# New North Atlantic Palaeo-Temperature Reconstruction from Terrestrial Sedimentary Archives: Interrogating Climates of the Past with Tools of Today

Matthew Allison

A thesis submitted to the University of Birmingham

for the degree of DOCTOR OF PHILOSOPHY

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

August 2025

# *Declaration of Authorship*

I, Matthew S. Allison, declare that this thesis titled, "New North Atlantic Palaeo-Temperature Reconstruction from Terrestrial Sedimentary Archives: Interrogating Climates of the Past with Tools of Today" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Date: 29$^{th}$ August 2025

# *Abstract*

This thesis investigates the preservation, reliability, and application of lipid biomarker archives for reconstructing past climates, with a focus on the late Miocene–early Pliocene Barmur Group (Tjörnes Peninsula, NE Iceland). First, an updated stratigraphic and structural model of the Barmur Group is developed using drone-based virtual outcrop mapping, refining the composite thickness to 442 m. This revised framework places further constraints on the age of the first opening of the Bering Strait, placing it c. 5.0 Ma, establishing a robust foundation for subsequent palaeoenvironmental reconstructions. Second, the Gaussian Process Emulator (GPE) - OPTiMAL, is developed further. The novel Contextualised Distance ($CD$) metric is introduced to quantify the analogue status of ancient biomarker distributions relative to modern calibration datasets, demonstrating that the isoGDGT system appears to satisfy the uniformitarian assumption, despite transient deviations during greenhouse climates. This GPE framework is applied to a 200 Ma record of isoGDGT data, the largest to date. Third, the Barmur Group lignites are interrogated to test the resilience of brGDGTs, a terrestrial biomarker, under high geothermal gradients. Burial and thermal modelling, integrated with vitrinite reflectance and supporting biomarker indices, reveal systematic diagenetic alteration of recovered brGDGTs yet preservation of interpretable palaeoclimatic signals. Finally, a combined machine learning workflow is employed in which a Gaussian Process Classifier predicts depositional setting of brGDGT assemblages, enabling targeted palaeotemperature reconstructions with a depositionally-specific GPE. These models reduce predictive uncertainty relative to conventional MBT/CBT indices and provide realistic estimates of warmest month air temperature and in situ water temperature, consistent with independent palaeoclimate proxies and regional records. Together, this thesis advances the methodological foundations of biomarker-based palaeothermometry and demonstrates the potential of Iceland's Barmur Group to bridge laboratory-scale experiments with geological archives of climate change.

*Skeifárfoss - Iceland. Best enjoyed with a pastry after a long morning's fieldwork.*

# *Acknowledgements*

I would like to start by thanking my supervisory team, Steve Jones and James Bendle. I signed up to a Msci project many years ago because it looked interesting and also offered a chance to travel to Iceland. That decision set me firmly on the trajectory to completing this PhD, a trajectory that has been full of adventure, travel, learning, and at times quite a lot of effort! Both my supervisors offered me a tremendous degree of trust and freedom during the course of my PhD, for which I am very grateful. When I reflect on the last few years of academic effort, they were present to support, nudge in the correct direction, and encourage me to arrive at this thesis. Thank you both.

I started my PhD during COVID. The project had many opportunities to be a lonely one, disconnected from those embarking on similar academic exploits. However, I feel very fortunate in the company I shared these last few years. Between fellow PhD students and post docs, office/desk/lab-mates, I have never been short of people to talk to and enjoy the company of. Some of my fonder memories of my PhD will be of Friday morning coffee, the yearly Burn's Night knees-up, and the general camaraderie that emerged from this shared experience. Thank you to everyone who has put up with me for the last few years, and thank you for enduring my need to ask questions such as "can you laminate a slice of ham?" or "if this computer was a type of cheese, what should it be?".

Thank you to my lovely wife Bryony, I hope very shortly to stop being a student! You are a wonderfully supportive person and a first-rate encourager and enabler. The PhD has been a long slog and, at times, a consuming project to commit myself to. Thank you for your patience and support throughout this chapter of my life, I love you.

This work was funded by the Natural Environment Research Council (NERC) via the DTP - Central England NERC Training Alliance (CENTA). Thank you for taking a punt on me as a PhD student! Thank you also to Diet Coke, your contribution to this thesis is somewhat immeasurable.

# Contents

# List of Figures

# List of Tables

# List of Equations

$$TEX_{86} = \frac{GDGT\ 2 + GDGT\ 3 + cren'}{GDGT\ 1 + GDGT\ 2 + GDGT\ 3 + cren'}$$

$$TEX_{86} = 0.015 \times SST + 0.28$$

$$CD_n = \sqrt{\sum_{f=1}^{6} \left(\sigma_f \left(d_f - a_f^{(n)}\right)\right)^2}$$

$$MBT' = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIb + IIc + IIIa + IIa' + IIb' + IIc' + IIIa'}$$

$$CBT = -log(\frac{Ib + IIb + IIb'}{Ia + IIa + IIa'})$$

$$MBT'_{5ME} = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIb + IIc + IIIa}$$

$$CBT_{5ME} = -log(\frac{Ib + IIb}{Ia + IIa})$$

$$CPI = \frac{1}{2}(\frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{24} + C_{26} + C_{28} + C_{30} + C_{32}} + \frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{26} + C_{28} + C_{30} + C_{32} + C_{34}})$$

$$Geohopanoid\ Index = \frac{\beta\beta}{\alpha\beta + \beta\beta}$$

$$k = Ae^{\frac{-E_a}{RT}}$$

# List of Abbreviations

- **NVZ** = Northern Volcanic Zone
- **TFZ** = Tjörnes Fraction Zone
- **GOR** = Grímsey Oblique Rift
- **ER** = Eyjafjarðaráll Rift
- **HFF** = Húsavík-Flatey Fault
- **KB** = Kolbiensey Ridge
- **NADW** = North Atlantic Deep Water
- **AMOC** = Atlantic Meridional Overturning Circulation
- **ES21** = Eiríksson and Símonarson (2021)
- **SfM** = Structure from Motion
- **DJI** = Da-Jiang Innovations™
- **GPS** = Global Positioning System
- **DEM** = Digital Elevation Model
- **SST** = Sea Surface Temperature [°C]
- **WMT** = Warmest Month air Temperature
- **MAAT** = Mean Annual Air Temperature
- **ISWT** = In Situ Water Temperature
- **MMCO** = Middle Miocene Climate Optimum
- **O/M T** = Oligocene/Miocene Transition
- **E/O T** = Eocene/Oligocene Transition
- **MECO** = Middle Eocene Climate Optimum
- **EECO** = Early Eocene Climate Optimum
- **PETM** = Palaeocene-Eocene Thermal Maximum
- **K-Pg** = Cretaceous-Palaeogene boundary
- **OAE** = Ocean Anoxic Event
- **T-OAE** = Torcian-OAE
- **ODP** = Ocean Drilling Project
- **isoGDGT** = isoprenoidal Glycerol Dialkyl Glycerol Tetraethers
- **$TEX_{86}$** = index (TetraEther indeX of tetraethers containing 86 carbon atoms
- **BIT** = Branched and Isoprenoid Tetraether index (Hopmans et al., 2004)
- **MI** = Methane Index (Zhang et al., 2011)
- **$\Delta$RI** = Ring Index (Zhang et al., 2016)
- **OPTiMAL** = Optimised Palaeothermometry from Tetraethers via MAchine Learning
- ***CD*** = Contextualised Distance
- **$D_{Nearest}$** = Screening metric, as defined in Dunkley Jones et al. (2020)
- **brGDGT** = branched Glycerol Dialkyl Glycerol Tetraethers
- **MBT** = Methylation of Branched Tetraethers
- **CBT** = Cyclisation of Branched Tetraethers
- **CPI** = Carbon Preference Index
- **$UK'_{37}$** = C37 alkenone unsaturation index
- **k–Ar** = Potassium – Argon
- **$^{40}Ar/^{39}Ar$** = $^{40}$Argon and $^{39}$Argon
- **$\delta^{18}O$** = Ratio of oxygen isotopes $^{18}$O and $^{16}$O relative to a global standard
- **TOM** = Total Organic Matter
- **TOC** = Total Organic Carbon
- **TLE** = Total Lipid Extract
- **PTFE** = Polytetrafluoroethylene
- **NMR** = Nuclear Magnetic Resonance
- **HPLC-MS** = High Performance Liquid Chromatography - Mass Spectrometry
- **GC-MS** = Gas Chromatography - Mass Spectrometry
- **GC-FID** = Gas Chromatography - Flame

Ionisation Detection

- **RMSE** = Root Mean Squared Error
- **95$^{th}$%** = 95$^{th}$ percentile error
- **GP** = Gaussian Process
- **GPC** = Gaussian Process Classification
- **GPE** = Gaussian Process Emulator
- **GPR** = Gaussian Process Regression
- **RBF** = Radial Basis Function
- **ARD** = Automatic Relevance Determination
- **ODE** = Ordinary Differential Equations

# 1 | Overview: Thesis Objectives and Aims

This thesis strives to one aim: to investigate the palaeoclimatic signal that may or may not be archived within the brGDGT lipid biomarkers residing in the organic rich, lignite horizons of the Barmur Group, NE Iceland (Fig. 1.1). Biomarkers represent biologically derived compounds capable of persisting into the geological record. Observations of environmental dependencies that govern the production of modern biomarkers enables the reconstruction of past environments from equivalent biomarkers archived in the geological record. This study focusses on two microbial biomarker systems, isoprenoidal Glycerol Dialkyl Glycerol Tetraethers (isoGDGT) and branched Glycerol Dialkyl Glycerol Tetraethers (brGDGT). isoGDGTs are produced in the oceanic realm and subsequently archived in marine sediments, whereas brGDGTs are produced in terrestrial settings and found in a range of depositional environments including soils, peatlands, and lakes. In response to changes in the environment inhabited by microbial communities a range of lipid structures are synthesised to modulate the physical structure of the cell membrane. This changing record of lipid distributions can be interpreted as a record of variable environmental conditions. Following microbial cell death, the iso-/br-GDGTs are deposited in the surrounding sediment and eventually interned in the geological record. These archived lipids can be subsequently recovered, analysed, and used to reconstruct past environmental conditions, a series of measurements resulting a record of palaeoclimatic change.

However, whilst brGDGTs appear to possess the chemical stability to persist unaltered


Figure 1.1: Coastal exposure of the Barmur Group (NE Iceland) rising c. 50 m above the typical black beaches of Iceland.

throughout stretches of geological time, when attempting to reconstruct past climates it is crucial to understand how this (and other) biomarkers respond to forces such as temperature, pressure and time that dominate geological settings. To conduct such a study, one would need a package of organic (biomarker) rich sediment that had experienced enough heat to progress degradation of lipid biomarkers, but not so much that they are unrecoverable. A sedimentary package that was accessible to study, as opposed to buried at great depth. And lastly, a package understood in enough detail for stratigraphy, age, and burial history to be confidently modelled. The Barmur Group represents just such a geological curiosity. A relatively young sedimentary sequence on an island otherwise dominated by primary igneous geology, where sediments have received high levels of geothermal heating from shallow burial - associated with proximity to Iceland's Northern Volcanic Zone. The shallow marine sediments of the Barmur

Group are punctuated by horizons of lignite, a thermally immature coal rich in biomarkers.

Recent stratigraphic and age modelling (Chapter 2) provides the foundation on which accurate palaeoclimatic predictions can be made. An updated methodology/machine learning approach to the modelling and prediction of environmental information from biomarkers is developed and assessed (Chapter 3), providing the toolset with which to analyse the data generated from Barmur Group lignites. The updated stratigraphic framework for the Barmur Group supports burial/thermal modelling of the sedimentary sequence, groundtruthed in onshore drill-records (Chapter 4). Finally, application of the new machine learning approach to brGDGT data of the Barmur Group lignites in combination with burial, stratigraphic, and age models permits an holistic understanding of brGDGT preservation in geological archives (Chapter 4). Insights from this study enable, for the first time, a correlation between laboratory scale pyrolysis experiments and observations from the geological record. The Tjörnes Peninsula and Barmur Group sediments are described here as a "natural laboratory", one which has curated the conditions for accelerated geochemical processes and accessible by foot in coastal exposure. The Tjörnes Peninsula's unique history of sedimentological, palaeontological, radiometric, and resource exploration provide the necessary research platform to achieve the aim of this thesis.

# 2 | Updated stratigraphy and structure of the Barmur Group (Tjörnes beds, Iceland) and implications for the opening of the Bering Strait

*The contents of this chapter are the work of this author. Fieldwork and data collection were supported by Jonathan Hall and fieldwork funding was supported by the School of Geography, Earth and Environmental Sciences, University of Birmingham. The results of this work can be found published in Allison et al. (2026), co-author contributions are provided here in CRediT (Contributor Roles Taxonomy) format:* ***Matthew Allison****: conceptualization (lead), data curation (lead), formal analysis (equal), funding acquisition (equal), investigation (lead), methodology (equal), software (lead), visualization (lead), writing–original draft (lead);* ***Jonathan Hall****: funding acquisition (equal), investigation (equal), methodology (equal), writing–review & editing (equal);* ***Max Papadopoulos****: investigation (equal), methodology (equal);* ***Jón Eiríksson****: writing–review & editing (supporting);* ***Stephen Jones****: conceptualization (equal), funding acquisition (equal), supervision (lead), writing–review & editing (equal).*

## 2.1 Abstract

The Barmur Group (Tjörnes beds) sedimentary sequence onshore Iceland records the first opening of the Bering Strait oceanic gateway, and late Miocene–early Pliocene cooling that preceded intensification of the Northern Hemisphere glaciation. The stratigraphy of the Barmur Group is not known in detail because of difficulties in correlating coastal and inland exposures. To address this issue, we carried out remotely piloted vehicle (drone-based) mapping of the entire 11 km outcrop. We correlated the processed 'virtual outcrop' image with previously logged sections and constructed a 3D structural model. The modelled composite thickness of the Barmur Group is 442 m, which is 16% thinner than previously estimated. This discrepancy reflects earlier difficulties in correlating coastal and inland exposures and the recognition of faulting at the base of the succession. We recently showed that the bioevent signalling the first Pacific-to-Atlantic faunal migration through the Bering Strait occurred during magnetic chrons C3n.4n – C3n.3r (5.235 – 4.896 Ma); the new model suggests that it occurred near the chron boundary (c. 5.0 Ma). Thus, we show that a 3D virtual outcrop is a useful tool to quantify uncertainty when assembling a composite section. We also provide the complete workflow used to plan, acquire and process the virtual outcrop.

## 2.2 Introduction

The Barmur Group (informally Tjörnes beds) is a sedimentary succession of latest Miocene to early Pliocene age exposed onshore in NE Iceland (Fig. 2.1) (Hall et al., 2023). The succession is internationally important for reconstructing the opening of the Bering Strait oceanic gateway because it contains the only well-dated record of the initial migration of Cenozoic marine fauna from the Pacific to the Atlantic (Marincovich and Gladenkov, 1999; Eiríksson and Símonarson, 2021). The Bering Strait is a nar-

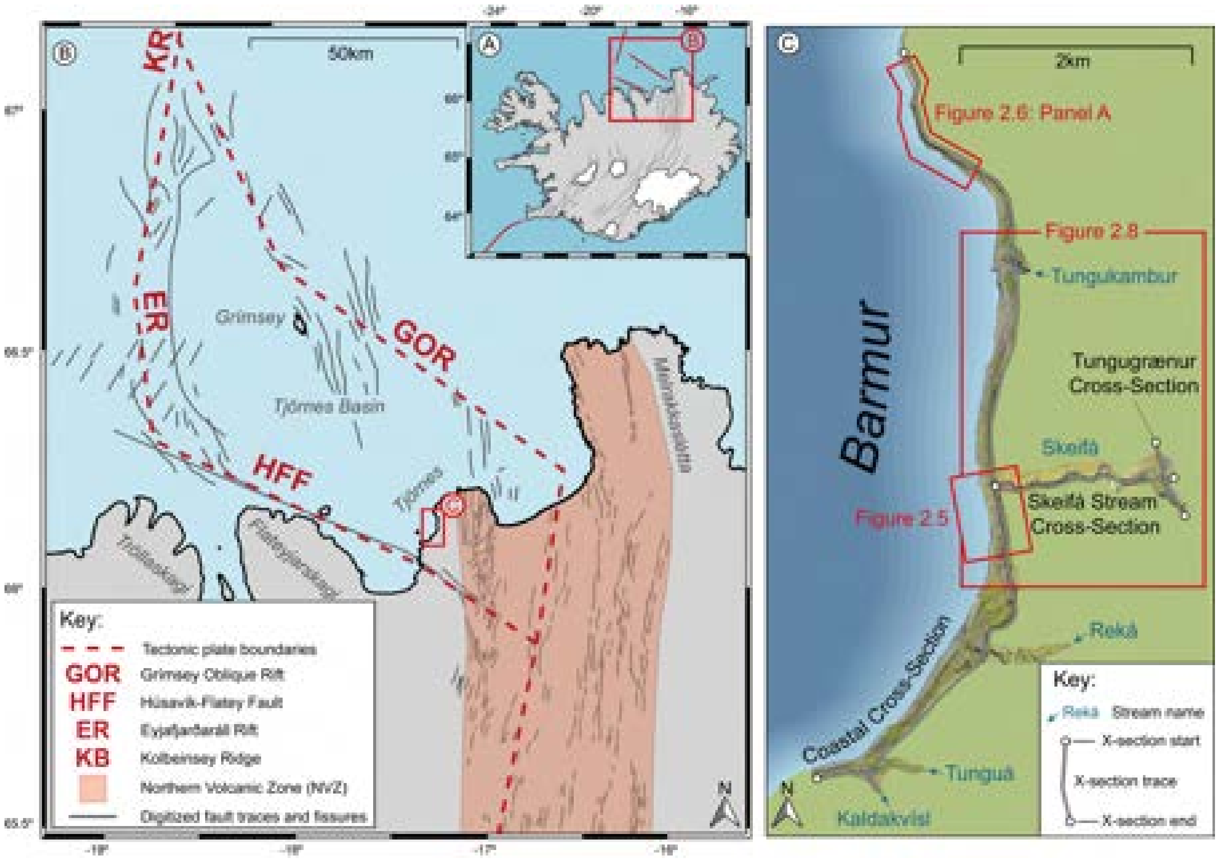


Figure 2.1: Location of the study area. (a) Iceland. (b) Regional structural overview of the Northern Volcanic Zone (NVZ) and the Tjörnes Fraction Zone (TFZ), which have generated the accommodation space for the deposition of the Barmur Group. Fault traces accessible through Landælingar Íslands. (c) Tjörnes Peninsula and extent of the drone-based, virtual outcrop model that covers the entire coastal extent of the Barmur Group. Traces of three cross-sections are shown: the coastal cross-section, the inland Tungugrænurr cross-section and the Skeifá Stream cross-section which connects the other two. The locations shown in Figures 2.5, 2.6 and 2.8 are marked with red boxes. Source: fault traces, National Land Survey of Iceland.

row, shallow gateway for flow between the Pacific and Arctic oceans. An open Bering Strait allows relatively fresh Pacific water into the Arctic and North Atlantic (Woodgate and Aagaard, 2005). Earth system modelling studies suggest that the opening of the Bering Strait after a period of closure might reduce the strength of the North Atlantic Deep Water (NADW), lower high-latitude Atlantic surface temperatures and promote growth of Northern Hemisphere ice sheets (Shaffer and Bendtsen, 1994; Hu et al., 2010; Brierley and Fedorov, 2016). These changes might also impact Southern Hemisphere climate, since the NADW is an integral component of the Atlantic Meridional

Overturning Circulation (AMOC) (Toggweiler and Key, 2001).

The Bering Strait was closed from the Late Cretaceous to the Late Cenozoic (Fiorillo, 2008; Símonarson and Eiríksson, 2021). During this period, distinct mollusc assemblages evolved in the northern Atlantic and Pacific oceans (Gladenkov et al., 2002). The subsequent mixing of these assemblages, as observed in the Barmur Group and a handful of other sedimentary successions bordering the Arctic, is taken to indicate the first Cenozoic opening of the Bering Strait (Fig. 2.2) (Bárdarson, 1925; Strauch, 1963; Beets, 1992; Marincovich, 2000; Gladenkov et al., 2002, 2004; Eiríksson and Sí-

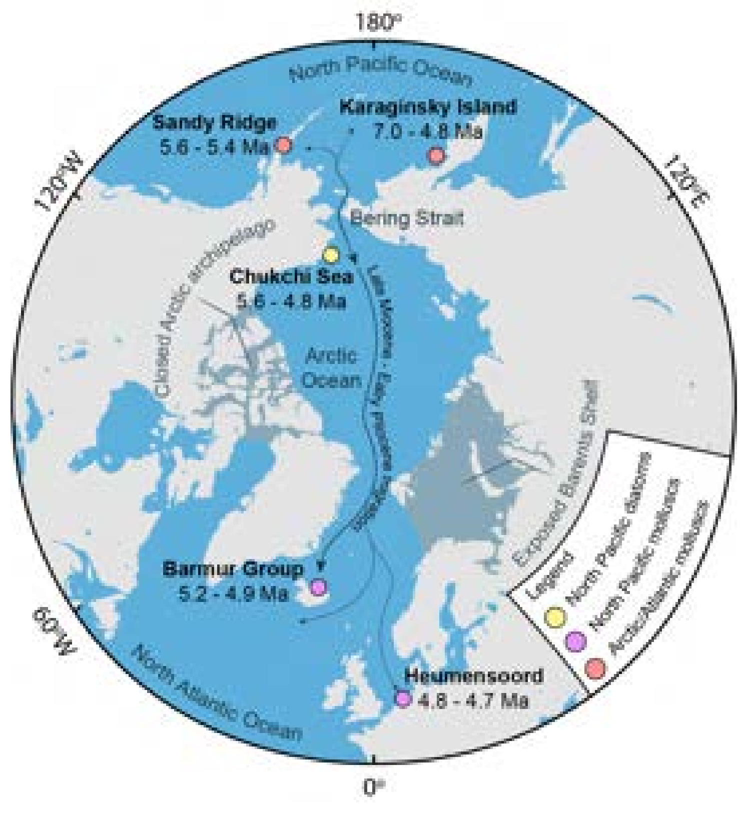


Figure 2.2: Summary of the late Neogene trans-Arctic molluscan migration model. A single-stage, bidirectional faunal exchange event (blue arrows) driven by the opening of the Bering Strait is predicted to have allowed Pacific molluscs to arrive in Iceland c. 5.2–4.9 Ma. Also shown are the late Miocene/early Pliocene Pacific diatoms in the Chukchi Sea and strontium dating of Pacific molluscs preserved at Heumensoord, Netherlands. Source: modified from Hall et al. (2023), Beets (1992), and Gladenkov et al. (2004); continental reconstruction at 5.0 Ma is shown following the tectonic reconstruction of Müller et al. (2019).

monarson, 2021). Until recently, the Pacific-to-Atlantic mixing bioevent within the Barmur Group was dated to the latest Pliocene or uppermost early Pliocene (c. 3.4–3.9 Ma) on the basis of K–Ar radiometric ages determined in the 1970's (Albertsson, 1978; Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Eiríksson and Símonarson, 2021), or 4.5 Ma on the basis of biostratigraphy (Verhoeven et al., 2011). The younger end of this range was used to support a hypothesis that the first opening of the Bering Strait triggered late Pliocene (3.5–3.0 Ma) expansion of Northern Hemisphere icesheets (Brierley and Fedorov, 2016). The older end of this range was used to hypothesize that formation of the East Greenland Current at 4.5 Ma and coeval development of sea ice in the Iceland Sea were triggered by the first opening of the Bering Strait (De Schepper et al., 2015; Clotten et al., 2019). Recently, we obtained the first $^{40}Ar/^{39}Ar$ radiometric ages for the Barmur Group and showed that the first opening of the Bering Strait occurred during magnetic chrons C3n.4n–C3n.3r (5.235–4.896 Ma; Hall et al. (2023)). This result does not support hypotheses of simple, direct links between the first opening of the Bering Strait gateway and Northern Hemisphere icesheet expansion, development of sea ice in the Arctic Ocean and Nordic Seas, or changes to ocean circulation in the Atlantic (Hall et al., 2023). Given the consequences of revising the Barmur Group age model, a main motivation of this study is to examine the uncertainty in the recent new age for the first opening of the Bering Strait in more detail.

The Barmur Group also contains a faunal and geochemical record of regional cooling that preceded intensification of the Northern Hemisphere glaciation (Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Verhoeven et al., 2011). Difficulties in establishing a reliable composite lithostratigraphic section mean that the age model for these records, and even their relative stratigraphic order, is uncertain, which limits their potential to reconstruct past environmental change. Estimates for the total thickness of the Barmur Group have ranged from c. 400 to 570 m (Bárdarson, 1925; Eiríksson, 1981; Buchardt and Símonarson, 2003; Eiríksson and Símonarson, 2021). This range arises because the composite stratigraphic section must be assem-

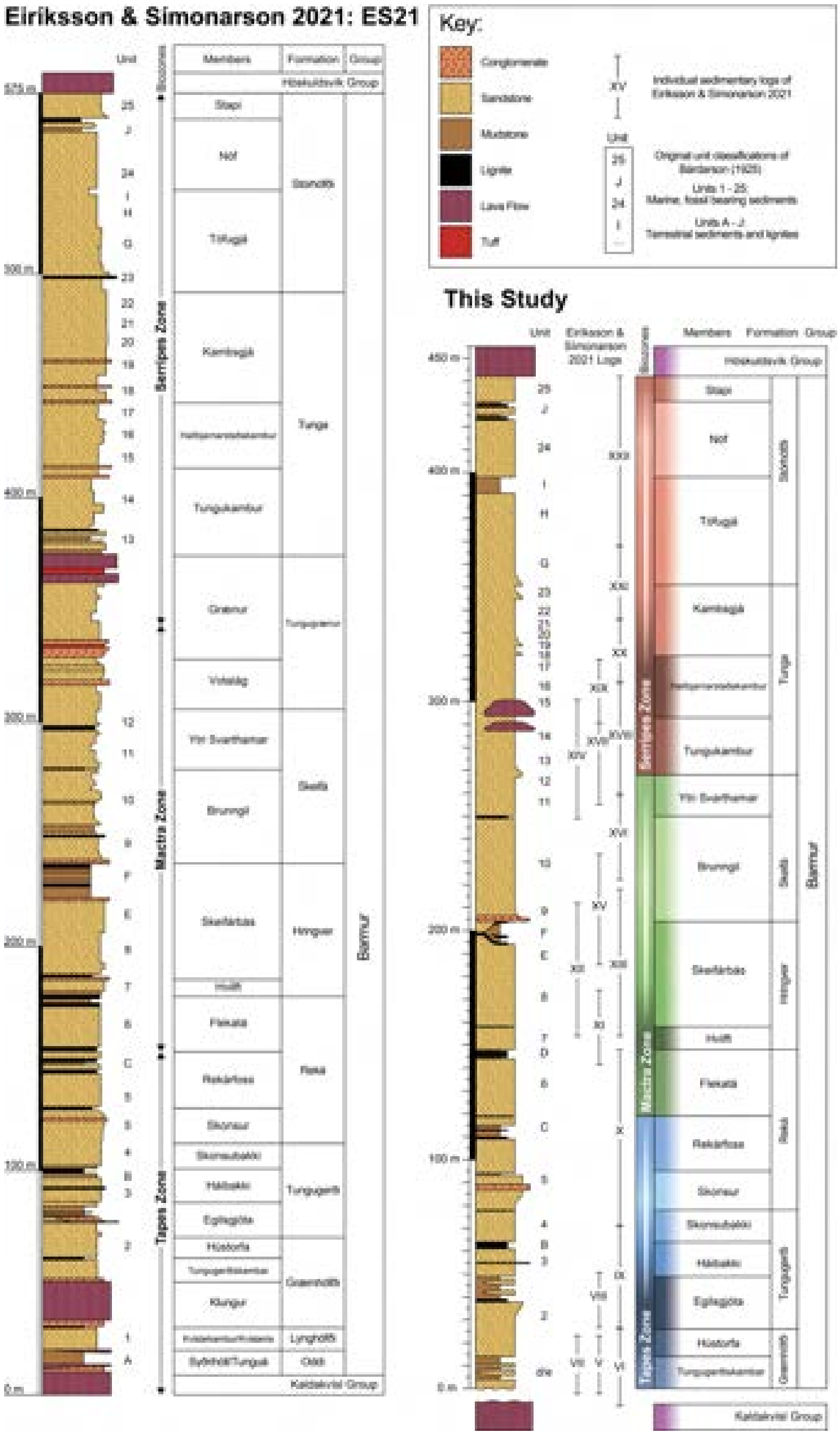


Figure 2.3: Comparison between the stratigraphic column of Eiríksson and Símonarson (2021) and this study. The new stratigraphic column is based on the coastal Barmur Group members as outlined in Eiríksson and Símonarson (2021). The Skeifá Lavas are shown schematically at the stratigraphic height predicted by our modelled stratigraphy. Source: left column, Eiríksson and Símonarson (2021).

bled from multiple logged sections that are not straightforward to correlate (Fig. 2.3). Recently, Eiríksson and Símonarson (2021) carried out a comprehensive lithological logging programme, formalized the stratigraphy of the Barmur Group and estimated a maximum total thickness of 570 m (Fig. 2.3). However, a formal structural framework within which to place the logged successions is still lacking. Consequently, it remains unclear whether the 22 logged successions represent the entire stratigraphy or whether unexposed stratigraphy exists between them.

A related issue is how to correlate the coastal and inland exposures. Most of the Barmur Group is exposed along a 7 km coastal cliff section. A second important exposure – the Tungugrænur exposure – occurs several kilometres inland in the Skeifá river valley (Fig. 2.1c). This exposure is significant because it contains the radiometrically dated Skeifá Lava, which crops out 17 m above the bioevent that signals the first opening of the Bering Strait oceanic gateway (Eiríksson and Símonarson, 2021). However, the inland and the coastal exposures are separated by c. 1.5 km of partly unexposed ground. The Skeifá Lava does not crop out along the coast, where the best magnetostratigraphic record is established (Doell, 1972; Gladenkov and Gurari, 1976; Eiríksson et al., 1990; Kristjánsson, 2004). This apparent lack of lateral continuity makes correlating the sedimentary units challenging. Thus, while the earliest Pliocene radiometric age of the Skeifá Lava is secure, the magnetostratigraphy-based uncertainty range for the faunal exchange event is less secure.

This study aims to increase the value of the Barmur Group environmental change archive by producing the first integrated structural and stratigraphic model. We carried out drone-based mapping of the entire Barmur Group succession, interpreted the resulting virtual outcrop alongside existing lithostratigraphic logs, and iteratively fitted the model to the stratigraphic observations. The new model yields a more reliable composite stratigraphic section. It also enables quantitative assessment of the relationship between the coastal and inland exposures. The results suggest a tightening of our previous estimate of the age range for the Pacific-to-Atlantic faunal exchange bioevent, which signals the first opening of the Bering Strait. The drone-based mapping also highlights the quality of Bárðarson's work, a century after his pioneering study of this rare sedimentary succession onshore Iceland (Bárdarson, 1925).

### 2.2.1 *Geological Background*

The Barmur Group is the lower part of a c. 1200 m thick sequence of latest Miocene to late Pleistocene sedimentary rocks known as the Tjörnes succession that crops out in the northwest of the Tjörnes Peninsula, NE Iceland (Fig. 2.1). The onshore Tjörnes succession is the eastern feather edge of the sedimentary fill of the Tjörnes Basin, which lies predominantly offshore (Rögnvaldsson et al., 1998; Richter and Gunnarsson, 2010), Fig. 2.1b. The deposition of the Tjörnes succession was accommodated by subsidence within the Tjörnes Fracture Zone, which acts as a transform system between two segments of the Mid-Atlantic Ridge plate spreading axis: the Eyjafjarðaráll Ridge offshore to the west, and the Northern Volcanic Zone onshore to the east (Einarsson, 2008). The Tjörnes Fracture Zone has two strands: the Húsavík–Flatey Fault to the south and the Grímsey Oblique Rift to the north. Although the Tjörnes succession accumulated between these major fault zones, it is cut by pervasive smaller faults, which reflect the active tectonic environment and complicate the assembly of a composite stratigraphic section (Eiríksson and Símonarson, 2021).

The Barmur Group overlies the Kaldakvísl Lava Group in the south and is overlain by the Höskuldsvík Lava Group in the north, both of which erupted in sub-aerial environments (Eiríksson and Símonarson, 2021). There is an erosional unconformity between the Kaldakvísl and Barmur Groups that spans 3–4 Ma (Hall et al., 2023). The Barmur Group comprises siltstones, sandstones, conglomerates, lignites and occasional basaltic lavas (Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Eiríksson and Símonarson, 2021). These sediments are interpreted to represent a predominantly marine foreshore to shoreface depositional environment, with multiple fluctuations in relative sea-level to give coastal plain or shallow marine environments (Verhoeven and Louwye, 2013; Eiríksson and Símonarson, 2021). Based on both sedimentology and radiometric dating, the long-term sediment accumulation rate is interpreted to be fairly consistent for the entire Barmur Group (Símonarson and Eiríksson, 2008; Hall et al., 2023). Throughout the Barmur Group, and within the transition to the overlying Höskuldsvík Group, there is abundant evidence for minor hiatuses consistent with the coastal depositional environment, but little evidence for major angular erosional unconformities (Eiríksson and Símonarson, 2021).

The Barmur Group is well exposed along a coastal section 7.2 km long, and more discontinuously along inland stream sections totalling 4 km (Fig. 2.1c). The composite stratigraphy of the Barmur Group was first described by Bárdarson (1925). He recognized 25 marine marker horizons characterized by abundant molluscs (labelled 1–25 from oldest to youngest) and 10 terrestrial marker horizons characterized by lignites (labelled A–J; Fig. 2.3). He also established the Tapes, Mactra and Serripes biozones based on mollusc assemblages within the marker beds. The Tapes and Mactra biozones contain predominantly Atlantic-type molluscs. Abundant Pacific-type molluscs appear suddenly at the base of the Serripes biozone. Thus, the Mactra/Serripes biozone boundary is the bioevent that signals the first opening of the Bering Strait. Bárðarson's original marker bed and biozone classifications have been largely retained in subsequent composite stratigraphies of the Barmur Group (Einarsson et al., 1967; Albertsson, 1978; Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Verhoeven et al., 2011; Eiríksson and Símonarson, 2021). Published palaeontological and palaeo-environmental records from the Barmur Group include Mollusca (Bárdarson, 1925), ostracods (Cronin, 1991), palynomorphs (Verhoeven et al., 2011; Verhoeven and Louwye, 2013), palaeobotany (Denk et al., 2011), foraminifera (Eiríksson and Símonarson, 2021) and stable oxygen isotopes (Buchardt and Símonarson, 2003).

| Stratigraphy | Minimum thickness ES21(m) | Maximum thickness ES21(m) | Thickness range (m) | Modelled thickness (m) | Model/ data fit (m) |
|---|---|---|---|---|---|
| **Stórhöfði Formation** | | | | | |
| Stapi Member | - | 11.1 | - | 11 | -0.1 |
| Nöf Member | - | 31.1 | - | 33 | 1.9 |
| Tófugjá Member | 13.0 | 38.3 | 25.3 | 47 | 8.7 |
| **Tunga Formation** | | | | | |
| Kambsgjá Member | 25.0 | 45.0 | 20.0 | 31 | -14.0 |
| Hallbjarnarstaðakambur Member | 8.5 | 22.5 | 14.2 | 27 | 4.5 |
| Tungukambur Member* | 16.5 | 23 | 6.5 | 25 | 2.0 |
| **Tungugrænur Formation** | | | | | |
| Grænur Member | - | 50.7 | - | - | - |
| Votalág Member | - | 22.1 | - | - | - |
| **Skeifá Formation** | | | | | |
| Ytri Svarthamar Member | 8.0 | 17.0 | 9.0 | 18 | 1.0 |
| Brunngil Member | 27.3 | 30.5 | 3.2 | 46 | 15.5 |
| **Hringver Formation** | | | | | |
| Skeifárbás Member* | 30.0 | 45.0 | 15.0 | 46 | 1.0 |
| Hvilft Member* | 4.3 | 12.9 | 8.6 | 10 | -2.9 |
| **Reká Formation** | | | | | |
| Flekatá Member | 8.6 | 19.4 | 10.8 | 29 | 9.6 |
| Rekárfoss Member | 24.8 | 24.8 | - | 24 | -0.8 |
| Skonsur Member | 17.1 | 17.1 | - | 17 | -0.1 |
| **Tungugerði Formation** | | | | | |
| Skonsubakki Member | - | 12.3 | - | 14 | 1.7 |
| Háibakki Member | - | 15.2 | - | 15 | -0.2 |
| Egilsgjóta Member | 14.7 | 18.0 | 3.3 | 23 | 5.0 |
| **Grænhöfði Formation** | | | | | |
| Hústorfa Member | 6.0 | 15.0 | 9.0 | 12 | -0.3 |
| Tungugerðiskambar Member* | 7.0 | 13.5 | 6.5 | 14 | 0.5 |
| Klungur Member | 9.2 | 28.8 | 21.6 | - | - |
| **Oddi Formation** | | | | | |
| Syðrihóll Member | 4.5 | 18.6 | 14.1 | - | - |
| Tunguá Member | 13.8 | 17.1 | 3.3 | - | - |
| **Lynghöfði Formation** | | | | | |
| Kvíslarós Member | 8.3 | 10.3 | 2.0 | - | - |
| Kvíslarkambur Member | 6.5 | 12.9 | 6.4 | - | - |
| **Totals and means** | 344.7 (387.0) | 484.5 (572.2) | 10.5† | 442 | 4.7† |

Table 2.1: Maximum and minimum thicknesses are logged thicknesses reported by Eiríksson and Símonarson (2021). The thickness range is the difference between the maximum and minimum logged thickness, and shows lateral variability. *Thicknesses of members are observed from base to top of the member in a single log; all other thickness are reconstructed from the lateral correlation of several logs. Model/data fit is the difference between the model thickness and the maximum logged thickness. Total thicknesses are for the conformable, modelled succession (base Tungugerðiskambar Member to top Stapi Member); the aggregate thickness of the entire logged succession is shown in parentheses. †Mean model/data fit.

The most recent and most comprehensive study of the Barmur Group was by Eiríksson and Símonarson (2021), hereafter termed ES21 (Fig. 2.3). They formally defined the lithostratigraphy of the Barmur Group for the first time (it was previously known informally as the Tjörnes beds) and divided it into 10 formations and 23 members on the basis of 22 lithological logs (Fig. 2.3). The aggregate thickness of the ES21 lithostratigraphy is 570 m, as measured between the top of the Kaldakvísl Group and the base of the Höskuldsvík Group (Fig. 2.3). In this study, we aim to model the apparently conformable package of Barmur Group sediments that is exposed along the coastline, from the base of the Tungugerðiskambar Member to the base of the Höskuldsvík Group. The Klungur Member that underlies this package consists of lava of same age as the Kaldakvísl Group lavas (Hall et al., 2023). The two basal formations (Lynghöfði, Oddi) occur in faulted contact with the Kaldakvísl Group and the Barmur Group, and may be lateral equivalents of the lower part of the conformable package of ES21. The aggregate maximum thickness of the conformable package in the ES21 lithostratigraphy is 525 m (Table 2.1).

The Tungugrænur exposure in the upper Skeifá river valley is of key importance in Bering Strait gateway reconstructions because it contains both the Skeifá Lavas (which have been radiometrically dated) and also the Mactra/Serripes biozone boundary bioevent that marks the Pacific-to-Atlantic faunal exchange. In ES21, the entire Tungugrænur exposure is assigned to the Tungugrænur Formation, of logged thickness c. 75 m. However, it is not straightforward to place the Tungugrænur Formation within the composite Barmur Group stratigraphy because the Skeifá Lavas are not exposed at the coast, and it is not clear from their facies whether the sedimentary units within the Tungugrænur Formation are lateral equivalents of those at the coast. ES21 proposed that the entire Tungugrænur Formation correlates with an erosional unconformity within the coastal succession. Therefore, in the ES21 composite stratigraphy, the Tungugrænur Formation is inserted between the Skeifá and the Tunga Formations (Fig. 2.3). Thus, the Mactra/Serripes biozone boundary occurs within a succession of relatively continuous deposition within the inland Tungugrænur exposure and correlates with a putative unconformity within the coastal exposure in ES21.

## 2.3 Investigative Methods

### 2.3.1 *Virtual Outcrop Building*

We aimed to produce a 3D ‘virtual outcrop’ model of the entire Barmur Group coastal and inland exposure. A remotely piloted vehicle (drone) was used to conduct automated Structure-from-Motion (SfM) surveying (Howell et al., 2021). Fieldwork was carried out during September 2021 and September 2022, and approximately 17,000 images of the Barmur Group sediments were acquired during 23 surveying days. We used a DJI Mavic 2 Pro drone and carried out survey planning using DJI’s iPad-hosted Ground Station Pro software (DJI, 2021). Owing to the large geographical extent of the survey, ground control points were not included in the survey workflow. The accuracy of the resulting SfM model is therefore equivalent to the positional accuracy of the drone’s internal GPS, i.e. c. 3 m (Kalacska et al., 2020), which is sufficient for the requirements of this study considering the 11 km extent of our final model. The drone imagery was processed using Agisoft Metashape (version 1.7.5) to produce an SfM photogrammetric model (Agisoft, 2022). This model has resolvable details at 5–10 cm scale across the entire exposure. The final virtual outcrop model was uploaded to the V3Geo cloud-based platform and is freely available in uninterpreted form (Allison et al., 2023a).

Field expedition planning and post-expedition processing were somewhat hampered by limited detailed examples. A fully documented methodology and entire workflow is provided in Appendix A.1.1. Our workflow is divided into three parts: (1) the project-planning; (2) the data-acquisition, including a detailed account of how the photogram-


Figure 2.4: Fig. 4. Virtual outcrop model and structural interpretation. (a) Bárðarson's (1925) original coastal section featuring the 25 marine horizons (Units 1–25) and 10 terrestrial horizons (Units A–J). (b) Modern analogue to Bárðarson's coastal section created using the virtual outcrop model. The 24 key horizons used in the creation of the coastal cross-section are highlighted in this panel. F1–F31, faults featured in the fitting the modelled stratigraphy to the virtual outcrop model. Exposures used in quantifying model fit are highlighted; orange lines represent the back-projection of cross-section stratigraphy onto the virtual outcrop model used in assessing fit. (c) Coastal cross-section presented as part of this study. See Allison et al. (2023b) for a high-resolution, online version of this figure (size 1.6 Gb). Both the virtual outcrop and cross-section are shown at 1:1 vertical exaggeration. Source: (a), modified from Bárdarson (1925).

metric survey was programmed and executed in the field; and (3) the data-processing stage, describing how the SfM virtual outcrop was built from the drone imagery.

### 2.3.2 *Stratigraphic and structural model building*

We aimed to generate a 3D integrated stratigraphical and structural model for the conformable sedimentary part of the Barmur Group (above the Klungur Member) that satisfies both ES21 and also the new virtual outcrop. We used Petroleum Experts Limited MOVE software (version 13.0) (Experts, 2022) to construct the stratigraphic/structural model and compare it quantitatively with our interpreted virtual outcrop. The final MOVE model that best fits the virtual outcrop defines both the best composite stratigraphic section and also the principal structures that deform the regional stratigraphy. The virtual outcrop was interpreted in terms of the ES21 stratigraphic framework. ES21 is based on 22 stratigraphic sections that were logged at centimetre scale (I–XXII in Fig. 2.3). These logs were compared with the virtual outcrop and used to identify a subset of stratigraphic units that are easily recognizable and can be traced for hundreds of metres or more. We interpreted 24 key horizons that span the entire stratigraphic column. These key horizons were traced in MOVE and used to constrain the stratigraphy as viewed in cross-section (Fig. 2.4) and are described in detail in Table 2.2. Multiple faults were also interpreted across the virtual outcrop.

The modelling strategy was to define an initial stratigraphic column spanning the 24 key horizons (we used that of ES21), to project the stratigraphy onto the virtual outcrop in 3D, and to iteratively add faults and adjust the model stratigraphic layer thicknesses until a satisfactory match between the model and the interpreted virtual outcrop was achieved. All these steps were carried out within MOVE, which is explicitly designed to carry them out interactively. The MOVE model can be projected onto the virtual outcrop in 3D, which allows direct comparison with the interpreted marker horizons and faults (Figs. 2.4 and 2.5). The corners where the three main river valley sections (Kaldakvísl,

| **ES21 Formation** | **ES21 Member** | **ES21 Bed** | **Bárðarson's Beds** | **Base or top** |
|---|---|---|---|---|
| Höskuldsvík Lava Group | | | | |
| Stórhöfði | Nöf | XXII:E | J | Base |
| | Nöf | XXII:D | 24 | Base |
| | Tófugjá | XXI:B | 23 | Base |
| Tunga | Hallbjarnarstaðakambur | XVIII:K | 15 | Base |
| | Tungukambur | XVIII:E | 14 | Base |
| Skeifá | Ytri Svarthamar | XVI:I | 11 | Base |
| | Brunngil | XVI:H | 10 | Base |
| | Brunngil | XIII:I | 9 | Base |
| Hringver | Skeifárbás | XIII:H | F | Base |
| | Skeifárbás | XIII:G | E | Base |
| | Skeifárbás | XIII:C | 8 | Base |
| | Hvilft | XI:C | 7 | Base |
| Reká | Flekatá | XI:B | D | Base |
| | Flekatá | X:G | 6 | Base |
| | Rekárfoss | X:F | C | Base |
| Tungugerði | Skonsubakki | IX:I | 4 | Base |
| | Háibakki | IX:H | B | Base |
| | Hústorfa | VI:N | 2 | Base |
| Grænhöfði | Tungugerðiskambar | VI:L | d/e | Top |
| | Tungugerðiskambar | VI:J | d/e | Base |
| | Klungur | VI:G | a/b | Top |
| Oddi | Tunguá | II:F | A | Base |
| Kaldakvísl Lava Group | | | | |

Table 2.2: Idealised stratigraphic model of the Barmur Group. Bed references are provided for both the Bárdarson (1925) classical scheme and the Eiríksson and Símonarson (2021) lithostratigraphic classification.

Skeifá and Hallbjarnarstaðaá) meet the coast were particularly useful for constraining the structure in 3D. The MOVE model can also be compared with the virtual outcrop along representative 2D cross-sections (Fig. 2.5). We show three cross-sections for illustrative purposes: a coastal section that trends roughly north–south along the entire coastal exposure (coastal cross-section, Fig. 2.1), an inland section containing the Skeifá Lava running parallel to the coastline (Tungugrænur cross-section, Fig. 2.1) and a section connecting the inland and coastal exposures oriented perpendicular to the coastline (Skeifá Stream cross-section, Fig. 2.1).

