lake sediment					0.003	0.120
USA 048	SD	lake sediment					0.008	0.271
ERROR							0.008	0.338

Table E5: Loss on Ignition (LOI) calculation

Appendix E6

3-OH FA STANDARDS				
Class	GC-MS run	Number	Concentration (ug/ml)	nC ₁₄ Peak Area
standard	1	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	4418902
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	13380954
standard		3-OH FA standard 3	100 ug/ml nC12, nC14, nC16	6341553
standard		3-OH FA standard 4	100 ug/ml nC12, nC14, nC16	10806632
standard		3-OH FA standard 5	100 ug/ml nC12, nC14, nC16	7046644
standard		3-OH FA standard 6	100 ug/ml nC12, nC14, nC16	8560890
standard	3	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	748808
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	6227530
standard		3-OH FA standard 3	100 ug/ml nC12, nC14, nC16	6230905
standard		3-OH FA standard 4	100 ug/ml nC12, nC14, nC16	4764053
standard	5	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	6979929
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	7881916
standard		3-OH FA standard 3	100 ug/ml nC12, nC14, nC16	8857724
standard		3-OH FA standard 4	100 ug/ml nC12, nC14, nC16	6857247
standard		3-OH FA standard 5	100 ug/ml nC12, nC14, nC16	8373416
standard		3-OH FA standard 6	100 ug/ml nC12, nC14, nC16	6625439
standard		3-OH FA standard 7	100 ug/ml nC12, nC14, nC16	7560434
standard		3-OH FA standard 8	100 ug/ml nC12, nC14, nC16	8124672
standard	6	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	8977826
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	6915829
standard	2	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	7790869
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	13976612
standard		3-OH FA standard 3	100 ug/ml nC12, nC14, nC16	7279142
standard		3-OH FA standard 4	100 ug/ml nC12, nC14, nC16	7986885
standard		3-OH FA standard 5	100 ug/ml nC12, nC14, nC16	7155803
standard	4	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	13544510
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	13035169
standard		3-OH FA standard 3	100 ug/ml nC12, nC14, nC16	12130924
standard		3-OH FA standard 4	100 ug/ml nC12, nC14, nC16	11102147
standard	7	3-OH FA standard 1	25 ug/ml nC12, nC14, nC16	16410696
standard		3-OH FA standard 2	25 ug/ml nC12, nC14, nC16	15283800
standard		3-OH FA standard 3	25 ug/ml nC12, nC14, nC16	19952930
standard	8	3-OH FA standard 1	100 ug/ml nC12, nC14, nC16	6051093
standard		3-OH FA standard 2	100 ug/ml nC12, nC14, nC16	18712454

Calculation of GC Response of 3-OH FA Standards						
GC-MS run	Concentration	Unit	Injection (ul)	Dilution (ml)	Mean nC ₁₄ Area	Concentration / Area
1	100	ug/ml	1	0.05	8425929.167	1.187E+05
2	100	ug/ml	1	0.05	8837862.2	1.131E+05
3	100	ug/ml	1	0.05	4492824	2.226E+05
4	100	ug/ml	1	0.05	12453187.5	8.030E+06
5	100	ug/ml	1	0.05	7657597.125	1.306E+05
6	100	ug/ml	1	0.05	7946827.5	1.258E+05
7	25	ug/ml	1	0.05	17215808.67	1.452E+06
8	100	ug/ml	1	0.05	12381773.5	8.076E+06

Table E6: 3-OH FA standards used for SW U.S. lake sediments and soils GC-MS runs, and calculation of response factor in preparation for calculation of 3-OH FA concentrations in samples.

Appendix E8

METADATA					STANDARD ERROR CALCULATION	
Sample	Class	GC-MS Run	Type	Method	Total Conc. (ng/g LOI)	Total Conc. (mg/g LOI)
					Error	Error
USA-041	sample	7	lake sediment	Wang, et al. (2012)		
USA-042	sample	7	lake sediment	Wang, et al. (2012)		
USA-043	sample	5	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-044	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-045	sample	7	lake sediment	Wang, et al. (2012)		
USA-046	sample	5	lake sediment	Wang, et al. (2012)		
USA-047	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-048	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-049	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-050	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-051	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-052	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-053	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-054	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-055	sample	7	lake sediment	Wang, et al. (2012)		
USA-056	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-057	sample	5	lake sediment	Wang, et al. (2012)		
USA-058	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-059	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-060	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-061	sample	3	lake sediment	Wang, et al. (2012)	1580.1098	1.5971
USA-062	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-063	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-064	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-065	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-066	sample	5	lake sediment	Wang, et al. (2012)	409.3409	0.4160
USA-041	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-042	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-043	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-044	sample	3	soil	Wang, et al. (2012)	1580.1098	1.5971
USA-045	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-046	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-047	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-048	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-049	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-050	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-051	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-052	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-053	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-054	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-055	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-056	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-057	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-058	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-059	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-060	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-061	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-062	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-063	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-064	sample	5	soil	Wang, et al. (2012)	409.3409	0.4160
USA-065	sample	6	soil	Wang, et al. (2012)	1.5224	0.0095
USA-066	sample	5	soil	Wang, et al. (2012)	1580.1098	1.5971
USA-041	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-042	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-043	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-044	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-045	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-046	sample	3	lake sediment	Low & Shaw	1580.1098	1.5971
USA-047	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-048	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-049	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-050	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-051	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-052	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-053	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-054	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-055	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-056	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-057	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-058	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-059	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-060	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-061	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-062	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-063	sample	1	lake sediment	Low & Shaw	11.1712	0.0192
USA-064	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-065	sample	2	lake sediment	Low & Shaw	278.4705	0.2865
USA-041	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-042	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-043	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-044	sample	4	soil	Low & Shaw	11.0112	0.0198
USA-045	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-046	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-047	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-048	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-049	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-050	sample	4	soil	Low & Shaw	11.0112	0.0198
USA-051	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-052	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-053	sample	8	soil	Low & Shaw		
USA-054	sample	3	soil	Low & Shaw	1580.1098	1.5971
USA-055	sample	3	soil	Low & Shaw	1580.1098	1.5971
USA-056	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-057	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-058	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-059	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-060	sample	8	soil	Low & Shaw		
USA-061	sample	8	soil	Low & Shaw		
USA-062	sample	8	soil	Low & Shaw		
USA-063	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-064	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-065	sample	1	soil	Low & Shaw	11.1712	0.0192
USA-066	sample	1	soil	Low & Shaw	11.1712	0.0192
STANDARD ERROR			lake sediment	Wang, et al. (2012)	881.2515	0.8890
			soil	Wang, et al. (2012)	508.0273	0.5607
			lake sediment	Low & Shaw	170.5165	0.1795
			soil	Low & Shaw	154.6783	0.1627

Table E8: Standard error calculation for 3-OH FA concentrations of SW U.S. lake sediments and soils extracted by both hydrolysis methods