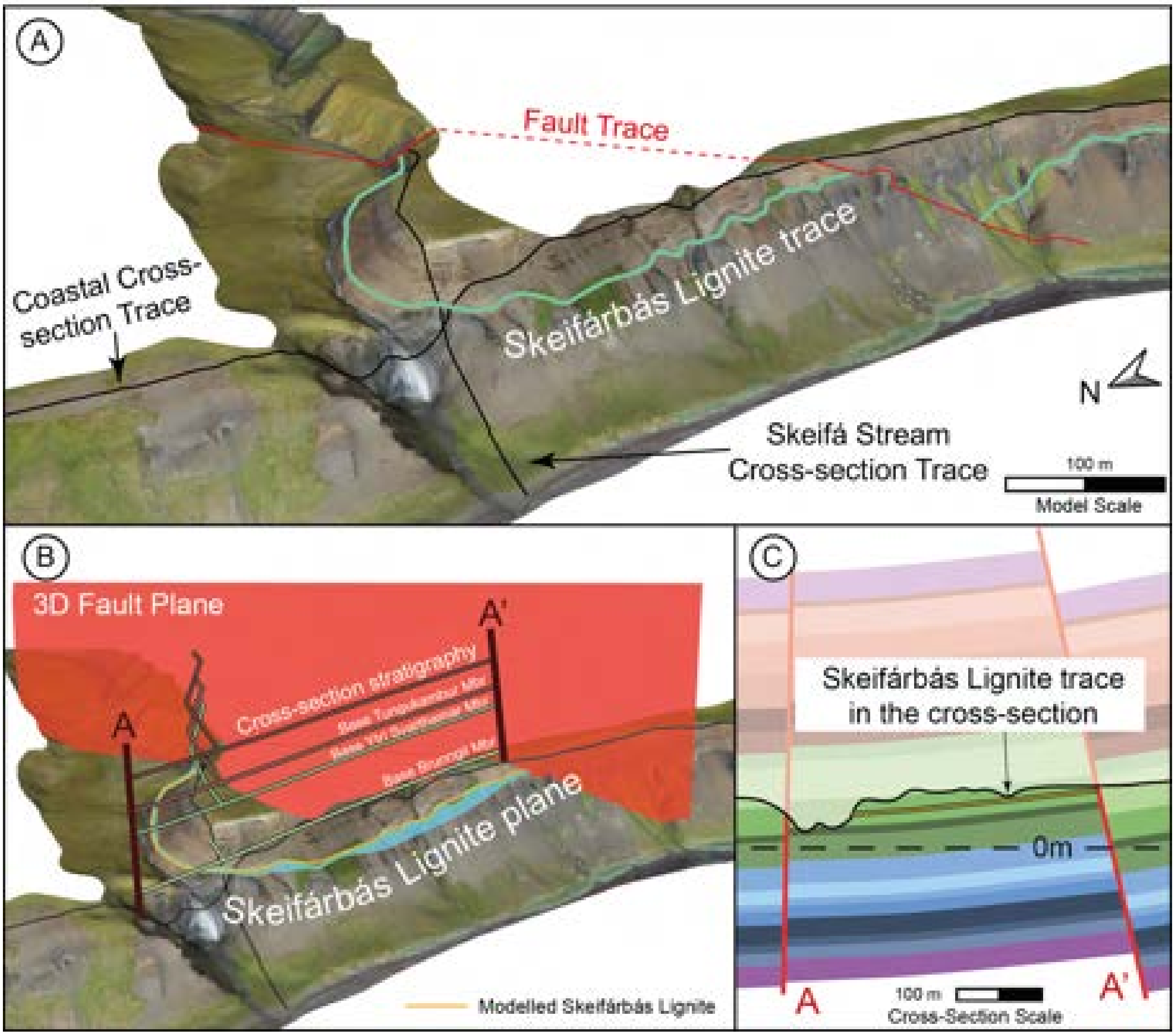


Figure 2.5: Example of the iterative and interactive model-building process that occurs within MOVE. All panels are shown at 1:1 vertical exaggeration and horizontal and vertical scales are equivalent. (a) Tracing of a key marker horizon (base of the Skeifárbás lignite, Bárðarson's Bed F) and faults on the virtual outcrop model. (b) Creation of planar structure from tracing of horizons around the 3D virtual outcrop and the generation of stratigraphy relative to those surfaces. Yellow line (modelled Skeifárbás lignite) represents the back-projection of the cross-section stratigraphy for model validation. (c) Cross-section constructed from tracing horizons in the virtual outcrop model and fitting the modelled stratigraphy to those traces. Source: modified from the stratigraphic framework of Eiríksson and Símonarson (2021), before iteratively adjusting the layer thickness to arrive at the stratigraphic interpretation of this study.

## 2.4 Results

### 2.4.1 *Model fit to virtual outcrop*

A comparison between the final model and the virtual outcrop along the coastal exposure is shown in Figure 2.4. By tracing the 24 key horizons across the virtual outcrop, we found that the logged sections V–XXII of ES21 overlapped to define the complete composite stratigraphy (Fig. 2.3). There is no evidence of any significant succession missing by non-exposure. We found that a single stratigraphic column of 442 m provides a satisfactory match to the entire coastal virtual outcrop of the Barmur Group (Fig. 2.4). We quantified the fit of the model to the stratigraphic data in four zones of good exposure that totalled 2.4 km of the 7 km-long coastal cross-section. Within these zones, modelled stratigraphic boundaries match their picked locations within $\pm$ 3 m (Fig. 2.4b). To achieve this fit, we altered the layer thicknesses relative to the ES21 model as shown in Table 1 and included normal faults that are distributed fairly evenly across the outcrop (Fig. 2.4b, F1–F31). We found that the faults that are most clearly visible in the virtual outcrop are not necessarily the faults with the greatest displacement. For this reason, not every fault observable in the field is included in the model (Fig. 2.4b, c). The fault planes in the model account for a total of 565 m of dip slip. In places where fault plane traces can be viewed in 3D in the virtual outcrop, the model fault surfaces tend to strike north–south. This orientation is comparable with the trend of faults within the Northern Volcanic Zone and Eyjafjarðaráll Rift spreading axes and the Tjörnes Fracture Zone (Metzger et al. (2013b); Fig. 2.1b). Thus, the larger faults observed within the Barmur Group probably accommodate some of the divergence between the European and Greenland/North American plates.

### 2.4.2 *Model fit to field observations*

Table 2.1 compares the thicknesses of the members in our model with their logged thicknesses. Some members were logged in multiple places, and maximum and minimum thicknesses are reported in ES21 (Table 2.1). The mean logged thickness range is 10.5 m, and the total range is 2.0–25.3 m (Table 2.1). Figure 2.6 shows the lateral thickness variation along four transects that span the entire coastal outcrop. Measured thickness variations over lateral distances of hundreds of metres to kilometres are 6.7, 8.3, 15.8, and 8.5 m. The similarity between the mean value of 9.8 m and the mean logged thickness range in Table 2.1 (10.5 m) suggests that the logged thickness range is a good indication of the lateral thickness variation, though it may also reflect lack of exposure in some cases. For most members, the difference between the modelled thickness and the maximum logged thickness is less than the logged thickness range (Table 2.1). This observation suggests that the new model fits the stratigraphic data adequately, considering the lateral thickness variations. Two exceptions are the Egilsgjóta and Brunngil Members, for which the model/data differences are 5.0 and 15.5 m, and the logged thickness ranges are 3.3 and 3.2 m, respectively. In both these cases, the minimum thickness estimates are not reliable because of lack of exposure. On average, the thickness of any model member is 4.7 m greater than its maximum logged thickness (Table 2.1). This difference is a measure of the effect of using the virtual outcrop to model the composite stratigraphy. The model represents the mean thicknesses of the members across their entire exposure in the virtual outcrop, while the logged thicknesses represent point measurements across a smaller footprint.

The mean local (tens-of-metres-scale) bedding orientation in the model is 030/10.5$^o$ NW (Fig. 2.7). The median bedding dip is 9.5$^o$, with an interquartile range of 7.5–12$^o$ and a total range of 4.5–18.5$^o$ (Fig. 2.7). The ranges compare well with point measurements made in the field of 10–15$^o$ (Bárdarson, 1925), 5–10$^o$ (F, 1963; Einarsson et al., 1967) and $< 10^o$ (Eiríksson and Símonarson, 2021). Our field measurements

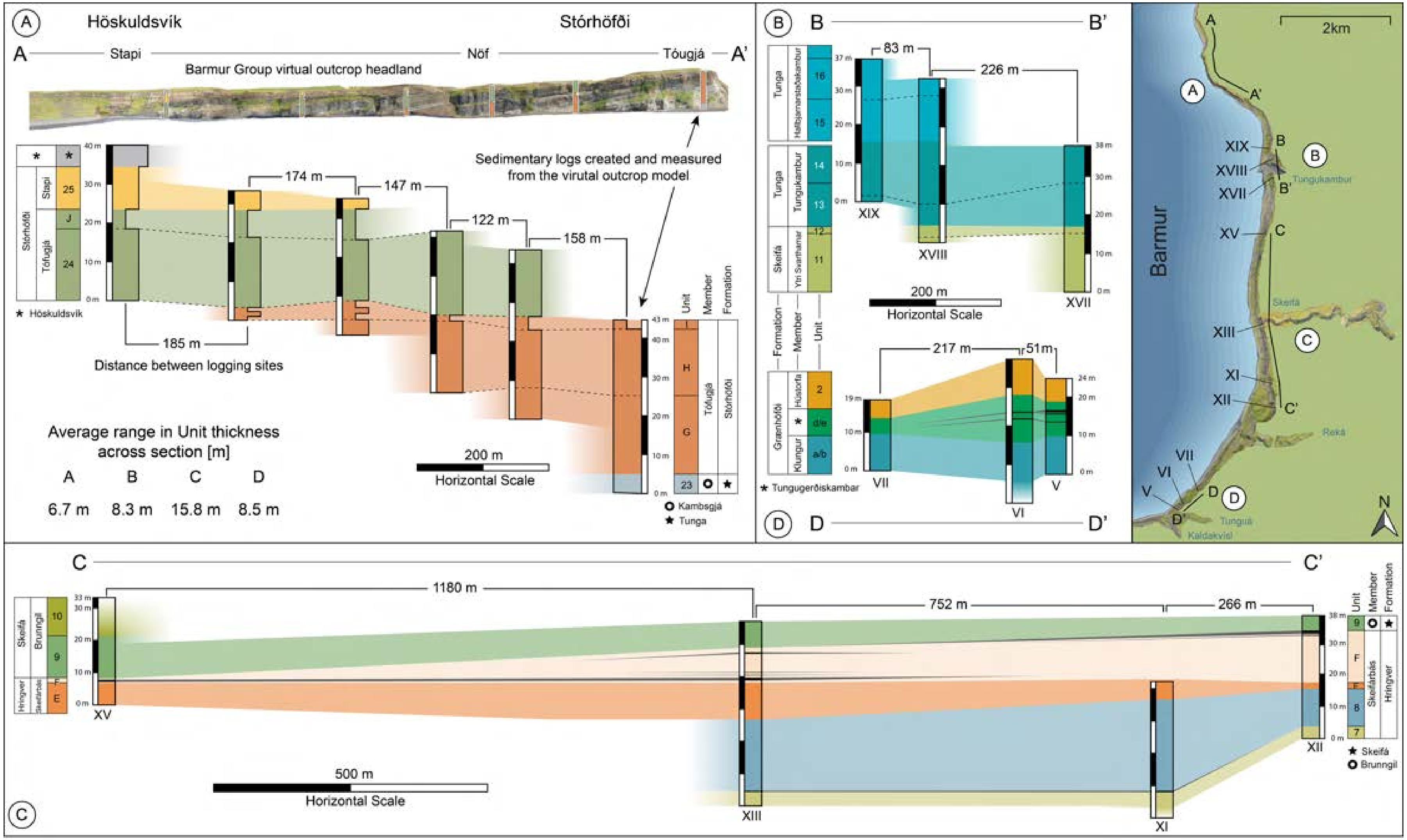


Figure 2.6: Lateral variation measured across the Barmur Group. (a) Lateral variation of the Stórhöfði cliff section (A–A'), measured directly from the virtual outcrop model. These sections are inaccessible without the use of remotely piloted vehicles. (b–d) Lateral variation measured from the virtual outcrop model at logged sections B–B' (b), C–C' (c) and D–D' (d) of Eiríksson and Símonarson (2021). The modelled stratigraphy (and corresponding sedimentary log from ES21 labelled with Roman numerals) and distance between adjacent stratigraphic measurements are shown. The lateral thickness variations are observed to range from 6.7 to 15.8 m across a scale of hundreds of metres. Source: modified from Eiríksson and Símonarson (2021).

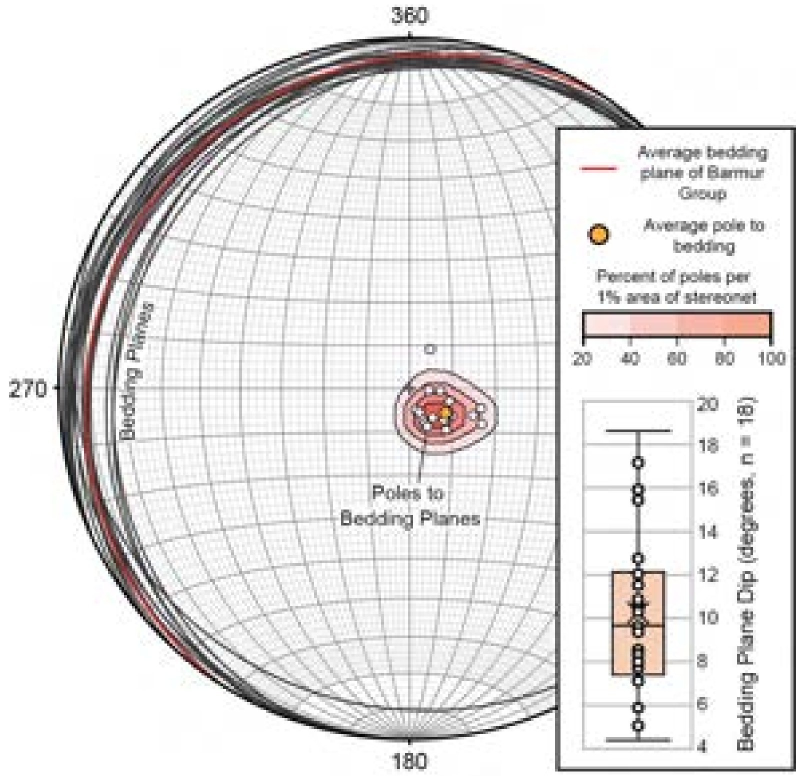


Figure 2.7: Bedding data extracted from key marker horizons (n = 20) across the virtual outcrop model. An average (mean) dip and strike of 030/10.5$^o$ NW was found to be representative of the entire section. Contours are shown for the bedding data, highlighting the percentage of poles to bedding residing within a 1% area of the stereonet.

agree with these published ranges. The range in local and point bedding orientations is probably associated with a combination of the pervasive regional faulting, and the variations in unit thickness on a lateral scale of hundreds of metres (Fig. 2.6). The mean dip of strata when averaged along the entire coastal cross-section in Figure 2.4 is 4.7$^o$ towards the north. This value is lower than the mean local dip because of the apparent dip effect. The azimuth of the coastal cross-section varies from northwest–southeast in the north to north–south in the centre to northeast–southwest in the south (Fig. 2.1). The corresponding apparent dips, based on our measured mean local bedding orientation, are predicted to be 10$^o$ in the north of the coastal section, 5$^o$ in the centre and 3$^o$ in the south. The consistent dip toward the northwest reflects the fact that the onshore outcrop of the Barmur Group lies toward the southeastern edge of the Tjörnes Basin depocentre (Fig. 2.1b).

### 2.4.3 *Correlation of coastal and inland exposure*

The new composite stratigraphy in Figure 2.3 was developed by fitting the model to the coastal exposure. We tested whether the same stratigraphy provides an adequate fit to the inland exposures along the Skeifá Valley (Fig. 2.8). A key marker horizon is a prominent lignite seam within the Skeifárbás Member (Table 2.2, base XIII:H, Bárðarson's Bed F), which is mapped across the coastal and inland exposures in ES21. In the coastal exposure at the mouth of the Tungukambur River, our model shows that the marker lignite seam occurs 75 m below the Mactra/Serripes biozone boundary (Fig. 2.8, Locality C). Inland, the same lignite marker is exposed (Fig. 2.8, Locality D) adjacent to the fault-bounded Tungugrænur Formation (Fig. 2.8, Locality E).

A model bedding surface was projected between the coastal and inland exposures of the marker lignite seam. We found no significant difference between the orientation of the projected surface and the orientation of the average bedding as measured along the coast (Fig. 2.8). This result suggests that the model of the coastal stratigraphy can adequately fit the inland stratigraphy. We therefore used the model to predict the relationship between the fault-bounded Tungugrænur Formation (Fig. 2.8, Locality E) and the rest of the Barmur Group, based on the position of the Mactra/Serripes biozone boundary, which is the only marker exposed on both sides of the fault. We estimate that the Tungugrænur Formation has been downthrown c. 30 m compared with the sediments immediately adjacent (Fig. 2.8, Locality D). In this correlation, the Tungugrænur Formation is laterally equivalent to the upper Skeifá Formation and the lower Tunga Formation.

### 2.4.4 *New composite stratigraphy*

The new model stratigraphic column has a total thickness of 442 m. We modelled the conformable part of the Barmur Group above the Klungur Member (i.e. above the basal

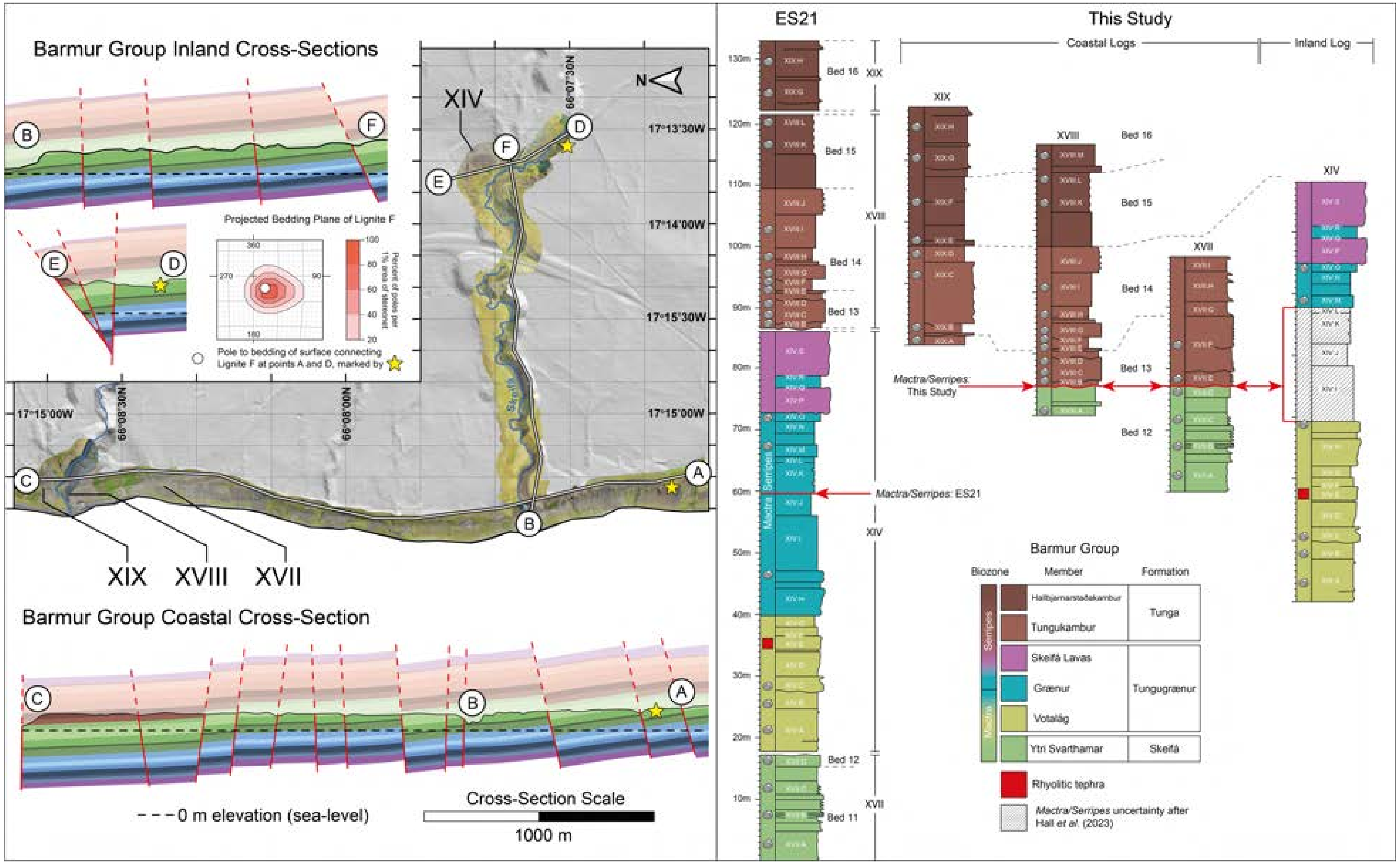


Figure 2.8: Comparison of the inland and coastal exposures of the Barmur Group, focusing on the Mactra/Serripes biozone boundary. All cross-sections are shown at 1:1 vertical exaggeration. Cross-sections have been created in MOVE: A–C, coastal exposure; D–E, Tungugrænurr exposure; F–B, Skeifá Stream exposure. Across the virtual outcrop model, Bárðarson's Bed F (Bed XIII:H, Eiríksson and Símonarson 2021) can be found in outcrop (yellow star). Modelled bedding surfaces between Bed F as exposed between inland and coastal exposures show no deviation from the average Barmur Group bedding orientation. Shown also are the sedimentary logs and stratigraphic interpretations of Eiríksson and Símonarson (2021) (logs XIX, XVII, XVIII and XIV) and the revised correlation of this study highlighting the uncertainty of the Mactra/Serripes boundary placement within the inland log. Crucially, we argue against any considerable unconformity in the coastal exposure at Locality C. Source: modified from Eiríksson and Símonarson (2021).

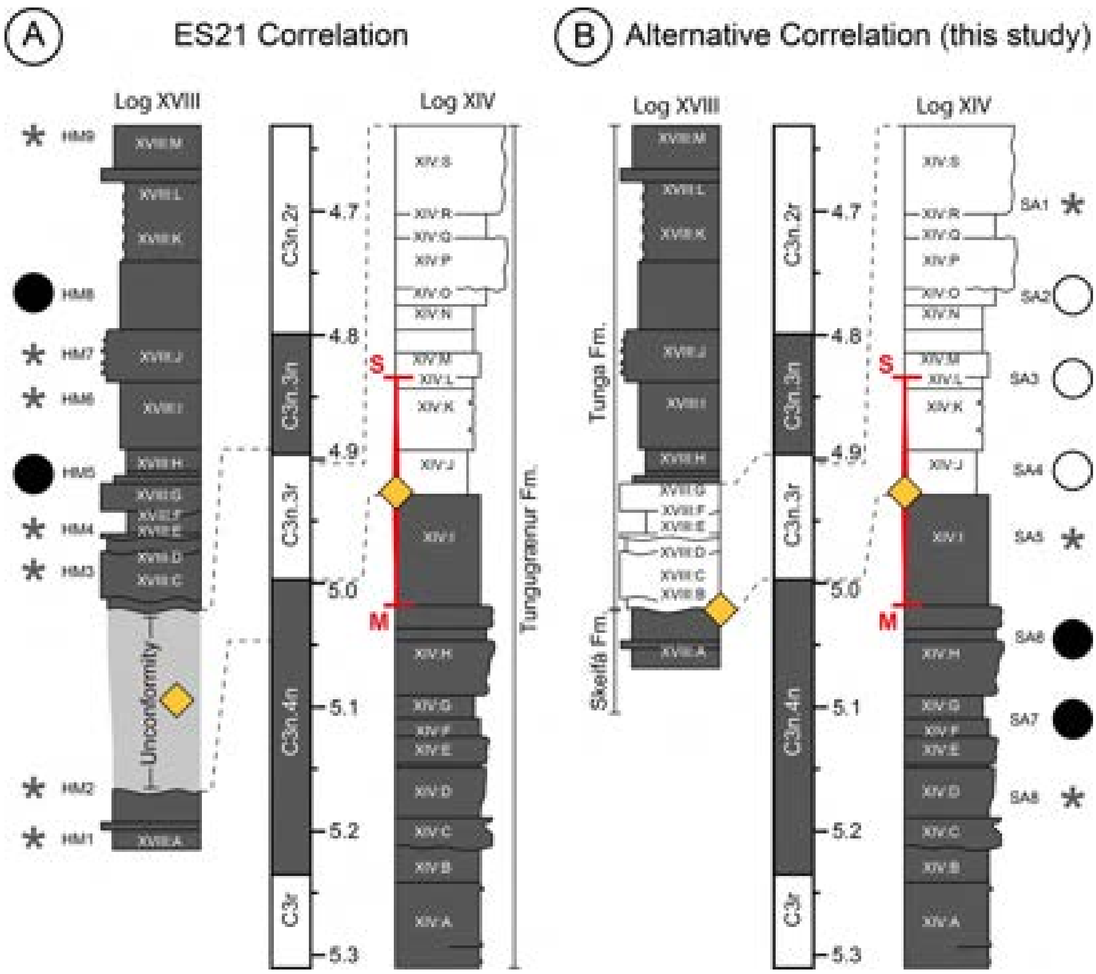


Figure 2.9: Palaeomagnetic interpretations of Eiríksson and Símonarson (2021) and this study. The palaeomagnetic record of Kristjánsson (2004) has been revised to highlight only reliable polarity results. The Mactra/Serripes (M/S) biozone boundary faunal uncertainty is highlighted in red. This study suggests that the Skeifá and Tunga formations are lateral equivalents of the Tungugrænurr Formation and tentatively assigns the base of the Tunga Formation to the subchron C3n.3r. The Geomagnetic Polarity Timescale is that of Ogg (2020). Source: modified from Eiríksson and Símonarson (2021), Kristjánsson (2004), and Ogg (2020).

fault zone). The comparable thickness of the published ES21 composite section is 525 m (Fig. 2.3). Thus, the new model stratigraphic column is 16% thinner than the ES21 stratigraphic column. The principal reason for this difference is that our model does not include the Tungugrænur Formation, which is predicted to be laterally equivalent to part of the coastal stratigraphy. The thickness of the Tungugrænur Formation on the ES21 composite section is 75 m. The difference between the model composite section and the ES21 composite section without the Tungugrænur Formation is 8 m. This remain-der is close to the mean lateral thickness variation. The sum of the maximum logged thicknesses of all the Barmur Group members is 572.5 m (Table 2.1), which is close to the equivalent thickness of 570 m as measured on the composite ES21 stratigraphic column (Fig. 2.3). However, the sum of the maximum logged thicknesses above the Klungur Member is 485 m (Table 2.1), which is less than the equivalent thickness of 525 m as measured on the composite ES21 stratigraphic column (Fig. 2.3), and closer to our model stratigraphic column of 442 m. This difference between the 525 and 485 m thickness estimates reflects the practicality of assembling a composite stratigraphic column from sparse logs of its constituent members and illustrates the advantage of fitting a stratigraphic model to a virtual outcrop.

## 2.5 Discussion

### 2.5.1 *Stratigraphic position of Skeifá Lava and age of Bering Strait opening*

The Skeifá Lava is significant because it has been dated to magnetic subchron C3n.3r (4.997–4.896 Ma) and occurs 17 m above the bioevent that marks the first opening of the Bering Strait (the Mactra/Serripes biozone boundary), which is assigned to subchron C3n.3r or C3n.4n (5.235–4.896 Ma; Hall et al. (2023)). The Skeifá Lava is only exposed within the Tungugrænur Formation, which is in faulted contact with the rest of the Barmur Group. It is therefore not straightforward to determine the stratigraphic position of the Tungugrænur Formation within the Barmur Group (Eiríksson and Símonarson, 2021). In ES21, the Tungugrænur Formation correlates with an erosional unconformity within the coastal succession and sits between the Skeifá and Tunga formations in the composite stratigraphy (Fig. 2.8). Here, we have presented an alternative model, in which no unconformity exists, and the Tungugrænur Formation is laterally equivalent to the upper part of the Skeifá Formation and the lower part of the Tunga Formation. In this section, we discuss the strengths and weaknesses of both hypotheses in terms of palaeomagnetic, sedimentological and structural evidence.

The upper Tungugrænur Formation, including the Skeifá Lavas, is reversely magnetized, and the lower Tungugrænur Formation, underlying and possibly including the Mactra/Serripes biozone boundary, is normally magnetized (Doell, 1972; Kristjánsson, 2004). The coastal succession above and below the Mactra/Serripes biozone boundary is reported to be entirely normally magnetized (Kristjánsson, 2004). In the ES21 stratigraphic model, an unconformity is proposed to explain the lack of evidence for reversely magnetized rocks within the coastal exposure that could correlate with the upper Tungugrænur Formation. However, the palaeomagnetic data from the Tungugrænur, Skeifá and Tunga formations are not of uniform quality. Of 19 magnetic determinations reported in Table 1 of Kristjánsson (2004; HM1 – 10, SA1 – 9), only seven are reported as reliable. In Figure 2.9, we have highlighted the reliable palaeomagnetic measurements. Notably, there are no reliable measurements around the biozone boundary in the coastal exposure. In Figure 2.9, we show an alternate to ES21 that fits the reliable palaeomagnetic data equally well: the Skeifá Formation and lower Tunga Formation (Tungukambur Member) could be reversely magnetized and could correlate with the upper Tungugrænur Formation, as supported by our model.

Eiríksson and Símonarson (2021) noted several pieces of sedimentological evidence in support of the ES21 composite stratigraphic model. The Mactra/Serripes biozone boundary is well exposed at the coast, and it is clear that the Skeifá Lavas are not present in the stratigraphy above it, as is observed within the Tungugrænur Formation. The Tungugrænur Formation contains a rhyolitic tephra unit (bed XIV:E, Figs. 2.8 and 2.9), which might be expected to have been deposited over a wide area, but which has no equivalent in the coastal exposure. Overall, no unequivocal lithological correlations can be made between the Tungugrænur Formation and the coastal exposure. There is no evidence for an angular unconformity near the Mactra/Serripes biozone boundary within the coastal exposure (Eiríksson and Símonarson, 2021). Nevertheless, the Ytri Svarthamar and Tungukambur members of the coastal Skeifá and Tunga formations contain several erosional surfaces, and the depositional environment is interpreted as a shallow shoreface environment characterized by strong, intermittent currents (Eiríksson and Símonarson, 2021). It is conceivable that the time equivalent to the Tungugrænur Formation could be missing, as shown in ES21. On the other hand, the coastal exposure and the Tungugrænur Formation exposure are separated by c. 2 km. There is clear evidence for lateral variability in sedimentary facies on scales of hundreds of metres to kilometres, both from the ES21 sedimentary logs and from our virtual outcrop (Fig. 2.6; Table 2.1). The lack of lavas in the coastal exposure could plausibly be explained by the limited spatial extent of the lava flows, and the lack of tuffs could plausibly be explained by the high-energy coastal depositional environment. We conclude that the sedimentological evidence is equally consistent with ES21 and our stratigraphic model.

Two scenarios could generate a coastal unconformity as proposed in ES21. One is that the Tungugrænur Formation accumulated in a laterally confined depocentre during a period of no deposition near the modern coast. However, the lithofacies within Skeifá, Tungugrænur and Tunga Formations have been interpreted in terms of a low-lying coastal plain sedimentary environment (Eiríksson and Símonarson, 2021). We showed above that the lateral changes in the thicknesses of individual Barmur Group units are typically on the metre scale (Fig. 2.6). It therefore seems unlikely that isolated deposition or erosion was associated with 50–100 m scale changes in topography at the depositional surface, comparable with the thickness of the Tungugrænur Formation. Another possibility is that the Tungugrænur Formation was deposited across the Tjörnes Basin, before being locally eroded away near the modern coastline. Such erosion could have been associated with faulting that occurred during and shortly after deposition of the Tungugrænur Formation. In this case, the total throw on the syndepositional faults would need to be 50–100 m. We searched for evidence of synsedimentary faulting in the virtual outcrop. Multiple faults are visible, but there is no clear evidence for any changes in the thicknesses of sedimentary units across any of these faults (Fig. 2.4). Thus, we interpret all the faulting as post-depositional. This interpretation is in line with our own field observations, and with published studies (Garcia et al., 2002; Eiríksson and Símonarson, 2021). We find no evidence that faulting with a combined throw of 50–100 m occurred during deposition of the Tungugrænur Formation.

In summary, the palaeomagnetic and sedimentological data are equally compatible with the ES21 and new model composite stratigraphies. The virtual outcrop does not provide positive evidence for the unconformity suggested in ES21, which may support the new model stratigraphy, although we acknowledge that absence of evidence is not evidence of absence. The conclusions of Hall et al. (2023) regarding the age of the Mactra/Serripes biozone boundary, which marks the first Bering Strait opening, were based on the Tungugrænur Formation only, and are therefore unaffected by the choice of composite stratigraphic model. The palaeomagnetic reversal record from the Tungugrænur Formation is reliable (Fig. 2.9), and it seems likely that the bioevent occurred close to the boundary between subchrons C3n.3r or C3n.4n, at c. 5.0 Ma. Unfortunately, the sparse occurrence of fossil mollusc beds within this formation makes it impossible to assign the bioevent to a single subchron. The coastal succession has a greater number of fossil mollusc beds at closer spacing, but the palaeomagnetic data are unreliable (Fig. 2.9). Future acquisition of reliable palaeomagnetic data from the coastal succession immediately surrounding the Mactra/Serripes biozone boundary could distinguish whether the Mactra/Serripes biozone boundary occurred in subchron C3n.3r or C3n.4n. Such data would also provide a more definite test between the ES21 and our new model stratigraphies. Irrespective of which hypothesis is preferred, our virtual outcrop and associated composite stratigraphy provides the most reliable means of correlating published and future records of past environmental change, most of which come from the coastal succession.

### 2.5.2 *Historical context and benefits of virtual outcrop*

The first thorough study of the Barmur Group was by Guðmundur Bárðarson (1925). Although the first written mention of the Barmur Group was in 1749, and several works in the late nineteenth and early twentieth centuries described the sediments and fossils, in all these studies the fossils from different horizons had been mixed together indiscriminately. Bárðarson was the first to develop a composite stratigraphy of the entire succession, to enable correlation of beds from place to place, and to document evolution of the mollusc assemblages. He achieved this through a combined stratigraphic and structural analysis along the main coastal section. Thus, his aims were the same as ours. It is interesting to compare what he achieved using field mapping techniques with our results using a virtual outcrop. Figure 2.4 compares Bárðarson's hand-drawn coastal section with the virtual outcrop and our model cross-section, and Figure 2.10 shows comparisons between Bárðarson's field photographs and our virtual outcrop.

To estimate the thickness of the succession, Bárðarson recognized that the local bedding dips were different from the mean regional dip of bedding, which he estimated to be about $5^o$ (we measure $4.7^o$ from the virtual outcrop). He estimated that the whole thickness of the Barmur Group would be about 520 m using Pythagoras, but recognized the need to adjust this value for the effects of faults, some with throws of up to 30 m. He concluded, "I think that one can without fear of exaggeration assume of at least abt. 450 m for the strata in question". This value is remarkably close to our model thickness of 442 m.

Bárðarson's view of the base of the Barmur Group is the same as ours. He interpreted a major unconformity between the Barmur Group and the Kaldakvísl Lava Group basement: "the basalt itself had been subjected to erosion for some considerable time during the period prior to the formation of the Crag strata". We estimated a time gap of 3–4 Myr using radiometric dating (Hall et al., 2023). He also recognized that parts of the lower

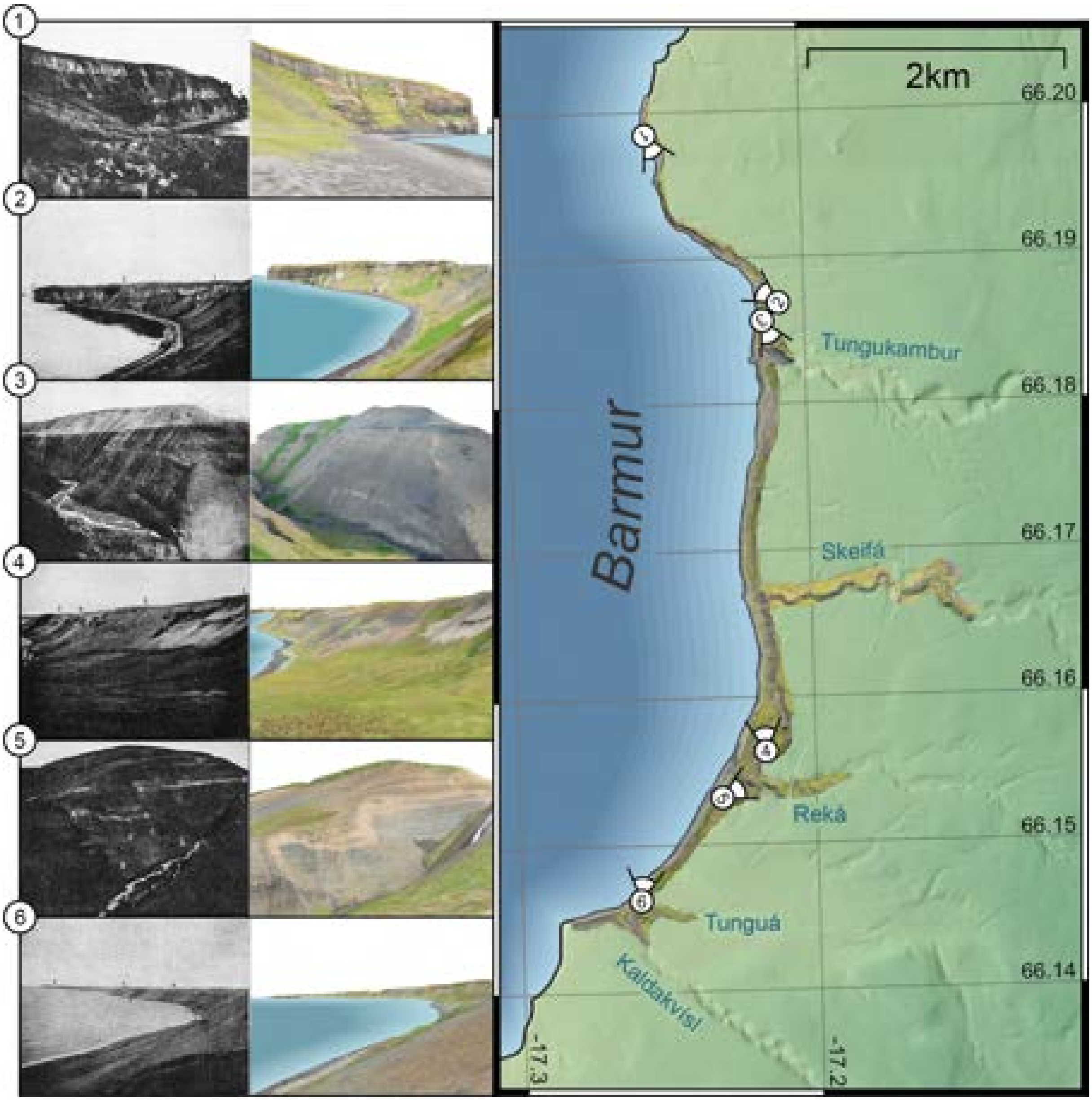


Figure 2.10: Comparison of Bárðarson's original (1925) photography with recreated snapshots of those same views in the virtual outcrop model nearly 100 years on (colour images). Source: modified from Bárdarson (1925).

Barmur Group are repeated by faulting: "Presumably, these strata with rounded stones belong together, although the stones here reach higher above the shore, the difference of level being possibly due to dislocation". Bárðarson did trace the units from the coastal exposure into the hills, but he did not directly address the problem of correlating the coastal and inland strata. In line with our new interpretation, he did not see any evidence of a major unconformity between the Mactra and Serripes biozones. Instead, he interpreted this boundary to occur within a transgressive cycle commencing with the formation of Bed 9, reaching its maximum at Bed 15, just above the Mactra/Serripes biozone boundary, and ending with the formation of Bed 23.

Thus, the main findings of the present study are not dissimilar to those of Bárðarson's classic study in Icelandic geology. What then is the value of our virtual outcrop? Published estimates of the composite thickness of the Barmur Group vary from 400 to 570 m (Bárdarson, 1925; Eiríksson, 1981; Buchardt and Símonarson, 2003; Eiríksson and Símonarson, 2021). This range of estimates has arisen because, without a formal structural–stratigraphical model, it has not been possible until now to confidently quantify the composite thickness for the entire outcrop. The coastal–inland correlation problem could in principle have been tackled using classical mapping and structural contouring techniques. Our approach of fitting a combined stratigraphic and structural model to a large virtual outcrop has allowed us to confidently determine total sedimentary thickness of the Barmur Group and systematically assess the natural variation in thickness of modelled members. Importantly, it is now straightforward to quantify how updates to any given part of the stratigraphy, including future new radiometric ages and palaeomagnetic datasets, will affect the age model for the entire Barmur Group.

The new virtual outcrop model offers a level of accessibility that has not previously been possible for the Barmur Group. Global events such as the COVD-19 pandemic highlight the need for high-quality resources to supplement and, in certain circumstances, replace in-person field trips and excursions (Fleming, 2022; Pugsley et al., 2022). Our model was commissioned partly as a teaching aid and fieldwork supplement during COVID-19. A strength of this model is its combination of large-scale, high-resolution and complete coverage of the entire Barmur Group, an important component of Iceland's thickest onshore sedimentary sequence. Of the eight Icelandic models currently accessible through the V3Geo website (`https://v3geo.com`), the Barmur Group model is one of the largest, the other being a lower-resolution model of an entire glacier (Gallagher and Askew, 2024). The Barmur Group virtual outcrop model provides a holistic geological model from start to end for one of Iceland's thickest onshore sedimentary sequences.

## 2.6 Conclusion

We aimed to resolve uncertainty in the composite stratigraphic section, and thus the age model, of the Barmur Group (Tjörnes Peninsula, Iceland) to increase the value of this sedimentary archive of late Miocene to early Pliocene palaeo-environmental change. We carried out drone-based mapping of the entire Barmur Group outcrop and created a virtual outcrop model. We have provided the complete workflow used to plan, acquire and process the virtual outcrop, to improve access to this burgeoning new technology in Earth science teaching and research. We integrated the virtual outcrop with the lithostratigraphic scheme of Eiríksson and Símonarson (2021) and produced the first integrated structural and stratigraphic model of the Barmur Group. We found that a single composite stratigraphic section of 442 m thickness adequately represents the stratigraphy across the entire outcrop of the Barmur Group. The new composite section is 16% thinner than the previous composite section, partly because we now account for repetition of stratigraphy by faulting at the base of the Barmur Group, and partly because we have improved the correlation between the coastal and inland exposures. The model suggests an alternative interpretation of the stratigraphy encompassing the inland Tungugrænur Formation, which contains both the radiometrically dated Skeifá Lavas and also the bioevent that signifies the first opening of the Bering Strait oceanic gateway – the Mactra/Serripes biozone boundary. This new correlation suggests that the first opening of the Bering Strait probably occurred during magnetic subchrons C3n.4n to C3n.3r and most likely at the transition between these two subchrons at 5.0 Ma. This new correlation could be tested in future by radiometric dating of a rhyolitic tephra that underlies the Mactra/Serripes biozone boundary within the Tungugrænur

Formation coupled with a more reliable palaeomagnetic dataset.

## 2.7 Data Availability

All data generated or analysed during this study are included in this published article and its supplementary material, including the details of the data acquisition and processing methods and batch processing files for Agisoft Metashape software. The complete uninterpreted virtual outcrop model is available in Allison et al. (2023a). The interpreted virtual outcrop model based on the combined stratigraphic/structural model presented here (i.e. a high-resolution version of Fig. 2.4) is available in Allison et al. (2023b).

# 3 | Tracing non-Uniformitarian behaviour across 200 million years of GDGT-based temperature reconstructions

*The contents of this chapter are the work of this author. The compilation of the 200 Ma record of isoGDGT data presented in this chapter was undertaken by Jonathan Hall. The author builds off the work of published model/code of OPTiMAL (Dunkley Jones et al., 2020) to produce the narrative, text and results of this chapter, shortly to be submitted for publication. The co-author contributions for this future publication are provided here in CRediT (Contributor Roles Taxonomy) format:* ***Matthew Allison****: Conceptualisation (equal), Methodology (lead), Software (equal), Formal analysis (lead), Investigation (equal), Visualisation (lead), Writing-original draft (lead)* ***Tom Dunkley Jones****: Conceptualisation (equal), Methodology (supporting), Supervision, Writing-review & editing (equal)* ***Sarah Greene****: Conceptualisation (equal), Software (equal), Writing-review & editing (equal)* ***Jonathan Hall****: Methodology (supporting), Investigation (equal), Data curation (lead), Writing-review & editing (equal)* ***Bella Duncan****: Writing-review & editing (equal)* ***Hana Ishii****: Writing-review & editing (equal)* ***Stephen Jones****: Conceptualisation (equal), Methodology (supporting), Supervision, Writing-review & editing (equal)* ***Ilya Mandel****: Methodology (equal), Writing-review & editing (equal)* ***James Bendle****: Conceptualisation (equal), Methodology (supporting), Supervision, Writing-review & editing (equal)*.

## 3.1 Abstract

Investigations of geological records must address the uniformitarian assumption: that the same processes operate in the present as in the past. Here we test this assumption on isoGDGT microbial membrane lipids, a widely used system for reconstructing Earth surface temperatures across the past 200 million years. Using a Gaussian Process emulator approach (OPTiMAL - Dunkley Jones et al. (2020)) we generate a single metric, quantifying the distance between an ancient isoGDGT assemblage and the modern marine core-top calibration dataset (Seki et al., 2014; Tierney and Tingley, 2015), scaled with respect to temperature-dependence within the system. The new measure, Contextualized Distance (*CD*) metric, provides a spatial map of the relative importance of calibration data for any past temperature estimate. *CD* values are calculated for a compilation of $>$12,000 ancient isoGDGT assemblages, spanning the last 200 Ma. This analysis demonstrates ancient isoGDGT assemblages are increasingly distant from the modern calibration data in warmer deep-time greenhouse climates but return to be more similar to modern in the cooler climates of the Early Cretaceous. We interpret that the isoGDGT system appears to obey the uniformitarian principle and can be used with care for reconstructing surface temperature for the past 200 million years.

## 3.2 Introduction

Organic biomarker palaeo-thermometers are a cornerstone of palaeoceanographic research that enable the reconstruction of surface temperatures over several hundred million years of Earth history (Brassell et al., 1986; Müller et al., 1998; Volkman et al., 1998; Schouten et al., 2002; Belt et al., 2007; Herbert et al., 2010). Their success stems from the worldwide occurrence of microbial-derived lipids that are commonly preserved within sedimentary deposits. Each palaeo-thermometer is based on a relationship established in modern environments, where both the biomarker composition produced by living microbes and the temperature of the environment they inhabit can be measured to create a calibration dataset. A fundamental assumption underpinning this approach is uniformitarianism – the principle that the biosynthetic pathways governing biomarker production have remained consistent between modern organisms and their ancient counterparts. Yet multiple lines of evidence raise the possibility that microbial populations have evolved over tens to hundreds of millions of years in a way that could invalidate the uniformitarian assumption and lead to systematic errors in palaeo-temperature estimates (Taylor et al., 2013; Schouten et al., 2013; Tierney and Tingley, 2014; Hurley et al., 2016).