298

Table E9: GDGT data from SW U.S. lake sediments and soils

Appendix E10

U.S. ENVIRONMENTAL CONDITIONS																															
METADATA				GENERAL CLIMATIC CONDITIONS										SOIL ENVIRONMENTAL CONDITIONS										LACUSTRINE CONDITIONS							
Sample	Transect	Latitude	Longitude	MAAT	T _{JJA}	T _{SON}	T _{DJF}	T _{MMM}	Seasonality (T _{JJA} -T _{DJF})	MAP	VPD	Elevation	Climate Regime	Precipitation Regime	Temperature Regime	ΔT (T _{soil} -MAAT)	Soil pH	T _{soil}	SMO	Vegetation	Soil Type	T _{water}	Water pH	Depth	CA/SA	% Catchment	Lake sed. pH				
USA 001	East Coast	44.17	-69.41	7.67	16.78	15.08	-1.65	0.48	18.43	1234.89	4.69	79.00	Snow	Fully Humid	Warm Summer	0.43	5.20	8.10	22.45	northern hardwoods	podzols										
USA 002	East Coast	44.62	-69.01	6.86	16.64	14.65	-2.99	-0.84	19.64	1158.67	4.76	34.00	Snow	Fully Humid	Warm Summer	1.24	5.04	8.10	23.97	Nottheastern spruce-fir forest	podzols										
USA 003	East Coast	45.00	-68.65	6.54	16.27	14.16	-3.22	-1.05	19.49	1086.90	4.91	66.00	Snow	Fully Humid	Warm Summer	1.36	4.28	7.90	38.86	Nottheastern spruce-fir forest	podzols										
USA 004	East Coast	44.09	-70.28	7.52	17.13	15.06	-2.20	0.08	19.34	1184.56	5.21	12.00	Snow	Fully Humid	Warm Summer	1.58	5.06	9.10	53.51	northern hardwoods	podzols										
USA 005	East Coast	43.73	-70.69	6.84	16.63	14.21	-2.91	-0.57	19.53	1248.90	5.18	113.00	Snow	Fully Humid	Warm Summer	2.36		9.20	36.10	northern hardwoods	podzols										
USA 006	East Coast	43.08	-71.15	8.83	18.12	15.93	-0.59	1.86	18.71	1176.08	5.77	90.00	Snow	Fully Humid	Warm Summer	1.17	4.11	10.00	19.38	northern hardwoods	podzols										
USA 007	East Coast	42.33	-71.48	9.64	18.79	16.54	0.46	2.77	18.32	1208.15	6.07	100.00	Snow	Fully Humid	mixed	0.46	3.84	10.10	45.43	Appalachian oak forest	podzols										
USA 008	East Coast	42.06	-71.79	9.22	18.10	16.11	0.30	2.36	17.79	1239.73	5.37	221.00	mixed	Fully Humid	Warm Summer	-0.02	3.56	9.20	30.45	Appalachian oak forest	podzols										
USA 009	East Coast	41.78	-72.28	9.65	18.41	16.38	0.81	2.99	17.60	1253.27	5.55	80.00	Snow	Fully Humid	Warm Summer	1.45	6.12	11.10	26.95	Appalachian oak forest	podzols										
USA 010	East Coast	41.75	-72.96	9.01	18.13	15.78	-0.09	2.20	18.22	1356.74	5.69	249.00	Snow	Fully Humid	mixed	0.89	4.13	9.90	17.33	Appalachian oak forest	podzols										
USA 012	East Coast	40.98	-74.21	10.57	19.48	17.22	1.57	3.99	17.91	1290.95	6.14	143.00	mixed	Fully Humid	Hot Summer	0.93	5.34	11.50	42.55	Appalachian oak forest	cambisols										
USA 014	East Coast	40.39	-75.28	11.15	20.13	17.67	2.07	4.76	18.06	1258.66	6.25	168.00	mixed	Fully Humid	Hot Summer	1.05	4.33	12.20	28.01	Appalachian oak forest	lvisols										
USA 015	East Coast	40.19	-75.80	11.03	19.73	17.55	2.19	4.64	17.54	1223.44	5.85	172.00	Warm Temperate	Fully Humid	Hot Summer	0.67	5.36	11.70	21.62	Appalachian oak forest	acrisols										
USA 016	East Coast	39.97	-76.51	11.86	20.89	18.31	2.69	5.56	18.20	1131.27	6.58	82.00	Warm Temperate	Fully Humid	Hot Summer	0.94	6.47	12.80	29.27	Appalachian oak forest	lvisols										
USA 017	East Coast	39.65	-76.76	11.88	20.54	18.13	3.01	5.84	17.53	1208.71	6.47	186.00	Warm Temperate	Fully Humid	Hot Summer	-0.38	4.61	11.50	20.60	Appalachian oak forest	acrisols							-1			
USA 018	East Coast	38.76	-77.12	14.15	22.74	20.51	5.08	8.27	17.66	1130.48	7.87	21.00	Warm Temperate	Fully Humid	Hot Summer	-0.35	4.76	13.80	13.77	Oak-hickory-pine forest	acrisols										
USA 019	East Coast	38.33	-77.14	14.14	22.45	20.29	5.39	8.44	17.06	1130.22	7.29	46.00	Warm Temperate	Fully Humid	Hot Summer	-0.64	4.90	13.50	15.28	Oak-hickory-pine forest	acrisols										
USA 020	East Coast	37.81	-77.12	14.92	23.01	20.67	6.30	9.67	16.71	1148.72	8.21	17.00	Warm Temperate	Fully Humid	Hot Summer	-0.32	4.14	14.60	14.35	Oak-hickory-pine forest	acrisols										
USA 021	East Coast	37.37	-77.55	14.50	22.77	20.34	5.70	9.19	17.08	1152.77	7.99	78.00	Warm Temperate	Fully Humid	Hot Summer	-0.10	5.05	14.40	20.62	Oak-hickory-pine forest	acrisols										
USA 022	East Coast	36.85	-77.93	14.61	22.67	20.22	6.16	9.38	16.51	1149.85	7.63	100.00	Warm Temperate	Fully Humid	Hot Summer	0.09	5.38	14.70	13.81	Oak-hickory-pine forest	acrisols										
USA 023	East Coast	36.44	-78.37	14.76	22.95	20.47	6.12	9.52	16.82	1127.11	8.36	111.00	Warm Temperate	Fully Humid	Hot Summer	-0.06	5.08	14.70	10.48	Oak-hickory-pine forest	acrisols										
USA 024	East Coast	35.95	-78.61	15.28	23.25	20.66	6.79	10.42	16.46	1174.36	8.11	94.00	Warm Temperate	Fully Humid	Hot Summer	-0.18	5.99	15.10	18.64	Oak-hickory-pine forest	acrisols										
USA 025	East Coast	35.46	-78.91	16.00	23.88	21.38	7.52	11.24	16.36	1195.21	8.70	110.00	Warm Temperate	Fully Humid	Hot Summer	-0.70	4.36	15.30	8.59	Oak-hickory-pine forest	acrisols										
USA 026	East Coast	34.67	-78.60	16.97	24.36	22.17	9.02	12.34	15.34	1260.67	8.85	17.00	Warm Temperate	Fully Humid	Hot Summer	-0.67	3.70	16.30	6.18	Oak-hickory-pine forest	podzols										
USA 027	East Coast	34.39	-79.00	16.94	24.33	21.97	9.05	12.40	15.28	1230.09	8.55	28.00	Warm Temperate	Fully Humid	Hot Summer	0.16	4.07	17.10	9.57	Oak-hickory-pine forest	acrisols										
USA 028	East Coast	34.33	-79.28	17.11	24.60	22.19	9.09	12.54	15.50	1207.92	8.70	40.00	Warm Temperate	Fully Humid	Hot Summer	-0.21	4.34	16.90	7.57	Oak-hickory-pine forest	acrisols										
USA 029	East Coast	33.95	-79.98	17.48	24.84	22.56	9.53	12.99	15.31	1252.42	9.25	61.00	Warm Temperate	Fully Humid	Hot Summer	-0.58	4.99	16.90	11.52	Oak-hickory-pine forest	acrisols										
USA 030	East Coast	33.52	-80.50	17.97	25.10	22.99	10.12	13.67	14.98	1231.76	9.20	66.00	Warm Temperate	Fully Humid	Hot Summer	-1.07	4.76	16.90	5.49	southern mixed forest	acrisols										
USA 031	East Coast	33.06	-81.09	18.39	25.29	23.25	10.76	14.27	14.53	1160.86	9.53	39.00	Warm Temperate	Fully Humid	Hot Summer	-0.59	4.01	17.80	8.67	southern mixed forest	acrisols										
USA 032	East Coast	32.87	-81.96	18.31	25.34	23.28	10.55	14.08	14.79	1152.67	9.59	54.00	Warm Temperate	Fully Humid	Hot Summer	-0.51	4.14	17.80	5.36	southern mixed forest	acrisols										
USA 033	East Coast	32.08	-82.14	19.10	25.74	23.84	11.77	15.06	13.97	1200.09	9.31	71.00	Warm Temperate	Fully Humid	Hot Summer	-0.70	4.92	18.40	13.87	southern mixed forest	acrisols										
USA 035	East Coast	30.21	-82.59	20.32	25.88	24.62	13.95	16.81	11.92	1344.17	9.55	68.00	Warm Temperate	Fully Humid	Hot Summer	-1.32	3.70	19.00	42.40	southern mixed forest	acrisols										
USA 036	East Coast	29.11	-81.62	21.34	26.17	25.57	15.74	17.87	10.43	1347.60	10.54	19.00	Warm Temperate	Fully Humid	Hot Summer	-0.54	4.09	20.80	13.16	southern mixed forest	regosols										
USA 037	East Coast	28.35	-81.64	22.68	26.89	26.44	17.76	19.63	9.13	1292.28	10.39	33.00	Warm Temperate	Fully Humid	Hot Summer	-1.38	6.01	21.30	9.57	southern mixed forest	podzols										
USA 039	East Coast	26.94	-81.31	23.04	26.67	26.59	18.71	20.20	7.96	1270.30	10.42	11.00	Warm Temperate	Fully Humid	Hot Summer	-0.94	4.49	22.10	3.04	Everglades and forest	podzols										
USA 040	East Coast	27.60	-82.55	23.01	27.04	26.85	18.30	19.85	8.74	1276.14	9.19	-4.00	Equatorial	Monsoonal	N/A	-1.01	7.09	22.00	8.32	Cypress savanna	podzols										
USA 041	Southwest	29.90	-95.15	21.02	27.25	25.89	13.72	17.24	13.53	1473.73	9.53	35.00	Warm Temperate	Fully Humid	Hot Summer	-0.92	5.17	20.10	2.84	woodland	vertisol	24.6	7.7	25	598.28	0.17	7.48				
USA 042	Southwest	35.71	-101.56	15.25	24.19	21.56	3.40	9.85	18.78	479.42	12.73	877.00	And	Steppe	Cold And	1.15	8.00	16.40	6.52	grassland	kastanozem	0	0	13	1997.23	0.05	7.37				
USA 043	Southwest	36.86	-105.10	6.88	14.00	12.32	-0.97	2.15	14.97	504.06	7.86	2500.00	Snow And	Fully Humid	Cold And	-0.68	6.38	6.20	14.21	woodland	lvisols	0	0	0.2	57.11	1.75	8.8				
USA 044	Southwest	34.12	-106.89	14.38	22.79	19.80	4.63	10.29	18.16	245.24	15.01	1382.00	And	Steppe	Cold And	2.52	8.09	16.90	6.39	shrubsteppe	yermosol	31.2	8.36	1	6944.31	0.01	7.51				
USA 045	Southwest	33.18	-107.20	16.78	25.28	22.17	7.22	12.45	18.06	251.32	16.93	1371.00	And	Desert	Cold And	1.72	8.50	18.50	1.05	shrubsteppe	yermosol	29	8.02	1	1978.99	0.05	7.62				
USA 046	Southwest	32.91	-107.31	16.66	24.90	22.04	7.38	12.50	17.51	275.67	17.22	1294.00	And	Desert	Cold And	2.84	8.36	19.50	0.20	shrubsteppe	yermosol	25	8.14	23	700.71	0.14	7.52				
USA 047	Southwest	32.72	-109.71	17.99	25.81	23.62	9.04	13.49	16.77	253.46	18.57	978.00	And	Steppe	Cold And	3.01	7.63	21.00	3.80	shrubsteppe	yermosol	25.5	8.42	21	357.23	0.28	8.14				
USA 048	Southwest	32.76	-109.70	18.06	25.94	23.72	9.04	13.56	16.90	246.12	18.76	956.00	And	Steppe	Cold And	3.24	7.76	21.30	19.44	shrubsteppe	yermosol	26.9	8.68	3	1402						

Appendix E11

METADATA						3-OF FA FRACTIONAL ABUNDANCE																				3-OH FA PROXIES												
Sample	Class	Type	Method	Core Depth (cm)	Age	<i>n</i> C ₁₀	<i>i</i> C ₁₁	<i>a</i> C ₁₁	<i>n</i> C ₁₁	<i>i</i> C ₁₂	<i>u</i> C ₁₂	<i>n</i> C ₁₂	<i>i</i> C ₁₃	<i>a</i> C ₁₃	<i>n</i> C ₁₃	<i>i</i> C ₁₄	<i>u</i> C ₁₄	<i>n</i> C ₁₄	<i>i</i> C ₁₅	<i>a</i> C ₁₅	<i>n</i> C ₁₅	<i>i</i> C ₁₆	<i>u</i> C ₁₆	<i>n</i> C ₁₆	<i>i</i> C ₁₇	<i>a</i> C ₁₇	<i>n</i> C ₁₇	<i>i</i> C ₁₈	<i>u</i> C ₁₈	<i>n</i> C ₁₈	RAN ₁₅	RAN ₁₇	3-OH FA ₁₅	RIN	RIN ₁₇	BR	RAN	
ESW1	sample	lake sediment	Low n' Slow	0	2021	6.887	0.427	0.135	0.262	0.369	0.827	7.657	2.463	0.658	0.459	2.034	0.148	13.830	12.585	7.437	0.765	2.066	0.341	15.701	10.023	5.873	0.919	1.543	0.185	6.404	9.719	6.387	1.435	0.176	0.596	10.902	0.863	0.267
ESW2	sample	lake sediment	Low n' Slow	1	2020	6.839	0.375	0.129	0.269	0.361	0.727	7.835	2.604	0.699	0.502	2.055	0.173	14.058	10.538	7.354	0.806	2.051	0.329	17.135	9.880	5.773	1.300	1.440	0.213	6.556	9.129	4.441	1.394	0.176	0.530	7.601	0.782	0.252
ESW3	sample	lake sediment	Low n' Slow	2	2019	7.550	0.412	0.161	0.280	0.407	0.871	8.478	2.768	0.740	0.525	2.017	0.200	13.869	9.937	7.373	0.751	2.062	0.343	16.661	9.614	5.555	1.256	1.386	0.209	6.577	9.822	4.423	1.410	0.180	0.511	7.654	0.758	0.247
ESW4	sample	lake sediment	Low n' Slow	3	2017	7.069	0.412	0.138	0.283	0.348	0.813	7.844	2.680	0.767	0.542	2.107	0.156	14.263	10.171	7.707	0.785	1.977	0.429	16.663	9.474	5.591	1.261	1.779	0.195	6.546	9.812	4.432	1.414	0.179	0.524	7.510	0.781	0.257
ESW5	sample	lake sediment	Low n' Slow	4	2016	6.342	0.377	0.127	0.208	0.334	0.815	7.654	2.564	0.728	0.437	1.839	0.172	13.315	9.869	7.374	0.751	1.974	0.388	16.885	10.460	5.840	1.174	1.026	0.214	9.133	9.820	4.974	1.666	0.174	0.509	8.910	0.761	0.252
ESW6	sample	lake sediment	Low n' Slow	5	2015	8.075	0.448	0.139	0.304	0.450	0.965	8.697	2.778	0.777	0.509	1.966	0.222	13.726	10.001	7.408	0.707	1.977	0.325	15.874	8.937	5.587	1.504	1.402	0.198	7.022	10.498	3.714	1.526	0.183	0.496	5.941	0.742	0.247
ESW7	sample	lake sediment	Low n' Slow	6	2014	5.574	0.349	0.123	0.233	0.353	0.740	7.815	2.575	0.755	0.462	1.985	0.176	14.541	9.965	7.906	0.769	2.173	0.413	17.754	10.172	5.238	1.422	1.508	0.213	6.786	10.277	3.683	1.632	0.171	0.525	7.152	0.779	0.253
ESW8	sample	lake sediment	Low n' Slow	7	2012	9.121	0.442	0.173	0.303	0.397	1.041	9.097	2.679	0.752	0.493	1.957	0.174	13.151	8.926	6.801	0.727	1.968	0.353	16.284	9.692	5.672	1.118	1.022	0.241	7.416	9.353	5.075	1.526	0.185	0.469	8.671	0.701	0.232
ESW9	sample	lake sediment	Low n' Slow	8	2011	3.380	0.252	0.093	0.173	0.273	0.549	6.131	2.214	0.620	0.414	2.439	0.211	17.114	8.576	7.029	0.939	2.367	0.445	19.281	10.668	6.246	1.232	1.140	0.284	7.931	7.484	5.070	1.498	0.160	0.493	8.661	0.741	0.247
ESW10	sample	lake sediment	Low n' Slow	9	2009	4.507	0.293	0.115	0.178	0.286	0.646	6.535	2.196	0.664	0.398	1.913	0.164	13.406	12.251	7.083	0.868	2.227	0.429	18.819	10.630	6.223	1.161	1.076	0.268	7.661	8.156	5.360	1.671	0.172	0.577	9.156	0.840	0.263
ESW11	sample	lake sediment	Low n' Slow	10	2007	6.564	0.284	0.096	0.177	0.280	0.661	7.570	2.340	0.616	0.414	1.823	0.153	14.311	10.643	7.719	0.618	1.766	0.353	17.977	9.953	5.567	1.116	1.651	0.215	7.131	12.480	4.987	1.487	0.156	0.514	8.916	0.765	0.251
ESW12	sample	lake sediment	Low n' Slow	11	2004	7.805	0.321	0.110	0.222	0.282	0.754	7.775	2.502	0.683	0.452	1.866	0.186	13.641	9.605	7.345	0.720	1.954	0.357	17.412	10.187	5.933	1.024	1.002	0.234	7.626	10.197	5.795	1.510	0.167	0.489	9.950	0.737	0.248
ESW13	sample	lake sediment	Low n' Slow	12	2001	7.466	0.311	0.118	0.237	0.379	0.892	8.203	2.636	0.783	0.529	2.113	0.195	14.265	10.012	7.918	0.781	1.969	0.357	16.990	9.659	5.603	1.168	1.104	0.187	6.123	10.133	4.797	1.479	0.181	0.505	8.269	0.764	0.259
ESW14	sample	lake sediment	Low n' Slow	13	1998	7.103	0.367	0.131	0.256	0.352	0.875	8.653	2.572	0.742	0.487	2.067	0.168	14.264	9.294	7.538	0.685	1.865	0.343	17.485	9.425	5.281	1.330	1.514	0.222	6.982	11.009	3.971	1.525	0.176	0.480	7.088	0.719	0.239
ESW15	sample	lake sediment	Low n' Slow	14	1995	7.331	0.384	0.141	0.259	0.389	0.665	8.949	2.544	0.712	0.500	2.177	0.163	14.784	9.617	8.118	0.785	1.815	0.319	17.582	8.635	4.600	1.435	1.215	0.208	6.672	10.346	3.206	1.424	0.177	0.459	6.018	0.692	0.233
ESW16	sample	lake sediment	Low n' Slow	15	1991	5.328	0.326	0.119	0.222	0.346	0.528	8.303	2.321	0.672	0.482	2.196	0.134	15.593	9.487	9.002	0.790	1.727	0.350	19.088	8.540	4.335	0.979	1.751	0.214	7.169	11.402	4.429	1.394	0.167	0.461	8.727	0.704	0.244
ESW17	sample	lake sediment	Low n' Slow	16	1988	7.572	0.401	0.144	0.275	0.388	0.916	8.636	2.123	0.580	0.429	1.840	0.163	14.328	8.468	8.539	0.714	1.708	0.319	19.495	8.432	4.484	1.471	1.467	0.222	6.887	11.951	3.047	1.352	0.160	0.415	5.730	0.645	0.230
ESW18	sample	lake sediment	Low n' Slow	17	1985	7.029	0.369	0.129	0.251	0.335	0.767	7.094	2.041	0.568	0.496	2.048	0.165	14.851	9.149	7.713	0.758	1.791	0.340	19.097	9.073	4.643	1.346	1.188	0.183	6.977	10.177	3.449	1.145	0.161	0.461	6.738	0.685	0.223
ESW19	sample	lake sediment	Low n' Slow	18	1983	4.918	0.272	0.101	0.193	0.266	0.723	6.849	1.813	0.516	0.388	1.689	0.123	15.211	10.140	6.378	0.750	1.728	0.276	19.213	10.668	4.601	1.396	1.696	0.200	9.892	8.504	3.296	1.329	0.137	0.481	7.642	0.678	0.197
ESW20	sample	lake sediment	Low n' Slow	19	1981	5.664	0.304	0.122	0.200	0.315	0.854	6.821	2.028	0.613	0.467	1.842	0.143	15.499	10.385	6.319	0.662	1.710	0.296	18.802	10.052	4.073	1.226	2.494	0.174	9.934	9.548	3.324	1.313	0.148	0.500	8.202	0.691	0.191
ESW21	sample	lake sediment	Low n' Slow	20	1980	6.136	0.293	0.109	0.187	0.260	0.681	7.763	1.960	0.511	0.455	1.913	0.161	14.108	9.195	6.035	0.765	1.877	0.319	20.340	10.024	3.707	1.496	2.535	0.240	8.927	7.886	2.478	1.123	0.156	0.466	6.700	0.638	0.172
ESW22	sample	lake sediment	Low n' Slow	21	1978	5.613	0.300	0.107	0.180	0.261	0.662	7.180	2.001	0.515	0.338	1.691	0.120	13.499	9.500	5.591	0.735	1.865	0.317	21.596	10.556	3.743	1.305	2.583	0.238	9.505	7.610	2.867	1.522	0.151	0.480	8.067	0.646	0.166
ESW23	sample	lake sediment	Low n' Slow	22	1976	6.178	0.325	0.134	0.228	0.311	0.957	8.385	2.598	0.671	0.514	2.196	0.187	13.789	9.682	5.155	0.782	1.848	0.413	18.987	9.803	3.731	1.167	2.757	0.251	8.951	6.590	3.197	1.305	0.182	0.500	8.400	0.665	0.164
ESW24	sample	lake sediment	Low n' Slow	23	1973	8.254	0.497	0.198	0.306	0.385	1.065	10.869	2.980	0.714	0.603	2.508	0.202	16.133	8.957	4.682	0.861	1.500	0.304	17.417	7.506	2.926	1.103	2.135	0.172	7.903	6.870	2.653	1.184	0.177	0.418	6.805	0.553	0.135
ESW25	sample	lake sediment	Low n' Slow	24	1971	6.622	0.410	0.157	0.202	0.269	0.740	8.157	2.368	0.569	0.506	2.028	0.149	13.690	8.812	4.695	0.817	1.720	0.359	18.361	9.236	3.680	1.534	3.292	0.253	11.374	5.744	2.398	1.124	0.168	0.459	6.019	0.608	0.149
ESW26	sample	lake sediment	Low n' Slow	25	1969	10.678	0.475	0.181	0.299	0.375	1.043	12.373	2.519	0.610	0.519	1.973	0.137	15.147	7.887	4.071	0.741	1.495	0.282	16.210	7.882	3.279	1.225	1.821	0.210	8.568	5.493	2.677	1.174	0.159	0.371	6.435	0.495	0.124
ESW27	sample	lake sediment	Low n' Slow	26	1967	7.435	0.380	0.136	0.230	0.350	0.545	10.949	2.323	0.601	0.502	2.152	0.113	16.166	9.392	4.496	0.795	1.653	0.304	17.891	8.495	3.481	1.453	2.210	0.198	7.751	5.658	2.396	1.195	0.157	0.427	5.848	0.565	0.138
ESW28	sample	lake sediment																																				