Here we focus on the isoGDGT (isoprenoidal Glycerol Dialkyl Glycerol Tetraethers) palaeo- thermometer. It is one of the most widely used lipid biomarker methods for reconstruction of Sea-Surface Temperature (SST) back to the Jurassic (c. 200 million years) (Schouten et al., 2002; Kim et al., 2008; Seki et al., 2014; Tierney and Tingley, 2015; Trommer et al., 2009; Inglis et al., 2015; O'Brien et al., 2017; Stickley et al., 2004). In response to SST variation, archaeal and bacterial producers of isoGDGTs modulate the relative abundance of ring structures within the carbon chain, with more rings resulting in a more thermally stable membrane (Schouten et al., 2002). Upon microbial cell death or predation, the isoGDGTs produced in the water column fall to the sea floor where a proportion of these robust lipids are c into the geological record. Based on a modern calibration dataset from marine sediment core-tops, the relative abundance of six isoGDGT structures (GDGT-0, -1, -2, -3, crenarchaeol (cren) and its regioisomer (cren')) can be correlated with SST (Schouten et al., 2002, 2007b,a). The original isoGDGT SST palaeo-thermometer was the $TEX_{86}$ index (TetraEther indeX of tetraethers containing 86 carbon atoms), formulated using four of the six measured isoGDGT compounds (Schouten et al., 2002).

Twenty years of research has produced variations on the original $TEX_{86}$ formulation, recognition of ocean-basin or region-specific deviations (Kim et al., 2008; Liu et al., 2009; Kim et al., 2010; Trommer et al., 2009; Fietz et al., 2016) and the documentation of factors that confound the primary temperature signal including (but not limited to) variation in the composition of microbial communities (Elling et al., 2015; Polik et al., 2018) water depth of lipid production (Kim et al., 2015; Taylor et al., 2013), post-depositional thermal maturation of sediment-hosted lipids (Schouten et al., 2004), and exogenous (e.g. from non-marine or in-sediment methanotrophic or methanogenic bacteria) sources of isoGDGTs (Inglis et al., 2015; Weijers et al., 2006b). A variety of indices have been developed to identify ancient isoGDGT assemblages affected by these factors, and partially mitigate the uncertainty they impart of SST estimates (including BIT (Hopmans et al., 2004), MI (Zhang et al., 2011), %GDGT – 0 (Blaga et al., 2009; Damsté et al., 2012), $\Delta RI$ (Zhang et al., 2016), fcren' (O'Brien et al., 2017) and GDGT 2/3 ratio (Taylor et al., 2013)). All these screening criteria, however, make the same uniformitarian assumption that underpins the isoGDGT palaeo-thermometer - that the biosynthetic processes acting in the present are the same as those of the past. In other words, they can screen out, in Rumsfeldian terms, the 'known unknowns' (Rumsfeld, 2002) – the biasing influences that we are aware of - but not the 'unknown unknowns' - such as the potential for a fundamental evolution of archaeal/bacterial lipid production over tens to hundreds of millions of years.

We have developed an objective method for assessing applicability of a calibration dataset to an ancient biomarker population. Crucially the method does not assume uniformitarianism, and can therefore be used to test this foundational assumption. We used a machine learning approach to isoGDGT based palaeo-thermometry that uses a Gaussian Process (GP) emulator to tackle this problem (Dunkley Jones et al., 2020). This approach was selected because it is agnostic about the form of isoGDGT temperature sensitivity and makes no assumptions about the underlying statistical regression model. As such, it uses all six components of the isoGDGT assemblage data, unlike the reductive $TEX_{86}$ index for which hidden variation in isoGDGT composition can exist behind a seemingly conserved number – for example that $TEX_{86}$ = 0.6 in the Jurassic is analogous to $TEX_{86}$ = 0.6 in the Holocene - is one of the untested assumptions of $TEX_{86}$-based SST estimation. Instead, we present a new metric, termed *CD* metric (Contextualised Distance), which captures any influencing factor acting on the isoGDGT system, known or unknown. We apply this GP approach to the largest compilation of ancient isoGDGT assemblages to data (n = 12,256) spanning the Holocene to the Early Jurassic (c. 200 Ma), making predictions of SST and presenting trends in the novel *CD* metric. The analysis presented in this study provides a streamlining of the traditional multi-index screening process to isoGDGT analysis and allows, for the first time, disparate influencing factors to be captured in one metric and directly compared. The *CD* metric enables the identification and quantification of data-points which behave in a non-uniformitarian manner - data presenting with non-modern, temperature weighted isoGDGT distributions.

## 3.3 Methods

Modern, marine core-top sediments have demonstrated the correlation between higher SSTs and the relative number of cyclopentane moieties in the archived isoGDGTs, which is the foundation of the isoGDGT SST palaeo-thermometer, first proposed as the $TEX_{86}$ index (TetraEther indeX of tetraethers containing 86 carbon atoms, Eq. 2.1, Eq. 2.2) (Schouten et al., 2002).

$$TEX_{86} = \frac{GDGT\ 2 + GDGT\ 3 + cren'}{GDGT\ 1 + GDGT\ 2 + GDGT\ 3 + cren'} \qquad \text{(Eq. 2.1)}$$

$$TEX_{86} = 0.015 \times SST + 0.28 \qquad \text{(Eq. 2.2)}$$

Where $GDGT\ 1$, $GDGT\ 2$, $GDGT\ 3$, $cren'$ are four of six isoGDGT structures (GDGT-0, -1, -2, -3, crenarchaeol (cren) and its regioisomer (cren')). The constants of Eq. 2.2 represent the constants of the linear regression used to correlate between $TEX_{86}$ and SST in $^{o}$C. More recent developments in isoGDGT analysis have utilized more sophisticated means by which to produce regressions through calibration data (Tierney and Tingley, 2015; Dunkley Jones et al., 2020). OPTiMAL is a Gaussian Process emulator with a squared exponential kernel with automatic relevance determination (ARD), designed to predict SST from ancient assemblages of isoGDGTs (Dunkley Jones et al., 2020). The model uses a calibration dataset of modern isoGDGT assemblages (n = 854), sampled from the top of marine sediment cores (Seki et al., 2014; Tierney and Tingley, 2015), as previously described by Dunkley Jones et al. (2020). A key limiting factor of a GP model is that the resulting predictions are constrained to the range of the calibration dataset, c. -3–+30 $^{o}$C (i.e. the temperature range of the modern ocean). This SST range is very close to the long-term minimum of such ranges throughout the Cenozoic–Mesozoic history of Earth's climate, the result being that deep time archives are inevitably confronted by SSTs that extend beyond the upper limit of modern calibration data of c. 30 $^{o}$C. The novel *CD* (Contextualized Distance) metric presented here extends the scope of OPTiMAL to enable useful information to be extracted from ancient isoGDGT assemblages which exist outside the observable distribution of isoGDGTs in the modern ocean.

When testing for non-uniformitarian behaviour of data through time, it might seem sufficient to investigate a Euclidean measure of distance between two data points in the isoGDGT space, however the GP emulator indicates certain dimensions carry far less temperature information than others (Table 3.1). During the machine learning stage of the GP, correlation length-scales are generated for each of the six isoGDGT dimensions, providing a first order estimate of the amount of SST ”information” stored within a given dimension. Should the end goal be to predict SST from isoGDGT assemblages - analysis should be most sensitive to the dimensions that carry this signal and less

| GDGT Fraction | Length-scale | Rank |
|---|---|---|
| GDGT-3 | 0.0047 | 1 |
| GDGT-1 | 0.0282 | 2 |
| GDGT-2 | 0.0478 | 3 |
| cren' | 0.0642 | 4 |
| GDGT-0 | 0.0911 | 5 |
| cren | 28.892 | 6 |

Table 3.1: The OPTiMAL derived length-scales for the six isoGDGT dimensions. The smaller the length-scale the more sensitive the given dimension is to encoding temperature information. The $TEX_{86}$ index does incorporate the four most temperature sensitive dimensions, but even within that subset of the isoGDGT space, order of magnitude differences exist in the learnt length-scales, implying a wide range of temperature sensitivity, contrary to the largely equal weighting of the $TEX_{86}$ formulation.

sensitive to the dimensions that do not. This is achieved by scaling the isoGDGT dimensions by their respective length-scales, encoding the machine learnt temperature sensitivity into the relative positions of the data. The *CD* metric is then a Euclidean measure of distance between data-points made in this length-scale adjusted space (Eq. 2.3).

$$CD_n = \sqrt{\sum_{f=1}^{6} \left(\sigma_f \left(d_f - a_f^{(n)}\right)\right)^2} \qquad \text{(Eq. 2.3)}$$

$d$ is a single isoGDGT data-point with six dimensions $f$, one for each lipid fraction, and $a^{(n)}$ is a list of isoGDGT data-points of size $n$. $\sigma$ are the GP learnt, correlation length-scales for each isoGDGT fraction. Eq. 2.3 results in a list of *CD* values (length $n$) describing the weighted distance between the data-point $d$ and each data-point in list $a$. In this study, we use Eq. 2.3 to investigate two modes of operation. **Mode 1**: $d$ represents an ancient data-point and $a$ represents the list of modern, calibration isoGDGT data. **Mode 2**: $d$ represents a modern/calibration data-point and $a$ represents a list of ancient isoGDGT data.

For any given pairing of calibration and ancient data a *CD* value connects them, resulting in a *CD* matrix the size of the two datasets, in this study: 854 × 12,256. For a single ancient isoGDGT assemblage, a list of 854 *CD* value exists (mode 1 operation). The smallest number in this list is the $D_{Nearest}$ metric described by Dunkley Jones et al. (2020), which represents the closest modern calibration data-point to the ancient data-point; the remaining 853 points, are the distances from the ancient to all other calibration data-points. Ancient data with $D_{Nearest} \leq 0.5$ is considered to have a modern calibration data point sufficiently nearby for a robust temperature prediction to be made (Dunkley Jones et al., 2020). In mode 2 operation, a single calibration data-point is compared to a compilation of ancient data, as an example, for the Holocene epoch (isoGDGT data n = 556) a *CD* matrix results of size 854 × 556. In both mode 1 and 2 operation, we investigate the median *CD* value, $CD_{Median}$. Any force, regardless of source, that perturbs the temperature sensitive isoGDGT system away from the modern will be recorded as an increase in the value of $CD_{Median}$, in this manner this novel metric has the scope to capture the ”unknown unknowns” acting within the temperature sensitive space of isoGDGT assemblages through time.

Sampling bias is of concern in this study and the temporal sampling rates are summarised in Table 3.2. The average age gap between samples increases with age; during the Cenozoic it is 6.07 Ka and during the Mesozoic it is 15.66 Ka. The palaeolatitudinal distribution of data is shown in Figure 3.2. Within a given epoch, the percentage of data which fail $D_{Nearest} \leq 0.5$ is summarised in Table 3.2.

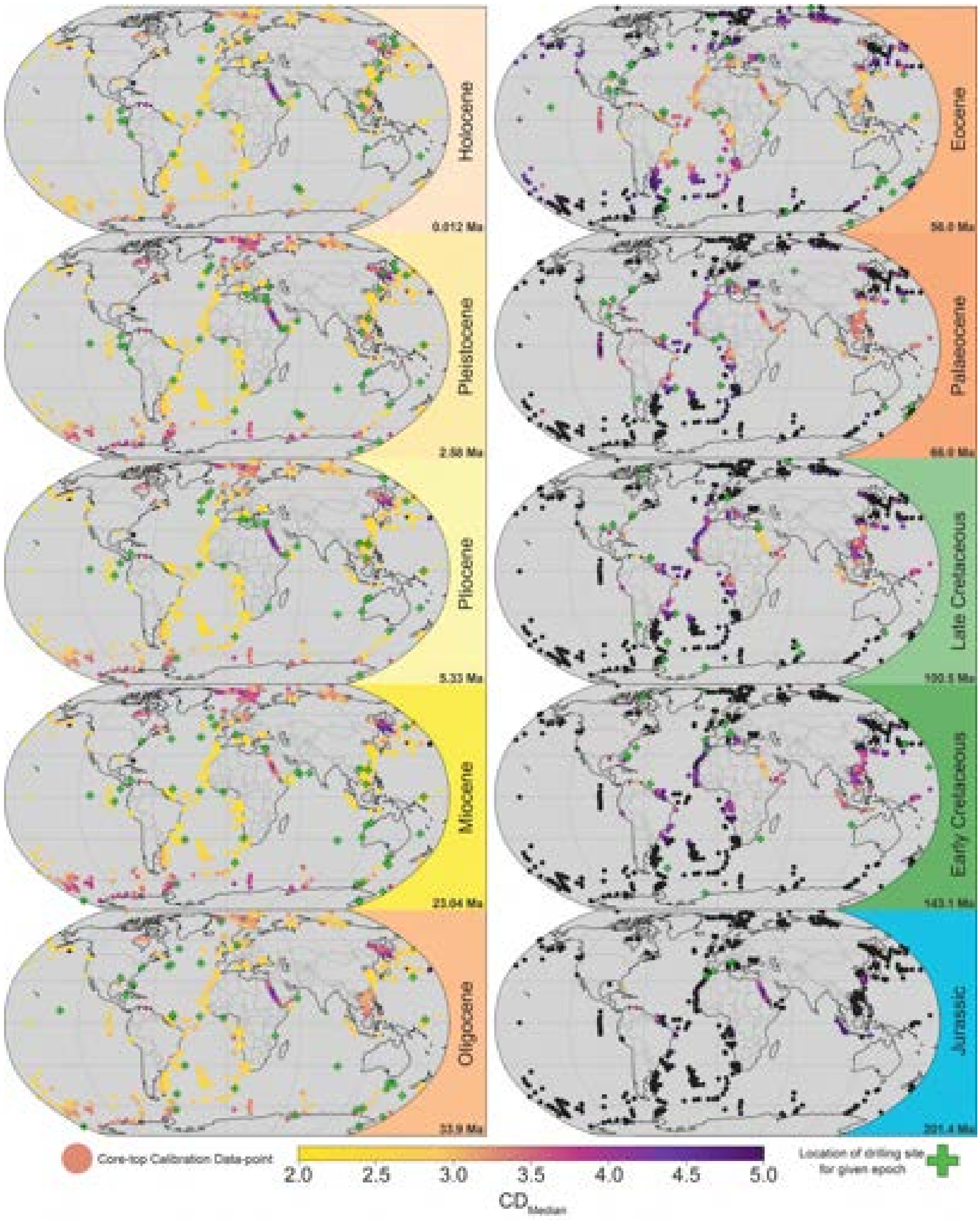


Figure 3.1: Calibration maps. Modern core-top calibration isoGDGT data (n=854) coloured by the $CD_{Median}$ values for a given epoch of ancient isoGDGT assemblages. Bright yellow colours highlight regions of the calibration dataset that are analogous to a given epoch and dark purple colours highlight data-points which are not. Black = median $CD_{Median} > 5$. Green crosses locate the sampling/drilling locations for the ancient isoGDGT data within a given epoch. Approximate basal ages for each epoch are shown in Ma.

| Epoch | No. Samples | Average age gap between samples [ka] | % Fail $D_{Nearest}$ Screening |
|---|---|---|---|
| Holocene | 556 | 0.02 | 12.6 |
| Pleistocene | 3132 | 0.82 | 33.2 |
| Pliocene | 2031 | 1.35 | 39.6 |
| Miocene | 2370 | 7.47 | 50.9 |
| Oligocene | 936 | 11.61 | 41.9 |
| Eocene | 1502 | 14.72 | 71.0 |
| Paleocene | 350 | 28.57 | 65.7 |
| Late Cretaceous | 881 | 38.68 | 87.4 |
| Early Cretaceous | 453 | 99.18 | 70.4 |
| Early Jurassic | 45 | 17.37 | 100.0 |

Table 3.2: The number of ancient isoGDGT assemblages compiled for each epoch (Holocene - Early Jurassic) and the average age spacing between sample, shown in Ka. The average age gap between samples through the Cenozoic is 6.07 Ka and through the Mesozoic is 15.66 Ka. Also highlighted are the percentage of data within an epoch that fail the $D_{Nearest} \leq 0.5$ screening test. Across all 12,256 ancient data-points, 48.5% of data fail the $D_{Nearest}$ screening test.

## 3.4 Results

### 3.4.1 *The Uniformitarian Assumption - Investigating the Modern isoGDGT System*

The *CD* metric is a unit-less measure, where higher values corresponds to a greater difference between the two datasets being compared. When calculated, a matrix of *CD* values is generated connecting all points across the modern calibration dataset (n = 854) and the compiled ancient dataset (n = 12,256). When considering a single data-point from either the modern or ancient datasets, a list of *CD* values results and we choose to analyse the median of this list, $CD_{Median}$. We hypothesise that uniformitarian behaviour of the isoGDGT system will result in transient perturbations of the $CD_{Median}$ metric across the many discrete episodes of climatic change across Earth's history (e.g. Paleocene-Eocene Thermal Maximum - PETM, Early Eocene Thermal Maximum - EECO etc.). In contrast, should the isoGDGT system display non-uniformitarian behaviour through the geological record, we hypothesise that this will be reflected in a long term, persistent increase in $CD_{Median}$ values with time. We investigated the relationship between the 854 modern, calibration data-points (Seki et al., 2014; Tierney and Tingley, 2015) and each epoch of ancient isoGDGT data (mode 2 operation, see methods) from the Holocene to the Jurassic, Fig. 3.1.

The majority of the calibration data-points show increasing $CD_{Median}$ values with epoch age, observed first in high latitude data and last in equatorial data. This result is expected, as modern, cold, high-latitude SSTs have limited or no analogues in past warm climates — even during the Miocene and Pliocene. In contrast, modern tropical systems, at least in the temperature sensitive isoGDGT space, resemble a broader range of ancient ocean conditions in warmer periods. There are exceptions to the trend of increasing $CD_{Median}$ values with age. First, the calibration data-points for most of the Red Sea (Trommer et al., 2009) and Cariaco Basin (Venezuelan coastline) (Turich et al., 2013) display decreasing $CD_{Median}$ values across the majority of the compiled record (Holocene - Cretaceous). This reinforces the previous conclusions that the Red Sea, although somewhat anomalous in the modern system, is of particular importance in understanding deep-time warm climates (Dunkley Jones et al., 2020). The relationship between the Red Sea and Cretaceous isoGDGT signals has been highlighted previously in the literature (Kim et al., 2010; Trommer et al., 2009; O'Brien et al., 2017) but is abundantly clear when the $CD_{Median}$ values are analysed, Fig. 3.1. Second, the Oligocene epoch shows a minor reversal in the long-term trend of increasing $CD_{Median}$ values with age, especially in polar calibration data. This result supports wider studies of the Oligocene epoch, characterized as a "one cold pole" climate, and is regarded as having cooler global mean temperature than either the Eocene or Miocene (Zachos et al., 2001). By the Jurassic, $CD_{Median} > 5$ for almost all calibration data, the implication being that there are no good modern analogues for isoGDGT assemblages preserved in the sampled Jurassic sediments.

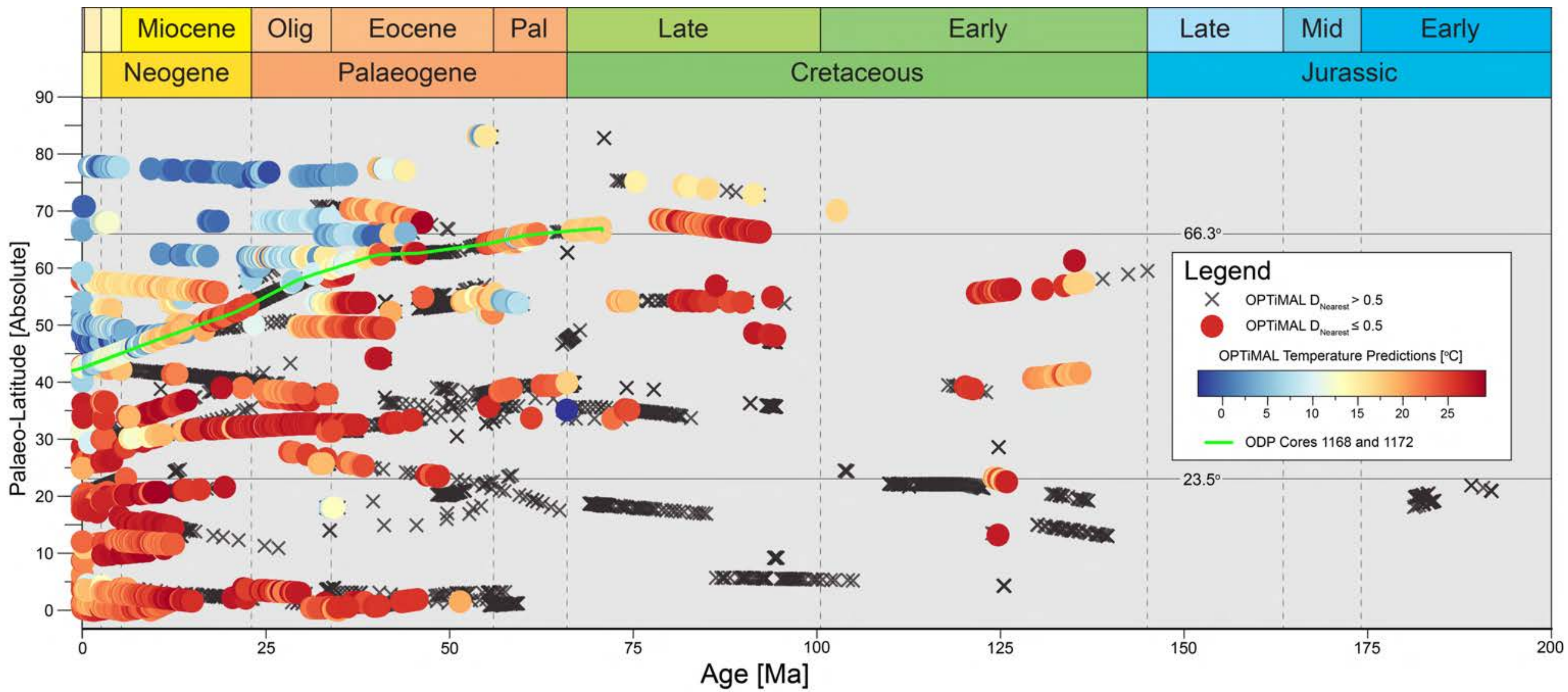


Figure 3.2: OPTiMAL Sea Surface Temperature (SST) predictions from the modern day to the Jurassic across a compilation of 12,256 isoGDGT assemblages. Filled circles = OPTiMAL $D_{Nearest} \leq 0.5$ and coloured by predicted SST, crosses = $D_{Nearest} > 0.5$. Green line marks the near continuous, Cenozoic isoGDGT record of ODP cores 1168 and 1172 (Fig. 3.4). Olig = Oligocene, Pal = Paleocene.

### 3.4.2 *Seeking Non-Uniformitarian Behaviour - 200 Ma Record of isoGDGT Assemblages*

OPTiMAL was used to predict SST on the 200 Ma record of compiled isoGDGT assemblages (Fig. 3.2) and to calculate the $CD_{Median}$ values between each ancient isoGDGT assemblage and the modern calibration data (mode 1 operation, see Chapter 3.3). The data compilation presented here is the biggest to date (n = 12,256), 56% larger than the next largest isoGDGT database, PhanSST (Judd et al., 2022). In Figures 3.2–3.4 we keep the original quality control metric of OPTiMAL (Dunkley Jones et al., 2020) ($D_{Nearest} \leq 0.5$) and only show OPTiMAL SST estimates for data which pass this test, whilst indicating failing data by a cross. Fig. 3.1 demonstrates that high-latitude calibration data contribute less to SST predictions as age increases, reflecting the reduced relevance of modern polar calibration data for past climates. In contrast, Fig. 3.2 illustrates the rising failure rate of SST predictions (increasing number of crosses) made on low palaeo-latitude data over time, highlighting how palaeo-tropical environments increasingly shift into a 'non-analogue to modern' climate space in warmer periods of deep time. Whilst a near uninterrupted, global record of isoGDGTs exists from the Holocene to the Late Cretaceous, low reliability data ($D_{Nearest} \leq 0.5$) begins to dominate the record from the mid-Eocene. Across the entire 200 Ma record, 48.5% of the compiled data (n = 5953) fails the $D_{Nearest}$ test and would be considered "unreliable" for SST prediction.

Analysis of the $CD_{Median}$ values seeks to extract information from all data-points, including those which would fail screening tests such as $D_{Nearest}$. Figure 3.3 compares the global $CD_{Median}$ record against a composite record of deep ocean, benthic foraminiferal $\delta^{18}$O spanning the Holocene–Jurassic (Judd et al., 2022; Westerhold et al., 2020). We choose $\delta^{18}$O as an independent proxy of oceanic temperature with which to compare against the $CD_{Median}$ record and its GP imbued temperature sensitivity. The resulting aggregate signal of 3.3A is representative of data from all latitudes and therefore may

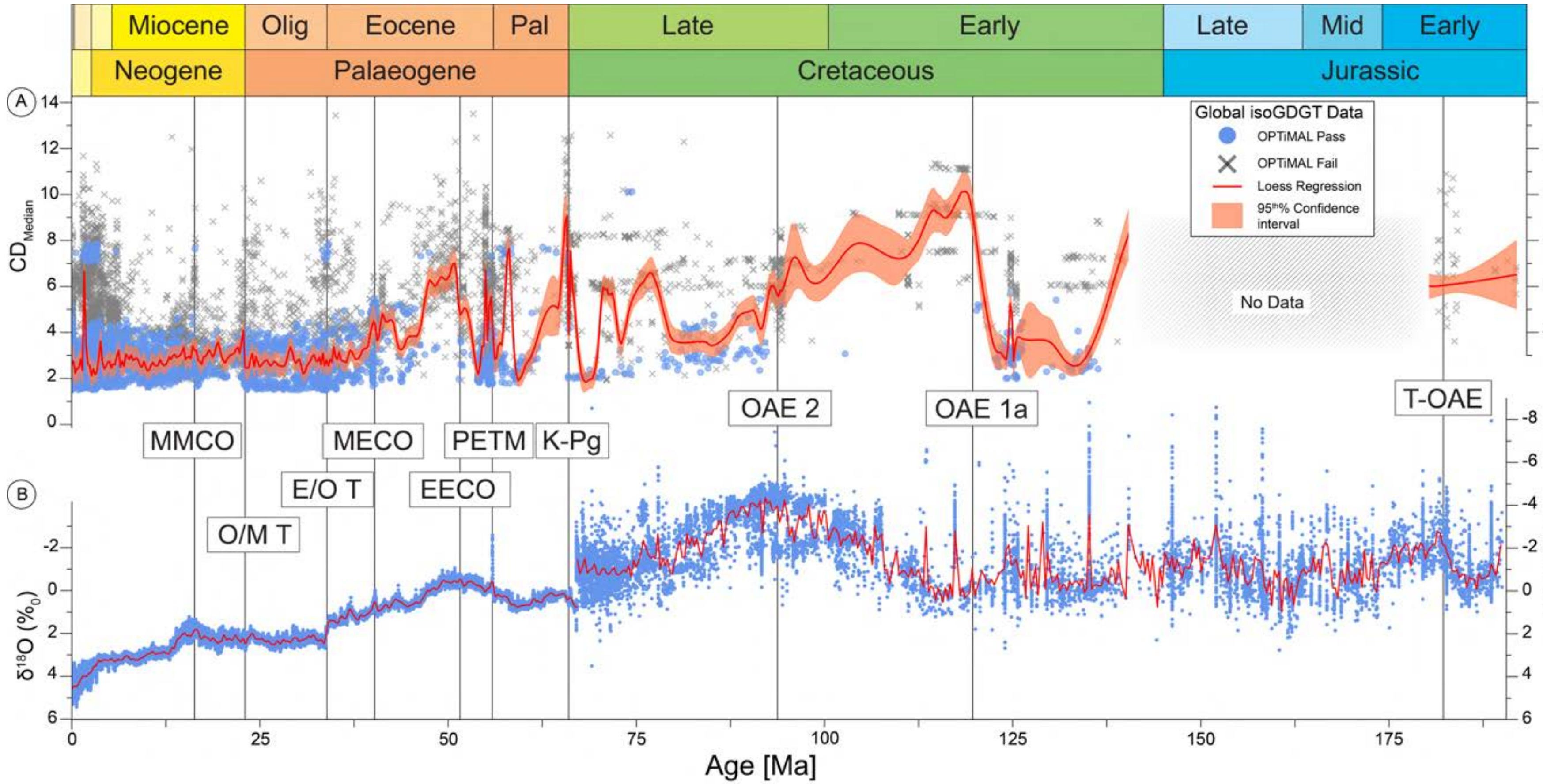


Figure 3.3: A) OPTiMAL calculated $CD_{Median}$ values for the compilation of 12,256 isoGDGT assemblages. Filled circles = OPTiMAL $D_{Nearest} \leq 0.5$, crosses = $D_{Nearest} > 0.5$. Red line = Loess regression. 10 global climate events: Middle Miocene Climate Optimum (MMCO), Oligocene/Miocene Transition (O/M T), Eocene/Oligocene Transition (E/O T), Middle Eocene Climate Optimum (MECO), Early Eocene Climate Optimum (EECO), Paleocene–Eocene Thermal Maximum (PETM), Cretaceous Palaeogene Boundary (K-Pg), Oceanic Anoxic Events 2 and 1a (OAE 2 and 1a), Torcian OAE (T-OAE). B) A compilation of two benthic foraminiferal $\delta^{18}O$ records: Cenozoic record Westerhold et al. (2020) and Mesozoic record Judd et al. (2022). Red line = Loess regression. $\delta^{18}O$ is a widely used stable isotope proxy for oceanic temperature. Olig = Oligocene, Pal = Paleocene.

record any latitudinal bias of the underlying data. This bias could result in decoupling between the $CD_{Median}$ record and the bottom water $\delta^{18}$O record. Despite this however, the $CD_{Median}$ data archive a record of the major global Cenozoic climate events. Perturbations in $CD_{Median}$ values can be attributed to any force influencing the temperature sensitive isoGDGT system. Whilst an individual $CD_{Median}$ value is arbitrary, the relative change from one data-point to the next is informative. $CD_{Median}$ values begin in the Holocene at c. 2 whilst the oldest data in the Jurassic show $CD_{Median}$ values c. 7. However, as far back as the Early Cretaceous low $CD_{Median}$ values are observed, implying that a high fraction of the ancient isoGDGT assemblages of deeper time records bear resemblance to those produced in specific modern day oceanic settings (e.g. the Red Sea, Fig. 3.1). This observation, that the ancient isoGDGT record periodically returns to quasi-modern/observable isoGDGT assemblages, may indicate that the isoGDGT system satisfies uniformitarian assumption. The notion of stable temperature dependency within the isoGDGT system appears supported by transient perturbations of the $CD_{Median}$ record and for robust temperature predictions ($D_{Nearest} \leq 0.5$) to be made on lipids as far back as the Early Cretaceous. However, it is possible that the course of geological time shifts an ancient isoGDGT distributions from outside to within the feasible range of the modern, temperature weighted isoGDGT space. This is a challenging behaviour to detect and one which OPTiMAL cannot offer insight into. It can however test explicitly for periods of non-uniformitarian behaviour, such as periods of increased $CD_{Median}$ values in the Greenhouse world of the early Cenozoic and late Cretaceous, which we attribute here to changes in the Earth's climate, moving the isoGDGT system transiently outside the range of the modern calibration dataset.

### 3.4.3 ***Detecting Non-Analogue Palaeoenvironments***

Finally, we apply the new $CD_{Median}$ metric to the sediment cores of ODP 1168 and 1172 (mode 1 operation). These cores were selected as they provide a near continuous

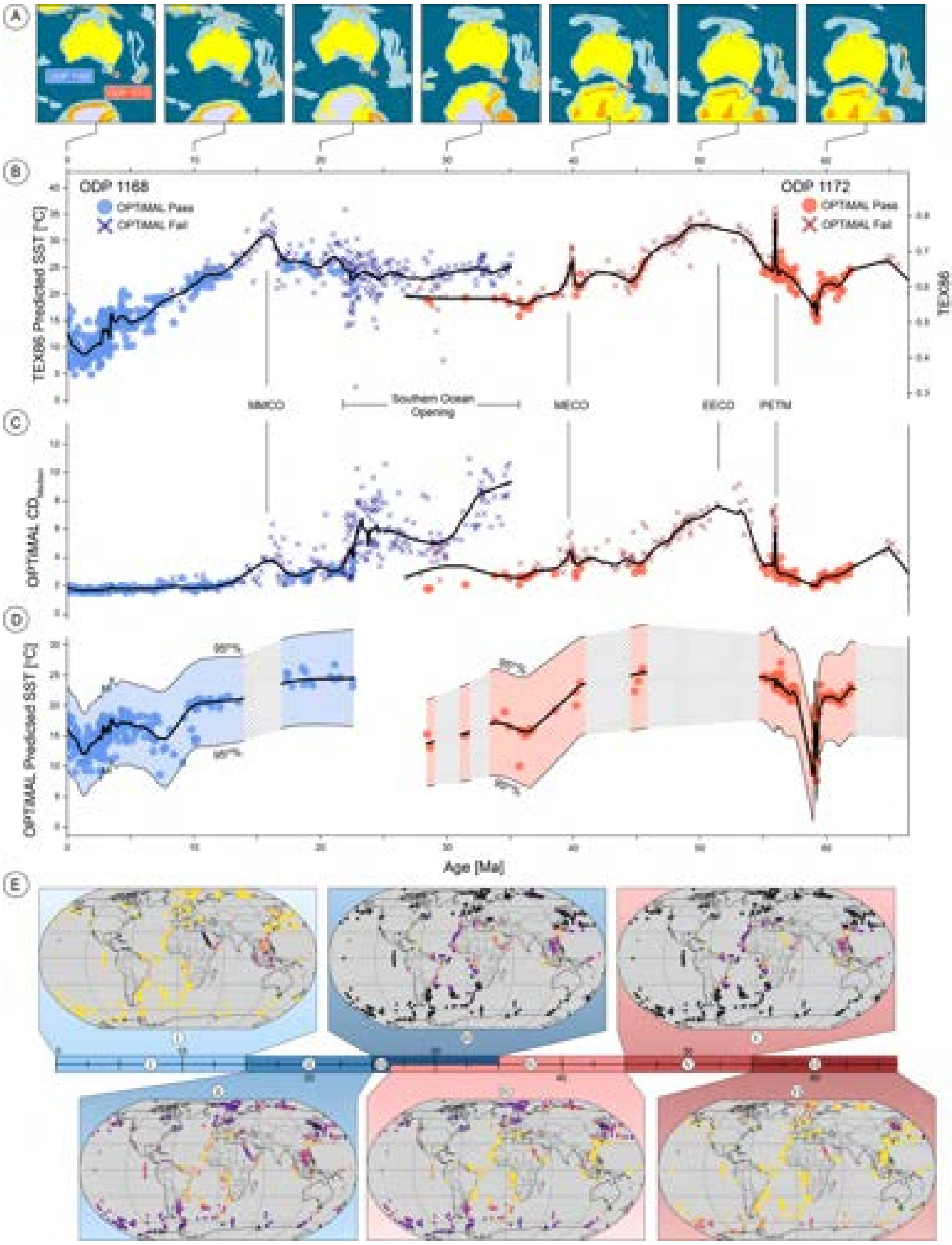


Figure 3.4: ODP 1168 and 1172 isoGDGT record. A) A palaeogeographic reconstruction Cao et al. (2017) of the core locations every 10 Ma, made using the program: GPlates. Orange = mountains, yellow = landmass, light-blue = shallow ocean, dark-blue = deep ocean B) $TEX_{86}$ record spanning ODP 1168 (blue) and 1172 (red). Filled circles = OPTiMAL $D_{Nearest} \leq 0.5$, crosses = $D_{Nearest} > 0.5$. Black line = Loess regression. C) OPTiMAL $CD_{Median}$ values. Marked are four palaeoclimate events affecting these records; Middle Miocene Climate Optimum (MMCO), Middle Eocene Climate Optimum (MECO), Early Eocene Climate Optimum (EECO), Paleocene–Eocene Thermal Maximum (PETM) D) OPTiMAL sea surface temperature predictions with $95^{th}$% bounds, Loess smoothing is applied over a span of 0.1 Ma. Hatched zones mark gaps in the predictive record where $D_{Nearest} > 0.5$. E) Calibration maps through six time slices (I–VI) of the ODP cores. See Fig. 3.1 for explanation.

record of isoGDGT data through the Cenozoic (Stickley et al., 2004; Sluijs et al., 2011; Houben et al., 2019; Bijl et al., 2021) (Fig. 3.2, green line). ODP 1168 and 1172 exist respectively on the western (enclosed basin) and eastern (open Pacific Ocean) continental margins of Tasmania. ODP 1168 records the final stages of opening of the Australian Bight and the formation of the Southern Ocean as Australia progressively separates from Antarctica (Stickley et al., 2004) (Figure 3.4A). Figure 3.4B–C provide a direct comparison between the classic $TEX_{86}$ SST proxy and $CD_{Median}$ values calculated as part of the new approach presented here. It is notable that despite the simplicity of the $TEX_{86}$ formulation, it performs well in evidencing the four major climate events acting on these ODP cores (Miocene Climate Optimum - MCO, Middle Eocene Climate Optimum - MECO, Early Eocene Climate Optimum - EECO, and the Paleocene–Eocene Thermal Maximum - PETM). However, $TEX_{86}$ appears largely unaffected by the oceanographic changes associated with the opening of the Southern Ocean, implying this developing ocean exerts no pressure on the isoGDGT palaeo-thermometer. In contrast, $CD_{Median}$ values calculated for ODP 1168 evidence a substantial change between c. 22–35 Ma from more to less non-analogue conditions over this time period, with particularly rapid change close to the Eocene–Oligocene and Oligocene–Miocene boundaries. The overlapping record of the neighbouring ODP 1172 core, at very similar latitude, indicates the changes in isoGDGT assemblages are restricted to within the opening Australian Bight and not some heretofore unrecognised climate event. This perturbation in the $CD_{Median}$ record is as large as the most severe hot-house climate events (i.e. EECO or PETM), implying that non-thermal, oceanographic changes can have just as big an impact on the temperature-sensitive component of isoGDGT assemblages as thermal, climate events – remembering that $CD$ values are explicitly scaled relative to temperature sensitivity in the modern calibration dataset. Restructuring of the Southern Ocean plankton ecosystem has been reported through this time period (c. 33.6 Ma), attributed to the first major Antarctic glaciation in the earliest Oligocene (Houben et al., 2013). Until now, the ability to directly compare disparate forces acting on isoGDGT temperature dependency in a single metric was not possible, this is however a strength of $CD_{Median}$ metric.

Figure 3.4D plots the OPTiMAL SST predictions from ODP 1168 and 1172, highlighting the sparseness of the record once data failing OPTiMAL's screening test ($D_{Nearest} \leq 0.5$) are removed. Large gaps exist in the SST record during global hot-house climate events, likely due to the ancient SSTs exceeding the upper limit of the calibration data (c. 30 $^{o}$C). Fig. 3.4E presents the modern calibration data coloured by $CD_{Median}$ values associated with time-slices of the ODP core data. These are similar plots to Fig. 3.1, however the ancient data used is restricted to the two ODP cores and the indicated time slices.

Fig. 3.4E III and IV reinforce the geographical restriction of the effects of the Southern Ocean opening to the ODP 1168 record. Regarding uniformitarianism in the isoGDGT system, Fig. 3.4E I–VI demonstrate that whilst perturbations exist across climate or geographical events, the encoding of temperature information appears conserved across this Cenozoic record. The $CD_{Median}$ metric enables testing of fundamental questions and presents a streamlining of the traditional isoGDGT analysis, condensing information normally spread over a range of isolated indices/metrics into one, all-inclusive measure, capable of tracking both the known unknowns and unknown unknowns.

## 3.5 Discussion

The quasi-modern isoGDGT assemblages observed in deep time settings, such as the Early Cretaceous (Fig. 3.3), may indicate uniformitarian behaviour in isoGDGT production through time. The modern ocean, and the resulting calibration dataset, must represent a naturally curated subset of the wider isoGDGT system that has operated with largely conserved systematics through Earth's history. Modern day sampling sites such as the Cariaco Basin (Turich et al., 2013) and the Red Sea (Fietz et al., 2016), which at times experience decreasing $CD_{Median}$ values with age, prove vital in extending the observable range of the isoGDGT system. It has been suggested that in the Cariaco Basin, stratified production of lipids at depth alters the archived isoGDGT record (Turich et al., 2013). In the Red Sea, the unusual character of the isoGDGT assemblages has been attributed to a distinct population of lipid producers residing in the northern Red Sea (Fietz et al., 2016). It is a non-trivial challenge to curate a calibration dataset that is neither over-cleaned nor diluted by non thermally-sensitive data, but the $CD_{Median}$ metric provides an agnostic metric to aid in this challenge. It would seem proactive to extend the calibration dataset with the currently under-represented, low-latitude locations of Indian coastal waters, eastern Africa, northern Australia, Madagascar, Timur Sea, Arafura Sea, and the Caribbean. The upper end of the modern SST range is an important reference for the typically hotter past ocean states, especially prior to the Oligocene.

Due to the formulation of the $CD_{Median}$ record, it is an aggregate signal of all forces acting on the production and archiving of isoGDGTs. Long timescale (100's Myr) perturbations in this metric might reflect diagenetic or evolutionary forces acting on the system, whereas shorter term (10's Myr) perturbations may reflect changes in microbial community composition, global climate or geographical events. The traditional suite of indices and metrics developed to screen isoGDGT data (e.g. BIT (Hopmans et al., 2004), MI (Zhang et al., 2011), $\Delta RI$ (Zhang et al., 2016) etc.) will prove important in deconvolving the $CD_{Median}$ signal into its constituent parts. But crucially, when the sum of these known components cannot fully explain the $CD_{Median}$ signal, the influence of unknown components, and importantly the quantified scale of this influence, will at least be constrained. Experimental/culture studies represent an obvious target for $CD_{Median}$ analysis however, results are often reported in the form of the reductive $TEX_{86}$ index (Schouten et al., 2007a; Wuchter et al., 2004) in contrast to the full suite of isoGDGTs. The $CD_{Median}$ metric would facilitate investigation of the relationship between high temperature ($> 30$ °C), culture derived isoGDGTs and isoGDGTs produced during hothouse climate events. To this extent, archaeal strains derived from locations such as the Red Sea and the Cariaco Basin should be the focus of future studies. Extending the observable/measurable subset of the isoGDGT system would enable a greater percentage of the geological isoGDGT record to be accessed for useful prediction.

## 3.6 Conclusion

We revisit and update the Gaussian Process emulator, OPTiMAL, to extend its capabilities beyond just the prediction of Sea Surface Temperatures (SST) from isoGDGT assemblages. We present the largest compilation of isoGDGT records to date (12,256 ancient isoGDGT assemblages) and, using OPTiMAL, we provide a predictive SST record from the present day to the Early Jurassic (c. 200 Ma). In addition to the SST predictions we provide a new metric, Contextualised Distance ($CD$) metric, which we use to assess the uniformitarian assumption on the isoGDGT data archived in the geological record. This new, agnostic metric streamlines the traditional isoGDGT workflow, split over 10's of individual metrics, into one single metric capable of tracking any force affecting the production and archiving of temperature information within isoGDGT assemblages. For the first time, we visualise the spatial and temporal variation in the analogue status of the modern ocean in predicting past SST through every epoch from the present day to the Jurassic. $CD_{Median}$ values are applied to the compiled ancient isoGDGT data, revealing a record of the major global climate events in the Cenozoic and the long-term stability of this climate archive. Whilst perturbations exist in the isoGDGT record, the ability for this system to return to quasi-modern isoGDGT assemblages through a 200 Ma record suggests that temperature dependency within isoGDGT production through time satisfies the uniformitarian assumption.

## 3.7 Code and Data Availability

The code used to generate all data and figures in this manuscript is available at the following link: `https://github.com/matthew-allison05?tab=repositories`. Data compiled as part of this study is visualised in Appendix A.2, Figure A.6, alongside a full reference list.

# 4 | Barmur Group Lignites: Insights into the preservation of organic biomarkers in high geothermal gradients

*The contents of this chapter are the work of this author. brGDGT analysis and integration was conducted at the National Environmental Isotope Facility (NEIF) - University of Bristol.*

## 4.1 Abstract

This study investigates the preservation of palaeoclimatic information within organic biomarkers - specifically brGDGTs - recovered from the lignite rich sediments of the Barmur Group (c. 4.5–7.5 Ma), exposed on the Tjörnes Peninsula, north-east Iceland. The high geothermal gradients present in N. Iceland present a unique opportunity to investigate the impacts of burial diagenesis on biomarker systems. Stratigraphic and burial models, groundtruthed with vitrinite reflectance data recovered from drill cores, predict a geothermal gradient of c. 130 $^{o}$Ckm$^{-1}$ and a maximum experience burial temperature of 96 $^{o}$C for the lowermost lignite horizon. A logarithmic loss in brGDGT abundance is observed with stratigraphic depth and correlated to independent assessments of diagenesis using *n*-alkane, Carbon Preference Index (CPI) and hopanoid stereoisomer ratios, confirming systematic thermal degradation trends with depth. A Gaussian Process Classifier (GPC) is developed to predict palaeodepositional environments from brGDGT assemblages, enabling targeted application of four depositional-specific Gaussian Process Emulator (GPE) models trained on compilations of global, modern brGDGT distributions. This machine learning framework outperformed conventional $MBT'_{5ME}$ calibrations, reducing predictive uncertainty and providing realistic estimates of Warmest Month air Temperature (WMT: 6.4–20.8 $^{o}$C) and In Situ Water Temperatures (ISWT: 6.0–21.1 $^{o}$C). Results indicate that brGDGTs recovered from the Barmur Group most reliably reconstruct WMT over MAAT supporting warm temperature bias in brGDGT distributions. This multi-proxy approach demonstrates the potential for robust palaeoclimatic reconstruction from thermally altered brGDGT archives when paired with predictive stratigraphic modelling and machine learning classification.