Appendix E12

METADATA						ENVIRONMENTAL DATA										
Sample	Class	Type	Method	Core Depth (cm)	Age	Alkalinity (μg L ⁻¹ CaCO ₃ pH)	NO ₃ ⁻ (μg/L)	NH ₄ ⁺ (μg/L)	TOTP (μg/L)	SRP (μg/L)	Winter SRP (μg/L)	SiO ₂ (μg/L)	CHLa (μg/L)	T _{water} (°C)	Water pH	
ESW4	sample	lake sediment	Low n' Slow	3	2017		268	69.2	26.7	2.89	7.76	2008	22.5	12.2	7.43	
ESW5	sample	lake sediment	Low n' Slow	4	2016		266	47.8	21.7	3.17	3.32	1499	20.1	12.2	7.42	
ESW6	sample	lake sediment	Low n' Slow	5	2015	436.8	338.4	41.2	21.2	2.68	9.1	1417.3	15.2	11.66	7.31	
ESW7	sample	lake sediment	Low n' Slow	6	2014	444.2	271.3	48.8	21.9	2.95	4.2	1517.4	14.1	12.3	7.33	
ESW8	sample	lake sediment	Low n' Slow	7	2012	472.6	279.3	36	15.3	2.58	7.2	1604.6	14.4	11.09	7.24	
ESW9	sample	lake sediment	Low n' Slow	8	2011	444.9	296.2	38.6	14.5	2.64	4.7	1568	21.7	11.91	7.27	
ESW10	sample	lake sediment	Low n' Slow	9	2009	451.5	307.6	35.5	22.1	3.31	10.3	1015.8	21.7	12.05	7.29	
ESW11	sample	lake sediment	Low n' Slow	10	2007	428.3	497.4	45.1	32.4	4.35	6.9	1295.2	20	12.17	7.26	
ESW12	sample	lake sediment	Low n' Slow	11	2004	447.8	476.6	16.4	31	5.16	9	1356.2	18.4	11.72	7.29	
ESW13	sample	lake sediment	Low n' Slow	12	2001	454.8	439.4	21.7	33.2	5.83	12.9	1143.3	24	11.93	7.38	
ESW14	sample	lake sediment	Low n' Slow	13	1998	436.6	689.3	22.4	39.4	8.5	18	1509.6	22	10.8	7.26	
ESW15	sample	lake sediment	Low n' Slow	14	1995	419.5	470.7	45.9	42.3	4.13	5.8	1478.7	21.5	11.68	7.22	
ESW16	sample	lake sediment	Low n' Slow	15	1991	394.9	594.2	62	33.3	4.27	13.7	933.6	26.8	11.22	7.28	
ESW17	sample	lake sediment	Low n' Slow	16	1988	416.3	375	62.2	30.1	4.62	13.1	1547.6	26.7	11.87	7.39	
ESW18	sample	lake sediment	Low n' Slow	17	1985	387.4	776.9	29.1	27.6	2.28	6.6	1238.3	18.6	11.53	7.33	
ESW19	sample	lake sediment	Low n' Slow	18	1983	411.2	420.7	34.1	27.9	2.64	8.2	939	19	11.68	7.5	
ESW20	sample	lake sediment	Low n' Slow	19	1981		383.4	43.1	27.3	6.75	12.4	1041.5	13.4	11.53		
ESW21	sample	lake sediment	Low n' Slow	20	1980		454.1	45.5	28.8	3.94	8.3	1127.9	15.7	11.17		
ESW22	sample	lake sediment	Low n' Slow	21	1978		474.6	39.3	28.6	4.04	10.6	1544.5	27.6	10.81	7.27	
ESW23	sample	lake sediment	Low n' Slow	22	1976	407.1	444.2	27.7	23.9	2.65	4.1	1510.4	18.2	11.31	7.14	
ESW24	sample	lake sediment	Low n' Slow	23	1973	418	305.2	44.7	28	2.15	6	1237.3	28	10.86		
ESW25	sample	lake sediment	Low n' Slow	24	1971	484.5	393.5	64.3	22.6	2.47	6.6	1816.6	18.5	11.01		
ESW26	sample	lake sediment	Low n' Slow	25	1969		416.1			1.22	5.8	1166.8	16.6	11.23		
ESW27	sample	lake sediment	Low n' Slow	26	1967	354.1	281.8	43.6		1.41	1.5	1065.4	17	10.45		
ESW28	sample	lake sediment	Low n' Slow	27	1965	342.7	339.6			1.34	1.6	1090.1	21.7	9.4		
ESW29	sample	lake sediment	Low n' Slow	28	1963		270.8	36.9		1.02	1	1287.6		9.93		
ESW30	sample	lake sediment	Low n' Slow	29	1957	264	214			1.53	1.6	1488.6		10.98		
ESW31	sample	lake sediment	Low n' Slow	30	1950	257.6	247.3			1.05	1.1	1934.3		10.28		

Table E12: Environmental data from Esthwaite Water. Data from [Maberly et al. \(2017\)](#), and data from 2014 to 2017 from [Feuchtmayr et al. \(2021\)](#).

Appendix E13

METADATA			TEMPERATURE CALIBRATIONS																pH CALIBRATIONS						
			3-OH FA CALIBRATION ESTIMATED MAAT								MBT EST. MAAT		ΔT (ESTIMATED T - MAAT)								PROXY EST. pH				
Sample	Transect	MAAT	Wang, et al. (2021)	Mt. Shennongjia (S)	U.S. Soils	Global Soils	Mt. Rungwe (R)	Mt. Majella (M)	SRM Combined	MAAT	MBT	Wang, et al. (2021)	Mt. Shennongjia (S)	U.S. Soils	Global Soils	Mt. Rungwe (R)	Mt. Majella (M)	SRM Combined	MBT	Soil pH	-log(RIN _{iso})	pH	-log(RIN _{iso})	pH	
pH (EST pH - pH)																									
USA 001	East Coast	7.67	7.02	7.13	11.01	12.43	12.68	6.91	10.12	8.50	7.40	-0.65	-0.55	3.33	4.76	5.01	-0.77	2.45	-1.10	5.20	5.80	4.70	0.60	-0.40	
USA 002	East Coast	6.86	4.79	5.73	9.98	11.76	11.42	6.62	8.99	8.20	9.20	-2.08	-1.14	3.12	4.89	4.55	-0.24	2.12	1.00	5.04	5.19	4.60	0.15	-0.20	
USA 003	East Coast	6.54	0.46	3.01	7.99	10.45	8.95	6.07	6.78	7.80	7.80	-6.08	-3.53	1.45	3.91	2.41	-0.46	0.24	0.00	4.28	5.50	5.00	1.22	-0.10	
USA 004	East Coast	7.52	2.74	4.44	9.04	11.14	10.25	6.36	7.95	8.40	11.50	-4.77	-3.07	1.53	3.62	2.74	-1.15	0.43	3.10	5.06	6.09	3.10	1.03	-2.20	
USA 005	East Coast	6.84	6.09	6.54	10.58	12.15	12.16	6.79	9.65	8.10	8.90	-0.75	-0.30	3.74	5.31	5.32	-0.05	2.81	0.80		5.34	6.10			
USA 006	East Coast	8.83	-2.34	1.25	6.71	9.60	7.36	5.72	5.36	9.60	14.50	-11.17	-7.58	-2.13	0.77	-1.47	-3.11	-3.47	4.90	4.11	4.32	3.70	0.21	-1.10	
USA 007	East Coast	9.64	10.77	9.48	12.73	13.57	14.82	7.38	12.04	10.20	19.10	1.13	-0.16	3.09	3.93	5.18	-2.26	2.40	8.90	3.84	5.57	5.80	1.73	0.50	
USA 008	East Coast	9.22	18.58	14.38	16.32	15.93	19.26	8.37	16.01	9.80	14.30	9.36	5.16	7.10	6.71	10.04	-0.84	6.80	4.50	3.56	4.78	4.00	1.22	-0.80	
USA 009	East Coast	9.65	8.08	7.79	11.49	12.75	13.29	7.04	10.66	10.20	12.60	-1.57	-1.86	1.85	3.11	3.64	-2.60	1.02	2.40	6.12	6.60	5.70	0.48	-0.20	
USA 010	East Coast	9.01	5.67	6.28	10.39	12.02	11.92	6.74	9.44	9.90	13.30	-3.33	-2.73	1.38	3.02	2.91	-2.27	0.43	3.40	4.13	4.66	4.00	0.53	-1.50	
USA 012	East Coast	10.57	17.00	13.39	15.60	15.45	18.36	8.17	15.21	11.00		6.44	2.82	5.03	4.88	7.80	-2.39	4.64		5.34	5.69		0.35		
USA 014	East Coast	11.15	7.63	7.51	11.29	12.62	13.03	6.98	10.43	12.30	12.10	-3.53	-3.65	0.13	1.46	1.88	-4.17	-0.72	-0.20	4.33	3.76	4.00	-0.57	-2.10	
USA 015	East Coast	11.03	1.60	3.73	8.52	10.79	9.60	6.22	7.36	12.40	14.00	-9.43	-7.30	-2.51	-0.23	-1.42	-4.81	-3.66	1.60	5.36	5.54	4.60	0.18	-1.50	
USA 016	East Coast	11.86	10.22	9.13	12.48	13.40	14.51	7.31	11.75	13.10	17.60	-1.64	-2.73	0.62	1.54	2.64	-4.55	-0.11	4.50	6.47	6.80	6.60	0.33	-0.70	
USA 017	East Coast	11.88	14.73	11.96	14.55	14.76	17.07	7.89	14.05	12.60	14.60	2.85	0.08	2.67	2.88	5.19	-3.99	2.17	2.00	4.61	4.34	3.80	-0.27	-2.00	
USA 018	East Coast	14.15	16.31	12.95	15.28	15.24	17.97	8.09	14.85	14.30	11.80	2.16	-1.20	1.13	1.09	3.81	-6.07	0.70	-2.50	4.76	4.43	5.20	-0.33	0.60	
USA 019	East Coast	14.14	9.70	8.81	12.24	13.24	14.21	7.25	11.49	14.80	18.00	-4.44	-5.34	-1.90	-0.90	0.07	-6.90	-2.65	3.20	4.90	5.14	3.40	0.24	-1.90	
USA 020	East Coast	14.92	14.71	11.95	14.54	14.76	17.06	7.88	14.04	15.10	15.60	-0.20	-2.96	-0.37	-0.16	2.14	-7.03	-0.87	0.50	4.14	4.75	3.70	0.61	-1.60	
USA 021	East Coast	14.50	13.02	10.89	13.76	14.25	16.10	7.67	13.18	15.00	20.30	-1.48	-3.61	-0.74	-0.25	1.60	-6.83	-1.32	5.30	5.05	4.40	3.90	-0.65	-1.60	
USA 022	East Coast	14.61	-1.73	1.64	6.98	9.79	7.71	5.80	5.67	15.00	21.80	-16.34	-12.97	-7.62	-4.82	-6.90	-8.81	-8.94	6.80	5.38	5.28	4.00	-0.10	-1.50	
USA 023	East Coast	14.76	13.74	11.34	14.10	14.46	16.51	7.76	13.55	15.20	19.20	-1.02	-3.42	-0.67	-0.30	1.74	-7.00	-1.22	4.00	5.08	5.56	4.10	0.48	-2.70	
USA 024	East Coast	15.28	19.84	15.17	16.90	16.31	19.98	8.53	16.66	15.70	18.70	4.56	-0.11	1.62	1.03	4.70	-6.74	1.38	3.00	5.99	5.39	3.90	-0.60	-1.60	
USA 025	East Coast	16.00	14.02	11.52	14.22	14.55	16.66	7.80	13.69	16.10	21.70	-1.99	-4.49	-1.78	-1.46	0.66	-8.21	-2.32	5.60	4.36	4.78	4.00	0.42	-1.50	
USA 026	East Coast	16.97	27.25	19.82	20.30	18.55	24.19	9.48	20.43	17.00	23.60	10.27	2.84	3.33	1.57	7.21	-7.50	3.45	6.60	3.70	4.09	4.20	0.39	-0.90	
USA 027	East Coast	16.94	16.81	13.27	15.51	15.39	18.25	8.15	15.11	17.20	21.40	-0.12	-3.66	-1.43	-1.54	1.32	-8.78	-1.82	4.20	4.07	5.95	4.70	1.88	-0.60	
USA 028	East Coast	17.11	21.83	16.42	17.81	16.91	21.11	8.79	17.67	17.30	21.80	4.72	-0.69	0.71	-0.20	4.00	-8.32	0.56	4.50	4.34	4.22	4.10	-1.22	-1.20	
USA 029	East Coast	17.48	17.52	13.72	15.84	15.61	18.66	8.24	15.47	17.60	22.50	0.04	-3.76	-1.64	-1.87	1.18	-9.24	-2.01	4.90	4.99	4.31	3.90	-0.68	-1.20	
USA 030	East Coast	17.97	14.05	11.53	14.24	14.56	16.68	7.80	13.70	18.20	19.30	-3.92	-6.44	-3.73	-3.41	-1.29	-10.17	-4.27	1.10	4.76	3.66	3.80	-1.10	-1.80	
USA 031	East Coast	18.39	28.44	20.56	20.85	18.91	24.86	9.63	21.03	18.40	16.30	10.04	2.17	2.46	0.51	6.47	-8.77	2.64	-2.10	4.01	5.89	4.60	1.88	-0.70	
USA 032	East Coast	18.31	21.00	15.90	17.43	16.66	20.64	8.68	17.25	18.70	21.50	2.69	-2.42	-0.88	-1.65	2.32	-9.63	-1.07	2.80	4.14	3.93	6.00	-0.21	0.50	
USA 033	East Coast	19.10	21.63	16.29	17.72	16.85	20.99	8.76	17.56			2.52	-2.81	-1.38	-2.26	1.89	-10.34	-1.54		4.92	3.84		-1.08		
USA 035	East Coast	20.32	30.80	22.04	21.94	19.62	26.21	9.93	22.23	10.48		10.48	1.73	1.62	-0.70	5.89	-10.39	1.92		3.70	3.91		0.21		
USA 036	East Coast	21.34	31.71	22.61	22.36	19.90	26.73	10.04	22.70	10.37		10.37	1.28	1.02	-1.44	5.39	-11.29	1.36		4.09	4.53		0.44		
USA 037	East Coast	22.68	10.24	9.14	12.49	13.40	14.51	7.32	11.76	-12.44	-13.54	-10.19	-9.28	-8.17	-15.36	-10.92			6.01	5.64		-0.37			
USA 039	East Coast	23.04	22.67	16.94	18.20	17.16	21.58	8.89	18.09	-0.37		-0.37	-6.10	-4.84	-5.88	-1.46	-14.15	-4.95		4.49	3.12		-1.37		
USA 040	East Coast	23.01	26.96	19.63	20.17	18.46	24.02	9.44	20.28			3.95	-3.37	-2.84	-4.55	1.01	-13.57	-2.73		7.09	5.58		-1.51		
USA 041	Southwest	21.02	24.42	18.04	19.01	17.69	22.58	9.12	18.99	21.02	9.27	3.40	-2.98	-2.02	-3.33	1.56	-11.91	-2.04	-11.76	5.17	4.57	0.29	-0.60	-4.88	
USA 042	Southwest	15.25	17.23	13.53	15.70	15.52	18.49	8.20	15.33	15.25	-1.07	1.98	-1.72	0.45	0.27	3.24	-7.05	0.08	-16.32	8.00	7.87	6.80	-0.13	-1.20	
USA 043	Southwest	6.88	-6.83	-1.56	4.64	8.24	4.81	5.15	3.07	6.88	-7.73	-13.71	-8.44	-2.23	1.37	-2.07	-1.73	-3.81	-14.61	6.38	5.87	2.90	-0.51	-3.48	
USA 044	Southwest	14.38	-31.35	-16.94	-6.63	0.83	-9.14	2.03	-9.42	14.38	-4.56	-45.73	-31.32	-21.01	-13.55	-23.51	-12.34	-23.80	-18.94	8.09	8.38	7.28	0.29	-0.81	
USA 045	Southwest	16.78	-0.50	2.41	7.55	10.16	8.41	5.95	6.29	16.78	-8.62	-17.28	-14.37	-9.23	-6.62	-8.37	-10.83	-10.48	-25.39	8.50	7.52	4.83	-0.98	-3.67	
USA 046	Southwest	16.66	9.45	8.65	12.13	13.17	14.07	7.22	11.36	16.66	-13.28	-7.20	-8.00	-4.53	-3.49	-2.59	-9.44	-5.29	-29.94	8.36	7.45	4.03	-0.91	-4.33	
USA 047	Southwest	17.99	8.76	8.22	11.81	12.96	13.67	7.13	11.01	17.99	-11.00	-9.23	-9.78	-6.19	-5.03	-4.32	-10.86	-6.98	-28.99	7.63	7.44	5.73	-0.19	-1.90	
USA 048	Southwest	18.06	-10.08	-3.60	3.15	7.26	2.96	4.74	1.42	18.06		-28.14	-21.66	-14.91	-10.80	-15.10	-13.33	-16.65		7.76	8.84		1.08		
USA 049	Southwest	20.47	7.32	7.52	11.15	12.52	12.86	6.95	10.28	20.47	-16.73	-13.15	-13.16	-9.33	-7.95	-7.62	-13.53	-10.19	-37.20	7.29	7.88	0.74	0.59	-6.55	
USA 050	Southwest	21.91	21.54	16.23	17.68	16.82	20.94	8.73	17.52	21.91	-2.30	-0.37													