## 4.2 Introduction

Iceland's geological record is overwhelmingly volcanic, with approximately 90% of the island formed from primary volcanic activity. Of the remaining 10%, the sediments found on the Tjörnes Peninsula represent the thickest onshore sedimentary sequence on the island, Fig. 4.1. The Barmur Group comprises a c. 440 m thick sedimentary sequence dated to between c. 4.3–6.0 Ma (Hall et al., 2023; Allison et al., 2026). Originally studied by Bárdarson (1925), the Barmur Group (formerly the "Tjörnes Beds") is well known for its mollusc assemblages, defining three key biozones: *Tapes*, *Mactra*, and *Serripes*. These shallow marine to intermittently terrestrial sediments capture environmental changes linked to shifts in ocean circulation and gateway connectivity

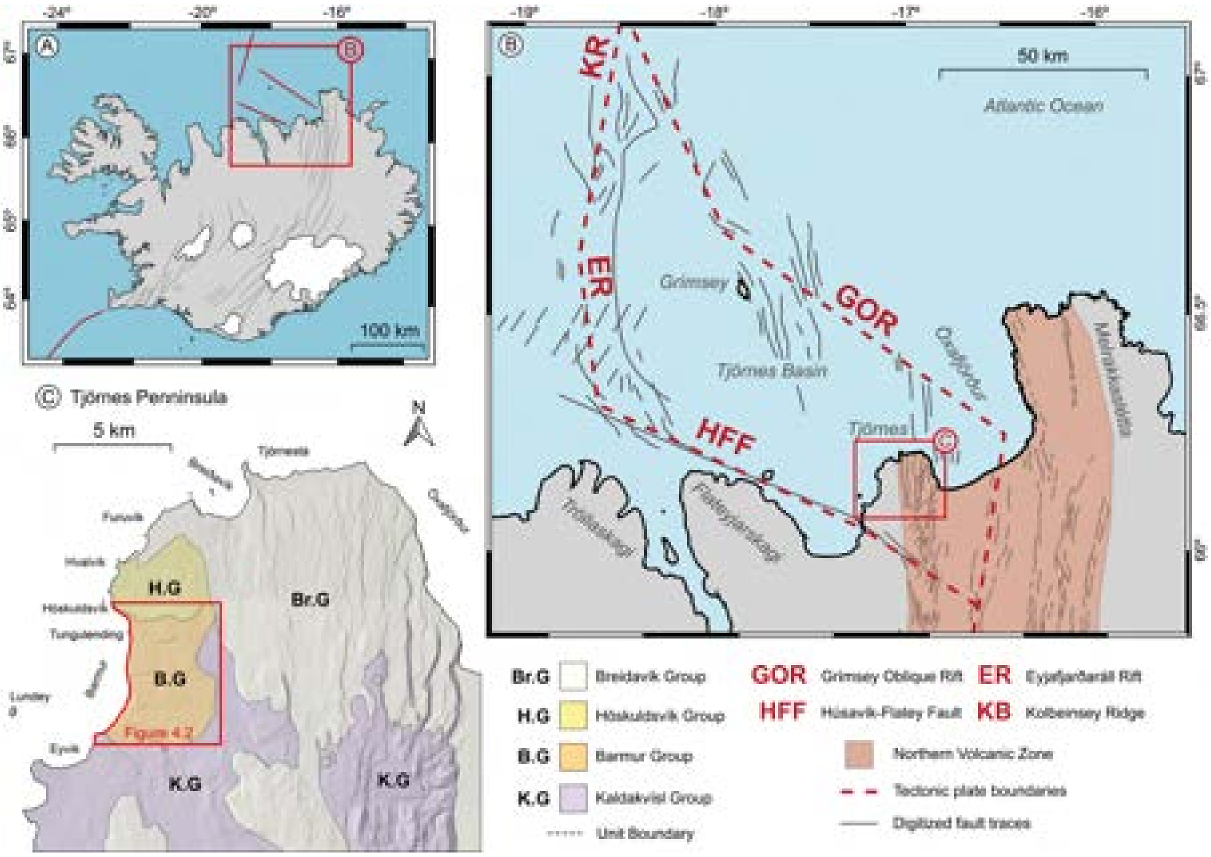


Figure 4.1: Regional map of study area. A) Iceland. B) NE Iceland and the Tjörnes Fracture Zone. C) The geology of the Tjörnes Peninsula: Kaldakvísl Group, Barmur Group, Höskuldvík Group, and the Breiðavík Group. Source: B) fault traces, C) DEM, National Land Survey of Iceland.

(Bárdarson, 1925; Símonarson and Eiríksson, 2008; Eiríksson and Símonarson, 2021; Hall et al., 2023). The Barmur Group rests unconformably atop the Kaldakvísl lavas and is overlain by subaerial lava flows of the Höskuldvík Group, dated using $^{40}Ar/^{39}Ar$ at 5.06 ± 0.16 Ma (Hall et al., 2023). Above the Höskuldvík Group is the c. 480 m thick Breiðavík Group, which comprises alternating lava and sedimentary units interspersed with diamictites marking the onset of lowland glaciation in Iceland (Eiríksson, 1981, 1985; Eiríksson et al., 1994; Eiríksson and Símonarson, 2021).

The Barmur Group consists of alternating marine (siltstones, sandstones and conglomerates) and terrestrial sediments (siltstones, lignites and sandstones) occasionally interrupted by basaltic lavas flows (Eiríksson and Símonarson, 2021), Fig. 4.2. The terrestrial lignite horizons are predominantly found in the *Tapes* and *Mactra* biozones, and interpreted as periods of emergence in a tidal flat - coastal swamp depositional environment (Eiríksson and Símonarson, 2021; Hall, 2021). A return to regional uplift during the upper *Serripes* biozone produced sedimentation typical of estuarine and paludal depositional environments (Buchardt and Símonarson, 2003). The Barmur Group sediments have been the subject of study for a century, with palaeoclimatic reconstructions drawing on molluscs (Bárdarson, 1925; Strauch, 1972), oxygen isotopes (Gladenkov et al., 1975), $\delta^{18}$O from marine shells (Buchardt and Símonarson, 2003), pollen and dinoflagellate assemblage analysis (Verhoeven, 2012; Verhoeven and Louwye, 2013) and brGDGT (Branched Glycerol Dialkyl Glycerol Tetraethers) Mean Annual Air Temperature (MAAT) reconstruction (Hall, 2021). A range of hydrocarbon and geothermal studies have also been undertaken across the wider peninsula, including investigations into the Öxarfjörður gas seeps on the east of the Tjörnes Peninsula (Ólafsson et al., 1993; Ármannsson et al., 1998), which are hypothesised to originate from the Barmur Group lignites buried at depth (Richter and Gunnarsson, 2010). In 2002, drill cores were taken at the coastal exposure above the Skeifá waterfall (Fig. 4.2) to recover unweathered material from lignite horizons and provide source rock analysis for lignites beds D and F (Richter and Bojesen-Koefoed, 2002).

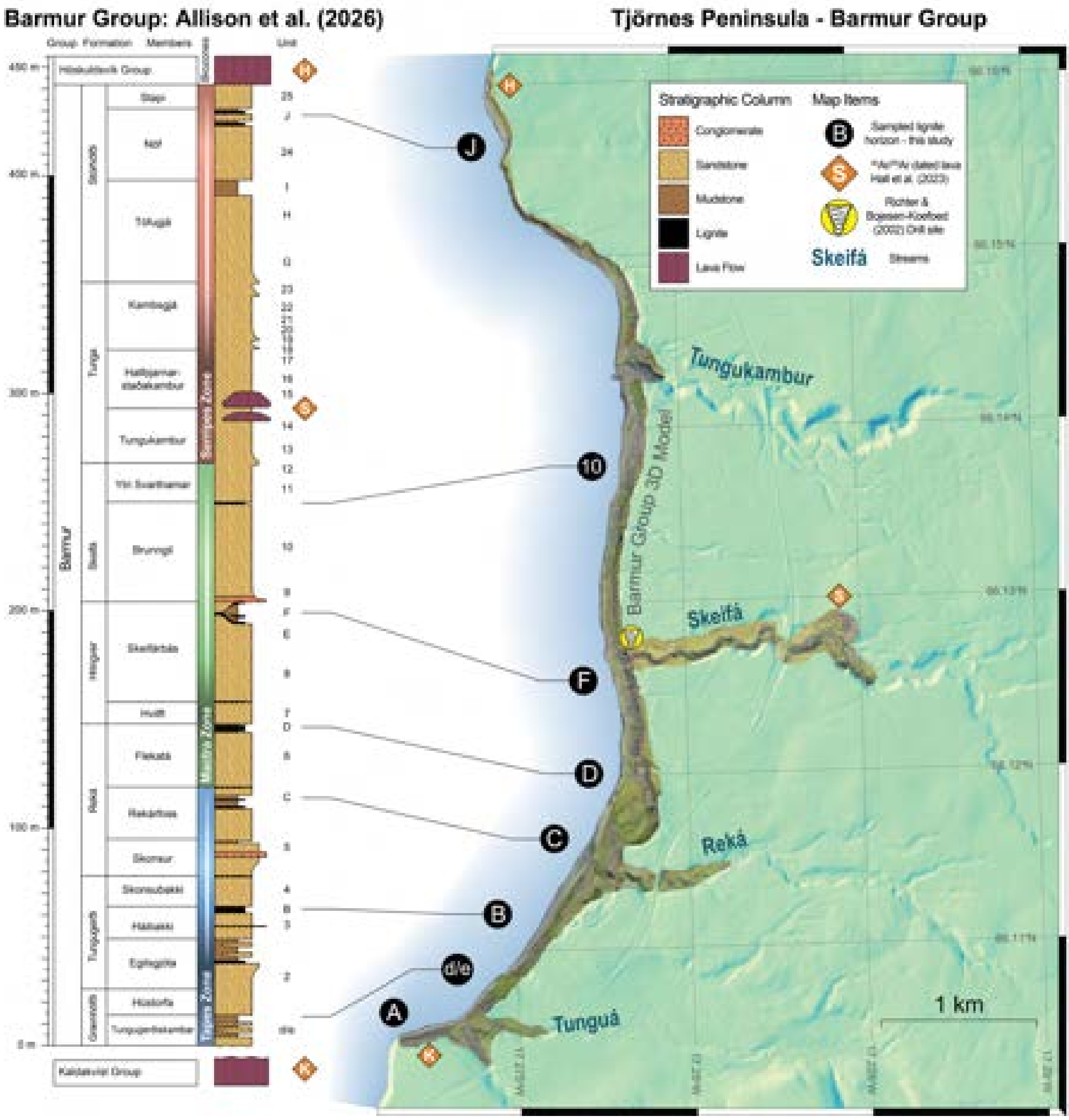


Figure 4.2: Local map of the Barmur Group sediments exposed along the coastal and stream-section exposures. A base-map of a 2 m DEM (Digital Elevation Model) is used. The coastal 3D photogrammetry model and modelled stratigraphic column of Allison et al. (2026) are shown alongside the lignite sampling locations of this study. $^{40}Ar/^{39}Ar$ ages of Hall et al. (2023) and drilling site of Richter and Bojesen-Koefoed (2002) are highlighted. Source: DEM, National Land Survey of Iceland.

The focus of this study are the terrestrial biomarkers, brGDGTs, expanding on published records from the Barmur Group lignites (Hall, 2021). Produced by bacterial communities, the relative abundance of the 15 lipid isomers (Fig. 4.3) have been demonstrated to correlate with environmental variables such as environmental temperature and pH (Weijers et al., 2007, 2011). Recent studies have suggested that brGDGT distributions may be best predicted by warm-biased temperature measures, such as Warmest Month Air Temperature (WMT) (Crampton-Flood et al., 2020; Raberg et al., 2021, 2022). brGDGT distributions have also been shown to vary as a function of depositional environment (Tierney et al., 2010). In this study we utilise the brGDGT

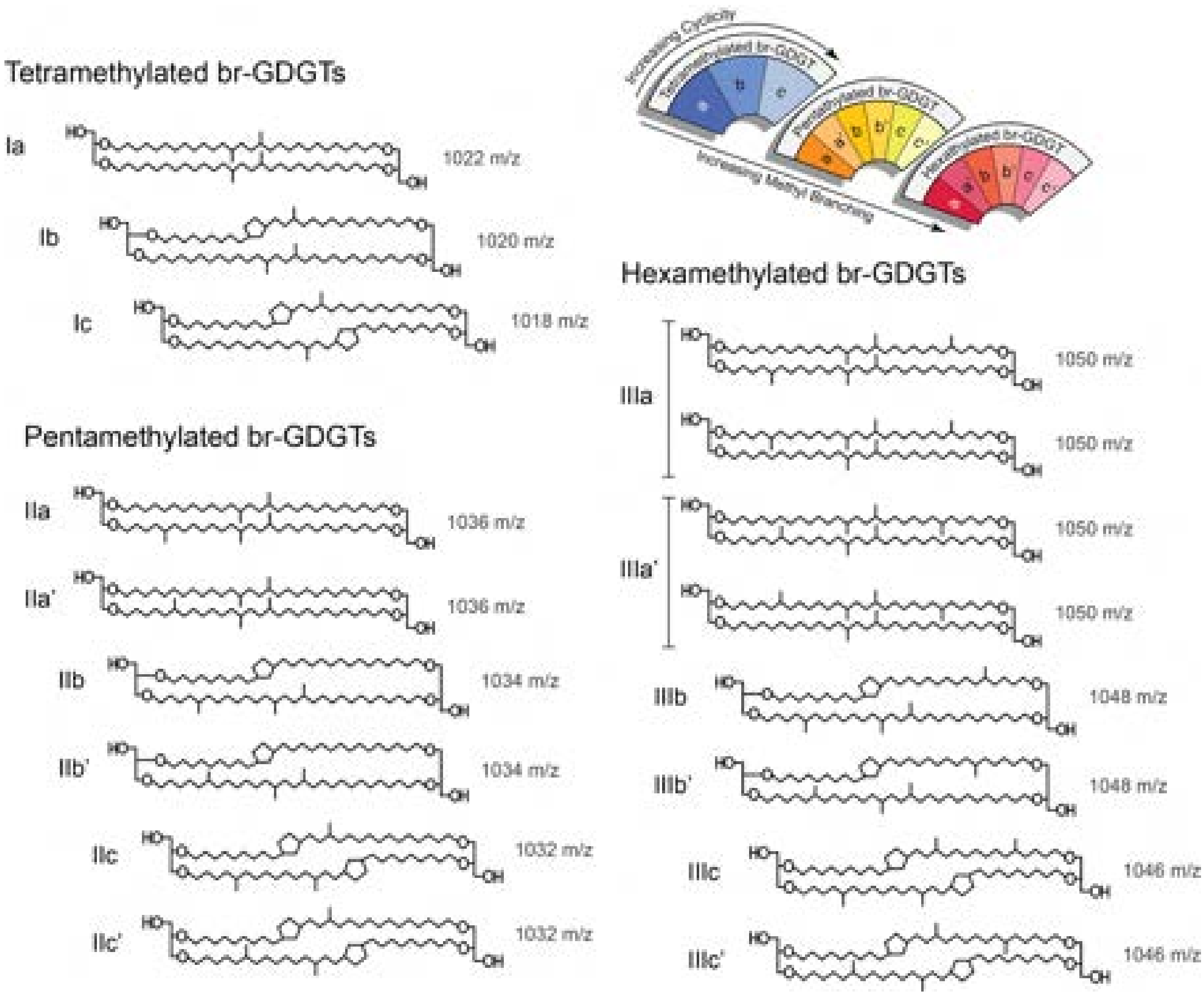


Figure 4.3: 15 brGDGT isomers split into the three categories of tetramethylated, pentamethylated and hexamethylated brGDGTs. Mass/charge (m/z) ratio for HPLC-MS analysis are shown alongside each structure.

biomarker as a potential archive of palaeoclimatic information for the reconstruction of past air temperature (MAAT, and WMT) and In Situ Water Temperature (ISWT). Traditional approaches utilising the $MBT'_{5ME}$ climate proxy (De Jonge et al., 2014) to reconstruct the palaeoclimate from the Barmur Group lignites resulted in an unrealistically wide range of air temperatures c. -12–19 $^{o}$C, attributed to thermal degradation of the lipids from burial diagenesis (Hall, 2021). Here we trial the Gaussian Process Emulator (GPE) approach of OPTiMAL (Dunkley Jones et al., 2020), but trained on a calibration dataset of global, modern brGDGT data (Raberg et al., 2022). Before making temperature predictions using the GPE, palaeodepositional environments are predicted using a Gaussian Process Classifier (GPC), enabling depositionally specific GPE models to be applied. The use of a Gaussian Process for both classification and prediction facilitates an agnostic and flexible approach to brGDGT analysis, providing a means to make meaningful palaeoclimatic predictions from the potentially degraded brGDGT data of the Barmur Group.

The exploratory history at the Tjörnes Peninsula provides a unique opportunity to combine a detailed study of biomarkers with a burial model and thermal history of the sediments and lignite bearing horizons. For any study of organic biomarkers preserved in ancient sediment, the impact of diagenesis represents an inescapable challenge. Laboratory studies cannot replicate geological timescales, and the average geothermal gradient of sedimentary basins (c. 25 $^{o}$Ckm$^{-1}$ - Kolawole and Evenick (2023)) restricts high temperature environment to deep burial settings. However, proximity of the Tjörnes Peninsula to the Atlantic spreading axis of the Northern Volcanic Zone (NVZ) results in a high heat flux, in the range of from 100–450 mWm$^{-2}$ (Davies and Davies, 2010; Elders and Fridleifsson, 2010). This geological curiosity presents the opportunity to study organic (biomarker) rich sediments which have experience high heat flux with relatively minimal burial (< 2 km), with vitrinite reflectance data from the Barmur Group lignites (Richter and Bojesen-Koefoed, 2002) enabling the thermal history of the Tjörnes Peninsula to be modelled. These investigations are complemented by investigations of *n*-Alkane and hopanoid compounds recovered from the lignites, these biomarkers having previously been used to estimate the effects of thermal maturation using the Carbon Preference Index (CPI: n-alkanes) (Cooper and Bray, 1963; Allan and Douglas, 1977; Nelson et al., 1998) and Geohopanoid Index (Mackenzie et al., 1980; Peters and Moldowan, 1991). The combination of traditional source rock analysis, burial modelling, and study of organic biomarkers provides the opportunity to better understand the effects of diagenesis and thermal degradation on the brGDGT system.

### 4.2.1 *Geological Background - Barmur Group*

The basal Kaldakvísl Group forms a widespread basement to the overlying geology at the Tjörnes Peninsula. Subaerial lavas flows are interbedded with thin, red coloured sediments (Eiríksson and Símonarson, 2021) and thickness estimates of the Group range from 500 m to > 1000 m (Sæmundsson, 1974). Radiometric $^{40}Ar/^{39}Ar$ investigations date the youngest lava flows at 9.78 ± 0.13 Ma (Hall, 2021), preceding an unconformity of c. 3–4 Ma before the onset of Barmur Group sedimentation. $^{40}Ar/^{39}Ar$ ages of dykes across North Iceland indicate that the North Atlantic spreading axis is predicted to have jumped to its modern position c. 8–8.5 Ma (Garcia et al., 2003). The palaeo-rift, located westward at Skagafjördur, and the NVZ are predicted to have operated synchronously until c. 2.8 Ma, at which point the NVZ operated as the sole spreading axis in Northern Iceland (Garcia et al., 2003). The Kaldakvisl Group is a physical manifestation of this rift relocation and associated volcanism, during which time the Tjörnes Fracture Zone (TFZ) developed as a trans-tensional fault system to accommodate plate motion and transfer spreading from the onshore NVZ to the offshore Kolbeinsey Ridge system (Metzger et al., 2013a).

The Barmur Group is a collection of shallow marine to occasionally sub-aerial sediments, accommodated by the developing basin within the TFZ (Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Verhoeven and Louwye, 2013; Eiríksson and Símonarson, 2021). Sedimentological observations and radiometric dating indicate a relatively consistent long-term accumulation rate, with a small rate increase towards the top of Barmur Group (Símonarson and Eiríksson, 2008; Eiríksson and Símonarson, 2021; Hall et al., 2023). Minor hiatuses in sedimentation exist throughout the Group, consistent with the coastal depositional environment, but no major angular unconformities are evidenced within the sedimentary package (Eiríksson and Símonarson, 2021; Hall et al., 2023; Allison et al., 2026). The original classification of the Barmur Group lithostratigraphy by Bárdarson (1925) defined 25 marine (Units 1–25) and 10 terrestrial (Units A–J) marker horizons. Three biozones were defined from mollusc assemblages: the basal *Tapes* and overlying *Mactra* biozones are rich in Atlantic fauna, in sharp contrast to the *Serripes* biozone, which is rich in Pacific fauna derived from the opening of the Bering Strait (Buchardt and Símonarson, 2003; Símonarson and Eiríksson, 2008; Verhoeven, 2012; Eiríksson and Símonarson, 2021; Allison et al., 2026). The most recent lithostratigraphical framework provided for the Barmur Group defines 10 formations and 23 members on the basis of 22 lithological logs (Eiríksson and Símonarson, 2021) and 3D modelling of this framework predicts a total sedimentary thickness of c. 440 m (Allison et al., 2026)

Intra-sedimentary age constraints for the Barmur Group arise from $^{40}Ar/^{39}Ar$ dating of the Skeifá Lava (5.06 $\pm$ 0.16 Ma - Hall et al. (2023), Fig. 4.2) and interpretation of palaeomagnetic data (Doell, 1972; Kristjánsson, 2004). Some uncertainty exists in the palaeomagnetic data reported from the Barmur Group, resulting in two subtly different interpretations of the *Mactra*-*Serripes* biozone boundary (Allison et al., 2026). The sediments are capped by the subaerial lava flows of the Höskuldvík Group, the youngest of these flows yielding an $^{40}Ar/^{39}Ar$ age of 4.43 $\pm$ 0.12 Ma (Hall et al., 2023), Fig. 4.2. The lignite horizons at the centre of this study are semi-regularly spaced throughout the *Tapes* and *Mactra* biozones, but due to changes in regional uplift, only one lignite horizon (Unit J) exists at the top of the *Serripes* biozone. Representative of emergent periods in the depositional history of the Barmur Group, these lignites are interpreted as forming in a swamp/coastal-lagoon setting (Eiríksson and Símonarson, 2021). A pilot study of these lignites confirmed the presence of a range of organic biomarkers often used in the reconstruction of past environment and climate however, not all lignite horizons were identified and sampled as part of this study (Hall, 2021).

The lava dominated Höskuldvík Group is overlain by the more sediment rich Breiðavík Group (Eiríksson, 1981, 1985). A depositional hiatus is predicted between these two groups, but its precise location and duration vary depending on the interpretation. Radiometric $^{40}Ar/^{39}Ar$ dating and palaeomagnetic data indicate a hiatus of c. 1.6 Ma between the Höskuldvík and Breiðavík groups (Hall et al., 2023). However, dynoflagellate cysts imply a hiatus exists within the basal formations of the Breiðavík Group spanning c. 0.8 Ma (Verhoeven, 2012). Within the Breiðavík Group, 14 diamictite horizons mark a cyclical pattern of sedimentation, interpreted as the repeated growth and retreat of lowland ice-sheets across the Tjörnes Peninsula (Eiríksson et al., 1994; Eiríksson, 2008; Geirsdóttir, 2011; Verhoeven, 2012; Eiríksson and Símonarson, 2021). Whilst basal, angular unconformities are observable for each Breiðavík Group cycle, the effect of these unconformities is predicted to be minimal due to the completeness of the preserved cycles (Eiríksson, 1981). Overall, the Breiðavík Group spans an estimated thickness of c. 480 m and a maximum age range of c. 0–3.5 Ma (Kristjánsson et al., 1988; Eiríksson et al., 1990; Verhoeven and Louwye, 2013; Eiríksson and Símonarson, 2021; Hall et al., 2023). Uplift and rotation of the Tjörnes Peninsula to its current form is predicted to have taken place within the last 1 Ma, with the west of peninsula uplifted > 0.5 km and further east (Öxarfjörður) downthrown > 1.5 km (Ólafsson et al., 1993; Gunnarsson, 1996; Eiríksson and Símonarson, 2021).

To facilitate burial and thermal modelling of the Barmur Group lignites, we explore the three stratigraphic/age models posited in Hall et al. (2023). All three models are the same during the deposition of the Kaldakvísl and Barmur groups and they share the eventual uplift beginning c. 1.0 Ma. The differences between the three models are outlined as follows: Model 1: Assumes a near continuos deposition of sediment at Tjörnes. Deposition of the Höskuldvík Group spans c. 2.6–4.4 Ma. Model 2: Places a depositional hiatus (c. 1.6 Ma) in between the Höskuldvík and Breiðavík groups, inline with radiometric evidence. Model 3: Follows the dynoflagellate evidence of Verhoeven (2012). Deposition of the Breiðavík Group begins at c. 3.55 Ma and a hiatus is inserted between the basal Furuvík and Hörgi formations lasting c. 0.8 Ma. Burial modelling of these three stratigraphic models aim to explore a feasible range of thermal histories for the sediments found on the west of the Tjörnes Peninsula, providing a first order estimate of the impact of diagenesis on the brGDGTs - and associated palaeoclimatic record - recovered from Barmur Group lignites.

### 4.2.2 *Organic Biomarkers in Terrestrial Settings*

brGDGTs are the bacterial equivalents of the archaeal-derived isoGDGT (isoprenoidal Glycerol Dialkyl Glycerol Tetraethers) system of lipids (Schouten et al., 2002). Production of brGDGTs has been attributed to bacterial communities following stereochemistry analysis of the glycerol moiety within the brGDGT structure, finding it is inverse to that of isoGDGTs (Weijers et al., 2006a), a stereochemical arrangement attributed to the membrane lipids of bacteria not archaea (Koga et al., 1998). brGDGTs have been found in a range of depositional environments including soils (Weijers et al., 2007; Crampton-Flood et al., 2020), peatlands (Naafs et al., 2017), lakes (Tierney and Russell, 2009; Tierney et al., 2010; Damsté et al., 2009), and some marine environments (Peterse et al., 2009). 15 brGDGT structures are commonly recognised and form the focus of the analysis presented here, however the number of brGDGT structures has grown in the c. 20 years of research into this lipid family (De Jonge et al., 2014). Initial investigation of this widespread and ubiquitous family of lipids described 9 brGDGTs, used to demonstrate the relationship of brGDGT fractional abundance to MAAT and soil pH (Weijers et al., 2007, 2011). Subsequent studies have identified additional brGDGT isomers, notably using NMR (Nuclear Magnetic Resonance) (Damsté et al., 2000) and refinements in HPLC-MS (High Performance Liquid Chromatography Mass Spectrometer) methodologies (De Jonge et al., 2013) expanding the commonly recognised number brGDGTs to 15, Fig. 4.3.

Attempts to use the distribution of brGDGT isomers to predict MAAT and pH were initially formulated in a similar manner to the isoGDGT $TEX_{86}$ index, with a effective but reductive linear regression. The MBT/CBT index (Methylation of Branched Tetraethers / Cyclisation of Branched Tetraethers) demonstrated correlations with MAAT ($r^2 = 0.62$) and pH ($r^2 = 0.70$) (Weijers et al., 2007). Following the original MBT and CBT formulation, new penta- and hexa-methylated brGDGTs were identified in peat samples previously obscured by co-elution (De Jonge et al., 2013). Due to strong pH dependencies observed within hexamethylated brGDGT isomers, revisions to the MBT/CBT index were made to deconvolve pH signals previously included in the co-eluting brGDGT response (De Jonge et al., 2014). The MBT index was renamed MBT' and updated indices were published ($MBT'_{5ME}$ and $CBT_{5ME}$) addressing the identified pH bias, Eq. 3.1.1 – Eq. 3.1.4.

$$MBT' = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIb + IIc + IIIa + IIa' + IIb' + IIc' + IIIa'} \quad \text{(Eq. 3.1.1)}$$

$$CBT = -log(\frac{Ib + IIb + IIb'}{Ia + IIa + IIa'}) \quad \text{(Eq. 3.1.2)}$$

$$MBT'_{5ME} = \frac{Ia + Ib + Ic}{Ia + Ib + Ic + IIa + IIb + IIc + IIIa} \quad \text{(Eq. 3.1.3)}$$

$$CBT_{5ME} = -log(\frac{Ib + IIb}{Ia + IIa}) \quad \text{(Eq. 3.1.4)}$$

Lipid distributions have also been demonstrated to vary systematically with depositional environment (De Jonge et al., 2014; Naafs et al., 2017; van Bree et al., 2020; Crampton-Flood et al., 2020; Raberg et al., 2022). The key depositional categories considered as part of this study are soils, peats, lacustrine and marine settings. Methodologies have been established to classify palaeodepositional environments from brGDGT distributions utilising a range of statistical approaches, notably Véquaud et al. (2022); Martínez-Sosa et al. (2023b). The reasons for such distinct depositional trends in brGDGTs distributions has been attributed to redox conditions (Weber et al., 2018), microbial community composition (Weber et al., 2018; De Jonge et al., 2019; van Bree et al., 2020), and organic matter source (Crampton-Flood et al., 2020). Despite decades of research into brGDGTs, the source organisms of these lipids have remained only partly resolved, with early culture experiments producing only a limited subset of the more than 15 known brGDGT isomers from just two strains of Acidobacteria (Sinninghe Damsté et al., 2011; Halamka et al., 2021). Recently, a culture study of the Acidobacteria *Solibacter usitatus* synthesised a structurally diverse set of brGDGTs (Ia, Ib, IIa, IIb, and IIIa) with demonstrable responses to temperature, pH and, $O_2$ conditions (Halamka et al., 2023). In addition, identification of the radical SAM protein Tes (tetraether synthase) and its catalysing role in the conversion of diether to tetraether lipids provides insights into the mechanistic pathways for brGDGT synthesis (Zeng et al., 2022). However, uncertainties still remain regarding the identity of the dominant brGDGT-producing taxa and the full suite of environmental controls on their biosynthesis and export to the geological record. For that reason, attempts to use brGDGT distributions for palaeoclimatic study must also classify the palaeodepositional environment to maximise the confidence in and utility of any resulting prediction.

Whilst the effects of diagenesis are a near geological surety, experimental efforts to constrain the effects of thermal degradation on this lipid system are rare. Schouten et al. (2004) conducted a hydrous pyrolysis experiment to assess how thermal maturation impacts isoGDGTs (via the $TEX_{86}$ index), and Schouten et al. (2013) released further data from this original experiment, reporting the response of brGDGTs (via the MBT/CBT index). The MBT index was found to gradually decrease and the CBT index gradually increase with pyrolysis temperature, CBT increasing markedly at c. 240 °C (Schouten et al., 2013). The BIT (Branched Isoprenoidal Tetraether) index was found to remain relatively stable until c. 240 °C where the index progressively decrease towards total degradation of the iso- and brGDGTs by 300 °C (Schouten et al., 2004). The recovery of isoGDGTs from deep geological archives (c. 1.5 km below sea-floor - Littler et al. (2011); Naafs and Pancost (2016)), evidences the longevity of these lipids in geological settings and instructs the need for a good understanding of the effects of thermal degradation on GDGTs. Bijl et al. (2025) provides the most recent best practises for the extraction and analysis of GDGT data, in which they recommend an assessment of thermal maturation via vitrinite reflectance, RockEval pyrolysis, or an assessment of the hopane composition and stereochemical configuration. This study complements all three of these recommendations with the addition of burial modelling to predict the thermal history of any biomarker bearing horizon.

Many permutations of the MBT/CBT index exist for brGDGTs recovered from a variety of sedimentary basins and depositional regimes, for example: MAAT and pH in peatlands ($MBT_{peat}$ and $CBT_{peat}$, Naafs et al. (2017)). A similar view of these calibrations is taken here as in the previous chapter on the topic of the isoGDGT $TEX_{86}$ index: there is value and limitation in these formulations. These regressions are often linear and regularly reduce the dimensionality of the underlying data. Whilst dependencies within calibration data can and have been demonstrated, it is crucial to utilise an approach that has the capacity to capture the complexity, subtlety and unknowns present in brGDGT fractional abundance data. Here, we use a Gaussian Process to both classify the palaeodepositional environment and predict environmental temperature from ancient brGDGT distributions. These Gaussian models are trained on modern brGDGT distributions from soil, peat, lacustrine or marine depositional settings.

In addition to the investigation of brGDGTs in this study, *n*-alkanes and hopanes have been identified and analysed using Gas Chromatography Mass Spectrometry (GC-MS) and Gas Chromatography Flame Ionisation Detection (GC-FID). These established biomarkers are often abundant in organic rich samples and proxies/indices already exist which may provide information on the effects of thermal maturation. Long chain (> $C_{25}$), odd-numbered n-alkanes are saturated hydrocarbons produced in plant waxes (Eglinton and Hamilton, 1967), as such are often abundant in organic rich, lignite horizons. The Carbon Preference Index (CPI) is a ratio of even to odd carbon number molecular chains and commonly used to indicate sediment maturity (Cooper and Bray, 1963). Studies have demonstrated that the CPI value decreases with increasing thermal maturity (Allan and Douglas, 1977; Nelson et al., 1998). It is defined here as:

$$CPI = \frac{1}{2}\left(\frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{24} + C_{26} + C_{28} + C_{30} + C_{32}} + \frac{C_{25} + C_{27} + C_{29} + C_{31} + C_{33}}{C_{26} + C_{28} + C_{30} + C_{32} + C_{34}}\right) \qquad \text{(Eq. 3.2)}$$

Geohopanoids are the diagenic degradation products of biohopanoids - pentacyclic triterpenoids produced by a wide range of bacteria. During this diagenetic transformation, the biologically-derived 17$\beta$,21$\beta$(H)-hopanoid is replaced by the more thermally stable 17$\beta$,21$\alpha$(H) and 17$\alpha$,21$\beta$(H) stereoisomers (Rohmer et al., 1984; Pearson et al., 2007). It is possible to use the ratio of $\beta\beta$ to $\alpha\beta + \beta\beta$ to estimate thermal maturation of a system (Rohmer et al., 1984; Farrimond et al., 1998; Pearson et al., 2007). Decreasing $\beta\beta$ to $\alpha\beta + \beta\beta$ values are correlated with increasing thermal maturation.

$$Geohopanoid\ Index = \frac{\beta\beta}{\alpha\beta + \beta\beta} \qquad \text{(Eq. 3.3)}$$

Lastly, this study uses the Contextualised Distance ($CD$) metric, defined in Chapter 2, to assess any unknown forces acting on a brGDGT system. The median $CD$ value ($CD_{Median}$) is chosen and analysed here. In contrast to the isoGDGT system, brGDGT dependency on depositional environment complicates the usage of the $CD$ metric. This unit-less and relative metric is dependent on the correlation length-scales learnt during the Gaussian model training - different training data will produce different $CD$ metrics and as such void direct, model-model comparison. Here, the $CD_{Median}$ value is assessed within depositional groups as outlined in Chapter 3.3. It is hypothesised that an indiscriminate force, such as the diagenetic thermal degradation of brGDGTs, will act on all records of $CD_{Median}$ - resulting in a coherent signal across depositional groups. This study aims to use a multi-biomarker approach alongside a Gaussian Process workflow and basin/thermal modelling to enable a more holistic understanding of the effects of thermal degradation on brGDGTs recovered from geological samples. It is hypothesised that if thermal degradation of brGDGTs has occurred, this will be captured by decreasing CPI and Geohopanoid indices and increases in the value of $CD_{Median}$.

## 4.3 Methods

### 4.3.1 *Stratigraphic and Burial Modelling*

For the purposes of this study and and burial/thermal modelling, the stratigraphic thickness and age of the Barmur, Höskuldvík, and Breiðavík groups are required. Chapter 2 outlines revisions to the stratigraphic model of the Barmur Group using drone photogrammetry and structural modelling (Allison et al., 2026), the results of which are used here. Above the Barmur Group, reconciliation of magnetic polarity records (Eiríksson et al., 1990; Kristjánsson et al., 1988), radiometric ages (Hall et al., 2023), and biostratigraphic analysis (Verhoeven and Louwye, 2013) result in the three stratigraphic models of Hall et al. (2023). These three models are used to investigate a range of possible thermal histories for the Tjörnes Peninsula using an industry-standard burial maturation scheme as described in Rowley and White (1998) and Jones et al. (2019). A stratigraphic/age model and a selection of prescribed physical constants are prerequisites for the thermal history of the basin to be calculated. A predictive output of this model is vitrinite reflectance, which can be directly compared to the actual measured vitrinite reflectance data of Richter and Bojesen-Koefoed (2002). In matching the predicted and observed reflectance data, it is possible to constrain the model parameters and estimate the thermal history of all the Barmur Group lignite horizons.

The burial model ingests a summarised stratigraphy detailing the basal depth [m], age [Ma] and composition (sandstone, basalt, shale etc.) of geological layers. The model time-steps and deposits the geology accordingly, accounting for the effects of compaction through each step. Uplift is handled by considering the depth and age that the uppermost unaffected layer would have after compaction, running the sedimentation model in reverse - removing layers during the model time-steps but continuing to model kinetic reactions throughout the event. Due to the uplift, erosion and tilting that has occurred on the Tjörnes Peninsula an erosional event is required however, the magnitude of this event varies depending on where on the Peninsula is being modelled. The exact timing and rate of uplift is unknown, with Eiríksson and Símonarson (2021) suggesting that uplift of the Peninsula occurred within the last 1 Ma. A single uplift/erosion event is permitted in this model, and here an uplift (760 m) and erosion event is modelled prior to the deposition of the Grasafjall Formation (c. 0.8 Ma - Eiríksson and Símonarson (2021)). This event is sufficient to expose the lignite-bearing Unit F at c. 30 m above sea-level, replicating the modern coastline and borehole logs of Richter and Bojesen-Koefoed (2002) from the Skeifá Waterfall, Fig. 4.2. At the base of the succession the stratigraphic position of Unit d/e is secure however, Unit A is exposed in a fault block and its exact stratigraphic position is uncertain (Eiríksson and Símonarson, 2021; Allison et al., 2026). Due to faulting lowering the stratigraphic position of Unit A, it is included here as a separate horizon at the base of the stratigraphic section.

For predictions of vitrinite reflectance, kinetic parameters for thermal maturation reactions are used, as constrained from inverse modelling by Rowley and White (1998).

Following the construction of a stratigraphic/age model, a series of model runs are completed with varying heat flux inputs [$mWm^{-2}$]. The computed vitrinite reflectance data are then compared to the drill reports of Richter and Bojesen-Koefoed (2002) until a satisfactory match is found. Two key stratigraphic uncertainties exist which could impact burial modelling: uplift timing and stratigraphic thickness. Kinetic reactions will occur most vigorously at depth where temperatures are highest, as such the model will be most sensitive to variations during peak burial - just prior to the final uplift event. Here, uplift timing is varied between 0.5–1.0 Ma and the sedimentary thickness is varied by $\pm$ 100 m prior to uplift, representative of a generous upper limit of the published erosional unconformity estimates for sedimentation on the Tjörnes Peninsula (Eiríksson and Símonarson, 2021; Allison et al., 2026). The outputs of these models are compared to estimate model uncertainty. Full details of the stratigraphic/age model, input parameters and physical constants are found in Appendix A.3.2.

### 4.3.2 *Organic Extractions*

brGDGT distributions presented here are the culmination of several field seasons at the Tjörnes Peninsula. A total of 14 lignite seams were sampled across 8 of the Units described by Bárdarson (1925) (A, d/e, B, C, D, F, 10, J). Whilst Unit 10 was not originally classified as a terrestrial unit (Bárdarson, 1925), a small lignite seam is present at the boundary between Units 10 and Bed 11 (Eiríksson and Símonarson, 2021) and was sampled as part of this study. Two lignite seams were sampled from Units B and J (denoted $\text{Unit}_{\text{Lower}}$ and $\text{Unit}_{\text{Upper}}$) and three lignites seams sampled from Units C and F (denoted $\text{Unit}_{\text{Lower}}$, $\text{Unit}_{\text{Middle}}$, and $\text{Unit}_{\text{Upper}}$). The rough locations of these sampling localities are detailed in Fig. 4.2 and precise GPS locations reported in Table 4.1. Once a lignite seam was identified in the field, a fresh surface was excavated and unweathered material exposed. For each seam, c. 100 g of material was collected using a geological hammer (cleaned with solvent before each sample) and stored in a pre-furnace

| Lignites | GPS Location | Stratigraphic Height (Barmur Group) | Max. Burial Depth [m] | Ave. Burial Temp. [$^{\circ}C$] | Vitrinite Reflectance [%] | Averge TOM of 5 samples [%] |
|---|---|---|---|---|---|---|
| Unit J | 66.15686 N 17.27308 W | 430 | 395.2 | 24.4 | 0.20 | 16.5 |
| Unit 10 | 66.13781 N 17.25700 W | 251 | 528.6 | 35.1 | 0.21 | 34.8 |
| Unit F | 66.12608 N 17.25829 W | 199 | 570.8 | 37.9 | 0.24 | 37.4 |
| Unit D | 66.11908 N 17.25652 W | 145 | 610.2 | 41.5 | 0.29 | 65.5 |
| Unit C | 66.11365 N 17.26240 W | 114 | 637.3 | 43.8 | 0.35 | 38.0 |
| Unit B | 66.11133 N 17.26972W | 63 | 675.9 | 47.2 | 0.47 | 27.1 |
| Unit d/e | 66.10702 N 17.27890 W | 14 | 725.5 | 52.3 | 0.61 | 16.2 |
| Unit A | 66.10553 N 17.28343 W | 0 | 739.3 | 53.9 | 0.65 | 43.4 |

Table 4.1: Barmur Group lignite beds. Exact GPS locations and stratigraphic heights are provided (using the stratigraphic model of Allison et al. (2026)). Modelled predictions of maximum burial depth, average burial heating and vitrinite reflectance are provided alongside measured Total Organic Matter (TOM).

envelope of aluminium foil. Foil envelopes were placed in a sterile Whirl-Pak™ bag and labelled with the corresponding GPS location and sample identifiers. Once the samples were transported from the field, they were weighed, freeze-dried and finally homogenised in a solvent cleaned TEMA disc mill, before storage in a freezer until solvent extractions could take place.

A standard solvent extraction protocol was used to extract 4 solvent fraction (N1–N4). A detailed protocol and equipment list can be found in Appendix A.3.1, but is summarised here. 5 g of homogenised lignite was extracted using a mixture of dichloromethane/ methanol (3:1). The solution was heated at 40 $^{\circ}$C for one hour and centrifuged at 3000 RPM to separate solid from liquid and the supernatant extracted. This process was repeated three times, or until the extracted supernatant was colourless. Extracts were recombined and concentrated to form the Total Lipid Extract (TLE). The TLE was further separated into 4 fractions using silica-gel column chromatography: N1 - Apolar fraction containing n-alkanes and hopanes, N2, N3, N4 - Polar fraction containing GDGTs. N1 compounds were identified using an GC-MS and analysed using a GC-FID at the University of Birmingham, and the N4 fraction was redissolved in 200 $\mu$l of n-hexane/i-propanol (99:1) and filtered using a 0.45 $\mu$m polytetrafluoroethylene (PTFE) syringe filter prior to analysis on a HPLC-MS at the University of Bristol. n-alkanes were identified using GC-MS by filtering for mass ion m/z 57. The C30 hopane was identified using GC-MS by filtering for mass ion m/z 191.1. Quantification of n-alkanes and the C30 hopane were made using GC-FID on the TIC with specific compounds identified by cross-referencing with the GC-MS results.

### 4.3.3 *Palaeodepositional and Palaeoclimatic Predictions*

Chapter 3 details much of the ethos and rationale behind utilising a Gaussian Process to study ancient archives of high dimensional (n > 3) data, such as the isoGDGTs. Using the brGDGT system for palaeoclimatic applications poses a similar set of challenges. OPTiMAL provides a flexible approach to addressing these challenges for isoGDGTs (Dunkley Jones et al., 2020) and can be easily modified to work with increased dimensionality of brGDGTs. A key difference between the branched and isoprenoidal systems is the observed dependency of brGDGT distributions on depositional environment. The training data used here are a subset of those compiled by Raberg et al. (2022), totalling 2289 samples splits across the sample types of soils (n = 1153), peats (n = 473), lacustrine (n = 539), and marine (n = 124). Four GPE models are trained on the four depositional categories, but before application it is necessary to classify the depositional environment of an ancient brGDGT assemblage in order to apply the most relevant GPE model. This is a necessity both in order to extract the maximum model performance, and because the measure of environmental temperature within each sample type is not necessarily consistent. In this study, brGDGT distributions from soil and peat samples are correlated with Mean Annual Air Temperature (MAAT [$^{\circ}$C]) and Warmest Month air Temperature (WMT [$^{\circ}$C]), lacustrine samples with in-situ Water Temperature [$^{\circ}$C], and marine samples with mean annual Sea Surface Temperature [$^{\circ}$C]. Published methods do exist with which to classify depositional environments from GDGT distributions, notably the Branched and Isoprenoid GDGT Machine Learning Classification Algorithm (BIGMaC) (Martínez-Sosa et al., 2023b). However, for this study a GPC framework is trialled for the assignment of depositional environment from ancient brGDGTs assemblages.

No classification scheme is necessarily right or wrong; all possess strengths, weaknesses and requirements. A GPC scheme requires labelled training data, but performs well with complex, non-linear decision boundaries and makes no assumption about the underlying functional form of the data, whilst providing both predictive capability and confidence estimates (Rasmussen, 2003). Here, a GPC with both an isotropic and anisotropic kernel is tested for classifying palaeodepositional environment (soil, peat, lacustrine or marine) from brGDGT data. Whilst more computationally demanding, anisotropic kernels result in correlation length-scales which can modulate the sensitivity of the model in favour of dimensions that carry predictive power. As such, the correlation length-scales provide a first order measure of dimensions which are more or less important when making a prediction of, for example, air temperature from the distribution of brGDGT isomers. In addition, the following kernels are also assessed: additive kernel, rational quadratic kernel, dot-product RBF (Radial Basis Function) kernel, and product kernel.

The Barmur Group lignites are consistently interpreted as originating from swamp/ coastal-lagoon setting (Verhoeven, 2012; Eiríksson and Símonarson, 2021; Hall, 2021). It is important to reflect on whether it is appropriate to apply a GPC with a small number of coarse depositional categories to the task of predicting the palaeodepositional environment of the Barmur Group lignites. The four depositional categories of soil, peat, lacustrine and marine would all be expected to operate in close proximity to the interpreted palaeodepositional environment of the lignites. However, it is also expected that the broadly similar lithology of the lignite horizons would originate from a broadly similar and stable depositional environment. Should this be true, it is expected that a

GPC approach will return a consistent classification - provided the training data faithfully captures depositional dependencies. The training data utilised here is categorised coarsely and so the resulting classifications made here reflect the training data structure (Raberg et al., 2022). As such, this study assumes discrete, coarse depositional categories. Importantly, should a sample be incorrectly or poorly classified, this will result in a brGDGT distribution that will be ill-represented by the training data used for palaeoclimatic prediction. These samples will be more likely to fail the subsequent data screening that forms part of the OPTiMAL (Dunkley Jones et al., 2020), ensuring confidence in the eventual temperature prediction.

To quantify the accuracy with which the GPC can predict depositional environment, 5% (n = 114) of the training data were reserved at random from the training stage of the GPC model. The remaining 95% (n = 2175) were used for training and to make predictions of sample type on the reserved data, for which the correct answer is known. Via random sampling of the whole dataset, the distribution of sample types are, on average, preserved across both the reserve and training data. This process was repeated 50 times and the results averaged to representatively assess the model performance. All GPC models were trained on the full suite of 15 brGDGT isomers and also subsets of brGDGTs with a fractional abundance $\geq$ 3% and $\geq$ 10%.