Appendix E14

METADATA			PROXY ESTIMATED MAAT					ΔT (CALIBRATION EST. MAAT - MAAT)				
Sample	Transect	MAAT	$(C_{12}+aC_{13}+aC_{17})/nC_B$	RIN ₁₇ *	RAN ₁₃	MLR	MBT _{SME}	RIN ₁₇ *	$(C_{12}+aC_{13}+aC_{17})/nC_B$	RAN ₁₃	MLR	MBT _{SME}
Yuqiao	China (Yang, et al. (2021))	12.20	18.49	11.42	17.83	16.11		-0.78	6.29	5.63	3.91	
Daihai	China (Yang, et al. (2021))	4.90	18.92	7.59	16.02	15.92		2.69	14.02	11.12	11.02	
Lake Caizi	China (Yang, et al. (2021))	15.90	18.77	14.06	19.39	17.47		-1.84	2.87	3.49	1.57	
Lake Chagan	China (Yang, et al. (2021))	5.30	18.41	7.35	17.94	18.40		2.05	13.11	12.64	13.10	
Lake Changdang	China (Yang, et al. (2021))	15.30	18.80	15.96	18.72	13.50		0.66	3.50	3.42	-1.80	
Lake Chao	China (Yang, et al. (2021))	16.00	19.58	12.05	19.63	17.92		-3.95	3.58	3.63	1.92	
Lake Dongping	China (Yang, et al. (2021))	13.90	19.31	17.64	18.23	9.81		3.74	5.41	4.33	-4.09	
Lake Hengshui	China (Yang, et al. (2021))	13.00	18.08	13.39	17.48	16.98		0.39	5.08	4.48	3.98	
Lake Hongu	China (Yang, et al. (2021))	16.60	19.30	15.76	18.65	14.76		-0.84	2.70	2.05	-1.84	
Lake Hongze	China (Yang, et al. (2021))	14.00	18.89	13.85	18.45	10.82		-0.15	4.89	4.45	-3.18	
Lake Liangzi	China (Yang, et al. (2021))	17.00	18.93	14.22	18.75	16.79		-2.78	1.93	1.75	-0.21	
Lake Longgan	China (Yang, et al. (2021))	16.70	18.91	14.53	19.26	16.57		-2.17	2.21	2.56	-0.13	
Lake Luoma	China (Yang, et al. (2021))	14.20	16.79	14.05	17.70	11.17		-0.15	2.59	3.50	-3.03	
Lake Shijiu	China (Yang, et al. (2021))	15.50	19.19	13.79	18.84	19.14		-1.71	3.69	3.34	3.64	
Lake Weishan	China (Yang, et al. (2021))	13.80	18.53	15.88	17.75	11.29		2.08	4.73	3.95	-2.51	
Lake Wuchang	China (Yang, et al. (2021))	16.00	18.98	13.61	19.69	17.51		-2.39	2.98	3.69	1.51	
Lake Yangcheng	China (Yang, et al. (2021))	15.84	18.51	14.76	17.85	14.33		-1.08	2.67	2.01	-1.51	
Lake Zhengzhu	China (Yang, et al. (2021))	8.30	18.52	9.21	17.93	18.87		0.91	10.22	9.63	10.57	
Wuliangshuai	China (Yang, et al. (2021))	6.10	18.26	13.99	18.47	13.82		7.89	12.16	12.37	7.72	
Yuqiao	China (Yang, et al. (2021))	10.90	18.03	9.05	18.75	19.99		-1.85	7.13	7.85	9.09	
USA 041	Southwestern U.S.	21.02	19.61	10.88	20.03	20.63	20.66	-10.15	-1.41	-1.00	-0.39	-0.36
USA 042	Southwestern U.S.	15.25	14.70	-1.79	18.19	15.28	16.32	-17.04	-0.55	2.94	0.03	1.07
USA 043	Southwestern U.S.	6.88	19.52	3.01	18.05	20.01	16.86	-3.86	12.65	11.18	13.13	9.99
USA 044	Southwestern U.S.	14.38	17.49	12.74	17.75	13.44	17.54	-1.64	3.11	3.37	-0.94	3.16
USA 045	Southwestern U.S.	16.78	12.31	-23.89	16.42	16.91		-40.66	-4.47	-0.36	0.13	
USA 046	Southwestern U.S.	16.66	16.19	8.97	17.70	13.70	17.07	-7.69	-0.46	1.05	-2.96	0.42
USA 047	Southwestern U.S.	17.99	19.24	15.44	17.76	17.80	20.45	-2.55	1.24	-0.23	-0.20	2.46
USA 048	Southwestern U.S.	18.06	17.96	15.09	18.75	18.49	19.57	-2.97	-0.11	0.69	0.42	1.51
USA 049	Southwestern U.S.	20.47	17.73	2.16	17.77	18.86	17.85	-18.31	-2.75	-2.71	-1.62	-2.63
USA 050	Southwestern U.S.	21.91	20.11	8.52	18.59	22.36		-13.39	-1.81	-3.32	0.45	
USA 051	Southwestern U.S.	20.76	18.55	4.75	18.15	19.61	18.32	-16.01	-2.21	-2.62	-1.15	-2.44
USA 052	Southwestern U.S.	21.70	18.69	7.15	17.77	20.14	18.11	-14.55	-3.02	-3.93	-1.56	-3.60
USA 053	Southwestern U.S.	6.08	20.11	6.52	17.44	24.33	15.96	0.44	14.03	11.36	18.25	9.88
USA 054	Southwestern U.S.	14.29	16.20	1.92	15.81	15.09	15.82	-12.37	1.92	1.52	0.81	1.53
USA 055	Southwestern U.S.	15.50	16.07	5.80	15.88	16.05	16.59	-9.69	0.57	0.38	0.55	1.10
USA 056	Southwestern U.S.	15.30	15.15	5.51	14.52	15.35		-9.79	-0.16	-0.78	0.05	
USA 057	Southwestern U.S.	16.12	17.54	5.95	18.17	18.56	16.71	-10.17	1.42	2.05	2.44	0.59
USA 058	Southwestern U.S.	17.81	18.75	9.61	19.38	19.60	19.41	-8.20	0.94	1.57	1.79	1.60
USA 059	Southwestern U.S.	17.77	18.83	7.67	19.12	17.20	17.47	-10.10	1.05	1.34	-0.58	-0.30
USA 060	Southwestern U.S.	18.18	18.21	7.39	18.79	18.30		-10.79	0.03	0.61	0.12	
USA 061	Southwestern U.S.	18.59	18.30	16.63	19.69	18.30	19.30	-1.96	-0.29	1.10	-0.29	0.70
USA 062	Southwestern U.S.	18.69		21.20			19.38		2.51			0.68
USA 063	Southwestern U.S.	20.22		21.20			19.86		0.98			-0.36
USA 064	Southwestern U.S.	20.26	20.28	14.48	19.11	20.76	17.64	-5.79	0.01	-1.16	0.49	-2.62
USA 065	Southwestern U.S.	20.67	20.19	7.59	20.18	20.47	18.56	-13.07	-0.48	-0.49	-0.20	-2.10
USA 066	Southwestern U.S.	20.15	18.84	11.49	20.01	19.75	19.75	-8.66	-1.31	-0.14	-0.40	-0.40

Table E14: Calibration data for Chapter 4

Appendix E15

A

Environmental Variable	R ²	p-value
MAAT	0.09	0.053
MAP	0.59	0.001 ***
VPD	0.29	0.001 ***
Soil pH	0.66	0.001 ***
TSoil	0.07	0.09
SMO	0.02	0.49
ΔT	0.41	0.001 ***
Seasonality	0.19	<0.01 **

	Df	Variance	F	Pr(>F)	
MAAT	1	0.4600	6.7091	0.002	**
MAP	1	3.0057	43.8353	0.001	***
VPD	1	0.1486	2.1672	0.078	.
pH	1	0.3267	4.7643	0.012	*
Tsoil	1	0.1390	2.0276	0.092	.
SMO	1	0.1463	2.1330	0.097	.
Tjja	1	0.0677	0.9867	0.379	
Tdjf	1	0.0718	1.0473	0.341	
Residual	53	3.6342			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

B

	Df	Variance	F	Pr(>F)	
MAP	1	2.8937	40.9843	0.001	***
VPD	1	0.6157	8.7200	0.001	***
pH	1	0.2519	3.5675	0.021	*
Tdjf	1	0.2142	3.0342	0.039	.
Residual	57	4.0245			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

C

Table E15: Statistics associated with PCA and RDA in Chapter 3. **A** Results of permutation analyses used to fit environmental variables to PCA (Figure 3.9). **B** Constrained variance of 3-OH FA proxy distributions amongst U.S. soils explained by environmental variables with significance indicated (Figure 3.10). **C** RDA results following forward selection of significant environmental descriptors on 3-OH FA proxy constrained variance (Figure 3.10).

Appendix E16

3-OH FAs	R^2	p -value	Significance
n10	0.35829247	0.001	***
i11	0.3748584	0.001	***
a11	0.05593344	0.005	**
n11	0.38115354	0.001	***
i12	0.25180285	0.001	***
n12	0.33824821	0.001	***
i13	0.42696983	0.001	***
a13	0.24570564	0.001	***
n13	0.06799457	0.007	**
i14	0.12835617	0.001	***
n14	0.48075236	0.001	***
i15	0.04560976	0.003	**
a15	0.17934673	0.001	***
n15	0.11647083	0.001	***
i16	0.63036666	0.001	***
n16	0.22635586	0.001	***
i17	0.51678946	0.001	***
a17	0.3753597	0.001	***
n17	0.37775294	0.001	***
i18	0.10483256	0.001	***
n18	0.37496381	0.001	***

Table E16: Results of permutation tests used to fit 3-OH FA fractional abundance distributions to NMDS of global soils in Figure 3.14.

Appendix E17

A				
	PC1	PC2	r2	Pr(>r)
tmean	-0.64032	-0.76811	0.2697	0.047 "
tmean_jja	-0.71277	-0.70140	0.2275	0.075 .
tmean_djf	-0.58918	-0.80800	0.2841	0.034 "
pmean	-0.17126	-0.98523	0.3126	0.024 "
water_temp	-0.19014	-0.98176	0.0629	0.533 .
water_ph	-0.98675	0.16224	0.1740	0.153 .
water_depth	0.21480	0.97666	0.0063	0.925 .
ca_sa	-0.42114	0.90700	0.1438	0.189 .
seasonality	0.44400	0.89603	0.2508	0.048 "
sed_ph	0.99999	-0.00382	0.1142	0.258 .
SDT[Lake]	0.86284	-0.50548	0.0542	0.546 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

B				
	PC1	PC2	r2	Pr(>r)
tmean	-0.55854	0.82948	0.3600	0.009 **
tmean_jja	-0.51968	0.85436	0.3987	0.007 **
tmean_djf	-0.58666	0.80983	0.3207	0.013 "
pmean	-0.98402	0.17806	0.0482	0.572 .
vpdmean	0.20983	0.97774	0.0941	0.359 .
soil_temp	-0.28461	0.95864	0.2677	0.041 "
soil_moisture	-0.43004	0.90281	0.0267	0.759 .
delta_t	0.99254	0.12193	0.1777	0.112 .
seasonality	0.71600	-0.69810	0.1395	0.206 .
soil_ph	0.44240	-0.89682	0.1615	0.168 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

C				
	Df	Variance	F	Pr(>F)
tmean	1	1.8948	2.3757	0.017 "
tmean_jja	1	1.0832	1.3581	0.188 .
tmean_djf	1	0.6737	0.8446	0.575 .
pmean	1	1.6347	2.0495	0.049 "
water_temp	1	1.0062	1.2616	0.261 .
water_ph	1	1.3485	1.6907	0.083 .
water_depth	1	0.5415	0.6789	0.730 .
ca_sa	1	0.8262	1.0359	0.399 .
sed_ph	1	0.9615	1.2055	0.277 .
X.ΔT.Lake.	1	0.6607	0.8284	0.544 .
Residual	13	10.3688		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

D				
	Df	Variance	F	Pr(>F)
tmean	1	0.9589	1.2080	0.283 .
tmean_jja	1	0.7982	1.0056	0.431 .
tmean_djf	1	0.7608	0.9584	0.501 .
pmean	1	1.3071	1.6466	0.068 .
vpdmean	1	1.3660	1.7208	0.058 .
soil_temp	1	1.5056	1.8967	0.036 "
soil_moisture	1	1.2581	1.5849	0.125 .
soil_ph	1	1.1070	1.3945	0.185 .
Residual	10	7.9382		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

E				
	Df	Variance	F	Pr(>F)
tmean	1	1.50205	8.6679	0.001 ***
tmean_jja	1	0.65909	3.8034	0.015 "
tmean_djf	1	0.02591	0.1495	0.964 .
pmean	1	0.75879	4.3787	0.010 **
water_temp	1	0.40763	2.3523	0.098 .
water_ph	1	0.20303	1.1716	0.273 .
water_depth	1	0.46152	2.6633	0.062 .
ca_sa	1	0.45888	2.6481	0.067 .
sed_ph	1	0.49654	2.8654	0.062 .
X.ΔT.Lake.	1	0.77381	4.4654	0.023 "
Residual	13	2.25276		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

F				
	Df	Variance	F	Pr(>F)
tmean	1	1.9463	7.7793	0.001 ***
tmean_jja	1	0.7131	2.8501	0.038 "
tmean_djf	1	0.7088	2.8328	0.042 "
pmean	1	0.2845	1.1370	0.321 .
vpdmean	1	0.2409	0.9628	0.409 .
soil_temp	1	0.1545	0.6173	0.640 .
soil_moisture	1	0.0827	0.3305	0.890 .
soil_ph	1	0.1165	0.4655	0.788 .
Residual	15	3.7529		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table E17: Statistics associated with PCA and RDA in Chapter 4. **A** Results of permutation analyses used to fit environmental variables to PCA for SW lake sediments (Figure 4.10 A), and **B** for soils (Figure 4.10 B). **C** Constrained variance of 3-OH FA fractional abundance distributions explained by environmental variables with significance indicated for SW lake sediments (Appendix B4 A), and **D** for SW soils (Appendix B4). **E** Constrained variance of 3-OH FA proxy distributions explained by environmental variables with significance indicated for SW lake sediments (Figure 4.15 A), and **F** for SW soils (Figure 4.15 B).