For the final task of predicting environmental temperature from brGDGT distributions, a modified version of the GPE - OPTiMAL (Dunkley Jones et al., 2020) - is used. The strengths of a GPE approach are highlighted in Chapter 3, and considering the increased dimensionality and unknown producer of brGDGTs, an agnostic predictor which makes no assumptions about data structure or temperature dependencies is desirable for this study. In contrast to the single isoGDGT training model, here multiple GPE models are trained and assessed. Recent studies suggest that brGDGTs distributions are better reflected in warm bias temperature records (Crampton-Flood et al., 2020; Raberg et al., 2021, 2022). In the case of soil and peat type sample, predictive models for both MAAT and WMT are assessed. For water temperature estimates, lacustrine type samples are assessed using predictive models of ISWT and marine type samples using predictive models of SST. An additional assessment is included as part of this study to investigate the effects of making predictions using the entire suite of 15 brGDGT isomers versus smaller subsets defined by fractional abundances $\geq$ 3% and $\geq$ 10%. These three models are designed to investigate the amount of temperature information stored within the most abundant fractions. It is pragmatic to assess whether a performative model can be constructed on the data most reliably recovered from the geological record.

Complicating matters however, in the global training data it is observed that a large percentage of the relative abundance data is published as "zero", particularly for isomers containing greater degrees of cyclical structures and methyl branching. For example, in the global compilation (n = 2289) > 75% of data entries for fraction IIc' are listed as "zero". It is crucial that a machine learning process is trained on data containing dependencies for which prediction is desired - for example, a data entry of "zero" could result from instrumental insensitivity as opposed to environmentally driven factors. To assess this potential pitfall, all GPE models are trained twice, once with all data and once with data entries of zero are removed prior to training. The end result of these investigations are six training datasets for environmental temperature prediction: soil type (MAAT and WMT), peat type (MAAT and WMT), lacustrine type (IWST), and marine type (SST). For each of these six datasets, six GPE models were trained and assessed, summarised as follows GPE-Model 1: All 15 brGDGT isomers - the maximum amount of training data. GPE-Model 2: All 15 brGDGT isomers, but data entries of zero removed. GPE-Model 3: brGDGT isomers with fractional abundance $\geq$ 3% . GPE-Model 4: brGDGT isomers with fractional abundance $\geq$ 3%, but data entries of zero removed. GPE-Model 5: The same as GPE-Model 3, but for fractional abundance $\geq$ 10%. GPE-Model 6: The same as GPE-Model 4, but for fractional abundance $\geq$ 10%.

OPTiMAL is complemented by a data screening metric termed $D_{Nearest}$ (Dunkley Jones et al., 2020). $D_{Nearest} \leq 0.5$ was chosen qualitatively as a value which screens for data with an acceptable predictive error (Dunkley Jones et al., 2020). Due a change in training data (iso- versus brGDGT) and the four sample types explored in this study, a $D_{Nearest}$ screening value is defined more rigorously for each depositional environments. GPE predictions are observed to form a cumulative distribution curve which, when modelled, can be used to reliably define a $D_{Nearest}$ cut-off, the details of this are found in Appendix A.3.3. Similarly, the $CD_{Median}$ metric is calculated within and not across the four depositional categories, however trends in this metric can be compared across categories.

## 4.4 Results

### 4.4.1 *Stratigraphic and Burial Modelling*

The vitrinite reflectance data of Richter and Bojesen-Koefoed (2002) was found to be replicated by a modelled heat flux ranging from 250–300 mWm$^{-2}$, with the best model match at 275 mWm$^{-2}$. The depth of Units D and F in the drill core of Richter and Bojesen-Koefoed (2002) were within 1 m of the modelled depth using the stratigraphy of Allison et al. (2026). Richter and Bojesen-Koefoed (2002) reported the average vitrinite reflectance of Unit D and F as 0.28% and 0.26% respectively, which compare to a modelled vitrinite reflectance of 0.29% and 0.24% at a heat flux of 275 mWm$^{-2}$. Using three input heat flux values (250, 275 and 300 mWm$^{-2}$) to investigate the three stratigraphic models of Hall et al. (2023), resulted in a total of nine burial models. The uncertainties in the three stratigraphic models of the Tjörnes Peninsula appear to have a minimal impact on the results of the burial model. For a single heat flux value, models produced divergent results during the stratigraphic discrepancies between 2.0–4.5 Ma, but convergent results during all other times, Fig. 4.4. For all lignite bearing horizons

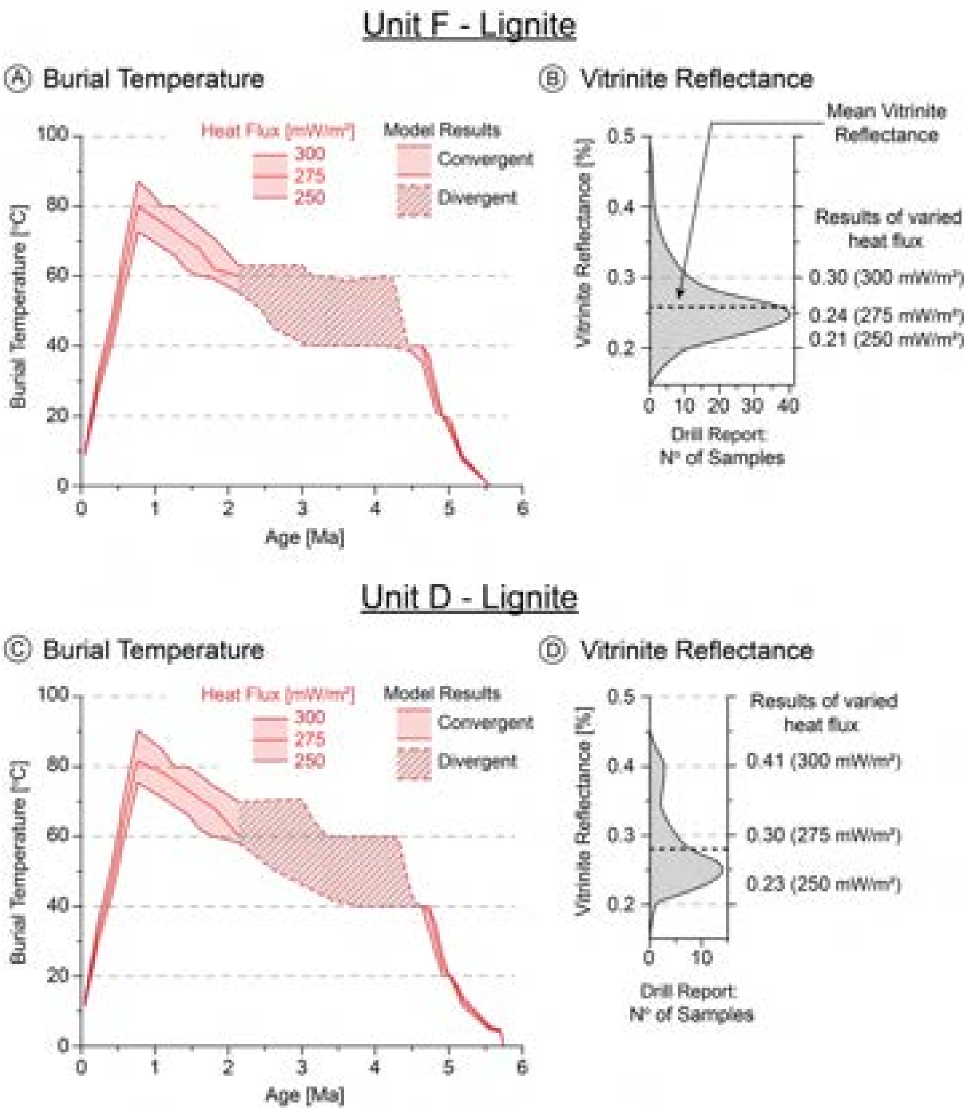


Figure 4.4: Modelled burial temperature and vitrinite reflectance for lignites of Unit F and Unit D. A) Burial temperature predictions for Unit F resulting from three heat flux inputs (250, 275, and 300 $mWm^{-2}$), tested on the three stratigraphic models of Hall et al. (2023). Nine models result which produced convergent results outside of the variations in stratigraphic models (divergent results = hashed region). B) Distribution curves of vitrinite reflectance from Richter and Bojesen-Koefoed (2002) alongside the vitrinite reflectance predictions from varying model parameters. A heat flux of 275 $mWm^{-2}$ was found to best replicate the average measured vitrinite reflectance (dashed line). C) Same as A) but for Unit D. D) Same as B) but for Unit D.

the difference in predicted maximum burial temperatures was < 0.6 °C and vitrinite reflectance varied < 0.01% across the three stratigraphic models. Peak heating is predicted to have occurred during maximum burial, just prior to the uplift event at c. 0.8 Ma. For Units D and F, the maximum predicted temperatures ranged from 68.3–87.4 °C depending on the modelled heat flux, Fig. 4.4. Burial of Units D and F is modelled to a depth of 610 m and 570 m respectively, before uplift to their modern positions as exposed on the Tjörnes coastline. Estimates of maximum burial temperature, burial depths and vitrinite reflectance for each lignite horizon are detailed in Table 4.1 and burial models for the three stratigraphic models are visualised in Fig. 4.5.

Sensitivity of the burial model was assessed by adjusting the final burial depth and timings to observe the effects on maximum burial temperatures and vitrinite reflectance using a heat flux of 275 $mWm^{-2}$. Uplift timings were adjusted between 0.5–1.0 Ma and the sensitivity of the model to burial depth was assessed by adding or subtracting up to 100 m of sediment to the stratigraphy prior to uplift. Between the three stratigraphic models, the results of the sensitivity testing were similar and varied < 0.5 $^{\circ}$C for temperature predictions and < 0.01% for vitrinite reflectance calculations, no one model proving more or less sensitive than the other. The maximum predicted burial temperature and vitrinite reflectance were minimised at - 100 m of sediment and maximised at + 100 m of sediment. For Unit D, maximum burial temperatures ranged from 68.5–91.6 $^{\circ}$C (ideal model: 80.0 $^{\circ}$C) and vitrinite reflectance varied from 0.23–0.44% (ideal model: 0.29%). For Unit F, maximum burial temperatures ranged from 63.3–86.0 $^{\circ}$C (ideal model: 75.2 $^{\circ}$C) and vitrinite reflectance varied from 0.21–0.33% (ideal model: 0.24%). The model sensitivity to changes in sedimentary thickness are akin to altering the heat flux of the ideal model by $\pm$ 50 $mWm^{-2}$) (heat flux range: 225–325 $mWm^{-2}$). The model was less sensitive to the timing of uplift than to changes in the sedimentary thickness. The difference in maximum burial temperature between uplift starting at 1 Ma versus 0.5 Ma was 0.4 $^{\circ}$C for Unit D and 1.4 $^{\circ}$C for Unit F. Vitrinite reflectance deviated by $\pm$ 0.01% from the idealised model, with the highest levels of vitrinite reflectance occurring when uplift was delayed, as the kinetic model could progress reactions further during delayed uplift.

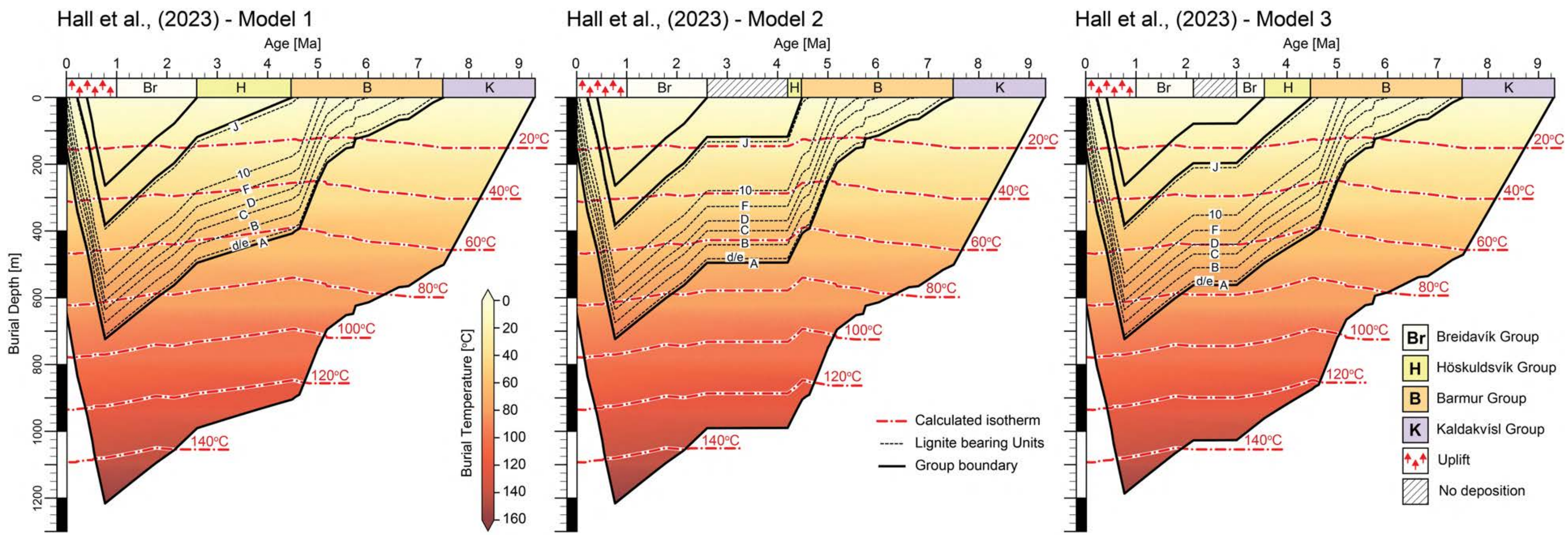


Figure 4.5: Burial models for the three stratigraphic models of Hall et al. (2023) using a heat flux input of 275 $mWm^{-2}$. Calculated isotherms are shown in red dashed lines (every 20 $°C$ and interpolated to produce the burial temperature colour map. 760 m of uplift begins at c. 0.8 Ma and raising Unit F to the modern day position at the Skeifá waterfall - c. 30 m above sea-level.

### 4.4.2 *Organic Biomarkers*

brGDGTs were recovered in 13 of the 14 sampled lignite seams and the absolute abundance of recovered brGDGTs broadly increased with stratigraphic height, with the exception of the Unit $J_{Lower}$ lignite. The raw abundance values have been corrected to provide a semi-quantitative form of analysis. Firstly, the total brGDGT response is summed across all fractions. Secondly, this total value is divided by the mass of Total Lipid Extract (TLE) analysed on silica gel, column chromatography. Finally, this TLE adjusted value is normalised between 1 (maximum response) and zero (no response). The absolute abundance of the total brGDGT response on HPLC-MS range from 5231–8.34$x10^8$. Fitting a logarithmic regression to the normalised and TLE adjusted brGDGT response and stratigraphic height of the lignite seams yields a moderate fit ($r^2 = 0.65$). The lowest levels of brGDGTs were recovered from the lowest stratigraphic lignites (Units A and d/e) - no brGDGTs were found in c. 40% of these samples (n = 14). Of samples containing brGDGTs, c. 70% of isomers were absent. The relative abundance of all 15 brGDGT isomers are summarised in Fig. 4.6. Whilst the lignite seams of Unit $J_{Lower}$ and Unit 10 produced a consistent brGDGT distribution signal across samples, the remaining seams produced a greater variety in brGDGT distribution, despite homogenisation of the samples. Trends in particular brGDGT isomers with stratigraphic height are not clear, but a reduction in the total number of recovered isomers with stratigraphic height is observable (linear regression: $r^2 = 0.74$). $MBT'_{5ME}$ and $CBT_{5ME}$ values range from 0.00–0.94 and 0.08–2.00 respectively. Using temperature calibrations such as $MAAT_{soil}$ and $MAAT_{peat}$ (Naafs et al., 2017), this $MBT'_{5ME}$ range translate to temperature predictions of -15.25–22.50 °C and -23.05–26.13 °C respectively.

*n*-alkanes and hopanes were analysed in 9 of the 14 lignites seams (Units $J_{Upper}$, 10, $F_{Upper}$, $F_{Lower}$, D, $C_{Lower}$, $B_{Upper}$, d/e and A). The total absolute abundance of all measured *n*-alkanes ($C_{19}$–$C_{35}$) recorded by GC-FID range between 3.6$x10^8$–8.2$x10^9$.

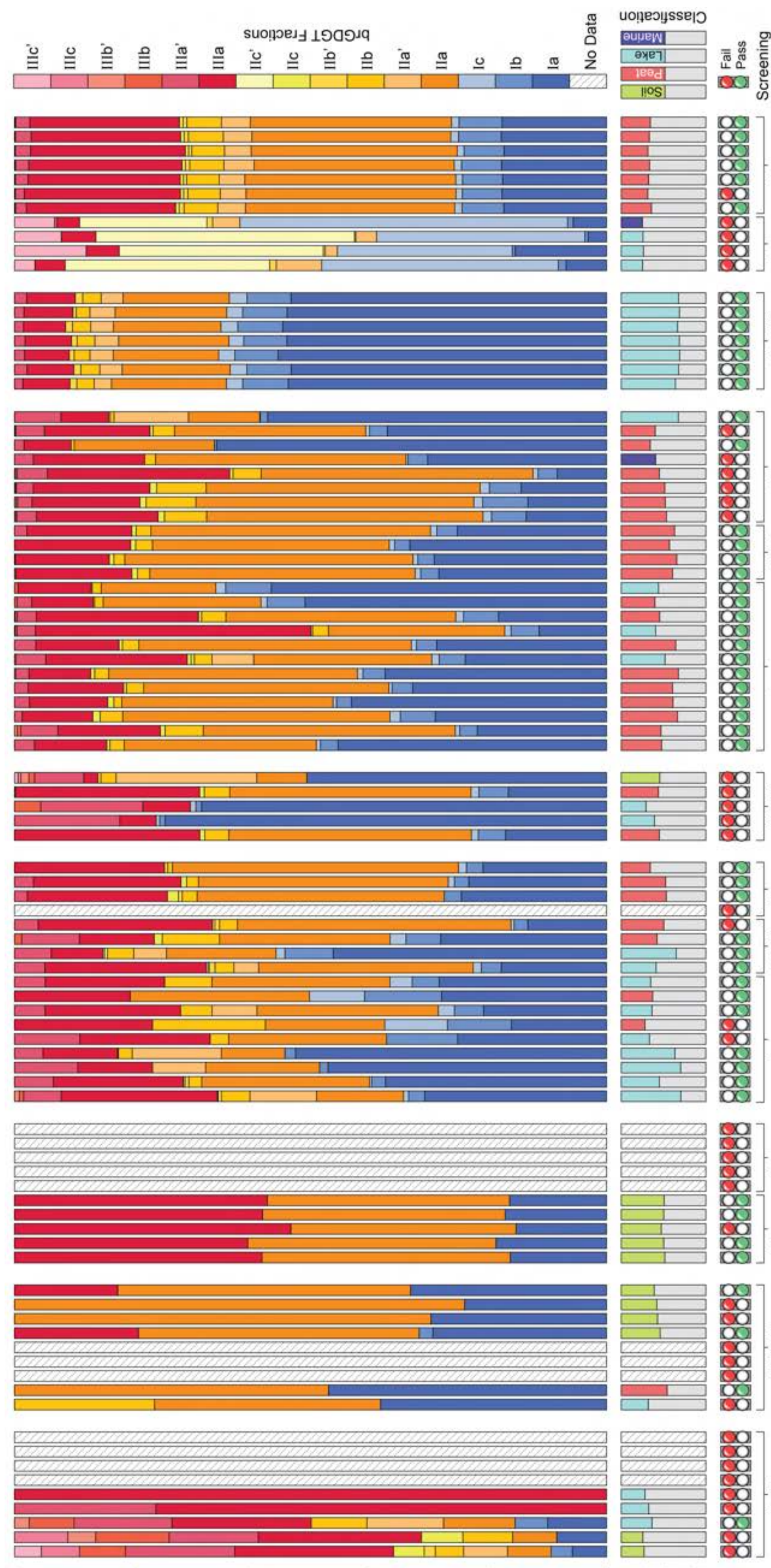


Figure 4.6: A) brGDGT relative abundance data for the Barmur Group lignites. B) Results of the Gaussian Process Classifier (GPC) for the prediction of palaeodepositional environment (soil, peat, lacustrine or marine) from the relative abundance brGDGT data. Classification confidence (0.0–1.0) is shown as a coloured bar. The remaining grey bar represents the sum of the other three classification types. C). Gaussian Process Emulator (GPE) screening metric $D_{Nearest}$. Unique cut-off values exist per sample type (see text) providing a means to screen data for palaeoclimatic prediction (green) or to reject data due to poor representation in the training dataset (red).

Unlike the brGDGTs, no relationship is observed between absolute abundance of n-alkanes and stratigraphic height (linear regression: $r^2 = 0.02$). n-alkane CPI ranged from 2.26–9.62, Fig. 4.7. If the outlier of Unit $J_{Upper}$ is removed from the analysis, the relationship between CPI values and stratigraphic height are modelled well by a positively correlated linear regression ($r^2 = 0.83$). The total absolute abundance of all measured hopanoids ($\alpha\beta$, $\beta\beta$, $\beta\beta$) recorded by the GC-FID range between $2.5x10^7$–$1.1x10^9$. Similar to the n-alkanes results, no relationship is observed between absolute abundance of hopanoids and stratigraphic height (linear regression: $r^2 = 0.02$). Geohopanoid Index values ranged from 0.32–0.95, with values exhibiting a sharp decrease towards the lowermost lignite - Unit A, Fig. 4.7. The relationship between the Geohopanoid Index and stratigraphic height is modelled well by a positive logarithmic correlation ($r^2 = 0.86$). All regressions and trends outlined above can be found visualised in Appendix A.3.3.

### 4.4.3 *Palaeodepositional Predictions*

Assessment of GPC models indicate an anisotropic RBF kernel utilising all 15 brGDGTs isomers is the most accurate predictor of depositional environment/sample type from brGDGT assemblages, with an overall classification accuracy of 90% and average F1 score of 89%. Within each depositional environment, the classification accuracy was found to be 91% for soils (n = 2905), 88% for peats (n = 1145), 87% for lacustrine (n = 1311) and 100% for marine (n = 339). The most consistent false positive result were found in lacustrine samples, with 12% of samples incorrectly classified as soils. Peat samples were most commonly misclassified as lacustrine samples (9% of samples) and soil samples were misclassified as either peat (4%) or lacustrine (4%) samples. Marine samples were never misclassified. Model accuracy for all tested kernels and input data are summarised in Table 4.2. Classification confidence is expressed as four values (between 0–1) which sum to 1, the largest fraction dictating the sample type and

| Kernel choice | Model accuracy | brGDGT isomers in training data |
|---|---|---|
| Anisotropic RBF | 89.9% | All |
| Isotropic | 89.0% | All |
| Anisotropic RBF | 88.9% | Ia, Ib, IIa, IIa',IIb, IIb', IIc', IIIa,IIIa' |
| Anisotropic RBF | 88.6% | Ia, Ib, IIa, IIa',IIb, IIb', IIIa ,IIIa' |
| Anisotropic RBF | 88.5% | Ia, Ib, IIa, IIa',IIb, IIb', IIIa, IIIa' |
| Additive | 88.0% | All |
| Anisotropic RBF | 87.8% | Ia, Ib, IIa, IIa', IIb', IIIa, IIIa' |
| Product Kernel | 87.8% | All |
| Dot-Product RBF | 85.1% | All |
| Rational Quadratic | 84.3% | All |
| Anisotropic RBF | 78.4% | Ia, IIa, IIIa |

Table 4.2: Gaussian Process Classifier (GPC) model assessment for predicting depositional environments from brGDGT distributions. Kernel choice, model accuracy and brGDGT isomers used in the trainining data are shown.

| Class | Precision | Recall | F1 Score | Model Accuracy |
|---|---|---|---|---|
| Soil | 90.88 | 91.12 | 91.02 | 89.89 |
| Peat | 88.12 | 89.45 | 87.78 | |
| Lacustrine | 86.65 | 87.79 | 87.22 | |
| Marine | 100.00 | 88.74 | 94.04 | |

| Class | Soil | Peat | Lacustrine | Marine |
|---|---|---|---|---|
| Soil | 90.9% | 9.1% | 11.6% | 0% |
| Peat | 3.9% | 88.1% | 0.4% | 0% |
| Lacustrine | 4.5% | 2.4% | 86.7% | 0% |
| Marine | 0.7% | 0.3% | 1.4% | 100% |

Table 4.3: Confusion matrix results for the best performing GPC model - anisotropic RBF kernel trained on all 15 brGDGT isomers. The grey diagonal highlights the correctly predicted depositional classes.

the confidence estimate. Of the correctly classified samples, the average classification confidence was 63% for soils (n = 2640), 64% for peats (n = 1009), 64% for lacustrine (n = 1136) and 66% for marine (n = 270). The average confidence of incorrectly classified samples was 51% for soils (n = 265), 50% for peats (n = 136), 49% for lacustrine (n = 175) and N/A for marine (n = 0). The anisotropic kernel was both accurate and confident in prediction, indicating it was not over-fitted. The confusion matrix for the anisotropic RBF kernel can be found in Table 4.3.

For the Barmur Group lignites, the GPC resulted in both palaeodepositional predictions and an estimation of classification confidence for 79 samples of the 92 analysed samples (13 samples recovering no brGDGTs). Soil classifications (n = 12) had an average

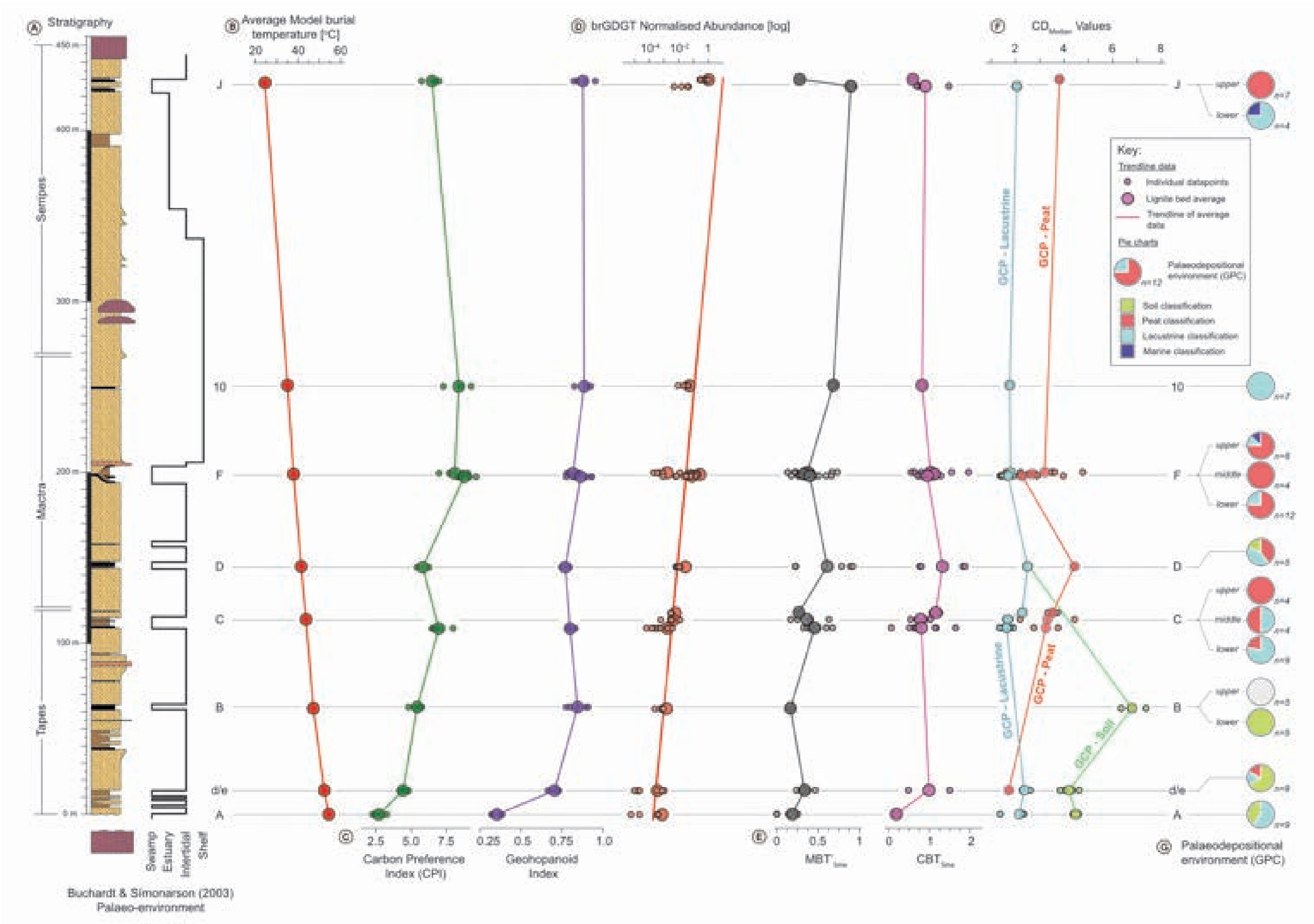


Figure 4.7: Summary of biomarker analysis for the Barmur Group lignites. A) Stratigraphic column (Allison et al., 2026) and palaeoenvironmental interpretation (Buchardt and Símonarson, 2003). B) Modelled average burial temperatures for all lignite horizons. C) Carbon Preference Index (CPI) and Geohopanoid Index, lower values representing more thermally mature samples. D) Normalised total abundance of all brGDGT data adjusted for Total Lipid Extract (TLE) quantities, (log scale) with a linear regression ($r^2 = 0.65$). E) $MBT'_{5me}$ and $CBT_{5me}$ indices. F) Contextualised Distance ($CD$) metric for soil, peat and lacustrine type samples classified via GPC (see text). Greater values indicate a greater distance between a given data-point and the global training dataset, when weighted for temperature sensitivity. G) Pie charts displaying the GPC predictions of palaeodepositional environment per lignite seam.

classification confidence of 43% and were predominantly found lower in the stratigraphy. Peat classifications (n = 36) had an average confidence of 47% with the number of peat predictions increasing with stratigraphic height. Lacustrine classifications (n = 29) has an average confidence of 49% and were distributed throughout the stratigraphy. Only 2 samples were classified as marine samples with a confidence of 25% and 41%. All classification results are summarised in Fig. 4.6 and Fig. 4.7, highlighting the potential changes in palaeodepositional environmental predicted across the stratigraphic section.

### 4.4.4 *Palaeoclimatic Predictions*

Investigations into GPE models trained on subsets of the 15 brGDGTs isomers revealed equally, and occasionally better, performing models can be trained using only the most abundant brGDGT fractions. By predicting on reserved training data, for which the actual temperature is known, the models were assessed based on the Root Mean Square Error (RMSE) and the value of the $95^{th}\%$ of predictive error. Reductions in RMSE and $95^{th}\%$ error represent better models, and the results for GPE-Models 1–6 are summarised in Tables 4.4 and 4.5. Investigations into the comparative performance of models trained on MAAT or WMT show that models trained on WMT preform consistently better than those trained on MAAT, Table 4.4. GPE-Models 3 and 4 resulted in the best performing models (smallest RMSE and $95^{th}\%$ ) across all depositional environments. GPE-Model 5 consistently performed worse than GPE-Model 3.

Removal of zero data (GPE-Models 2, 4, and 6) resulted in generally worse performance than their respective counterparts for both RMSE and $95^{th}\%$ error. The only exceptions to this observation are GPE-Model 4 for soil (WMT) and lacustrine (lWST) training datasets, which performed marginally better (< 0.05 $^{\circ}$C RMSE) than GPE-Model 3. GPE-Models 2, 4, and 6 experienced either parity or a reduction in the size of training data, which negatively impacts model performance. The most severe re-

| **Model** | **Soil - MAAT** | | **Soil - WMT** | | **Peat - MAAT** | | **Peat - WMT** | |
|---|---|---|---|---|---|---|---|---|
| | RMSE | $95^{th}\%$ | RMSE | $95^{th}\%$ | RMSE | $95^{th}\%$ | RMSE | $95^{th}\%$ |
| *brGDGT isomers* | *all* | | | | *all* | | | |
| GPE-Model 1 | 3.89 | 6.38 | 3.23 | 5.30 | 2.83 | 4.63 | 3.59 | 5.92 |
| GPE-Model 2 | 4.10 | 6.74 | 3.47 | 5.70 | 3.68 | 6.00 | 3.80 | 5.79 |
| *brGDGT isomers* | *Ia, Ib, IIa, IIa', IIb', IIIa, IIIa'* | | | | *Ia, IIa, IIIa* | | | |
| GPE-Model 3 | 3.83 | 6.29 | 3.36 | 5.52 | 3.57 | 5.87 | 2.67 | 4.39 |
| GPE-Model 4 | 3.91 | 6.43 | 3.20 | 5.27 | 3.55 | 5.82 | 3.00 | 4.93 |
| *brGDGT isomers* | *Ia, IIa, IIa'* | | | | *Ia, IIa* | | | |
| GPE-Model 5 | 4.21 | 6.92 | 3.85 | 6.32 | 3.31 | 5.45 | 2.80 | 4.60 |
| GPE-Model 6 | 4.39 | 7.23 | 3.74 | 6.15 | 3.19 | 5.25 | 3.00 | 4.93 |

Table 4.4: Model assessments for Gaussian Process Emulator (GPE) air temperature prediction (MAAT and WMT) - soil and peat type depositional environments. RMSE and $95^{th}\%$ error of predictions are used to quantify model performance (lower is better) and the brGDGT isomers used in the training data are shown. GPE-Models 1, 3, 5 contain all training data and GPE-Models 2, 4, 6 have had training data containing "zero" entries removes (see text). GPE-Models predicting WMT performed consistently better than models predicting MAAT. GPE-Model 4 (WMT) performed best for soil type samples and GPE-Model 3 (WMT) performed best for peat type samples.

duction in training data was experienced between GPE-Models 1 and 2 for soil and peat samples, with training data reducing in size by 75% and 97% respectively. The reduction is less marked between Models 3 and 4 (< 15%), a trend which continued between Models 5 and 6 (< 10%). The effect of zero data in the training datasets is unclear, any result likely masked behind the effect of large changes in the size of training data, see Chapter 4.5.2 for further discussion. As such, and for the purposes of this study, all palaeoclimatic predictions made in this study use GPE-Model 3. This model performs well in comparison to those tested and is only marginally and occasionally outperformed by GPE-Model 4. Prediction of MAAT are discussed here for comparative purposes, but this study recommends the use of WMT when training predictive model on brGDGT assemblage data.

Four GPE models are therefore developed here (GPE-Model 3: soil - WMT, peat - WMT, lacustrine - ISWT, and marine - SST), and following GPC palaeodepositional classification, the most applicable GPE is applied to a given sample for temperature prediction. Each of these models have a pre-determined screening metric ($D_{Nearest}$)

| **Model** | **Lacustrine - ISWT** | | **Marine - SST** | |
|---|---|---|---|---|
| | RMSE | $95^{th}$% | RMSE | $95^{th}$% |
| *brGDGT isomers* | *all* | | *all* | |
| GPE-Model 1 | 3.32 | 5.30 | 1.16 | 1.91 |
| GPE-Model 2 | 4.31 | 7.02 | 1.50 | 2.39 |
| *brGDGT isomers* | *Ia, Ib, IIa, IIa', IIb, IIb', IIIa, IIIa'* | | *Ia, Ib, IIa, IIa', IIb, IIb', IIc', IIIa, IIIa'* | |
| GPE-Model 3 | 3.22 | 5.30 | 0.91 | 1.50 |
| GPE-Model 4 | 3.20 | 5.26 | 1.09 | 1.78 |
| *brGDGT isomers* | *Ia, IIa, IIa', IIIa, IIIa'* | | *Ia, Ib, IIb, IIIa, IIIa'* | |
| GPE-Model 5 | 3.40 | 5.60 | 1.47 | 2.40 |
| GPE-Model 6 | 3.69 | 6.07 | 1.44 | 2.36 |

Table 4.5: Model assessments for Gaussian Process Emulator (GPE) water temperature prediction (ISWT and SST) - lacustrine and marine type depositional environments. RMSE and $95^{th}$% error of predictions are used to quantify model performance (lower is better). and the brGDGT isomers used in the training data are shown. GPE-Models 1, 3, 5 contain all training data and GPE-Models 2, 4, 6 have had training data containing "zero" entries removes (see text). GPE-Model 4 performs best for lacustrine samples and GPE-Model 3 performs best for marine samples.

adjusted to provide a similar threshold of screening despite the difference in training data defined as follows: soil (WMT) $\leq$ 0.57, peat (WMT) $\leq$ 0.57, lacustine (IWST) $\leq$ 0.67, marine (SST) $\leq$ 1.09. When applying the GPE-Models to the Barmur Group lignites, the data which failed screening is as follows: soil (WMT) - 50%, peat (WMT) - 28%, lacustrine - 46%, marine - 100%. Across all Barmur Group lignite samples, a total of 40 lignite samples failed the $D_{Nearest}$ screening metric and 52 passed, with failure rates broadly increasing with burial depth, Fig. 4.6. Across all lignite samples a weak, negative correlation is observed between the confidence of classification and the value of $D_{Nearest}$ (power law regression: $r^2 = 0.31$). This relationship is observed most strongly in the lacustrine classifications (power law regression: $r^2 = 0.52$) and most weakly in the peat classifications (power law regression: $r^2 = 0.08$).

The $CD_{Median}$ metrics calculated by the GPE-Models range from 2.55–7.37 for soil type lignites, 1.37–4.77 for peat type lignites, 1.35–2.60 for lacustrine type lignites. Due to only two lignites samples being classified as marine type and both samples failing data screening, they are excluded from this analysis. The hypothesis that $CD_{Median}$ values would increase with depth as the effects of catagenesis increase is not well supported,

Fig. 4.7. Whilst lacustrine $CD_{Median}$ values peak at the lowermost lignites of Units d/e and A, $CD_{Median}$ values for soil type lignites peak at Unit B and for peat type lignites at Unit D. The lowest $CD_{Median}$ values for both soil and peat type lignites are found in Units d/e and A.

Of data which pass the screening metrics, predictions of WMT range from 6.4–12.4 $^{o}$C and average 8.8 $^{o}$C in soil type lignites (n = 6), and range from 11.1–20.8 $^{o}$C and average 16.2 $^{o}$C in peat type lignites (n = 26). Predictions of MAAT were lower than WMT, ranging from -14.7–4.0 $^{o}$C with an average of -12.5 $^{o}$C in soil type lignites (n = 6), and ranging from -2.4–9.3 $^{o}$C with an average of 2.7 $^{o}$C in peat type lignites (n = 14). ISWT predictions range from 6.0–21.1 $^{o}$C and average 13.7 $^{o}$C in lacustrine type lignites (n = 20), Fig. 4.8. No samples classified as marine passed the $D_{Nearest}$ screening test, so any temperature predictions made on these samples are considered unreliable.

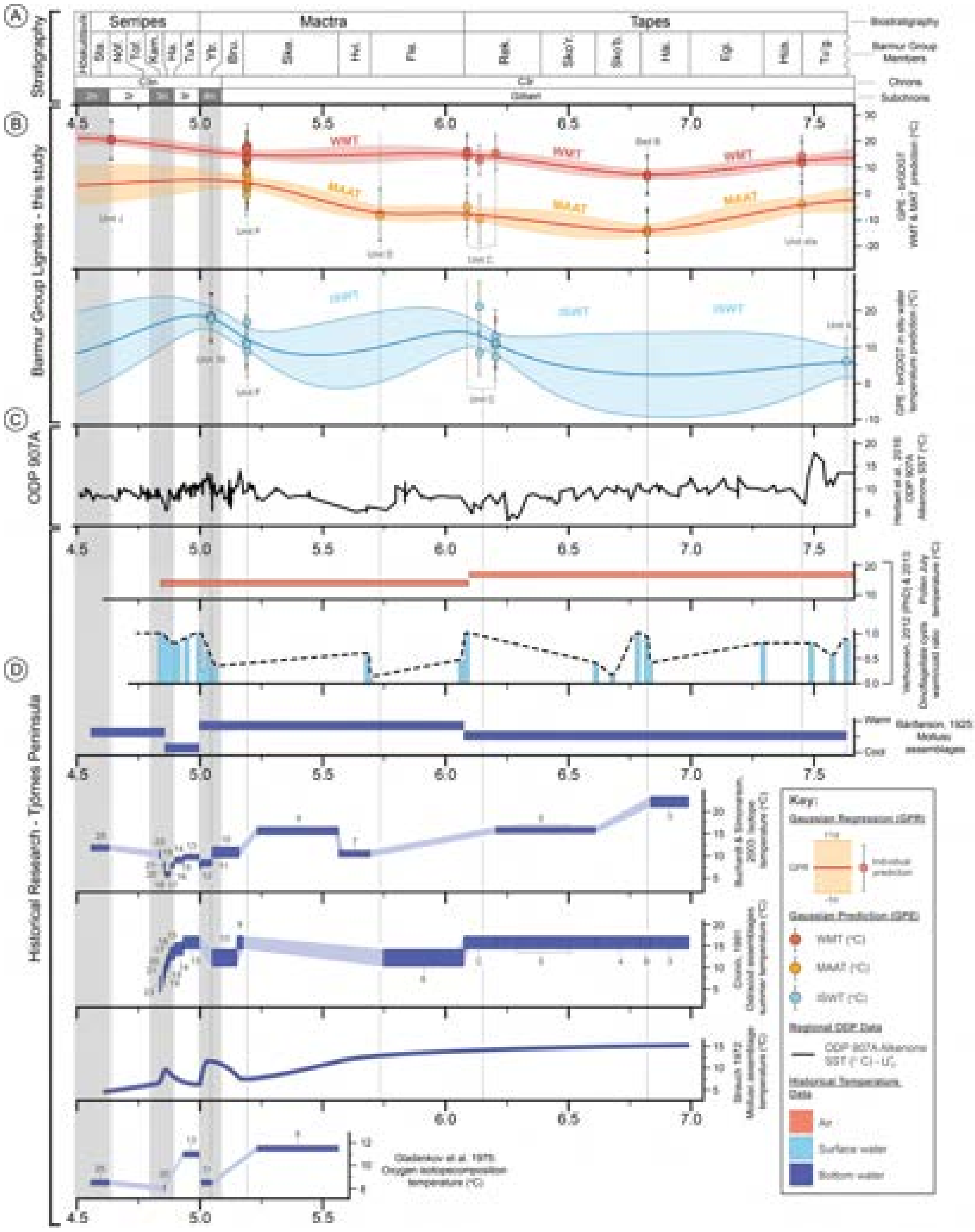


Figure 4.8: Compilation of all palaeoclimatic studies from the Barmur Group replotted on the stratigraphic/age model of Hall et al. (2023) & Allison et al. (2026). A) Barmur Group stratigraphy (Fig. 4.2 for abbreviations) and Chronostratigraphy (Ogg, 2020). B) The Gaussian Process Emulator (GPE) predictions of Warmest Month Air Temperature (WMT), Mean Annual Air Temperature (MAAT), and In Situ Water Temperature (ISWT) for the Barmur Group lignites. All GPE data plotted passed $D_{Nearest}$ screening. Gaussian Process Regression (GPR) - Matern kernel. C) Alkenone ($U^{k'}_{37}$) SST record from ODP Hole 907A. D) Historical palaeoclimate data for the Barmur Group. Air temperature prediction: Verhoeven and Louwye (2013). Surface water temperature predictions: Verhoeven (2012).Bottom water temperature predictions: Bárdarson (1925), Buchardt and Símonarson (2003), Cronin (1991), Strauch (1972), Gladenkov et al. (1975).

## 4.5 Discussion

### 4.5.1 *Stratigraphic and Burial Modelling*

Burial modelling indicates that a heat flux of 275 $mWm^{-2}$ provides an acceptable fit to observations, resulting in a geothermal gradient of c. 130 $^{\circ}Ckm^{-1}$. The nearby geothermal power-plant at Krafla (c. 50 km SE of the Barmur Group) reports elevated geothermal gradients of 200–300 $^{\circ}Ckm^{-1}$ associated with proximity to both the NVZ and a volcanic centre (Friðleifsson et al., 2003), and drilling at the Húsavík-Flatey Fault (c. 10 km south of the Barmur Group) measured temperatures of 110 $^{\circ}C$ at 1.2 km depth (Friðleifsson et al., 1994). From the onset of Barmur Group deposition it is likely that the heat flux at Tjörnes has decreased through time, as tectonic motion has transported the sedimentary basin away from the NVZ. The burial model used here assumes a constant heat flux through time, a simplistic but pragmatic assumption, which likely results in an overestimate of modern geothermal gradients. Improvements to this model would include the ability to vary the heat flux with time, increasing the complexity of the model but potentially yielding more representative results.

Burial modelling also offers an opportunity to connect the observations of laboratory experiments to those of geological scales. Using the Arrhenius equation it is possible to estimate kinetic reaction parameters from the results of Schouten et al. (2004) in order to provide a first order estimate of thermal degradation of GDGTs in geological settings. This can be compared to the results of GDGT analysis and burial modelling to assess whether the short timescale, high heat hydrous pyrolysis experiment of Schouten et al. (2004) is representative of lipids recovered from geological settings. The Arrhenius equation is defined in Eq. 3.4.

$$k = Ae^{\frac{-E_a}{RT}} \quad \text{(Eq. 3.4)}$$

k is the rate constant, the frequency of collisions resulting in reactions [$s^{-1}$]. A is the pre-exponential/frequency factor [$s^{-1}$]. $E_a$ is the activation energy [$Jmol^{-1}$]. R is the universal gas constant [8.314 $JK^{-1}mol^{-1}$] and T is the absolute temperature [K]. The following assumptions are made for the purposes of using this equation: That the application of Eq. 3.4 is a suitable representation of first order reactions, such as thermal degradation. That the response of isoGDGTs to thermal degradation is similar to that of brGDGTs, an assumption supported by the results published in Schouten et al. (2013). That increasing reaction temperatures only reduce the GDGT content, in contrast to the release of GDGTs observed in Schouten et al. (2004) between 200–240 $^{o}$C - these data are excluded from this analysis. Finally, that the reaction temperature (T) in laboratory scale experiments is equivalent to the average burial temperature in geological settings. The simplistic application of the Arrhenius equation here does not explore the Ordinary Differential Equations (ODE) necessary to calculate variable reaction temperatures (i.e. those of geothermal gradients).