Appendix E18

PERMUTATION FOR FIG. 4.4				PERMUTATION FOR FIG. 4.19				PERMUTATION FOR FIG.4.20			
3-OH FAs	R ²	p-value	Significance	3-OH FAs	R ²	p-value	Significance	3-OH FAs	R ²	p-value	Significance
n10	0.374908937	0.001	***	n10	0.692258325	0.001	***	n10	0.475419114	0.001	***
i11	0.556436003	0.001	***	i11	0.734733092	0.001	***	i11	0.442871695	0.001	***
a11	0.458399038	0.001	***	a11	0.666945869	0.001	***	a11	0.176886006	0.001	***
u11	0.112797845	0.057		n11	0.711901982	0.001	***	n11	0.456710668	0.001	***
n11	0.533426446	0.001	***	i12	0.654875598	0.001	***	i12	0.545348778	0.001	***
i12	0.519846316	0.001	***	n12	0.64404562	0.001	***	n12	0.527533821	0.001	***
u12	0.220997214	0.008	**	i13	0.47742334	0.001	***	i13	0.501581938	0.001	***
n12	0.239732897	0.003	**	a13	0.557489615	0.001	***	a13	0.536577124	0.001	***
i13	0.494587597	0.001	***	n13	0.449991259	0.001	***	n13	0.289139399	0.001	***
a13	0.515469517	0.001	***	i14	0.273216057	0.002	**	i14	0.712507851	0.001	***
n13	0.287698412	0.002	**	n14	0.619401632	0.001	***	n14	0.21051426	0.001	***
i14	0.543453177	0.001	***	i15	0.000258991	0.992		i15	0.026875793	0.002	**
u14	0.199953828	0.013	*	a15	0.151054844	0.026	*	a15	0.735949136	0.001	***
n14	0.077571316	0.131		n15	0.343887267	0.001	***	n15	0.246932265	0.001	***
i15	0.19863151	0.005	**	i16	0.26979271	0.001	***	i16	0.244626949	0.001	***
a15	0.582059478	0.001	***	n16	0.57005578	0.001	***	n16	0.351323466	0.001	***
n15	0.416750184	0.001	***	i17	0.274710141	0.001	***	i17	0.049120506	0.001	***
i16	0.249248629	0.003	**	a17	0.129194404	0.043	*	a17	0.502591693	0.001	***
u16	0.128571247	0.04	*	n17	0.45861245	0.001	***	n17	0.071021075	0.001	***
n16	0.456462955	0.001	***	i18	0.205491854	0.006	**	i18	0.054525744	0.001	***
i17	0.212033951	0.006	**	n18	0.514607639	0.001	***	n18	0.604945421	0.001	***
a17	0.466168865	0.001	***								
n17	0.157973926	0.029	*								
i18	0.469998232	0.001	***								
u18	0.246279628	0.008	**								
n18	0.073280713	0.165									

Table E18: Permutation tests for Chapter 4 NMDS ordination. Left: Fig.4.4. Middle: Fig.4.19. Right: Fig.4.20.

Appendix E19

PERMUTATIONS FOR FIG.5.1				PERMUTATIONS FOR FIG.5.4				PERMUTATIONS FOR APPENDIX C8				PERMUTATION FOR FIG.5.8			
3-OH FAs	R ²	p-value	Significance	3-OH FAs	R ²	p-value	Significance	3-OH FAs	R ²	p-value	Significance	3-OH FAs	R ²	p-value	Significance
n10	0.3781466	0.001 ***		n10	0.47137209	0.001 ***		n10	0.4490514	0.001 ***		n10	0.84905213	0.001 ***	
i11	0.74943566	0.001 ***		i11	0.65112348	0.001 ***		i11	0.49146643	0.001 ***		i11	0.85002387	0.001 ***	
a11	0.38279249	0.001 ***		a11	0.60340617	0.001 ***		a11	0.18628093	0.001 ***		a11	0.84386068	0.001 ***	
u11	0.08043521	0.139		u11	0.06026891	0.053 *		n11	0.51178603	0.001 ***		n11	0.85026627	0.001 ***	
n11	0.36641515	0.001 ***		n11	0.47431232	0.001 ***		i12	0.17592368	0.001 ***		i12	0.68050179	0.001 ***	
i12	0.0753946	0.141		i12	0.11514699	0.001 ***		n12	0.60993524	0.001 ***		n12	0.42952007	0.004 **	
u12	0.61421554	0.001 ***		u12	0.31020008	0.001 ***		i13	0.30611936	0.001 ***		i12	0.85300381	0.001 ***	
n12	0.48434349	0.001 ***		n12	0.41549329	0.001 ***		a13	0.49252348	0.001 ***		i13	0.63056289	0.001 ***	
i13	0.50399455	0.001 ***		i13	0.55520985	0.001 ***		n13	0.21589409	0.001 ***		a13	0.66557273	0.001 ***	
a13	0.07925674	0.159		a13	0.19860007	0.001 ***		i14	0.71086273	0.001 ***		n13	0.53112028	0.001 ***	
n13	0.09012666	0.115		n13	0.15063442	0.001 ***		n14	0.10425113	0.001 ***		i14	0.13229223	0.147	
i14	0.62562693	0.001 ***		i14	0.43808957	0.001 ***		i15	0.0663279	0.001 ***		n14	0.45543332	0.001 ***	
u14	0.44540576	0.001 ***		u14	0.25902002	0.001 ***		a15	0.77328515	0.001 ***		n14	0.03563545	0.6	
n14	0.04917754	0.344		n14	0.03933181	0.171		n15	0.14756405	0.001 ***		i15	0.47059288	0.002 **	
i15	0.14493179	0.026 *		i15	0.06009244	0.071 *		i16	0.23311071	0.001 ***		a15	0.59917154	0.001 ***	
a15	0.81956149	0.001 ***		a15	0.71701133	0.001 ***		n16	0.33247184	0.001 ***		n15	0.25164321	0.033 *	
u15	0.03481326	0.486		u15	0.02168281	0.353		i17	0.06763679	0.001 ***		i16	0.41788243	0.003 **	
n15	0.12975635	0.044 *		n15	0.01482345	0.483		a17	0.53117669	0.001 ***		u16	0.37552531	0.005 **	
i16	0.18610554	0.006 **		i16	0.10855077	0.005 **		n17	0.04435319	0.001 ***		n16	0.68791331	0.001 ***	
u16	0.03368054	0.471		u16	0.00293529	0.901		i18	0.10773737	0.001 ***		i17	0.75112595	0.001 ***	
n16	0.3330195	0.001 ***		n16	0.51302475	0.001 ***		n18	0.69726154	0.001 ***		a17	0.77951685	0.001 ***	
i17	0.24512177	0.002 **		i17	0.36578274	0.001 ***						n17	0.16108551	0.097	
a17	0.58491391	0.001 ***		a17	0.45602561	0.001 ***						i18	0.5608761	0.001 ***	
n17	0.28499251	0.001 ***		n17	0.1427795	0.001 ***						u18	0.13932704	0.13	
i18	0.40970656	0.001 ***		i18	0.55460527	0.001 ***						n18	0.51898278	0.001 ***	
u18	0.36515661	0.001 ***		u18	0.31504671	0.001 ***									
n18	0.22929016	0.002 **		n18	0.30786669	0.001 ***									

PERMUTATION FOR FIG.5.8			
Environmental	R ²	p-value	Significance
age	0.79851045	0.001 ***	
alkalinity	0.40187476	0.002 **	
NO3	0.11285035	0.202	
NH4	0.09556917	0.274	
top	0.02795094	0.686	
SRP	0.32190987	0.011 *	
wg_SRP	0.22176776	0.041 *	
sio2	0.03300127	0.656	
CHLa	0.02334384	0.723	
Twater	0.40218091	0.003 **	
Water pH1	0.00506631	0.939	

Table E19: Permutation tests for Chapter 5 NMDS ordination. From left to right: (i) 3-OH FA fractional abundance permutation Fig. 5.1. (ii) Fig.5.4, (iii) Appendix C8, (iv) Fig 5.8. (v) environmental variables permutation for Esthwaite Water lacustrine sequence, Fig.5.8