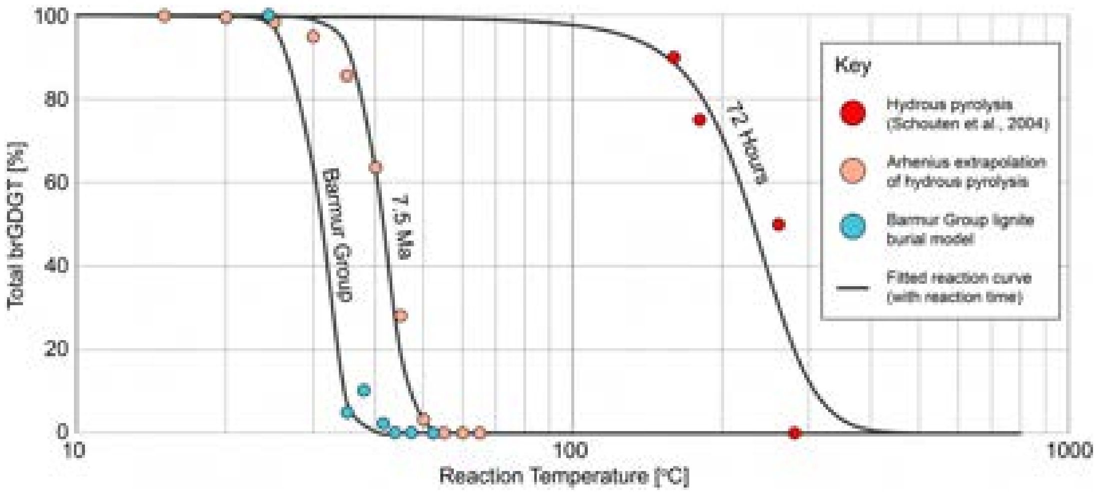


Figure 4.9: Arrhenius equation predictions for thermal degradation of brGDGTs across laboratory and geological timescales. Total brGDGT content represents the total abundance of all 15 brGDGT isomers when compared to a chosen control sample/threshold. **72 Hours** - loss of brGDGTs reported from hydrous pyrolysis (Schouten et al., 2004). **7.5 Ma** - extrapolation of Schouten et al. (2004) to the timescale of the depositional history of the Barmur Group. **Barmur Group**- the loss in brGDGTs measured in Barmur Group lignites, normalised to the absolute total abundance of brGDGTs recovered from Unit J.

The hydrous pyrolysis results of Schouten et al. (2004) result in a pre-exponential constant of 9.75x$10^{13}$ $s^{-1}$ and an activation energy of 173.21 kJmol$^{-1}$. This prediction of activation energy compares well to the experimental thermal degradation of organic rich matter (130–200 kJmol$^{-1}$ - Alwani et al. (2013); Ornaghi Jr et al. (2024)). These values can be used to extrapolate the progress of kinetic reactions across a range of timescales. For the purposes of comparison, total thermal degradation of GDGTs will be considered as when < 0.1% of the original material remains. Using a timescale of 7.5 Ma (i.e. the entire depositional history of the Barmur Group) the results of Schouten et al. (2004) predicts that total thermal degradation of GDGTs will take place at an average reaction temperature of 55 $^{o}$C. For Unit A lignites, absolute abundances of recovered brGDGTs are < 0.1% of the uppermost Unit J lignites. Burial modelling predicts an average burial temperature of 53 $^{o}$C for Unit A. Despite extrapolating a 72 hour laboratory experiment to timescales of several million years, this first order match in results supports the approach of hydrous pyrolysis at laboratory timescales as an analogue for diagenetic degradation of GDGTs. Whilst simplistic assumptions have been used in this application of Arrhenius equation, it outlines how GDGTs (iso- and br-) can be used to estimate the impact of diagenetic heating and even estimate geothermal gradients from the absolute abundance of GDGTs across a stratigraphic section. To further explore this concept it is recommended that more hydrous pyrolysis data be generated to ratify the single experiment of Schouten et al. (2004) and that results be published with all individual GDGT isomers, as the thermal sensitivity of individual isomers cannot be assessed by the publication of metrics such as $TEX_{86}$ or MBT/CBT.

### 4.5.2 *Gaussian Process Predictions*

The GPC approach to palaeodepositional classification using brGDGT distributions enabled the most appropriate GPE-Model to be selected for palaeotemperature prediction. Only two samples were classified as marine-type samples, both data-points also failing the $D_{Nearest}$ screening test, Fig. 4.6. Whilst the GPC can only return one of four classifications defined in the training data, it is a reassuring result that almost all data were classified as non-marine, matching the sedimentological interpretations of the lignite horizons. The correlation length-scales learnt during model training offer a first order assessment of the dimensions of the training dataset which carry the useful information when making predictions. Length-scale analysis indicates that the fractions IIIc', IIc, and IIIb were most important and IIIa, IIa, and Ia the least important when classifying depositional environment. Whilst a difference in training data exists, little similarity is observed between the analysis of correlation length-scales and the importance values of the random forest algorithm - BIGMaC (Martínez-Sosa et al., 2023a). The GPC presented here was assessed to be a more reliable predictor of lacustrine and marine environments but a worse predictor of soil and peat environments when compared to the performance of BIGMaC: Martínez-Sosa et al. (2023a) - Figure 5 vs. this study Table 4.3.

The classification confidence of the Icelandic lignite samples (c. 45%) was found to be lower than the average classification confidence (c. 60%) during the GPC model assessment. It is a recommendation of this study that further data screening could take place using the metric of classification confidence. If brGDGT assemblages show strong dependencies on depositional environment, then it is of importance to classify samples accurately before attempting palaeoclimatic reconstruction. A minimum threshold for classification confidence could improve the robustness of resulting predictions at the cost of an increase in unusable data.

Crampton-Flood et al. (2020) outlines a Bayesian approach to reconstructing palaeotemperature from brGDGT distributions in global soils, BayMBT. Unlike the Gaussian Process approach employed here, BayMBT assumes that the $MBT'_{5ME}$ formulation correctly captures the dependencies of brGDGT distributions on environmental temperature. On equivalent predictive targets (soil samples - MAAT), BayMBT achieved an

RMSE of 6.0 °C (temperature space: -25.6–29.9 °C), in comparison to GPE-Model 3 with an RMSE of 3.83 (temperature space: -17.7–28.0 °C). The most successful model was found when correlating brGDGT distributions with the average temperature of all months > 0 °C ($BayMBT_0$) (Crampton-Flood et al., 2020), achieved 3.8 °C (temperature space: 3.8–27.1 °C). In comparison, GPE-Model 4 (soil samples - WMT) achieved 3.2 °C (temperature space: 4.2–35.5 °C). This comparison highlights the competitiveness of Gaussian Process approaches for brGDGT studies. Notably, the best-performing models were achieved when dimensionality of the training data was reduced. Whilst MBT/CBT-derived formulations capture observable environmental dependencies, these assumptions may ultimately limit predictive performance. The fact that agnostic, data-driven GP models can outperform a Bayesian approach suggests that the priors underlying BayMBT may not be optimally specified. GPE-Model 3, utilising brGDGT isomers with fractional abundances $\geq$ 3%, represents a flexible and capable approach to reconstructing palaeoclimatic information from brGDGT data. Observations of model performance and $D_{Nearest}$ screening failures supports the notion that brGDGTs are best correlated with warm bias temperature measures such as WMT.

It is perhaps surprising that the most accurate temperature predictions came from GPE models trained on subsets of the 15 brGDGT isomers (GPE-Model 3), the isomers utilised were not consistent across depositional types, supporting the notion of depositionally specific dependencies (Table 4.4 and 4.5). It might be expected that reducing the dimensionality of the training dataset would hamper model performance, but this result suggests that noise within less abundant fractions may have negative impacts on the eventual model accuracy. GPE-Model 5 consistently performed worse than GPE-Model 3, implying a sweet-spot exists between dimensionality reduction and model performance. It is proposed here that this observed sweet-spot is driven in part by "zero" entries in the training dataset. It is notable that GPE-Models 1 and 3 share the same sized training dataset, but due to analysing naturally abundant fractions in GPE-Model 3 the amount of "zero" entries consequently decreases. Between GPE-Models 3 and 4, a small decrease in the size of training datasets is experienced (< 15%). GPE-Model 4 for soil and lacustrine type samples marginally outperforms the equivalent GPE-Model 3, a result which is speculatively attributed to the removal of zero data with little loss in the size of the training dataset. This inference is supported by observations of GPE correlation length-scales, which evidence a difference in learnt temperature dependencies between the two models. This result is not observed in peat models because the training datasets are similar enough that loss in any data negatively impacts model performance.

It is perhaps counter-intuitive that the most accurate predictions came from GPE models trained on subsets of the 15 brGDGT isomers (e.g. GPE-Model 3). The isomers retained were not consistent across depositional types, supporting the notion of depositionally/specific dependencies (Tables 4.4 and 4.5). Although one might expect reducing the dimensionality of the training dataset to impair model performance, these results suggest instead that noise from less abundant fractions may negatively impact model accuracy. Consistently poorer results from GPE-Model 5 relative to GPE-Model 3 imply the existence of a “sweet spot” between dimensionality reduction and model performance. Here, this sweet spot is attributed partly to the presence of “zero” entries in the training datasets. For example, GPE-Models 1 and 3 were trained on datasets of the same size, but the use of naturally more abundant fractions in GPE-Model 3 reduced the number of zero entries, resulting in improved model performance. Between GPE-Models 3 and 4, a small (< 15%) reduction in training dataset size was offset by fewer zero entries; GPE-Model 4 for soil and lacustrine samples marginally outperforming GPE-Model 3. This inference is supported by GPE correlation length-scales, which indicate differences in learnt temperature dependencies between the two models. In contrast, peat models did not display this pattern, as any further reduction in the training dataset - already free of most zero entries - had a negative impact on model performance. Details of the amount of zero entries in the training dataset and correlation length-scales produced by GPE-Models 3 and 4 are shown in Fig. 4.10.

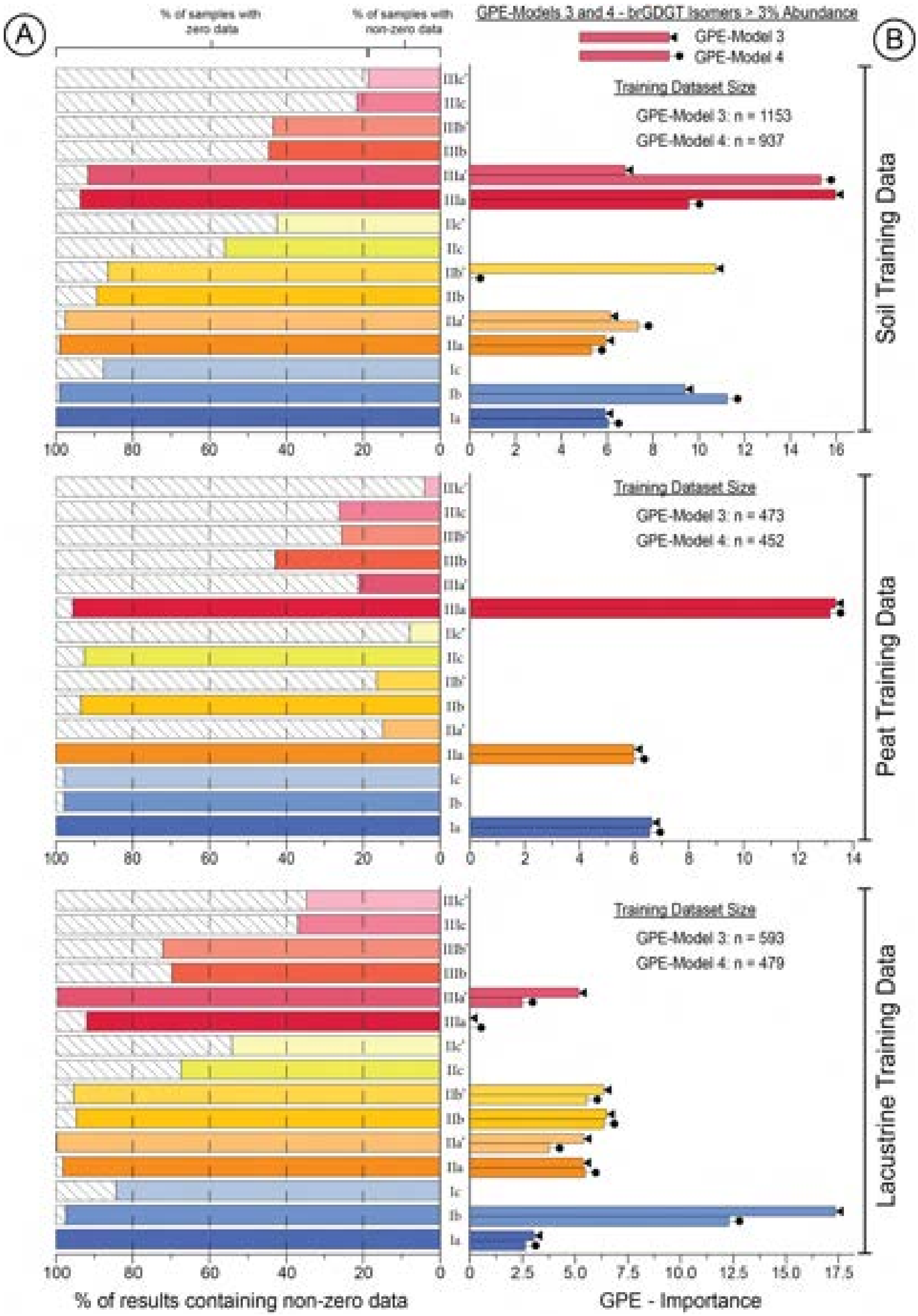


Figure 4.10: A) brGDGT training data for soil, peat and lacustrine type samples. Coloured bars represent the percentage of data containing non-zero entries, and the dashed bars represent the percentage of zero entries. B) GPE-Model 3 (triangles) and GPE-Model 4 (circles) importance (correlation length-scales) for brGDGT isomers > 3% abundance. The greater the importance, the more sensitive the model is to that dimension as a predictor of environmental temperature. Size of training datasets are shown.

The observed decrease in the diversity of brGDGT isomers with burial depth is interesting for two reasons. Firstly, that certain isomers may prove more susceptible to thermal degradation than others and secondly, that despite this degradation certain brGDGT samples continue to pass the GPE $D_{Nearest}$ screening metric. Schouten et al. (2004) recommends that $TEX_{86}$ based GDGT reconstructions occur only on sediments with a Geohopanoid Index > 0.5. It is concluded here, that the more flexible GPE approach to temperature reconstruction may extend the use of GDGT-based palaeothermometers to sediments with Geohopanoid Index values as low as 0.3, Fig. 4.7. Further study in geological and laboratory settings would be needed to better understand the thermal stability of this and other biomarker systems.

Excluding three outlier TOM measurements (Lignite D - 85% and 93%, and Lignite F - 6%), individual measurements of TOM range from 11–68% with an average value of 33% (n = 65). Small variations in TOM are observed between palaeodepositional classifications. Average TOM measurements were recorded of 31%, 32%, and 35% for palaeodepositional classifications of peat, soil, and lacustrine respectively. It might be expected that distinct depositional environments would present distinct trends in TOM, however this is not observed in the Barmur Group lignites. This observation may support a more nuanced approach to palaeodepositional classification, as opposed to the coarse classification scheme employed as part of this study. The challenge of appropriate classification begins with the training dataset. Correct labelling of distinct environments is crucial, but so is acknowledgment of when systems act as a discontinuous or not. It is expected that the lignite facies of the Barmur Group would originate from broadly similar depositional environments, as the geological product is broadly similar across the many lignite horizons. These sediments are commonly interpreted as originating from organic rich, waterlogged, but sub-aerially exposed environments (Eiríksson, 1985; Verhoeven, 2012; Eiríksson and Símonarson, 2021; Hall, 2021; Hall et al., 2023). The variation in classification results from the GPC do indicate a degree of variation in depositional environment, but it is important to consider the confidence of these classifications (4.6, small differences in classification confidence driving coarse classification results. Future studies may investigate more graded classifications than the four groups investigated here, but any investigation is beholden to the quality and or structure of the training data. Although brGDGT distributions demonstrate systematic variation across depositional settings, robust criteria for defining these settings and determining the most suitable classification method will form the basis for future studies.

The *CD* metrics prove a challenging set of signals to interpret. In contrast to the original hypothesis, the largest *CD* values were not found at the base of the succession implying that thermal degradation of the brGDGTs is not perhaps the most impactful force acting on brGDGT distributions. In the peat and lacustrine type lignite samples, the Unit D lignite records some of the highest values of the *CD* metric, this corresponds with some of the highest measured Total Organic Matter (TOM) results, Table 4.1. Weak correlations were observed between TOM and the *CD* metric. In soil type samples a negative correlation was observed (linear regression: $r^2 = 0.42$) and in peat type samples a positive correlation was observed (linear regression: $r^2 = 0.40$). No relationship was observed in lacustrine type samples (linear regression: $r^2 = 0.07$). The lack of obvious relationship between the *CD* metric and increasing burial is interpreted here as supporting the notion that catagenesis has a limited impact on the temperature information archived within brGDGT distributions. Whilst other forces may act on a sample over geological time to lessen its suitability for the reconstruction of palaeoclimate information, the omnipresent force of diagenesis does not appear principal amongst these forces.

### 4.5.3 *Palaeoclimatic Reconstruction of the Barmur Group*

Records of Northern Hemisphere, Miocene–Pliocene air temperature are uncommon. Ballantyne et al. (2010) used brGDGTs recovered from Arctic peat deposits to reconstruct Pliocene MAAT, reporting MBT-based temperature predictions of -0.6 °C, 19.1 °C warmer than modern. Similar levels of warming relative to modern (+9.7–16.7 °C) are reported from Pliocene–Miocene wood samples sourced from the Arctic Archipelago of Canada using hydrogen isotopes of ligninmethoxy groups ($\delta^2 H_{LM}$ proxy - Porter et al. (2022)). Fletcher et al. (2019) reconstructed mean summer air temperatures using brGDGT recovered from Pliocene sediments on Ellesmere Island, reporting temperatures of 14.1–17.4 °C. The WMT temperature range for the Barmur Group lignites (6.4–12.4 °C) corroborates regional estimates of air temperature during these time periods. Verhoeven and Louwye (2013) provides estimates of July air temperature from pollen records, whilst both records provide temperature estimates between 10–20 °C, the brGDGT record evidences a warming trend across the Group in contrast to the cooling trend of the pollen record, Fig. 4.8. Alkenone ($U_{37}^{k'}$) Sea Surface Temperature (SST) reconstructions from the Icelandic Sea (ODP Hole 907A) during the Miocene–Pliocene predict temperatures between 5–15 °C with an average of c. 10 °C (De Schepper et al., 2015; Herbert et al., 2016). Whilst sparse, the estimates for both WMT and IWST provided by the Barmur Group lignites are in agreement with the SST signal of the nearby ODP-Hole 907A. Between Units A and B a decrease in temperature is shared between the records, similarly with the rising temperature trend between Units C and 10. Detailed comparison of the brGDGT derived surface water temperature record with the historical records published at Tjörnes is challenging due to the sparseness of the brGDGT data. Multiple temperature records report a rise in water temperature during the C3n.4n subchron (Bárdarson, 1925; Strauch, 1972; Cronin, 1991; Buchardt and Símonarson, 2003; Verhoeven, 2012), Fig. 4.8. This temperature rise is observable in the lacustrine IWST record, 7 °C of warming is observed between Units F and Bed 10, but either side of these data-points the Gaussian regression is unconstrained, limiting further comparison. Importantly, the GPE temperature predictions presented here offer feasible/realistic estimates of both air and water temperature for the time period, in contrast to the more established MBT-derived predictions ranging as much as c. -23–+26 °C ($MAAT_{peat}$ (Naafs et al., 2017)). The GPE approach outlined in this study provides both a means to more confidently predict temperature and also effectively screen data using metrics such as $D_{Nearest}$, resulting in more robust temperature predictions.

## 4.6 Conclusion

This study demonstrates that the Barmur Group lignites of the Tjörnes Peninsula preserve a valuable - if diagenetically altered - record of brGDGT distributions, offering insight into pre-glacial, palaeoenvironmental conditions during the Miocene–Early Pliocene. Through detailed stratigraphic analysis, burial and thermal modelling, and geochemical characterisation the thermal history of the lignites is constrained, assessed and correlated with the preservation state of organic biomarkers. The geothermal gradient of c. 130 $^{o}$Ckm$^{-1}$, derived from inverse modelling against vitrinite reflectance data, supports the hypothesis that these shallowly buried sediments experienced elevated heat flux in the absence of deep burial, providing a rare setting to test the resilience of biomarker systems over geological time.

The integration of Gaussian Process Classifiers (GPC) and Gaussian Process Emulators (GPE) enables a more nuanced interpretation of brGDGT-derived palaeoclimatic signals. By first classifying palaeodepositional environments and then applying environment-specific GPEs, the traditional limitations of MBT/CBT indices are overcome. Temperature predictions generated through this method yield realistic estimates of both air and surface water temperatures, which are consistent with regional palaeoclimatic reconstructions and independent proxy records from the Barmur Group itself. This study supports the recent observations of warm temperature bias dependencies in brGDGT distributions (Crampton-Flood et al., 2020; Raberg et al., 2021, 2022) and concludes that brGDGT derived temperature models are better predictors of WMT over MAAT. Moreover, our comparison of GPE outputs with thermal maturity indices (CPI, Geohopanoid Index) and the *CD* metric indicates that, although thermal degradation has occurred, it does not substantially distort the brGDGT temperature signals, allowing reliable temperature reconstructions across the entire stratigraphic section. Data screening metrics, such as $D_{Nearest}$ facilitate the extraction of trustworthy palaeoclimatic information from samples which appear to have preserved an intact signal over the course of geological time. For the first time, a connection is made between laboratory and geological results by a first order kinetic reaction model used to demonstrate that thermal degradation via hydrous pyrolysis (Schouten et al., 2004) appears representative of thermal degradation of GDGTs in geological settings.

Overall, this chapter illustrates the viability of applying agnostic machine learning approaches to palaeoclimate reconstructions from complex and thermally altered sedimentary archives. The multi-proxy approach outlined here not only strengthens confidence in the palaeotemperature estimates from the Barmur Group lignites but also provides a transferable framework for future work on diagenetically impacted brGDGT datasets. The findings underline the importance of incorporating depositional environment classification and thermal history constraints into biomarker-based climate reconstructions, particularly in regions which experience elevated geothermal gradients such as Iceland.

## 4.7 Code and Data Availability

The code used to generate all data and figures in this manuscript is available at the following link: `https://github.com/matthew-allison05?tab=repositories`. Data generated as part of this study is presented in Appendix B.

# *REFERENCES*


Agisoft. Metashape professional, 2022. Version 1.7.5.

K. J. Albertsson. Um aldur jardlaga á Tjörnesi. *Náttúrufræðingurinn*, 48:1–8, 1978.

J. Allan and A. Douglas. Variations in the content and distribution of n-alkanes in a series of Carboniferous vitrinites and sporinites of bituminous rank. *Geochimica et Cosmochimica Acta*, 41(9):1223–1230, 1977.

M. S. Allison, J. R. Hall, and S. M. Jones. Barmur Group, Tjörnes Peninsula. https://v3geo.com/model/367, 2023a. V3Geo.

M. S. Allison, J. R. Hall, M. Papadopoulos, J. Eíriksson, and S. M. Jones. Barmur Group cross section high quality. Zenodo, 2023b.

M. S. Allison, J. R. Hall, M. T. Papadopoulos, J. Eiríksson, and S. M. Jones. Updated stratigraphy and structure of the Barmur Group (Tjörnes beds, Iceland) and implications for Bering Strait opening. *Geological Society, London, Special Publications*, 553(1):SP553–2023, 2026.

M. S. Alwani, H. A. Khalil, O. Sulaiman, M. N. Islam, and R. Dungani. An approach to using agricultural waste fibres in biocomposites application: thermogravimetric analysis and activation energy study. *BioResources*, 9(1):218–230, 2013.

H. Ármannsson, M. Ólafsson, G. O. Fridleifsson, W. G. Darling, and T. Laier. Organic gas in Öxarfjörður, NE-Iceland. *Proc. of the 9th International Symposium on Water-Rock Interaction - Taupo, New Zealand*, 1998.

A. P. Ballantyne, D. Greenwood, J. Sinninghe Damsté, A. Csank, J. Eberle, and N. Rybczynski. Significantly warmer Arctic surface temperatures during the Pliocene indicated by multiple independent proxies. *Geology*, 38(7):603–606, 2010.

G. Bárdarson. A stratigraphical survey of the Pliocene deposits at Tjörnes, in northern Iceland. K. danske vidensk. *Selsk. Biol. Medd. IV*, 5:118, 1925.

C. J. Beets. *Calibration of late Cenozoic marine strontium isotope variations and its chronostratigraphic and geochemical applications*. Vrije Universiteit, 1992.

S. T. Belt, G. Massé, S. J. Rowland, M. Poulin, C. Michel, and B. LeBlanc. A novel chemical fossil of palaeo sea ice: IP25. *Organic Geochemistry*, 38(1):16–27, 2007.

P. K. Bijl, J. Frieling, M. J. Cramwinckel, C. Boschman, A. Sluijs, and F. Peterse. Maastrichtian–Rupelian paleoclimates in the southwest Pacific–a critical reevaluation of biomarker paleothermometry and dinoflagellate cyst paleoecology at Ocean Drilling Program Site 1172. *Climate of the Past*, 17(6):2393–2425, 2021.

P. K. Bijl, K. K. Sliwinska, B. Duncan, A. Huguet, S. Naeher, R. Rattanasriampaipong, C. Sosa-Montes de Oca, A. Auderset, M. Berke, B. S. Kim, et al. Reviews and syntheses: Best practices for the application of marine GDGTs as proxy for paleotemperatures: sampling, processing, analyses, interpretation, and archiving protocols. *EGUsphere*, 2025:1–74, 2025.

C. I. Blaga, G.-J. Reichart, O. Heiri, and J. S. Sinninghe Damsté. Tetraether membrane lipid distributions in water-column particulate matter and sediments: a study of 47 European lakes along a north–south transect. *Journal of Paleolimnology*, 41:523–540, 2009.

S. Brassell, G. Eglinton, I. Marlowe, U. Pflaumann, and M. Sarnthein. Molecular stratigraphy: a new tool for climatic assessment. *Nature*, 320(6058):129–133, 1986.

C. M. Brierley and A. V. Fedorov. Comparing the impacts of Miocene–Pliocene changes in inter-ocean gateways on climate: Central American Seaway, Bering Strait, and Indonesia. *Earth and Planetary Science Letters*, 444:116–130, 2016. doi: 10.1016/j.epsl.2016.03.010.

B. Buchardt and L. Símonarson. Isotope palaeotemperatures from the Tjörnes beds in Iceland: evidence of Pliocene cooling. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 189(1-2):71–95, 2003. doi: 10.1016/s0031-0182(02)00594-1.

W. Cao, S. Zahirovic, N. Flament, S. Williams, J. Golonka, and R. D. Müller. Improving global paleogeography since the late Paleozoic using paleobiology. *Biogeosciences*, 14(23):5425–5439, 2017.

C. Clotten, R. Stein, K. Fahl, M. Schreck, B. Risebrobakken, and S. De Schepper. On the causes of Arctic sea ice in the warm Early Pliocene. *Scientific Reports*, 9(1):989, 2019. doi: 10.1038/s41598-018-37047-y.

J. Cooper and E. Bray. A postulated role of fatty acids in petroleum formation. *Geochimica et Cosmochimica Acta*, 27(11):1113–1127, 1963.

E. D. Crampton-Flood, J. E. Tierney, F. Peterse, F. M. Kirkels, and J. S. S. Damsté. BayMBT: A Bayesian calibration model for branched glycerol dialkyl glycerol tetraethers in soils and peats. *Geochimica et Cosmochimica Acta*, 268:142–159, 2020.

T. M. Cronin. Late neogene marine ostracoda from Tjörnes, Iceland. *Journal of Paleontology*, 65(5):767–794, 1991. doi: 10.1017/S002233600003777X.

J. S. S. Damsté, E. C. Hopmans, R. D. Pancost, S. Schouten, and J. A. Geenevasen. Newly discovered non-isoprenoid glycerol dialkyl glycerol tetraether lipids in sediments. *Chemical Communications*, (17):1683–1684, 2000.

J. S. S. Damsté, J. Ossebaar, B. Abbas, S. Schouten, and D. Verschuren. Fluxes and distribution of tetraether lipids in an equatorial African lake: constraints on the application of the TEX86 palaeothermometer and BIT index in lacustrine settings. *Geochimica et Cosmochimica Acta*, 73(14):4232–4249, 2009.

J. S. S. Damsté, J. Ossebaar, S. Schouten, and D. Verschuren. Distribution of tetraether lipids in the 25-ka sedimentary record of Lake Challa: extracting reliable TEX86 and MBT/CBT palaeotemperatures from an equatorial African lake. *Quaternary Science Reviews*, 50:43–54, 2012.

J. S. S. Damsté, W. I. C. Rijpstra, B. U. Foesel, K. J. Huber, J. Overmann, S. Nakagawa, J. J. Kim, P. F. Dunfield, S. N. Dedysh, and L. Villanueva. 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, 2018.

J. H. Davies and D. R. Davies. Earth's surface heat flux. *Solid earth*, 1(1):5–24, 2010.

C. De Jonge, E. C. Hopmans, A. Stadnitskaia, W. I. C. Rijpstra, R. Hofland, E. Tegelaar, and J. S. S. Damsté. Identification of novel penta-and hexamethylated branched glycerol dialkyl glycerol tetraethers in peat using HPLC–MS2, GC–MS and GC–SMB-MS. *Organic geochemistry*, 54:78–82, 2013.

C. De Jonge, E. C. Hopmans, C. I. Zell, J.-H. Kim, S. Schouten, and J. S. S. Damsté. Occurrence and abundance of 6-methyl branched glycerol dialkyl glycerol tetraethers in soils: Implications for palaeoclimate reconstruction. *Geochimica et Cosmochimica Acta*, 141:97–112, 2014.

C. De Jonge, D. Radujković, B. D. Sigurdsson, J. T. Weedon, I. Janssens, and F. Peterse. Lipid biomarker temperature proxy responds to abrupt shift in the bacterial community composition in geothermally heated soils. *Organic Geochemistry*, 137: 103897, 2019.

S. De Schepper, M. Schreck, K. M. Beck, J. Matthiessen, K. Fahl, and G. Mangerud. Early Pliocene onset of modern Nordic Seas circulation related to ocean gateway changes. *Nature communications*, 6(1):8659, 2015. doi: 10.1038/ncomms9659.

T. Denk, F. Grimsson, R. Zetter, and L. A. Símonarson. *Late Cainozoic floras of Iceland: 15 million years of vegetation and climate history in the northern North Atlantic*, volume 35. Springer Science & Business Media, 2011.

DJI. Ground station pro, 2021. Version 2.0.

R. R. Doell. Palaeomagnetic studies of Icelandic lava flows. *Geophysical Journal International*, 26(5):459–479, 1972. doi: 10.1111/j.1365-246X.1972.tb05763.x.

T. Dunkley Jones, Y. L. Eley, W. Thomson, S. E. Greene, I. Mandel, K. Edgar, and J. A. Bendle. Optimal: a new machine learning approach for GDGT-based palaeothermometry. *Climate of the Past*, 16(6):2599–2617, 2020.

G. Eglinton and R. J. Hamilton. Leaf Epicuticular Waxes: The waxy outer surfaces of most plants display a wide diversity of fine structure and chemical constituents. *Science*, 156(3780):1322–1335, 1967.

P. Einarsson. Plate boundaries, rifts and transforms in Iceland. *Jökull*, 58(12):35–58, 2008. doi: 10.33799/jokull2008.58.035.

T. Einarsson, D. M. Hopkins, and R. Doell. The stratigraphy of Tjörnes, northern Iceland, and the history of the Bering Land Bridge. *In: Hopkins, D.M. (ed.) The Bering Land Bridge*, 3:312–125, 1967.

J. Eiríksson. Lithostratigraphy of the upper Tjörnes sequence, north Iceland: the Breidavík Group. *Acta Naturalia Islandica*, 29:1–37, 1981.

J. Eiríksson. Facies analysis of the Breidavík Group sediments on Tjörnes, north Iceland. *Acta Naturalia Islandica*, 31:1–56, 1985.

J. Eiríksson. Glaciation events in the Pliocene–Pleistocene volcanic succession of Iceland. *Jökull*, 58:315–329, 2008.

J. Eiríksson and L. A. Símonarson. *Pacific-Atlantic Mollusc Migration: Pliocene Inter-Ocean Gateway Archives on Tjörnes, North Iceland*, volume 52. Springer Nature, 2021.

J. Eiríksson, A. I. Gudmundsson, L. Kristjánsson, and K. Gunnarsson. Palaeomagnetism of Pliocene-Pleistocene sediments and lava flows on Tjörnes and Flatey, North Iceland. *Boreas*, 19(1):39–55, 1990. doi: 10.1111/j.1502-3885.1990.tb00420.x.

J. Eiríksson et al. Growth of an intermittent ice sheet in Iceland during the late Pliocene and early Pleistocene. *Quaternary research*, 42(2):115–130, 1994.

W. A. Elders and G. Fridleifsson. The Science Program of the Iceland Deep Drilling Project (IDDP): a study of supercritical geothermal resources. In *Proceedings World geothermal congress*, page 9, 2010.

F. J. Elling, M. Könneke, M. Mußmann, A. Greve, and K.-U. Hinrichs. Influence of temperature, pH, and salinity on membrane lipid composition and TEX86 of marine planktonic thaumarchaeal isolates. *Geochimica et Cosmochimica Acta*, 171:238–255, 2015.

P. Experts. Move, 2022. Version 13.0.

S. F. Zur geologie von Tjörnes (Nordisland). Sonderveröff (in German). *Geol. Inst*, 8: 129 pp., 1963.

P. Farrimond, A. Taylor, and N. TelnÆs. Biomarker maturity parameters: the role of generation and thermal degradation. *Organic Geochemistry*, 29(5-7):1181–1197, 1998.

S. Fietz, S. L. Ho, C. Huguet, A. Rosell-Melé, and A. Martínez-García. Appraising GDGT-based seawater temperature indices in the Southern Ocean. *Organic Geochemistry*, 102:93–105, 2016.

A. R. Fiorillo. Dinosaurs of Alaska: implications for the Cretaceous origin of Beringia. *GSA Special Publications*, 442, 2008. doi: 10.1130/2008.442(15).

Z. Fleming. Using virtual outcrop models and Google Earth to teach structural geology concepts. *Journal of Structural Geology*, 156:104537, 2022. doi: 10.1016/j.jsg.2022.104537.

T. L. Fletcher, L. Warden, J. S. Sinninghe Damsté, K. J. Brown, N. Rybczynski, J. C. Gosse, and A. P. Ballantyne. Evidence for fire in the Pliocene Arctic in response to amplified temperature. *Climate of the Past*, 15(3):1063–1081, 2019.

G. Ó. Friðleifsson, J. Eiríksson, H. Hafliðason, K. Gunnarsson, T. K. Ólafsson, G, B. Larsen, S. Thórhallsson, and Á. E. Sveinbjörnsdóttir. ODP proposal for drillsites within the Tjörnes Fracture Zone sedimentary basin. Late Cainozoic paleoceanography and sedimentation history at the Arctic boundary. *Report 28.06.1994*, 1994.

G. Ó. Friðleifsson, H. Ármannsson, K. Árnason, I. Þ. Bjarnason, and G. Gíslason. Part i: Geosciences and Site Selection. In G. Ó. Friðleifsson, editor, *Iceland Deep Drilling Project, Feasibility Report*, volume OS-2003-007 of *Orkustofnun Report*, pages 1–104. Orkustofnun, Reykjavík, Iceland, 2003. URL https://rafhladan.is/bitstream/handle/10802/19265/OS-2003-007.pdf?sequence=1.

C. R. Gallagher and R. A. Askew. Hrútfjallstindar. https://v3geo.com/model/417, 2024. V3Geo.

S. Garcia, J. Angelier, F. Bergerat, and C. Homberg. Tectonic analysis of an oceanic transform fault zone based on fault-slip data and earthquake focal mechanisms: the Húsavık–Flatey Fault zone, Iceland. *Tectonophysics*, 344(3-4):157–174, 2002. doi: 10.1016/s0012-821x(03)00400-x.

S. Garcia, N. O. Arnaud, J. Angelier, F. Bergerat, and C. Homberg. Rift jump process in Northern Iceland since 10 ma from 40Ar/39Ar geochronology. *Earth and Planetary Science Letters*, 214(3-4):529–544, 2003.

Á. Geirsdóttir. Pliocene and Pleistocene glaciations of Iceland: A brief overview of the glacial history. *Developments in Quaternary Sciences*, 15:199–210, 2011.

A. Y. Gladenkov, A. E. Oleinik, L. Marincovich Jr, and K. B. Barinov. A refined age for the earliest opening of Bering Strait. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 183(3-4):321–328, 2002. doi: 10.1016/S0031-0182(02)00249-3.

A. Y. Gladenkov et al. Onset of connections between the Pacific and Arctic Oceans through the bering strait in the neogene. *Stratigraphy and Geological Correlation*, 12:175–187, 2004.

Y. Gladenkov and G. Gurari. Paleomagnetic characteristics of the Plio-Pleistocene in Iceland (Tjörnes Peninsula). *Reports of the USSR Academy of Science*, 230:1173–1174, 1976.

Y. Gladenkov, E. Krasnov, A. Ignatev, and V. Schegus. On the temperature boundaries of habitats of Late-Cenozoic mollusca in the North Atlantic. In *Academy of Sciences USSR*, volume 223, pages 176–177, 1975.

K. E. K. A. D. . K. Y. Gunnarsson, K. Auðlindakönnun í Öxarfirði. Reykjavík, Iceland: Orkustofnun OS-96001/JHD-01. page 46 pp, 1996.

T. A. Halamka, J. McFarlin, A. Younkin, J. Depoy, N. Dildar, and S. Kopf. Oxygen limitation can trigger the production of branched GDGTs in culture. *Geochemical Perspectives Letters*, 2021.

T. A. Halamka, J. H. Raberg, J. M. McFarlin, A. D. Younkin, C. Mulligan, X.-L. Liu, and S. H. Kopf. Production of diverse brGDGTs by Acidobacterium Solibacter usitatus in response to temperature, pH, and O2 provides a culturing perspective on brGDGT proxies and biosynthesis. *Geobiology*, 21(1):102–118, 2023.

J. Hall, M. Allison, M. Papadopoulos, D. Barfod, and S. Jones. Timing and consequences of Bering Strait opening: new insights from 40Ar/39Ar dating of the barmur group (Tjörnes beds), northern Iceland. *Paleoceanography and Paleoclimatology*, 38:1–21, 2023. doi: 10.1029/2022PA004539.

J. R. Hall. *Reconstructing Pliocene climatic, oceanic and tectonic interactions in the North East Atlantic, Nordic Seas and onshore Iceland by geochemical fingerprinting of marine and terrestrial sediments*. PhD thesis, University of Birmingham, 2021.

T. D. Herbert, L. C. Peterson, K. T. Lawrence, and Z. Liu. Tropical ocean temperatures over the past 3.5 million years. *science*, 328(5985):1530–1534, 2010.

T. D. Herbert, K. T. Lawrence, A. Tzanova, L. C. Peterson, R. Caballero-Gill, and C. S. Kelly. Late Miocene global cooling and the rise of modern ecosystems. *Nature Geoscience*, 9(11):843–847, 2016.

E. C. Hopmans, J. W. Weijers, E. Schefuß, L. Herfort, J. S. S. Damsté, and S. Schouten. 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, 2004.

A. J. Houben, P. K. Bijl, J. Pross, S. M. Bohaty, S. Passchier, C. E. Stickley, U. Röhl, S. Sugisaki, L. Tauxe, T. van de Flierdt, et al. Reorganization of Southern Ocean plankton ecosystem at the onset of Antarctic glaciation. *Science*, 340(6130):341–344, 2013.

A. J. Houben, P. K. Bijl, A. Sluijs, S. Schouten, and H. Brinkhuis. Late Eocene Southern Ocean cooling and invigoration of circulation preconditioned Antarctica for full-scale glaciation. *Geochemistry, Geophysics, Geosystems*, 20(5):2214–2234, 2019.

J. Howell, M. Chmielewska, C. Lewis, S. Buckley, N. Naumann, and J. Pugsley. Acquisition of data for building photogrammetric virtual outcrop models for the geosciences using remotely piloted vehicles (RPVs). 2021. doi: 10.31223/X54914.

A. Hu, G. A. Meehl, et al. Influence of Bering Strait flow and North Atlantic circulation on glacial sea-level changes. *Nature Geoscience*, 3:118–121, 2010. doi: 10.1038/ngeo729.

S. J. Hurley, F. J. Elling, M. Könneke, C. Buchwald, S. D. Wankel, A. E. Santoro, J. S. Lipp, K.-U. Hinrichs, and A. Pearson. Influence of ammonia oxidation rate on thaumarchaeal lipid composition and the TEX86 temperature proxy. *Proceedings of the National Academy of Sciences*, 113(28):7762–7767, 2016.

G. N. Inglis, A. Farnsworth, D. Lunt, G. L. Foster, C. J. Hollis, M. Pagani, P. E. Jardine, P. N. Pearson, P. Markwick, A. M. Galsworthy, et al. Descent toward the Icehouse: Eocene sea surface cooling inferred from GDGT distributions. *Paleoceanography*, 30(7):1000–1020, 2015.

S. M. Jones, M. Hoggett, S. E. Greene, and T. Dunkley Jones. Large Igneous Province thermogenic greenhouse gas flux could have initiated Paleocene-Eocene Thermal Maximum climate change. *Nature Communications*, 10(1):5547, 2019.

E. J. Judd, J. E. Tierney, B. T. Huber, S. L. Wing, D. J. Lunt, H. L. Ford, G. N. Inglis, E. L. McClymont, C. L. O'Brien, R. Rattanasriampaipong, et al. The PhanSST global database of phanerozoic sea surface temperature proxy data. *Scientific data*, 9(1):753, 2022.

M. Kalacska, O. Lucanus, J. Arroyo-Mora, E. Laliberte, K. Elmer, G. Leblanc, and A. Groves. Accuracy of 3D landscape reconstruction without ground control points using different UAS platforms. *Drones*, 4(2):13, 2020. doi: 10.3390/drones4020013.

J.-H. Kim, S. Schouten, E. C. Hopmans, B. Donner, and J. S. S. Damsté. Global sediment core-top calibration of the TEX86 paleothermometer in the ocean. *Geochimica et Cosmochimica Acta*, 72(4):1154–1173, 2008.

J.-H. Kim, J. Van der Meer, S. Schouten, P. Helmke, V. Willmott, F. Sangiorgi, N. Koç, E. C. Hopmans, and J. S. S. Damsté. New indices and calibrations derived from the distribution of crenarchaeal isoprenoid tetraether lipids: Implications for past sea surface temperature reconstructions. *Geochimica et Cosmochimica Acta*, 74(16): 4639–4654, 2010.

J.-H. Kim, S. Schouten, M. Rodrigo-Gámiz, S. Rampen, G. Marino, C. Huguet, P. Helmke, R. Buscail, E. C. Hopmans, J. Pross, et al. Influence of deep-water derived isoprenoid tetraether lipids on the TEX86H paleothermometer in the Mediterranean Sea. *Geochimica et Cosmochimica Acta*, 150:125–141, 2015.

Y. Koga, T. Kyuragi, M. Nishihara, and N. Sone. Did archaeal and bacterial cells arise independently from noncellular precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. *Journal of molecular evolution*, 46:54–63, 1998.

F. Kolawole and J. C. Evenick. Global distribution of geothermal gradients in sedimentary basins. *Geoscience Frontiers*, 14(6):101685, 2023.

L. Kristjánsson, H. Jóhannesson, J. Eiríksson, and A. Gudmundsson. Brunhes–Matuyama paleomagnetism in three lava sections in Iceland. *Canadian Journal of Earth Sciences*, 25(2):215–225, 1988.

L. Kristjánsson. A reconnaissance study of paleomagnetic directions in the Tjörnes beds, Northern Iceland. *Jökull Journal*, 54:57–63, 2004. doi: 10.33799/jokull2004.54.057.

K. Littler, S. A. Robinson, P. R. Bown, A. J. Nederbragt, and R. D. Pancost. *Nature Geoscience*, 4(3):169–172, 2011.

Z. Liu, M. Pagani, D. Zinniker, R. DeConto, M. Huber, H. Brinkhuis, S. R. Shah, R. M. Leckie, and A. Pearson. Global cooling during the Eocene–Oligocene climate transition. *Science*, 323(5918):1187–1190, 2009.

A. Mackenzie, R. Patience, J. Maxwell, M. Vandenbroucke, and B. Durand. Molecular parameters of maturation in the Toarcian shales, Paris Basin, France—I. Changes in the configurations of acyclic isoprenoid alkanes, steranes and triterpanes. *Geochimica et Cosmochimica Acta*, 44(11):1709–1721, 1980.

L. J. Marincovich. Central American paleogeography controlled Pliocene Arctic Ocean molluscan migrations. *Geology*, 28:551–554, 2000. doi: 10.1130/0091-7613(2000)28<551:CAPCPA>2.0.CO;2.

L. J. Marincovich and A. Y. Gladenkov. Evidence for an early opening of the Bering Strait. *Nature*, 397:149–151, 1999. doi: 10.1038/16446.

P. Martínez-Sosa, J. E. Tierney, L. C. Pérez-Angel, I. C. Stefanescu, J. Guo, F. Kirkels, J. Sepúlveda, F. Peterse, B. N. Shuman, and A. V. Reyes. Development and application of the Branched and Isoprenoid GDGT Machine learning Classification algorithm (BIGMaC) for paleoenvironmental reconstruction. *Paleoceanography and Paleoclimatology*, 38(7):e2023PA004611, 2023a.

P. Martínez-Sosa, J. E. Tierney, L. C. Pérez-Angel, I. C. Stefanescu, J. Guo, F. Kirkels, J. Sepúlveda, F. Peterse, B. N. Shuman, and A. V. Reyes. Development and application of the Branched and Isoprenoid GDGT Machine learning Classification algorithm (BIGMaC) for paleoenvironmental reconstruction. *Paleoceanography and Paleoclimatology*, 38(7):e2023PA004611, 2023b.

S. Metzger, S. Jónsson, G. Danielsen, S. Hreinsdóttir, F. Jouanne, D. Giardini, and T. Villemin. Present kinematics of the Tjörnes Fracture Zone, North Iceland, from campaign and continuous GPS measurements. *Geophysical Journal International*, 192(2):441–455, 2013a.

S. Metzger, S. Jónsson, G. Danielsen, S. Hreinsdóttir, F. Jouanne, D. Giardini, and T. Villemin. Present kinematics of the Tjörnes Fracture Zone, North Iceland, from campaign and continuous GPS measurements. *Geophysical Journal International*, 192:441–455, 2013b. doi: 10.1093/gji/ggs032.

P. J. Müller, G. Kirst, G. Ruhland, I. Von Storch, and A. Rosell-Melé. Calibration of the alkenone paleotemperature index U37K 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, 1998.

R. D. Müller, S. Zahirovic, et al. A global plate model including lithospheric deformation along major rifts and orogens since the Triassic. *Tectonics*, 38:1884–1907, 2019. doi: 10.1029/2018TC005462.

B. Naafs and R. Pancost. Sea-surface temperature evolution across Aptian oceanic anoxic event 1a. *Geology*, 44(11):959–962, 2016.

B. D. A. Naafs, G. N. Inglis, Y. Zheng, M. Amesbury, H. Biester, R. Bindler, J. Blewett, M. Burrows, D. Del Castillo Torres, F. M. Chambers, et al. Introducing global peat-specific temperature and pH calibrations based on brGDGT bacterial lipids. *Geochimica et Cosmochimica Acta*, 208:285–301, 2017.

C. R. Nelson, W. Li, I. M. Lazar, K. H. Larson, A. Malik, and M. L. Lee. Geochemical significance of n-alkane compositional-trait variations in coals. *Energy & fuels*, 12 (2):277–283, 1998.

C. L. O'Brien, S. A. Robinson, R. D. Pancost, J. S. S. Damsté, S. Schouten, D. J. Lunt, H. Alsenz, A. Bornemann, C. Bottini, and S. C. Brassell. Cretaceous sea-surface temperature evolution: Constraints from TEX86 and planktonic foraminiferal oxygen isotopes. *Earth-Science Reviews*, 172:224–247, 2017.

J. G. Ogg. *Geomagnetic polarity time scale*. Elsevier, 2020.

M. Ólafsson, G. Ó. Friðleifsson, J. Eiríksson, and H. Ármannsson. *On the origin of organic gas in Öxarfjörður, NE-Iceland*. Orkustofnun, 1993.

H. L. Ornaghi Jr, M. Faccio, and M. R. F. Soares. Thermal degradation kinetics of natural fibers: Determination of the kinetic triplet and lifetime prediction. *Polysaccharides*, 5(3):169–183, 2024.

A. Pearson, S. R. Flood Page, T. L. Jorgenson, W. W. Fischer, and M. B. Higgins. Novel hopanoid cyclases from the environment. *Environmental Microbiology*, 9(9): 2175–2188, 2007.

K. Peters and J. Moldowan. Effects of source, thermal maturity, and biodegradation on the distribution and isomerization of homohopanes in petroleum. *Organic geochemistry*, 17(1):47–61, 1991.

F. Peterse, J.-H. Kim, S. Schouten, D. K. Kristensen, N. Koç, and J. S. S. Damsté. Constraints on the application of the MBT/CBT palaeothermometer at high latitude environments (Svalbard, Norway). *Organic Geochemistry*, 40(6):692–699, 2009.

C. A. Polik, F. J. Elling, and A. Pearson. Impacts of paleoecology on the TEX86 sea surface temperature proxy in the Pliocene–Pleistocene Mediterranean Sea. *Paleoceanography and Paleoclimatology*, 33(12):1472–1489, 2018.

T. J. Porter, T. Anhäuser, J. Halfar, F. Keppler, A. Z. Csank, and C. J. Williams. Canadian Arctic Neogene temperatures reconstructed from hydrogen isotopes of lignin-methoxy groups from sub-fossil wood. *Paleoceanography and Paleoclimatology*, 37 (2):e2021PA004345, 2022.

J. H. Pugsley, J. A. Howell, et al. Virtual field trips utilizing virtual outcrop: construction, delivery and implications for the future. *Geoscience Communication*, 5:227–249, 2022. doi: 10.5194/gc-5-227-2022.

J. H. Raberg, D. J. Harning, S. E. Crump, G. de Wet, A. Blumm, S. Kopf, Á. Geirsdóttir, G. H. Miller, and J. Sepúlveda. Revised fractional abundances and warm-season temperatures substantially improve brGDGT calibrations in lake sediments. *Biogeosciences*, 18(12):3579–3603, 2021.

J. H. Raberg, G. H. Miller, Á. Geirsdóttir, and J. Sepúlveda. Near-universal trends in brGDGT lipid distributions in nature. *Science Advances*, 8(20):eabm7625, 2022.

C. E. Rasmussen. Gaussian processes in machine learning. In *Summer school on machine learning*, pages 63–71. Springer, 2003.

B. Richter and J. Bojesen-Koefoed. Niðurstöður efnagreininga á þroska lífrænna efna í Tjörnessetlögunum. *Reykjavík: Orkustofnun OS-2002/004*, 2002.

B. Richter and K. Gunnarsson. Overview of hydrocarbon related research in Tjörnes. Technical Report íSOR-2010/007, Iceland GeoSurvey, 2010. Prepared for National Energy Authority.

S. T. Rögnvaldsson, Gudmundsson, and R. Slunga. Seismotectonic analysis of the Tjörnes Fracture Zone, an active transform fault in north Iceland. *Journal of Geophysical Research: Solid Earth*, 103:30117–30129, 1998. doi: 10.1029/98JB02789.

M. Rohmer, P. Bouvier-Nave, and G. Ourisson. Distribution of hopanoid triterpenes in prokaryotes. *Microbiology*, 130(5):1137–1150, 1984.

E. Rowley and N. White. Inverse modelling of extension and denudation in the East Irish Sea and surrounding areas. *Earth and Planetary Science Letters*, 161(1-4): 57–71, 1998.

D. Rumsfeld. News briefing, washington dc, usa, 2002.

K. Sæmundsson. Evolution of the axial rifting zone in northern Iceland and the Tjörnes fracture zone. *Geological Society of America Bulletin*, 85(4):495–504, 1974.

S. Schouten, E. C. Hopmans, E. Schefuß, and J. S. S. Damste. Distributional variations in marine crenarchaeotal membrane lipids: a new tool for reconstructing ancient sea water temperatures? *Earth and Planetary Science Letters*, 204(1-2):265–274, 2002.

S. Schouten, E. C. Hopmans, and J. S. S. Damsté. The effect of maturity and depositional redox conditions on archaeal tetraether lipid palaeothermometry. *Organic Geochemistry*, 35(5):567–571, 2004.

S. Schouten, A. Forster, F. E. Panoto, and J. S. S. Damsté. Towards calibration of the tex86 palaeothermometer for tropical sea surface temperatures in ancient greenhouse worlds. *Organic Geochemistry*, 38(9):1537–1546, 2007a.

S. Schouten, C. Huguet, E. C. Hopmans, M. V. Kienhuis, and J. S. Sinninghe Damsté. Analytical methodology for TEX86 paleothermometry by high-performance liquid chromatography/atmospheric pressure chemical ionization-mass spectrometry. *Analytical chemistry*, 79(7):2940–2944, 2007b.

S. Schouten, E. C. Hopmans, and J. S. S. Damsté. The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Organic Geochemistry*, 54:19–61, 2013.

O. Seki, J. A. Bendle, N. Harada, M. Kobayashi, K. Sawada, H. Moossen, G. N. Inglis, S. Nagao, and T. Sakamoto. Assessment and calibration of TEX86 paleothermometry in the Sea of Okhotsk and sub-polar North Pacific region: Implications for paleoceanography. *Progress in Oceanography*, 126:254–266, 2014.

G. Shaffer and J. Bendtsen. Role of the Bering Strait in controlling North Atlantic ocean circulation and climate. *Nature*, 367:354–357, 1994. doi: 10.1038/367354a0.

J. S. Sinninghe Damsté, W. I. C. Rijpstra, E. C. Hopmans, J. W. Weijers, B. U. Foesel, J. Overmann, and S. N. Dedysh. 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, 2011.

A. Sluijs, P. K. Bijl, S. Schouten, U. Röhl, G.-J. Reichart, and H. Brinkhuis. Southern ocean warming, sea level and hydrological change during the Paleocene-Eocene thermal maximum. *Climate of the Past*, 7(1):47–61, 2011.

C. Stickley, H. Brinkhuis, K. McGonigal, G. Chaproniere, M. Fuller, D. Kelly, D. Nürnberg, H. Pfuhl, S. Schellenberg, J. Schönfeld, et al. Late Cretaceous–Quaternary biomagnetostratigraphy of ODP sites 1168, 1170, 1171, and 1172, Tasmanian Gateway. In *Proceedings of the Ocean Drilling Program, Scientific Results*, volume 189, pages 1–57. Texas A&M University College Station, Texas, 2004.

F. Strauch. Zur geologie von Tjörnes: (Nordisland). *Sonderveröffentlichungen des Geologischen Instituts der Universität Köln*, 7, 1963.

F. Strauch. Phylogenese, Adaptation und Migration einiger nordischer marine Molluskengenera (Neptunea, Panomya, Cyrtodaria und Mya) (in German). *Abh. Senckenberg*, (531):211 pp, 1972.

L. Símonarson and J. Eiríksson. Tjörnes-Pliocene and Pleistocene sediments and faunas. *Jökull*, 58:331–342, 2008. doi: 10.33799/jokull2008.58.331.

L. A. Símonarson and J. Eiríksson. Systematic overview of the molluscs and barnacles of the Quaternary Breiðavík Group North Iceland. In J. Eiríksson and L. Símonarson, editors, *Pacific-Atlantic Mollusc Migration: Pliocene Inter-Ocean Gateway Archives on Tjörnes, North Iceland*, volume 52 of *Topics in Geobiology*, pages 667–755. Springer Nature, 2021.

K. W. Taylor, M. Huber, C. J. Hollis, M. T. Hernandez-Sanchez, and R. D. Pancost. Re-evaluating modern and Palaeogene GDGT distributions: Implications for SST reconstructions. *Global and Planetary Change*, 108:158–174, 2013.

J. E. Tierney and J. M. Russell. Distributions of branched GDGTs in a tropical lake system: implications for lacustrine application of the MBT/CBT paleoproxy. *Organic Geochemistry*, 40(9):1032–1036, 2009.

J. E. Tierney and M. P. Tingley. A bayesian, spatially-varying calibration model for the TEX86 proxy. *Geochimica et Cosmochimica Acta*, 127:83–106, 2014.

J. E. Tierney and M. P. Tingley. A TEX86 surface sediment database and extended Bayesian calibration. *Scientific data*, 2(1):1–10, 2015.

J. E. Tierney, J. M. Russell, H. Eggermont, E. Hopmans, D. Verschuren, and J. S. Damsté. Environmental controls on branched tetraether lipid distributions in tropical East African lake sediments. *Geochimica et Cosmochimica Acta*, 74(17):4902–4918, 2010.

J. R. Toggweiler and R. M. Key. Thermohaline circulation. In *Encyclopedia of Ocean Sciences*, volume 6, pages 2941–2945. 2001. doi: 10.1006/rwos.2001.0111.

G. Trommer, M. Siccha, M. T. van der Meer, S. Schouten, J. S. S. Damsté, H. Schulz, C. Hemleben, and M. Kucera. Distribution of Crenarchaeota tetraether membrane lipids in surface sediments from the Red Sea. *Organic Geochemistry*, 40(6):724–731, 2009.

C. Turich, S. Schouten, R. C. Thunell, R. Varela, Y. Astor, and S. G. Wakeham. Comparison of TEX86 and U37K temperature proxies in sinking particles in the Cariaco Basin. *Deep Sea Research Part I: Oceanographic Research Papers*, 78:115–133, 2013.

L. G. van Bree, F. Peterse, A. J. Baxter, W. De Crop, S. Van Grinsven, L. Villanueva, D. Verschuren, and J. S. Sinninghe Damsté. Seasonal variability and sources of in situ brGDGT production in a permanently stratified African crater lake. *Biogeosciences Discussions*, 2020:1–36, 2020.

P. Véquaud, A. Thibault, S. Derenne, C. Anquetil, S. Collin, S. Contreras, A. T. Nottingham, P. Sabatier, J. P. Werne, and A. Huguet. FROG: A global machine-learning temperature calibration for branched GDGTs in soils and peats. *Geochimica et Cosmochimica Acta*, 318:468–494, 2022.

K. Verhoeven. *A reconstruction of the depositional history and palaeoenvironment of the Plio-Pleistocene of northeast Iceland*. PhD thesis, Ghent University, 2012.

K. Verhoeven and S. Louwye. Palaeoenvironmental reconstruction and biostratigraphy with marine palynomorphs of the plio–pleistocene in Tjörnes, Northern Iceland. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 376:224–243, 2013. doi: 10.1016/j.palaeo.2013.03.002.

K. Verhoeven, S. Louwye, J. Eiríksson, and S. De Schepper. A new age model for the Pliocene–Pleistocene Tjörnes section on iceland: its implication for the timing of North Atlantic–Pacific palaeoceanographic pathways. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 309:33–52, 2011. doi: 10.1016/j.palaeo.2011.04.001.

J. K. Volkman, S. M. Barrett, S. I. Blackburn, M. P. Mansour, E. L. Sikes, and F. Gelin. Microalgal biomarkers: a review of recent research developments. *Organic Geochemistry*, 29(5-7):1163–1179, 1998.

Y. Weber, J. S. Sinninghe Damsté, J. Zopfi, C. De Jonge, A. Gilli, C. J. Schubert, F. Lepori, M. F. Lehmann, and H. Niemann. Redox-dependent niche differentiation provides evidence for multiple bacterial sources of glycerol tetraether lipids in lakes. *Proceedings of the National Academy of Sciences*, 115(43):10926–10931, 2018.

J. W. Weijers, S. Schouten, E. C. Hopmans, J. A. Geenevasen, O. R. David, J. M. Coleman, R. D. Pancost, and J. S. Sinninghe Damsté. Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits. *Environmental Microbiology*, 8(4):648–657, 2006a.

J. W. Weijers, S. Schouten, O. C. Spaargaren, and J. S. S. Damsté. Occurrence and distribution of tetraether membrane lipids in soils: Implications for the use of the TEX86 proxy and the BIT index. *Organic Geochemistry*, 37(12):1680–1693, 2006b.

J. W. Weijers, S. Schouten, J. C. van den Donker, E. C. Hopmans, and J. S. S. Damsté. Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochimica et Cosmochimica Acta*, 71(3):703–713, 2007.

J. W. Weijers, P. Steinmann, E. C. Hopmans, S. Schouten, and J. S. S. Damsté. Bacterial tetraether membrane lipids in peat and coal: Testing the MBT–CBT temperature proxy for climate reconstruction. *Organic Geochemistry*, 42(5):477–486, 2011.

T. Westerhold, N. Marwan, A. J. Drury, D. Liebrand, C. Agnini, E. Anagnostou, J. S. Barnet, S. M. Bohaty, D. De Vleeschouwer, F. Florindo, et al. An astronomically dated record of Earth's climate and its predictability over the last 66 million years. *Science*, 369(6509):1383–1387, 2020.

R. A. Woodgate and K. Aagaard. Revising the Bering Strait freshwater flux into the Arctic Ocean. *Geophysical Research Letters*, 32:L02602, 2005. doi: 10.1029/2004GL021747.

C. Wuchter, S. Schouten, M. J. Coolen, and J. S. Sinninghe Damsté. Temperature-dependent variation in the distribution of tetraether membrane lipids of marine Crenarchaeota: Implications for TEX86 paleothermometry. *Paleoceanography*, 19(4), 2004.

J. Zachos, M. Pagani, L. Sloan, E. Thomas, and K. Billups. Trends, rhythms, and aberrations in global climate 65 Ma to present. *science*, 292(5517):686–693, 2001.

Z. Zeng, H. Chen, H. Yang, Y. Chen, W. Yang, X. Feng, H. Pei, and P. V. Welander. Identification of a protein responsible for the synthesis of archaeal membrane-spanning GDGT lipids. *Nature communications*, 13(1):1545, 2022.

Y. G. Zhang, C. L. Zhang, X.-L. Liu, L. Li, K.-U. Hinrichs, and J. E. Noakes. Methane Index: A tetraether archaeal lipid biomarker indicator for detecting the instability of marine gas hydrates. *Earth and Planetary Science Letters*, 307(3-4):525–534, 2011.

Y. G. Zhang, M. Pagani, and Z. Wang. Ring Index: A new strategy to evaluate the integrity of TEX86 paleothermometry. *Paleoceanography*, 31(2):220–232, 2016.

# A | Appendix A - Methodologies and Supporting Evidence

## A.1 Barmur Group Drone Photogrammetry Methodology

### A.1.1 *Drone Selection and Survey Design*

The information provided here is also available in the supplementary of Allison et al. (2026) and aims to outline, in detail, the methodology of how the Barmur Group virtual outcrop, SfM (Structure from Motion) model was generated (Allison et al., 2023). A DJI Mavic Pro 2 was used to collect aerial photography, with the onboard camera providing a 20 MP image from a 1 inch sensor. Unusually for a drone camera, it has aperture control between f/2.8–f/11, which provides ample flexibility in capturing images of different rock types in different light conditions. The camera is stabilized on a 3-axis gimbal, providing manual tilt control between –90$^o$ and 30$^o$ (from vertically down to 30$^o$ above the horizontal). The remote controller has a maximum transmission range of 6 km (unobstructed line of sight and no interference) and the drone is rated for flight in up to 38 km/hr winds. The drone has a surprisingly small form factor, making it an good candidate for remote geological fieldwork.

An iPad was used to control the drone and initiate pre-prepared surveys using DJI's Ground Station Pro (GS Pro) software. The automated surveys generated a substan-


Figure A.1: Real world drone performance. Data representative of 60 flights. RTH, Return To Home, an automated process by which the drone returns to its take-off location when on low battery.

tial number of images, so high capacity micro-SD cards were required as well as an external hard drive to allow for archiving and backup of data. SfM 3D models can be georeferenced by using either the onboard drone GPS or by using Ground Control Points (GCP). The Mavic Pro 2 has been shown to produce SfM models with an absolute positioning error of c. 3 m without the use of GCPs (Kalacska et al., 2020), deemed sufficient for this study, especially considering the large model area (7 km of coastline). Data was collected on the drone itself and used to provide real-world performance figures for the Mavic Pro 2, Fig. A.1. Whilst DJI publish performance figures, these are often representative of perfect flight conditions which in our experience over-promise flight performance by as much as a third. It was found that c. 5 hectares of surveying per hour was a realistic target over the course of this study.

The survey area (coastal exposure of the Barmur Group) was divided into smaller zones to ensure the drone was kept close to the operator and within line of sight at all times. The zones were designed in GS Pro to result in a survey flight time of c. 30 min-

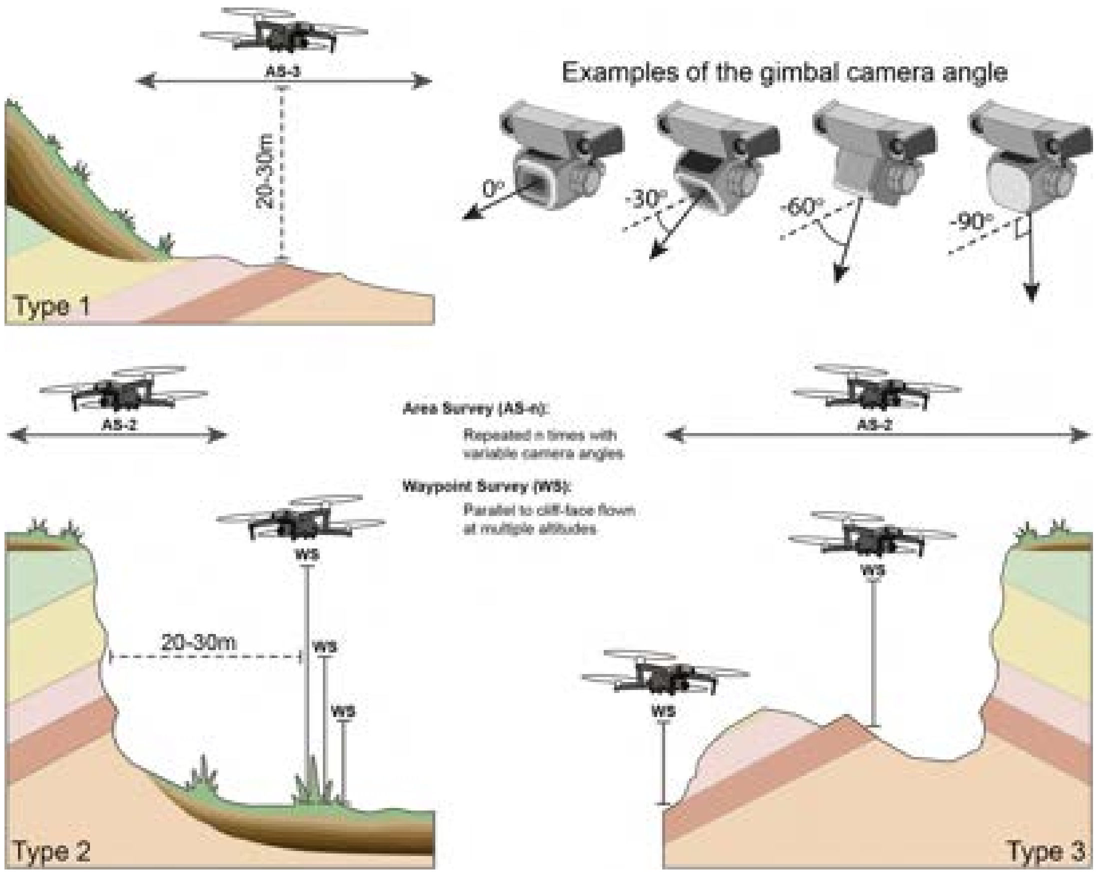


Figure A.2: 3 categories of outcrop and the proposed survey designs to best capture photogrammetric data. Type 1: Mostly horizontally exposed outcrop. Type 2: Mostly vertically exposed outcrop. Type 3: A mixture of Type's 1 and 2.

utes. 30 minutes was chosen as it represented roughly the flight time of a single drone battery. Two types of survey were planned for this study, termed here as "Area Surveys" (AS) and "Waypoint Surveys" (WS). These two surveys were designed to compliment each other to ensure that all topographical features were imaged thoroughly. Three, broad categories of topography were outlined and the appropriate survey type deployed to image them, Fig. A.2. Type 1: Mostly horizontal exposure, best captured by three overlapping Area Surveys utilising variable camera gimbal angles. Type 2: Mostly vertical exposure, best captured by two overlapping Area Surveys utilising variable camera gimbal angles and three waypoint surveys of variable altitudes. Type 3: A mixture of Type 1 and Type 2 exposure, best captured by two overlapping Area Surveys utilising variable camera gimbal angles and two non-overlapping waypoint surveys.

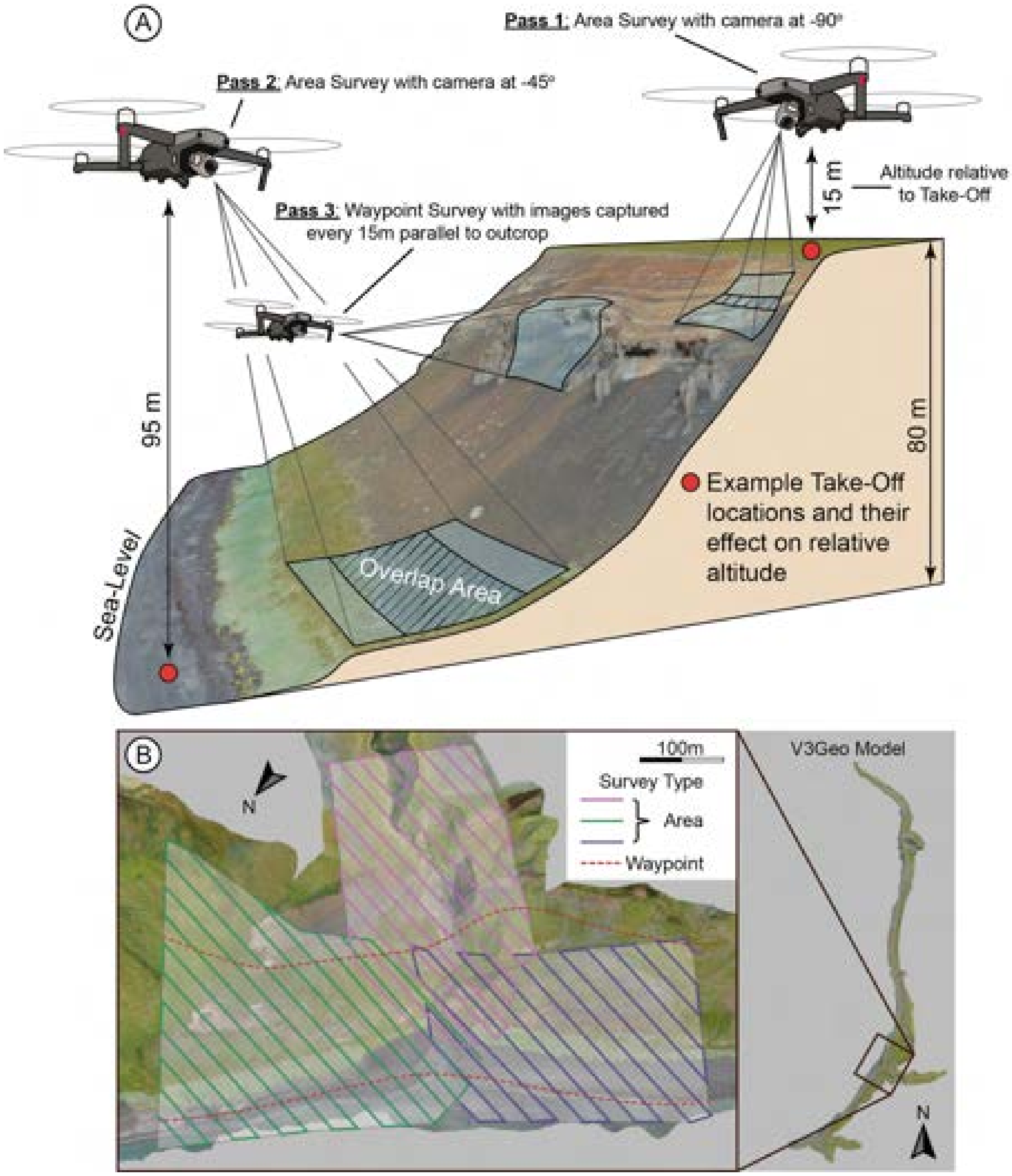


Figure A.3: a) Visual summary of data collection in the field. Highlighted is the variation in both image size and image overlap due to variable topography within the survey area and the dependent nature of relative altitude to the take-off location. b) An example of the survey design for a portion of the Barmur Group V3Geo model, highlighting overlapping Area Surveys and Waypoint Surveys to target prominent exposures

Type 3 exposure dominated the survey area and design for the majority of the imaged underlying the V3Geo model of the Barmur Group (15710 out of 17123 images, (Allison et al., 2023)). The target of this model was a 7 km stretch of coastal cliffs aligned roughly north – south. The geology crops out near the top of 40–70 m tall cliffs with the base of the cliffs often covered in scree and vegetation. In general, the main target of a given photograph was imaged from a distance of 20–30 m. This photographic distance resulted in visible centimetre-scale details across the entire section, sufficient overlap of neighbouring photographs for photogrammetric reconstruction and survey times small enough to image the entire Barmur Group within the time constraints of the field campaign. Areas rich in detail or consisting of complex geometry were imaged manually at distance c. 10 m.

A quirk of the Mavic Pro 2 is that when it takes-off, the altitude of the drone is set relative to this take-off location. Safe survey design required absolute topographical knowledge, and so for consistent survey altitudes to be flown the relative altitude of the drone needed correcting on a flight-by-flight basis. Should different survey zones be flown at variable absolute altitudes the resulting model fidelity would vary accordingly, the image sensor on the drone camera imaging less detail at greater heights. Survey zones were designed to overlap to allow for the eventual stichting together of all zones into one coherent model. Survey overlap served a secondary process in allowing for areas of prominent exposure to be captured by multiple surveys and a greater number of images, resulting higher quality SfM mode. An schematic example of a survey is shown in Figure A.3 and the actual survey parameters used in GSPro are shown in Figure A.4.

### A.1.2 *Camera Settings for SfM Models*

The camera provides the main means of data collection for SfM surveys and as such needs to be configured in an idealised manner. The desired outcome is one where

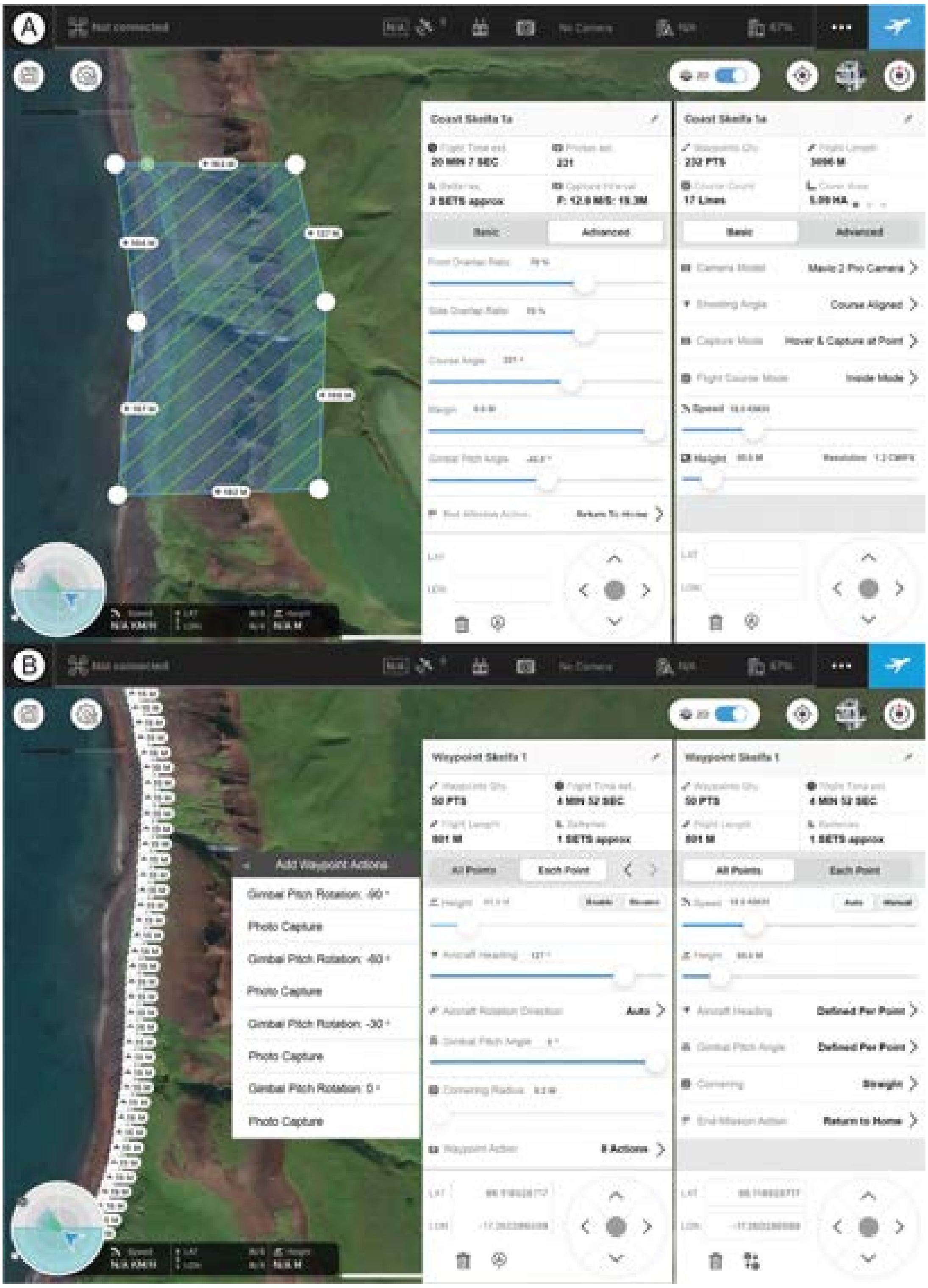


Figure A.4: Screenshots of DJI's Ground Station Pro software for autonomous survey planning. a) Area Survey b) Waypoint Survey with nodes spaced every 15 m. Key variables for both survey types are shown as they would be in the software.

as much of the image is in focus as possible, motion blur is minimised, exposure correctly balanced and the white balance consistent shot-to-shot. A digital camera offers a range of settings to control the above parameters, however they operate in a somewhat coupled manner, as outlined below. Digital cameras are primarily controlled by 4 settings. Aperture, shutter speed (exposure time), ISO (film speed) and white balance. The aperture represents the iris of the camera lens and dictates how much light enters the camera. It also controls the depth of field of an image, such that an open aperture (small f-number) sees greater background blur (low depth of field) and a softer image than a closed aperture (larger f-number), which produces an image that is sharper and more widely in focus (high depth of field). Shutter speed represents how long the sensor of the camera will gather light for. The greater (slower) the shutter speed, the brighter the image and the more motion blur that will be captured by the sensor. ISO represents how sensitive the sensor of the camera is to light. The higher the ISO the brighter and more noisy the image. White balance instructs a digital camera as to the qualities of the colour white in the imaged scene. White will look different under direct sunlight when compared to the same white under a fluorescent lightbulb. White balance can drastically change the colours of an image.

To create a sharp, mostly in focus image requires as small an aperture as possible (large f-number). To capture minimal motion blur the shutter speed needs to be as short (fast) as possible and to minimise noise it is desirable to keep the ISO as low as possible. However, all three of the end-member settings outlined above will result in very little light reaching the sensor and the image can end up under exposed. All images represent a compromise between the three settings outlined above.

The advantage of GS Pro and automated surveying, is that the drone can repeatedly stop and hover to take each image. Once hovering, the gimbal allows for a very steady camera platform, minimising motion blur and allowing for slower shutter speeds than for a drone in motion. Typical shutter speeds were found to be between 1/50–1/150 second. The slower shutter speeds without having to compromise on motion blur, allowed for the use of smaller apertures (greater f-numbers) resulting in sharper, more in focus images. In the same manner, the brighter images afforded by slower shutter speeds allowed for ISO values typically below 400, aiding in noise reduction. White balance should be manually adjusted to suit the daily changes in sunlight and weather. This will ensure colours are as consistent as possible across multiple images.

Weather and light play an important role in aerial photography. The ideal conditions results from an overcast day where the light would be described as ”flat”. Harsh, direct sunlight results in overexposed highlights, glare and severe shadows. The orientation of the exposure will also influence what time of day provides optimal lighting, for example a westward facing cliff section should not be surveyed in the morning as the camera will quickly become overwhelmed by the lighting conditions of having to image a target in shadow whilst still receiving direct sunlight on the camera sensor.

A final choice of image file can be made. Drones will typically save images in JPEG format, but it is possible to shoot in RAW. RAW offers greater flexibility when editing images. However, RAW produces much bigger file sizes and will need editing and exporting to a JPEG before use in SfM software. The survey of the Barmur Group produced c.17,000 JPEG images, the file size of each image ranged between c.7–16 MB, see Fig. A.1. It was not considered worthwhile to edit every RAW image for the small gains in image quality it would provide. Making careful adjustments of camera settings throughout the survey resulted in clear and high quality JPEG images, allowing for a straightforward import of the data into post-processing software, such as Agisoft Photoscan.

### A.1.3 *SfM Data Processing*

The workflow presented here is conducted using Agisoft Metashape Professional with the end goal of uploading the final model to V3Geo for hosting, as per Allison et al. (2023). Therefore, parts of the workflow are designed specifically to cater to the requirements of uploading to V3Geo's servers. Full details of V3Geo's submissions guidelines can be found at the following address: `https://v3geo.com/contribute`. If the overall survey area is large enough, it will be necessary to split the total model into small zones, each processed individually. Below is a step-by-step guide to creating a 3D photogrammetric model.

#### *A.1.3.1 Align Photos*

The first step with any Agisoft project is to load the photos into the project and align them. Agisoft guides the user through the software through the ”Workflow” menu. The first stage in the creation of a SfM model is to create a point cloud, representing a set of data points typically located within a Cartesian space. All survey images will be aligned in one step to produce a low density point cloud of the total survey area. The low density point cloud (and the associated images) can now be split into smaller regions, termed here as ”sub-zones”, each processed individually. Each sub-zones should be stored within their own ”chunk” in Agisoft. The reason for creating sub-zones is to ensure that memory requirements are not exceeded whilst processing within Agisoft and to match any eventual upload limits provided by the eventual destination/host of the model.

The low density point cloud produced in the alignment stage should bear a fuzzy resemblance to the entire surveyed area. It is important to align the entire dataset before the creation of sub-zones to ensure accurate meshing of the sub-zones following the creation of the final textured models. The low density point cloud should now to be sampled into the sub-zones by duplicating the entire dataset and deleting all cameras

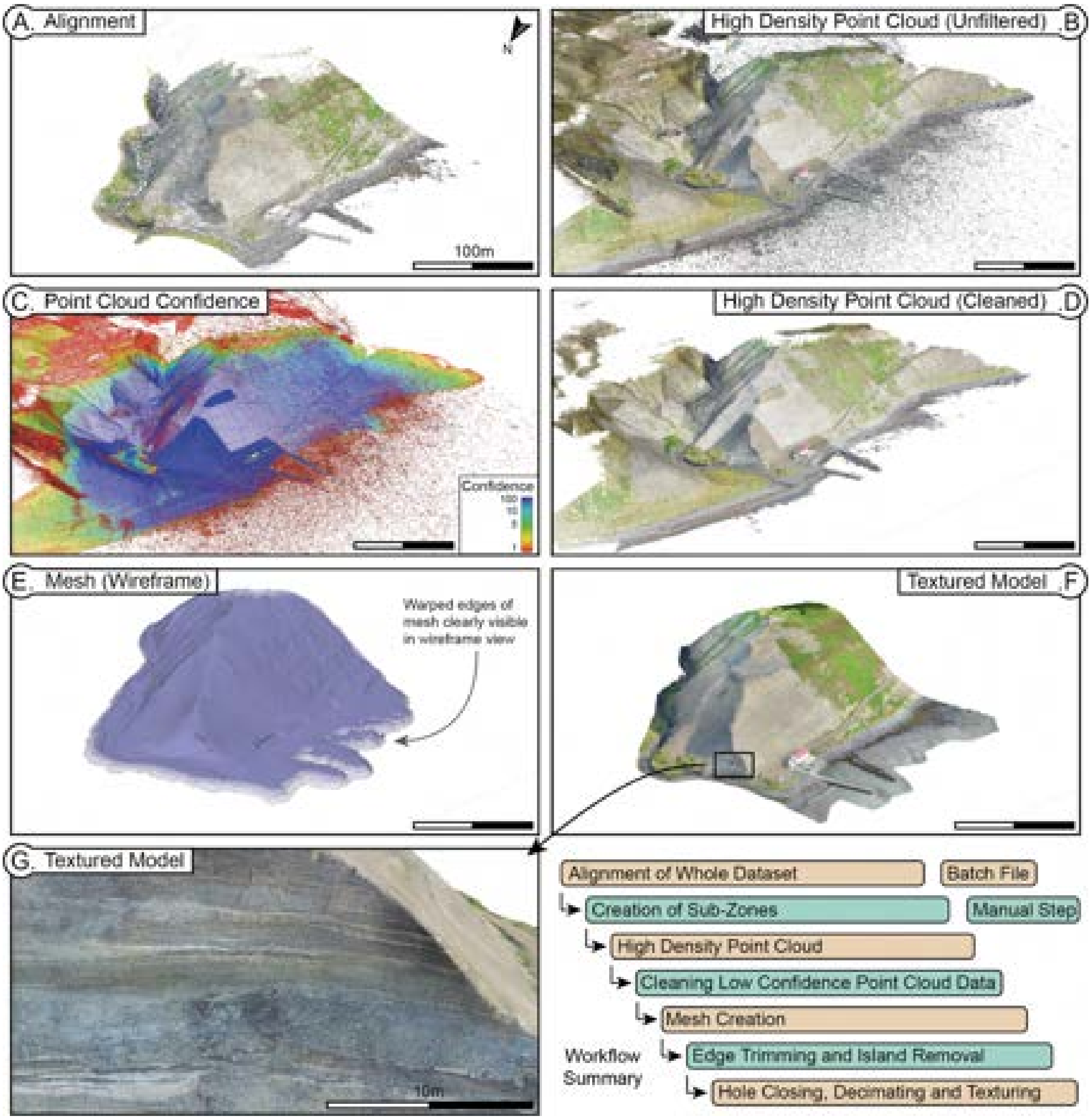


Figure A.5: Proposed workflow within Agisoft Metashape for a single sub-zone of the aligned model space. a) Alignment stage and low density point cloud b) A high density point cloud is constructed and c) low confidence points are d) cleaned (shown: cleaned below confidence = 2). e) The resulting mesh is then trimmed to removed warped edges and mesh islands are identified and removed. f) Holes in the mesh are closed, the mesh is decimated and textures applied for a finished model. g) a close-up of the final textured model highlighting the sedimentary structures visible in the model.

and point cloud data outside of the respective sub-zones, Fig. A.5a. A small amount of margin/overlap should be left between sub-zones, this is important to create clean edges for the eventual textured model mesh. Each of sub-zones will now be processed further.

#### *A.1.3.2 High Density Point Cloud*

The next step in the Agisoft workflow is to generate a high density point cloud for each sub-zone, Fig. A.5b. When building this dense point cloud, ensure that confidence values are calculated for each point. Once created, switch to a view of the point cloud where each point is coloured by its confidence value, Fig. A.5c. The actual exposure should be visible underneath a haze of low confidence noise. By using Agisoft's filtering options (Tools > Dense Point Cloud > Filter by Confidence), it is possible to view and delete these low confidence points, revealing the target exposure, Fig. A.5d. Experience revealed that points below a confidence of 2 offer very little to the end product. Filtering and deleting points between a confidence of 0–2 removes most of the noise produced when generating the high density point cloud. Points of confidence between 2–5 can often be deleted for gains in eventual textural quality. However, in areas of complicated geometry (e.g. craggy cliffs or overhangs) confidence will be naturally lower, and assessment may be required when deleting points between 2–5 confidence to determine if they represent useful information or noise. The fidelity of the model texture is also linked to the size of the model being processed, so breaking the sub-zones into smaller areas will increase the detail of the final product. As a reference, sub-zones for the Barmur Group model were composed of high density point clouds of c. 1,000,000 points.

#### *A.1.3.3 Mesh*

Once the cleaned, high density point clouds have been created, the clouds need meshing into a surface. The edges of the resulting mesh can become warped due to a lack of data at the edge of each sub-zone, Fig. A.5e. If appropriate overlap has been left between neighbouring sub-zones, then this warping can be trimmed to leave a clean surface which will mesh indistinguishably with adjacent sub-zones. This warping is particularly visible if the mesh is viewed as a wireframe, whereby the density of the frame drastically reduces as it approaches the warped edges.

Some part of the mesh may form ”islands”. These are floating and undesirable parts of the model created when the underlying point cloud data is inaccurate or too noisy to resolve the true geometry. These islands can be removed using Agisoft's Gradual Selection tool (Model > Gradual Selection > Connected Component Size). This tool allows for selection of the mesh based on a finite number of connected parts, making islands very easy to select and delete.

Once the mesh is trimmed, any holes in the mesh will need closing, this can be done with Agisoft's ”Close Hole's” feature. Now the mesh is trimmed and free of holes, it is decimating to achieve the desired file size of the upload destination. For example, V3Geo requests individual uploads do not exceed 300 MB. The mesh represents a series of triangle, or faces, which approximate the geometry of the modelled landscape. Higher face-counts result in higher fidelity models but at the cost of larger file sizes. The mesh can be decimated, to reduce the face-count and subsequent file size. It was found that decimating each sub-zone to a face count of 3,000,000 kept each sub-zone below V3Geo's 300 MB upload limit.

#### *A.1.3.4 Texture*

The final stage is to apply textures from the drone imagery to the blank mesh, Fig. A.5f. At this stage it is useful to explore all points of the models and check that the quality of the texturing is as desired. Depending on the destination of the model, there may be limits on the textural settings that can be applied. It is also sensible to combine each zone with its neighbouring zones and assess the quality of the boundary mesh, if done correctly, the seam should be invisible. Before export, it is important to check if the models need translating (x,y,z) relative to some known datum (e.g. sea-level) which is not quite at the correct altitude. This can be adjusted in the model export options. Where possible processing steps utilised batch files to standardise processing settings across sub-zones. Batch processing files are available through the supplementary of Allison et al. (2026) or via the following Github repository: `https://github.com/matthew-allison05/Agisoft_Metashape_SfM_Batchfiles`.

## A.2 Global isoGDGT Compilations and Updated OPTiMAL Code

The original OPTiMAL code repository and training data is available at the following repository: `https://doi.org/10.5281/zenodo.4293851` (Greene and Mandel, 2020). The data and code accompanying Chapter 3 of this thesis will form the supplementary of an imminently submitted science article. These data files will be made available through the following GitHub account: `https://github.com/matthew-allison05?tab=repositories`. The sampling sites of data compiled for this study are visualised in Figure A.6.

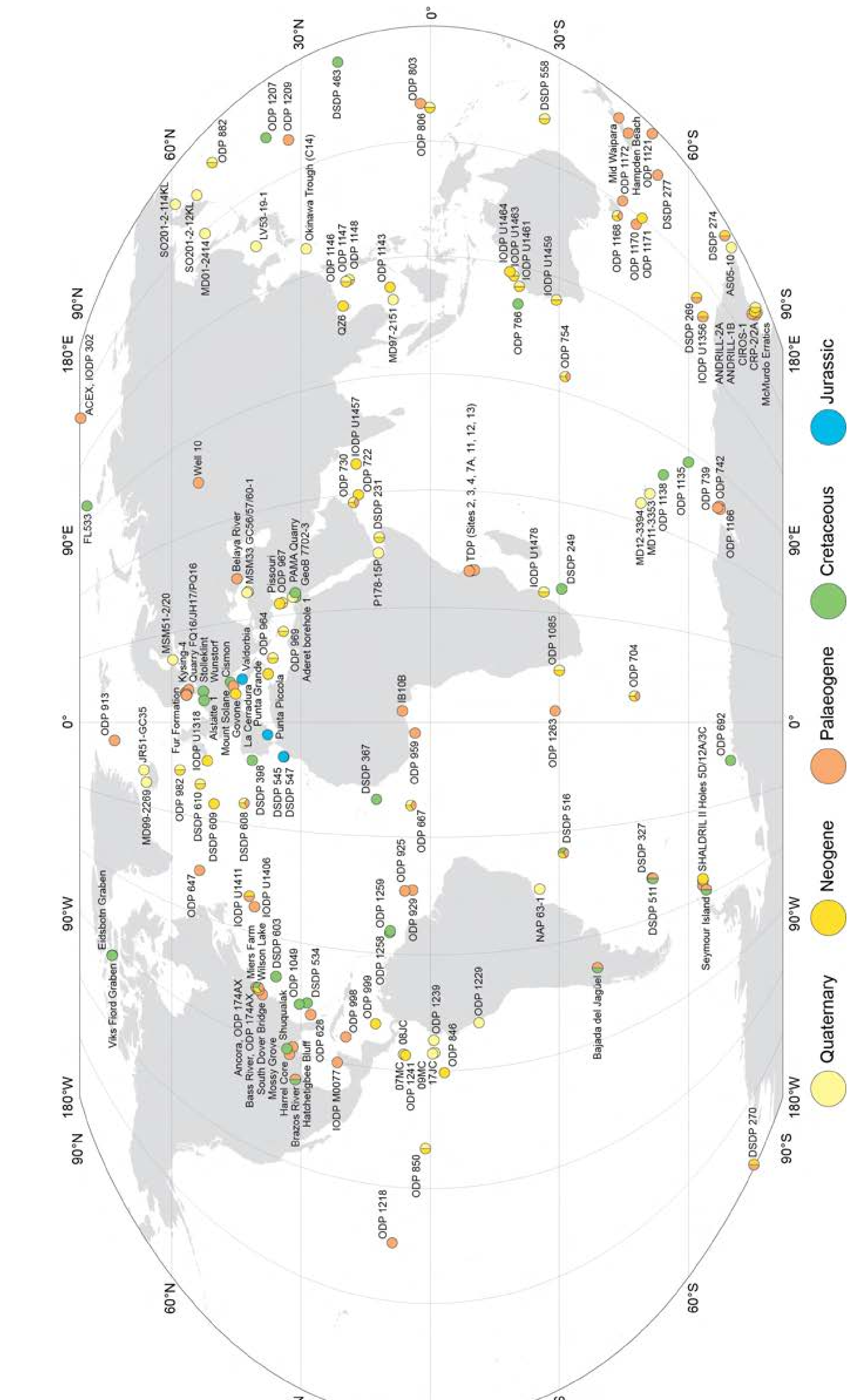


Figure A.6: Location of sampling sites for isoGDGT data compiled for Chapter 2

## A.3 Barmur Group Lignite Temperature Record

### A.3.1 *Organic Extraction Methodology*

Homogenising lignite samples

1. Homogenisation of freeze-dried lignite samples took place in a solvent-cleaned TEMA disc mill. Disc mill to be cleaned in between samples to limit cross-contamination.
2. Homogenised, powdered lignite samples to be stored in pre-weighed, furnace foil envelopes. Total dry lignite mass to be recorded

Organic Extraction - Total Lipid Extract

Ensure all solvents used in the process of organic extractions at High Performance Liquid Chromatography (HPLC) grade, and all glassware is clean and pre-furnaced.

1. Transfer 5 g of dry, homogenised lignite powder into 30 ml test tubes.
2. Add 20 ml of dichloromethan:methanol (DCM:MeoH) (3:1 v/v) and cap the test tube.
3. Mix the solution using a vortex mixer.
4. Leave the mixed solution on a heating block at 40 $^{\circ}C$ for 1 hour.
5. Centrifuge the heated samples for 3 minutes at 3000 RPM to separate the solution into the solid lignite powder and the Total Lipid Extract (TLE).
6. Pipette the TLE supernatant into a fresh test tube.
7. Repeat steps 2–6 at least three times or until the supernatant is clear.
8. Evaporate the TLE, either in a fume hood or using accelerated evaporation equipment.

9. Transfer the dry TLE to a fresh, pre-weighed 7 ml glass vial using 4 ml of DCM:MeOH (3:1).
10. Evaporate the solvent in the GC vial and weigh the solid TLE recovered from the sample.
11. Redissolve the TLE using DCM:MeOH (95:5 v/v) and split the solution into 2 aliquots, one for archiving purposes and the other for silica gel column chromatography. Ensure the TLE for the column does not exceed 10 mg.

Column Chromatography

Column chromatography is be conducted using disposable, glass Pasteur pipette filled with a small ball of glass wool, 2.5 cm of 99% deactivated silicon powder, and capped with c. 0.5 cm of sand. All contents of the column are to be pre-furnaced. Four fractions (N1 – N4) are to be collected from these columns. Four clean glass vials are to be prepared for each fraction in advance.

1. Pass 1 ml of *n*-hexane to the column and collect in a waste vial.
2. Transfer the TLE from the GC vial onto the top of the column. This can be done with either 1 ml of *n*-Hexane or by absorbing the TLE to a small amount of silica powder and transferring the dry silica to the top of the column.
3. N1 - aliphatic hydrocarbons: Pass 4 ml of *n*-hexane through the column and collect the N1 fraction.
4. N2 - aromatic hydrocarbons: Pass 2 ml of *n*-hexane:DCM (2:1 v/v) through the column and collect the N2 fraction.
5. N3 - aldehydes and ketones: Pass 4 ml of DCM through the column and collect the N3 fraction.
6. N4 - polar fraction: Pass 5 ml of DCM:MeOH (95:5 v/v) through the column and collect the N4 fraction.

7. Once all fractions are collected and solvents evaporated, transfer the contents to pre-weighed and labelled GC vials using 1 ml of the respective solvents.

N1–N3 were analysed at the University of Birmingham using GC-MS and GC-FID for compound identifications and quantification.

N4 Preparation for HPLC-MS Analysis

All equipment in this stage should be disposable or thoroughly washed in between contact with samples using *n*-Hexane and DCM.

1. Using a clean Hamilton syringe, add 200 µl of *n*-hexane:*i*-propanol (99:1 v/v) to re-dissolve the N4 fraction.
2. Draw back the syringe and collect the sample.
3. Carefully remove the needle whilst holding the sample in the syringe. Add a 0.45 µm PFTE syringe filter. Replace the needle and gentle pass the sample through the filter into a new GC vial containing a glass insert.
4. Dry down/evaporate the sample into the GC insert vial and cover with a septa screw cap.

N4 fractions, containing GDGTs, were analysed and integrated using HPLC-MS at the University of Bristol.

### A.3.2 *Burial Model Stratigraphy and Constants*

Table A.2 details the stratigraphic information used to generate the burial model of Chapter 3. For calculations of vitrinite reflectance, this model uses the physical constants constrained from inverse modelling from extensional basins (Rowley and White, 1998), these constants are: Frequency factor - $A_E$ = 4.00x10$^{10}$ [s$^-$1], Activation energy - $E_a$ = 242 [kJmol$^{-1}$], Standard deviation of activation energies - $\sigma$ = 4 [kJmol$^{-1}$]. Standard values are used for the physical properties of modelled layers. These properties are outlined in Table A.1.

| Layer type | Solid density [kgm$^{-3}$] | Initial porosity [%] | Thermal conductivity [Wm$^{-1}$K$^{-1}$] |
|---|---|---|---|
| Shale | 2.70x10$^3$ | 0.6 | 2.5 |
| Sandstone | 2.65x10$^3$ | 0.5 | 3.7 |
| Basalt | 3.00x10$^3$ | 0.1 | 2.3 |

Table A.1: Properties of density, porosity and thermal conductivity assumed in the burial model.

| Stratigraphy | Age [Ma] | Depth [m] | Shlale [%] | Sandstone [%] | Basalt [%] | Strat. Model |
|---|---|---|---|---|---|---|
| | Model Information | | | | | |
| Uplift Event | 0.0 | -760 | | | | |
| Miðlækur Mbr. – Bangastaðir Mbr. | 0.77 | 1 | | 70 | 30 | All |
| Svarthamar Mbr. – Búrfellsá Mbr. | 1.78 | 160 | | 30 | 70 | All |
| Hörgi Fm. – Fossgil Mbr. | 2.14 | 225 | | 100 | | All |
| Hiatus Model 3 | 3.00 | 226 | | 100 | | All |
| Furuvík Fm. | 2.60 | 310 | | | 100 | 1–2 |
| Hiatus Model 2 | 4.20 | 311 | | | | 2 |
| Furuvík Fm. | 3.55 | 310 | | 50 | 50 | 3 |
| Höskuldsvík Group | 4.49 | 430 | | | 100 | All |
| Stapi Mbr. | 4.56 | 441 | | 100 | | All |
| Bed J Lignites | 4.63 | 499 | 100 | | | All |
| Ytri Svarthamar Mbr. – Nöf Mbr. | 5.01 | 622 | | 100 | | All |
| Bed 10 Lignites | 5.02 | 623 | 100 | | | All |
| Brunngil Mbr. | 5.15 | 668 | | 100 | | All |
| Bed F Lignites | 5.18 | 678 | 100 | | | All |
| Hvilft Mbr. – Skeifárbás Mbr. | 5.56 | 724 | | 100 | | All |
| Bed D Lignites | 5.70 | 728 | 100 | | | All |
| Flekatá Mbr. | 5.75 | 753 | | 100 | | All |
| Bed C Lignites | 6.01 | 763 | 100 | | | All |
| Skonsubakki Mbr. – Rekárfoss Mbr. | 6.61 | 808 | | 100 | | All |
| Bed B Lignites | 6.80 | 811 | 100 | | | All |
| Hústorfa Mbr. – Háibakki Mbr. | 7.29 | 858 | | 100 | | All |
| Bed d/e Lignites | 7.31 | 860 | 100 | | | All |
| Tungugerðiskambar Mbr. | 7.50 | 872 | | 100 | | All |
| Kaldakvísl Group | 9.3 | 1350 | | | 100 | All |

Table A.2: Burial model stratigraphy table. Stratigraphy information represents the packages of sediment or lava used in the this model spanning. Age and depth are correlated with the basal age and depth of the corresponding stratigraphic package. Percentage shale, sandstone, basalt is used to characterise the physical properties of a package (e.g. heat conductivity or sediment compaction. The discrepancies between the three stratigraphic models (Strat. Model) are highlighted. Final uplift event is shown with a negative depth.

### A.3.3 *Gaussian Process Investigations - Supplementary*

In the original formulation of OPTiMAL (Dunkley Jones et al., 2020) $D_{Nearest} \leq 0.5$ was used to restrict the usage of low confidence data (i.e. large predictive standard deviation, $\sigma$). When $\sigma$ is plotted against the values of $D_{Nearest}$, data can be modelled adequately by a logistic curve. This curve has some qualitative properties associated with the training dataset, for example the upper and lower limits of this curve define the standard deviation and and lower noise threshold of the training dataset respectively. Fitting a logistics curve ($f(x)$) to the OPTiMAL predictions on isoGDGTs it is observed that the somewhat arbitrary cut-off value of $D_{Nearest}$ of $0.5$ results in $f(x) = 0.15$, that is 15% of the travel between the lower and upper bounds. Using this information, bespoke logistics curves are fit and $D_{Nearest}$ values generated for each GPE-Model 3 (soil, peat, lacustrine, and marine type training data), Fig. A.7.

A positive correlation is observed between the number of recovered brGDGT isomers and stratigraphic height, Fig. A.8. This relationship implies that the diversity of brGDGTs decreases with increasing burial, i.e. certain isomers are more thermally stable than others.

Weak to moderate correlations are observed between GPE $D_N earest$ values and GPC classification confidence, Fig. A.9. In lacustrine samples this relationship implies classification is more confident when the ancient data is represented well in the modern training data (low $D_N earest$ values).

Weak to moderate, linear correlations are observed between the average value of $CD_{Median}$ and TOM per lignite horizon, Fig. A.10. This relationship is observed most weakly in lacustrine type samples and most strongly in soil and peat type samples. The relationship could imply that the $CD_{Median}$ metric is contributed to more strongly by the TOM of the sediment. This would link the temperature sensitivity of brGDGTs in certain depositional environments to the TOM of that environment.

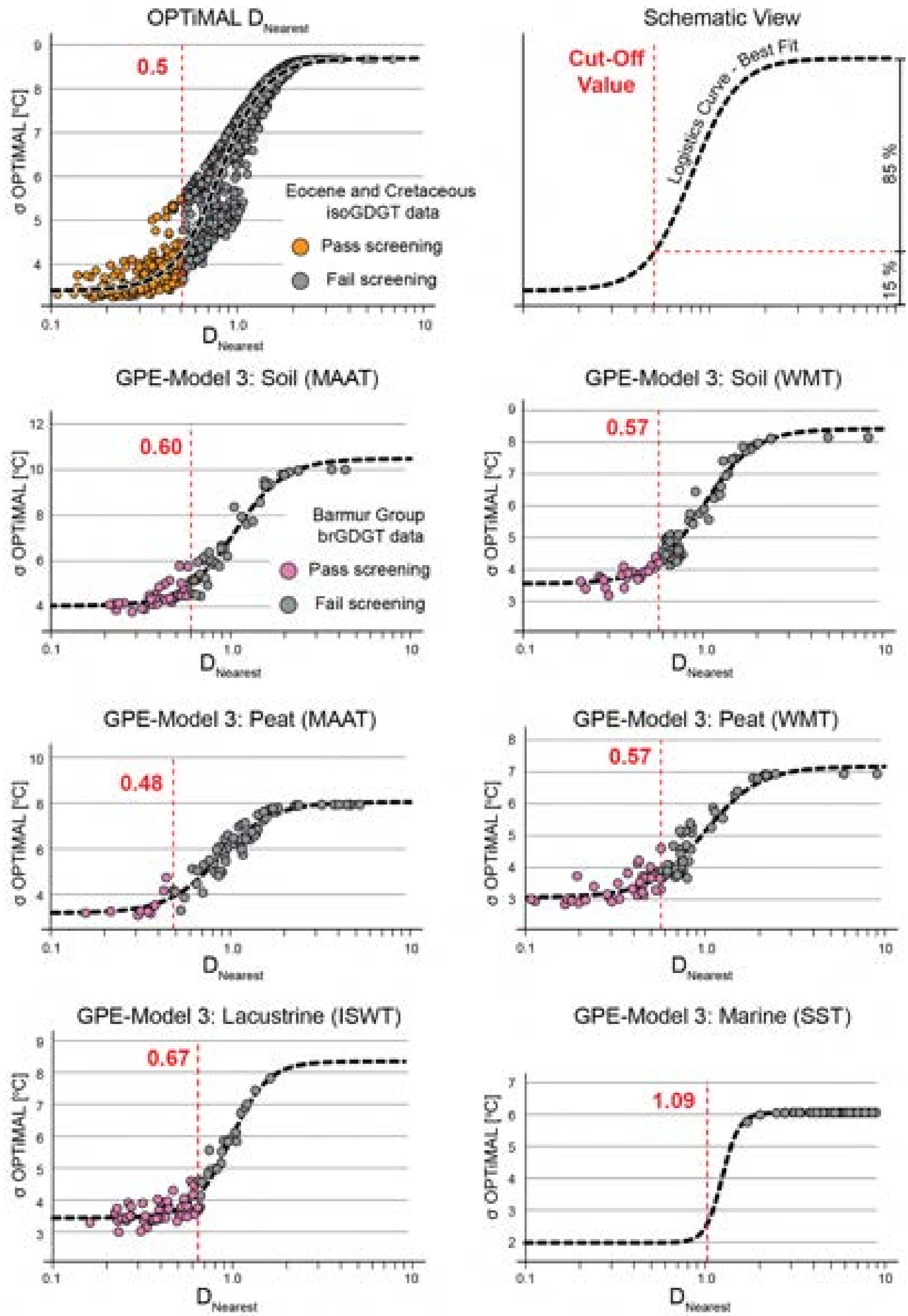


Figure A.7: $D_{Nearest}$ cut-off determination for brGDGT training data. Shown is the original $D_{Nearest} \leq 0.5$ of Dunkley Jones et al. (2020) and how a logistics curve can be fit to the modelled data. $D_{Nearest} \leq 0.5$ for isoGDGT data captures c. 15% of logistics curve between its lower and upper bounds. Using this threshold, curves can be fit to other training datasets and bespoke $D_{Nearest}$ cut-off determined.

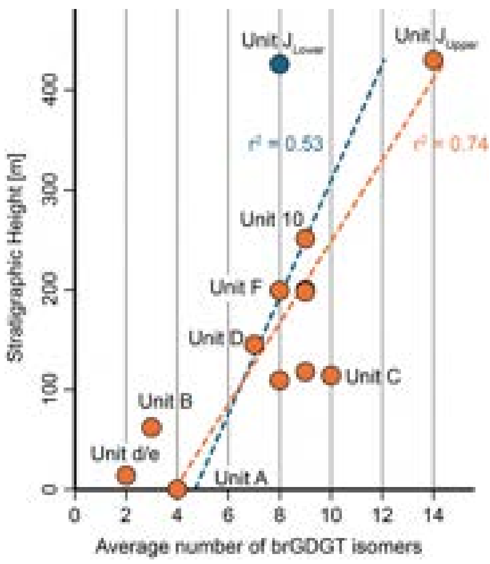


Figure A.8: The average number of brGDGT isomers (max 15) recovered from each lignite horizon against stratigraphic height. If all lignites beds are used, a linear regression of $r^2$ = 0.53 results. If the lignite samples of Unit $J_{Lower}$ is excluded from the analysis, this rises to $r^2$ = 0.74.

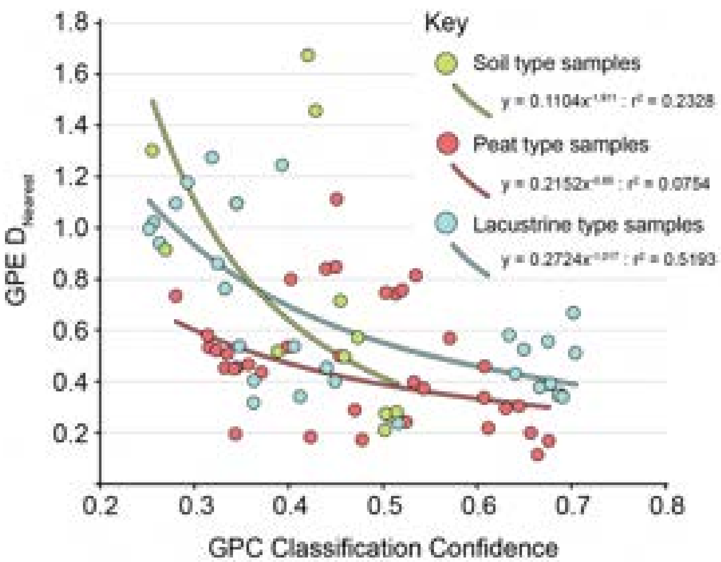


Figure A.9: The results of GPE $D_{Nearest}$ and GPC classification confidence for Barmur Group lignites. The lower the value of $D_{Nearest}$ the better represented a data-point is within the modern training data. No relationship is observed within soil types lignites, a weak relationship is observed with peat type lignites, and a moderate relationship observed in lacustrine type lignites.

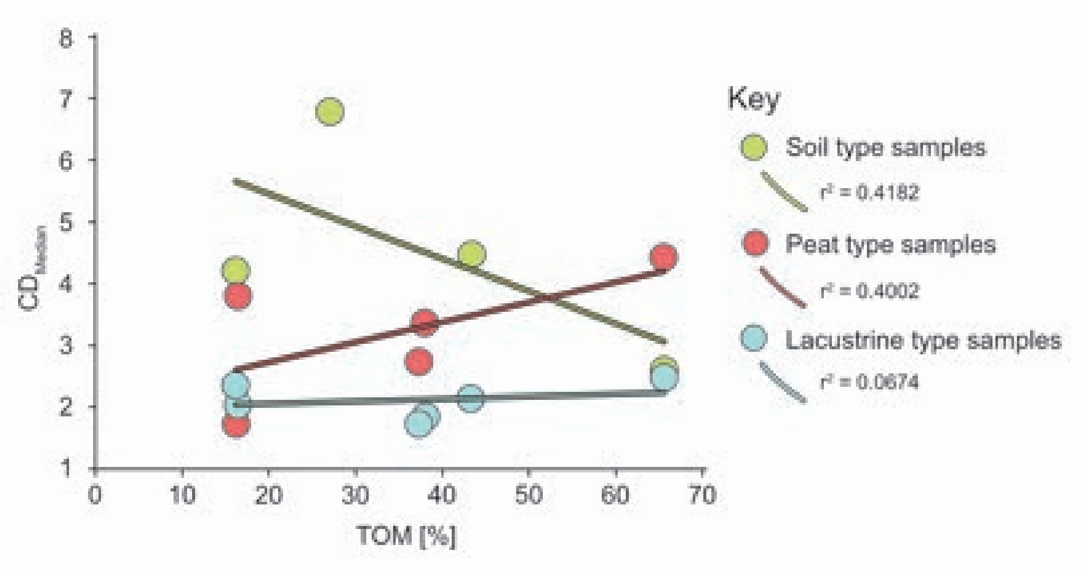


Figure A.10: Average Total Organic Matter (TOM) of sampled lignite horizons against the average $CD_{Median}$ value of soil, peat, and lacustrine type samples. A weak, linear relationship is observed in lacustrine type samples. Weak to moderate, linear relationships are observed in soil and peat type samples.

## *REFERENCES — Appendix A*


Allison, M. S., Hall, J. R. and Jones, S. M. (2023), 'Barmur Group, Tjörnes Peninsula', https://v3geo.com/model/367. V3Geo.

Allison, M. S., Hall, J. R., Papadopoulos, M. T., Eiríksson, J. and Jones, S. M. (2026), 'Updated stratigraphy and structure of the Barmur Group (Tjörnes beds, Iceland) and implications for Bering Strait opening', *Geological Society, London, Special Publications* **553**(1), SP553–2023.

Dunkley Jones, T., Eley, Y. L., Thomson, W., Greene, S. E., Mandel, I., Edgar, K. and Bendle, J. A. (2020), 'OPTiMAL: a new machine learning approach for GDGT-based palaeothermometry', *Climate of the Past* **16**(6), 2599–2617.

Greene, S. and Mandel, I. (2020), 'carbonatefan/OPTiMAL: First release'.
**URL:** *https://doi.org/10.5281/zenodo.4293851*

Kalacska, M., Lucanus, O., Arroyo-Mora, J., Laliberte, E., Elmer, K., Leblanc, G. and

Groves, A. (2020), 'Accuracy of 3D landscape reconstruction without ground control points using different UAS platforms', *Drones* **4**(2), 13.

Rowley, E. and White, N. (1998), 'Inverse modelling of extension and denudation in the East Irish Sea and surrounding areas', *Earth and Planetary Science Letters* **161**(1-4), 57–71.

# B | Appendix B - Data Tables

## B.1 Contents

Appendix B provides data tables of the following information:

- Table B.1 - Sample information for all the Barmur Group lignite data (bed identifiers, stratigraphic height, unique sample numbers, age model, GPS location, GPC and GPE results).
- Table B.2 - Barmur Group brGDGT data.
- Table B.3 - Barmur Group adjusted brGDGT signal data and TOM/TOC data.
- Table B.4 - Barmur Group *n*-Alkane data.
- Table B.5 - Barmur Group hopane data.

| Sample Information | | | | | | | GPC Results | | GPE Predictions [°C] | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bed | ES21 Log | ID | Stratigraphic Height [m] | Age [Ma] | Latitude [°] | Longitude [°] | palaeodepositional classification & Confidence [%] | | Prediction | -95$^{th}$% error | +95$^{th}$% error | $\sigma$ | $CD_{Median}$ | $D_{Nearest}$ | Pass/ fail |
| $J_{Upper}$ | XXII:G | 92 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.35 | 20.81 | 13.50 | 28.13 | 3.73 | 3.76 | 0.45 | Yes |
| $J_{Upper}$ | XXII:G | 91 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.33 | 20.84 | 13.50 | 28.17 | 3.74 | 3.74 | 0.45 | Yes |
| $J_{Upper}$ | XXII:G | 90 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.32 | 20.35 | 12.76 | 27.94 | 3.87 | 3.86 | 0.53 | Yes |
| $J_{Upper}$ | XXII:G | 89 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.34 | 20.52 | 12.95 | 28.10 | 3.86 | 3.82 | 0.51 | Yes |
| $J_{Upper}$ | XXII:G | 88 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.32 | 20.29 | 12.94 | 27.64 | 3.75 | 3.81 | 0.52 | Yes |
| $J_{Upper}$ | XXII:G | 87 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.31 | 20.00 | 12.37 | 27.63 | 3.89 | 3.87 | 0.58 | No |
| $J_{Upper}$ | XXII:G | 86 | 430 | 4.64 | 66.15686 | -17.27308 | Peat | 0.36 | 20.68 | 13.48 | 27.87 | 3.67 | 3.77 | 0.47 | Yes |
| $J_{Lower}$ | XXII:E | 85 | 426 | 4.65 | 66.15686 | -17.27308 | Marine | 0.25 | 12.51 | 0.66 | 24.37 | 6.05 | 3.19 | 2.71 | No |
| $J_{Lower}$ | XXII:E | 84 | 426 | 4.65 | 66.15686 | -17.27308 | Lacustrine | 0.26 | 6.72 | –4.87 | 18.32 | 5.92 | 2.14 | 1.02 | No |
| $J_{Lower}$ | XXII:E | 83 | 426 | 4.65 | 66.15686 | -17.27308 | Lacustrine | 0.26 | 6.22 | –5.25 | 17.69 | 5.85 | 2.07 | 0.94 | No |
| $J_{Lower}$ | XXII:E | 82 | 426 | 4.65 | 66.15686 | -17.27308 | Lacustrine | 0.25 | 5.84 | –5.59 | 17.27 | 5.83 | 1.98 | 0.99 | No |
| 10 | XVI:H | 81 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.68 | 18.32 | 11.66 | 24.98 | 3.40 | 1.73 | 0.39 | Yes |
| 10 | XVI:H | 80 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.69 | 18.46 | 11.97 | 24.95 | 3.31 | 1.75 | 0.35 | Yes |
| 10 | XVI:H | 79 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.67 | 18.25 | 11.66 | 24.84 | 3.36 | 1.77 | 0.38 | Yes |
| 10 | XVI:H | 78 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.69 | 18.06 | 11.52 | 24.60 | 3.34 | 1.75 | 0.35 | Yes |
| 10 | XVI:H | 77 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.69 | 18.51 | 11.99 | 25.03 | 3.33 | 1.79 | 0.34 | Yes |
| 10 | XVI:H | 76 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.68 | 18.30 | 11.66 | 24.94 | 3.39 | 1.73 | 0.39 | Yes |
| 10 | XVI:H | 75 | 251 | 5.04 | 66.13781 | -17.257 | Lacustrine | 0.64 | 17.58 | 10.82 | 24.34 | 3.45 | 1.78 | 0.43 | Yes |
| $F_{Upper}$ | XIII:H | 74 | 200 | 5.19 | 66.12608 | -17.25829 | Lacustrine | 0.68 | 11.36 | 3.81 | 18.91 | 3.85 | 1.80 | 0.56 | Yes |
| $F_{Upper}$ | XIII:H | 73 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.40 | 15.15 | 7.98 | 22.33 | 3.66 | 2.21 | 0.80 | No |
| $F_{Upper}$ | XIII:H | 72 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.34 | 17.21 | 10.50 | 23.92 | 3.42 | 1.61 | 0.45 | Yes |
| $F_{Upper}$ | XIII:H | 71 | 200 | 5.19 | 66.12608 | -17.25829 | Marine | 0.41 | 11.84 | –0.03 | 23.72 | 6.06 | 9.61 | 8.28 | No |
| $F_{Upper}$ | XIII:H | 70 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.45 | 17.84 | 6.33 | 29.35 | 5.87 | 4.77 | 1.11 | No |
| $F_{Upper}$ | XIII:H | 69 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.51 | 20.14 | 12.24 | 28.04 | 4.03 | 3.50 | 0.74 | No |
| $F_{Upper}$ | XIII:H | 68 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.52 | 21.30 | 13.06 | 29.54 | 4.20 | 3.47 | 0.76 | No |
| $F_{Upper}$ | XIII:H | 67 | 200 | 5.19 | 66.12608 | -17.25829 | Peat | 0.54 | 19.94 | 11.71 | 28.18 | 4.20 | 3.61 | 0.81 | No |
| $F_{Middle}$ | XIII:H | 66 | 199 | 5.19 | 66.12608 | -17.25829 | Peat | 0.63 | 12.81 | 6.81 | 18.82 | 3.06 | 2.82 | 0.29 | Yes |
| $F_{Middle}$ | XIII:H | 65 | 199 | 5.19 | 66.12608 | -17.25829 | Peat | 0.57 | 14.53 | 8.02 | 21.04 | 3.32 | 2.55 | 0.57 | Yes |
| $F_{Middle}$ | XIII:H | 64 | 199 | 5.19 | 66.12608 | -17.25829 | Peat | 0.66 | 11.09 | 5.41 | 16.77 | 2.90 | 2.53 | 0.20 | Yes |
| $F_{Middle}$ | XIII:H | 63 | 199 | 5.19 | 66.12608 | -17.25829 | Peat | 0.61 | 12.86 | 6.73 | 19.00 | 3.13 | 2.81 | 0.46 | Yes |
| $F_{Lower}$ | XIII:H | 62 | 198 | 5.19 | 66.12608 | -17.25829 | Lacustrine | 0.44 | 16.86 | 9.39 | 24.32 | 3.81 | 1.93 | 0.45 | Yes |
| $F_{Lower}$ | XIII:H | 61 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.40 | 17.44 | 11.00 | 23.89 | 3.29 | 1.37 | 0.53 | Yes |
| $F_{Lower}$ | XIII:H | 60 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.45 | 18.52 | 10.66 | 26.37 | 4.01 | 3.96 | 0.50 | Yes |
| $F_{Lower}$ | XIII:H | 59 | 198 | 5.19 | 66.12608 | -17.25829 | Lacustrine | 0.41 | 8.75 | 1.36 | 16.13 | 3.77 | 1.68 | 0.54 | Yes |
| $F_{Lower}$ | XIII:H | 58 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.64 | 12.92 | 7.17 | 18.67 | 2.93 | 2.34 | 0.30 | Yes |
| $F_{Lower}$ | XIII:H | 57 | 198 | 5.19 | 66.12608 | -17.25829 | Lacustrine | 0.52 | 10.69 | 4.80 | 16.57 | 3.00 | 1.46 | 0.24 | Yes |
| $F_{Lower}$ | XIII:H | 56 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.68 | 13.33 | 7.75 | 18.92 | 2.85 | 1.60 | 0.17 | Yes |
| $F_{Lower}$ | XIII:H | 55 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.61 | 13.16 | 7.22 | 19.09 | 3.03 | 2.25 | 0.34 | Yes |
| $F_{Lower}$ | XIII:H | 54 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.61 | 16.67 | 10.82 | 22.52 | 2.99 | 1.55 | 0.22 | Yes |
| $F_{Lower}$ | XIII:H | 53 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.66 | 14.14 | 8.35 | 19.92 | 2.95 | 2.17 | 0.11 | Yes |
| $F_{Lower}$ | XIII:H | 52 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.47 | 17.54 | 11.33 | 23.76 | 3.17 | 2.89 | 0.29 | Yes |
| $F_{Lower}$ | XIII:H | 51 | 198 | 5.19 | 66.12608 | -17.25829 | Peat | 0.48 | 16.80 | 10.89 | 22.72 | 3.02 | 1.43 | 0.17 | Yes |
| D | XI:B | 50 | 145 | 5.73 | 66.11908 | -17.25652 | Soil | 0.46 | 20.40 | 10.67 | 30.13 | 4.97 | 2.55 | 0.71 | No |
| D | XI:B | 49 | 145 | 5.73 | 66.11908 | -17.25652 | Peat | 0.44 | 17.45 | 7.21 | 27.69 | 5.23 | 4.42 | 0.84 | No |
| D | XI:B | 48 | 145 | 5.73 | 66.11908 | -17.25652 | Lacustrine | 0.29 | 10.13 | –3.05 | 23.30 | 6.72 | 2.44 | 1.18 | No |
| D | XI:B | 47 | 145 | 5.73 | 66.11908 | -17.25652 | Lacustrine | 0.39 | 11.20 | –2.30 | 24.69 | 6.88 | 2.57 | 1.24 | No |
| D | XI:B | 46 | 145 | 5.73 | 66.11908 | -17.25652 | Peat | 0.45 | 17.33 | 7.00 | 27.67 | 5.27 | 4.42 | 0.85 | No |
| $C_{Upper}$ | X:F | 45 | 118 | 6.09 | 66.11365 | -17.2624 | Peat | 0.34 | 14.26 | 6.95 | 21.57 | 3.73 | 3.73 | 0.19 | Yes |
| $C_{Upper}$ | X:F | 44 | 118 | 6.09 | 66.11365 | -17.2624 | Peat | 0.52 | 16.21 | 9.53 | 22.89 | 3.41 | 3.52 | 0.24 | Yes |
| $C_{Upper}$ | X:F | 43 | 118 | 6.09 | 66.11365 | -17.2624 | Peat | 0.53 | 15.93 | 9.41 | 22.46 | 3.33 | 3.34 | 0.40 | Yes |
| $C_{Upper}$ | X:F | 42 | 118 | 6.09 | 66.11365 | -17.2624 | | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| $C_{Middle}$ | X:F | 41 | 114 | 6.14 | 66.11365 | -17.2624 | Peat | 0.50 | 16.96 | 7.04 | 26.89 | 5.06 | 4.43 | 0.74 | No |
| $C_{Middle}$ | X:F | 40 | 114 | 6.14 | 66.11365 | -17.2624 | Peat | 0.42 | 12.82 | 7.01 | 18.62 | 2.96 | 2.19 | 0.18 | Yes |
| $C_{Middle}$ | X:F | 39 | 114 | 6.14 | 66.11365 | -17.2624 | Lacustrine | 0.65 | 21.09 | 14.13 | 28.05 | 3.55 | 1.59 | 0.52 | Yes |
| $C_{Middle}$ | X:F | 38 | 114 | 6.14 | 66.11365 | -17.2624 | Lacustrine | 0.41 | 8.24 | 2.11 | 14.37 | 3.13 | 1.77 | 0.34 | Yes |
| $C_{Lower}$ | X:F | 37 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.35 | 10.91 | 4.06 | 17.77 | 3.50 | 1.64 | 0.54 | Yes |
| $C_{Lower}$ | X:F | 36 | 109 | 6.20 | 66.11365 | -17.2624 | Peat | 0.37 | 15.61 | 8.66 | 22.55 | 3.54 | 2.76 | 0.43 | Yes |
| $C_{Lower}$ | X:F | 35 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.36 | 10.73 | 4.80 | 16.66 | 3.03 | 1.44 | 0.32 | Yes |
| $C_{Lower}$ | X:F | 34 | 109 | 6.20 | 66.11365 | -17.2624 | Peat | 0.28 | 20.58 | 11.90 | 29.25 | 4.42 | 3.74 | 0.73 | No |
| $C_{Lower}$ | X:F | 33 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.33 | 19.98 | 10.52 | 29.45 | 4.83 | 1.91 | 0.76 | No |
| $C_{Lower}$ | X:F | 32 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.63 | 10.41 | 3.59 | 17.23 | 3.48 | 1.70 | 0.58 | Yes |
| $C_{Lower}$ | X:F | 31 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.70 | 13.00 | 5.62 | 20.38 | 3.76 | 1.61 | 0.67 | Yes |
| $C_{Lower}$ | X:F | 30 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.45 | 11.28 | 4.49 | 18.08 | 3.47 | 1.77 | 0.40 | Yes |
| $C_{Lower}$ | X:F | 29 | 109 | 6.20 | 66.11365 | -17.2624 | Lacustrine | 0.70 | 7.23 | 0.69 | 13.77 | 3.34 | 1.35 | 0.51 | Yes |
| $B_{Upper}$ | IX:H | 28 | 63 | 6.81 | 66.11133 | -17.26972 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| $B_{Upper}$ | IX:H | 27 | 63 | 6.81 | 66.11133 | -17.26972 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| $B_{Upper}$ | IX:H | 26 | 63 | 6.81 | 66.11133 | -17.26972 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| $B_{Upper}$ | IX:H | 25 | 63 | 6.81 | 66.11133 | -17.26972 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| $B_{Upper}$ | IX:H | 24 | 63 | 6.81 | 66.11133 | -17.26972 | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| $B_{Lower}$ | IX:H | 23 | 62 | 6.82 | 66.11133 | -17.26972 | Soil | 0.51 | 7.46 | 0.02 | 14.90 | 3.80 | 6.82 | 0.27 | Yes |
| $B_{Lower}$ | IX:H | 22 | 62 | 6.82 | 66.11133 | -17.26972 | Soil | 0.50 | 7.29 | –0.05 | 14.63 | 3.74 | 6.70 | 0.27 | Yes |
| $B_{Lower}$ | IX:H | 21 | 62 | 6.821 | 66.11133 | -17.26972 | Soil | 0.47 | 6.79 | –1.48 | 15.06 | 4.22 | 7.37 | 0.57 | No |
| $B_{Lower}$ | IX:H | 20 | 62 | 6.821 | 66.11133 | -17.26972 | Soil | 0.50 | 6.44 | –0.70 | 13.59 | 3.65 | 6.34 | 0.21 | Yes |
| $B_{Lower}$ | IX:H | 19 | 62 | 6.821 | 66.11133 | -17.26972 | Soil | 0.51 | 7.41 | 0.17 | 14.65 | 3.69 | 6.70 | 0.28 | Yes |
| d/e | VI:L | 18 | 14 | 7.45 | 66.10702 | -17.2789 | Soil | 0.39 | 12.35 | 4.24 | 20.46 | 4.14 | 3.84 | 0.52 | Yes |
| d/e | VI:L | 17 | 14 | 7.45 | 66.10702 | -17.2789 | Soil | 0.42 | 16.45 | 1.04 | 31.87 | 7.86 | 4.62 | 1.67 | No |
| d/e | VI:L | 16 | 14 | 7.45 | 66.10702 | -17.2789 | Soil | 0.43 | 14.73 | 0.09 | 29.37 | 7.47 | 4.33 | 1.46 | No |
| d/e | VI:L | 15 | 14 | 7.45 | 66.10702 | -17.2789 | Soil | 0.46 | 11.69 | 3.93 | 19.45 | 3.96 | 4.05 | 0.50 | Yes |
| d/e | VI:L | 14 | 14 | 7.45 | 66.10702 | -17.2789 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| d/e | VI:L | 13 | 14 | 7.45 | 66.10702 | -17.2789 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| d/e | VI:L | 12 | 14 | 7.45 | 66.10702 | -17.2789 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| d/e | VI:L | 11 | 14 | 7.45 | 66.10702 | -17.2789 | Peat | 0.54 | 14.54 | 7.05 | 22.04 | 3.82 | 1.73 | 0.37 | Yes |
| d/e | VI:L | 10 | 14 | 7.45 | 66.10702 | -17.2789 | Lacustrine | 0.32 | 8.78 | –4.94 | 22.51 | 7.00 | 2.60 | 1.27 | No |
| A | I:B | 9 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| A | I:B | 8 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| A | I:B | 7 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| A | I:B | 6 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.34 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| A | I:B | 5 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.28 | 7.73 | –4.27 | 19.72 | 6.12 | 2.28 | 1.10 | No |
| A | I:B | 4 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.33 | 6.84 | –2.91 | 16.60 | 4.98 | 2.34 | 0.86 | No |
| A | I:B | 3 | 0 | 7.63 | 66.10553 | -17.28343 | Lacustrine | 0.36 | 5.99 | –0.74 | 12.72 | 3.43 | 1.38 | 0.40 | Yes |
| A | I:B | 2 | 0 | 7.63 | 66.10553 | -17.28343 | Soil | 0.26 | 13.33 | –1.22 | 27.87 | 7.42 | 4.60 | 1.30 | No |
| A | I:B | 1 | 0 | 7.63 | 66.10553 | -17.28343 | Soil | 0.27 | 11.93 | –0.74 | 24.60 | 6.46 | 4.36 | 0.91 | No |

Table B.1: Barmur Group lignite sample information. Results of the Gaussian Process Classification (GPC) for palaeodepositional classification. Gaussian Process Emulator prediction of environmental temperature, see Chapter 3 text. 95$^{th}$% predictive error is included alongside the standard deviation ($\sigma$), values of $CD_{Median}$ and $D_{Nearest}$. Data which pass $D_{Nearest}$ screening are marked "Yes".

| Sample Information | | brGDGT [%] | | | | | | | | | | | | | | | Total brGDGT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bed | ID | Ia | Ib | Ic | IIa | IIa' | IIb | IIb' | IIc | IIc' | IIIa | IIIa' | IIIb | IIIb' | IIIc | IIIc' | |
| $J_{Upper}$ | 92 | 0.176374649 | 0.072742271 | 0.012736321 | 0.339819116 | 0.048754508 | 0.058333991 | 0.006487393 | 0.006067662 | 0 | 0.252762589 | 0.024018795 | 0.001167408 | 0.000429998 | 0.00024073 | 6.45683E-05 | 606297456.7 |
| $J_{Upper}$ | 91 | 0.177838502 | 0.072077645 | 0.013037309 | 0.335996702 | 0.048345685 | 0.059032758 | 0.007214013 | 0.00596019 | 0 | 0.252372104 | 0.026702065 | 0.001119919 | 0.000303107 | 0 | 0 | 833720710 |
| $J_{Upper}$ | 90 | 0.173074739 | 0.067524921 | 0.011556261 | 0.348129271 | 0.045019856 | 0.055168304 | 0.005536763 | 0.005274143 | 0 | 0.26165399 | 0.025596574 | 0.00102542 | 0.000215906 | 0.000187579 | 3.62722E-05 | 563914779.8 |
| $J_{Upper}$ | 89 | 0.17729435 | 0.067591353 | 0.01236969 | 0.337912049 | 0.051072783 | 0.057783105 | 0.007183071 | 0.00585056 | 0 | 0.259026835 | 0.02223071 | 0.001095549 | 0.000287115 | 0.000245891 | 5.69386E-05 | 656521390.3 |
| $J_{Upper}$ | 88 | 0.175485776 | 0.067718177 | 0.011450968 | 0.356368492 | 0.043035313 | 0.055110658 | 0.00552397 | 0.005367132 | 0 | 0.257086605 | 0.021301222 | 0.001020631 | 0.000263952 | 0.000212831 | 5.42732E-05 | 587395709.2 |
| $J_{Upper}$ | 87 | 0.176448568 | 0.06708917 | 0.010708228 | 0.354578313 | 0.043677155 | 0.054685559 | 0.007056495 | 0.005424236 | 0.000117706 | 0.263729978 | 0.015034376 | 0.000996665 | 0.000231156 | 0.000188113 | 3.428E-05 | 488375829.9 |
| $J_{Upper}$ | 86 | 0.173443092 | 0.070598013 | 0.012299733 | 0.353525247 | 0.046789144 | 0.056969095 | 0.008168511 | 0.006090679 | 0.000179746 | 0.251957474 | 0.018327104 | 0.001101778 | 0.000287855 | 0.000199726 | 6.28024E-05 | 680390356.5 |
| $J_{Lower}$ | 85 | 0.056109661 | 0.009747901 | 0.553471175 | 0 | 0.045973585 | 0 | 0.009765693 | 0 | 0.214969065 | 0.037758607 | 0.004633066 | 0 | 0 | 0 | 0.067571246 | 553214.393 |
| $J_{Lower}$ | 84 | 0.030958874 | 0.005613539 | 0.35179801 | 0 | 0.035356986 | 0 | 0.001899155 | 0 | 0.436843928 | 0.058491557 | 0 | 0 | 0 | 0 | 0.079037951 | 403348.262 |
| $J_{Lower}$ | 83 | 0.154386403 | 0.004856786 | 0.295399557 | 0 | 0.020453628 | 0 | 0.003319647 | 0 | 0.344828594 | 0.055744631 | 0 | 0 | 0 | 0 | 0.121010754 | 138347.242 |
| $J_{Lower}$ | 82 | 0.068143248 | 0.01339238 | 0.399618304 | 0 | 0.076617899 | 0 | 0.011355648 | 0 | 0.34542591 | 0.051064465 | 0 | 0 | 0 | 0 | 0.034382146 | 1242800.388 |
| 10 | 81 | 0.532955869 | 0.074302166 | 0.03002789 | 0.179285073 | 0.037200582 | 0.030956136 | 0.012717153 | 0 | 0 | 0.082051211 | 0.02050392 | 0 | 0 | 0 | 0 | 25562885.69 |
| 10 | 80 | 0.53959597 | 0.0750518 | 0.026896999 | 0.188456159 | 0.042717871 | 0.022961869 | 0.00561602 | 0 | 0 | 0.082806275 | 0.015897036 | 0 | 0 | 0 | 0 | 43956031.16 |
| 10 | 79 | 0.546877711 | 0.07596085 | 0.028577191 | 0.181554696 | 0.038230239 | 0.030826051 | 0.011929197 | 0 | 0 | 0.070785209 | 0.015258856 | 0 | 0 | 0 | 0 | 24522332.24 |
| 10 | 78 | 0.539886644 | 0.07293184 | 0.025064473 | 0.186305754 | 0.040099695 | 0.029904893 | 0.009662481 | 0 | 0 | 0.078339514 | 0.017804707 | 0 | 0 | 0 | 0 | 33450148.29 |
| 10 | 77 | 0.554821898 | 0.073061515 | 0.027300411 | 0.177995358 | 0.039393122 | 0.026679666 | 0.00836451 | 0 | 0 | 0.075372736 | 0.017010784 | 0 | 0 | 0 | 0 | 34478110.89 |
| 10 | 76 | 0.532704371 | 0.074715659 | 0.028461134 | 0.182321401 | 0.037797999 | 0.03197686 | 0.011791923 | 0 | 0 | 0.079122349 | 0.021108304 | 0 | 0 | 0 | 0 | 18172654.76 |
| 10 | 75 | 0.537530127 | 0.077052405 | 0.027522491 | 0.194136244 | 0.029222835 | 0.029325281 | 0.011659972 | 0 | 0 | 0.079564274 | 0.013986371 | 0 | 0 | 0 | 0 | 42084023.69 |
| $F_{Upper}$ | 74 | 0.572301318 | 0.012588486 | 0.000891022 | 0.120407668 | 0.125799816 | 0.006346909 | 0.002608091 | 0 | 0 | 0.080495329 | 0.07856136 | 0 | 0 | 0 | 0 | 395434.831 |
| $F_{Upper}$ | 73 | 0.370170674 | 0.030117881 | 0.006592459 | 0.3228857 | 0 | 0.036219614 | 0 | 0.005118721 | 0 | 0.178446235 | 0.048296807 | 0.001544019 | 0.00060789 | 0 | 0 | 1235944.969 |
| $F_{Upper}$ | 72 | 0.658379215 | 0.00398004 | 0 | 0.236641327 | 0 | 0.005311291 | 0 | 0 | 0 | 0.079685156 | 0.016002971 | 0 | 0 | 0 | 0 | 80567.791 |
| $F_{Upper}$ | 71 | 0.302140772 | 0.033095006 | 0.003553982 | 0.422942517 | 0 | 0.018378794 | 0 | 0 | 0 | 0.188227391 | 0.031661538 | 0 | 0 | 0 | 0 | 284536.617 |
| $F_{Upper}$ | 70 | 0.083102502 | 0.032424199 | 0.008807888 | 0.458983549 | 0 | 0.047627643 | 0 | 0.005358876 | 0 | 0.307902151 | 0.05110749 | 0.002889602 | 0.001796101 | 0 | 0 | 10767680.09 |
| $F_{Upper}$ | 69 | 0.144129372 | 0.053744767 | 0.015699513 | 0.462587223 | 0 | 0.083692252 | 0 | 0.011595794 | 0 | 0.19679915 | 0.028703481 | 0.001404836 | 0.001643612 | 0 | 0 | 56894657.19 |
| $F_{Upper}$ | 68 | 0.132848473 | 0.076371202 | 0.01481922 | 0.469499783 | 0 | 0.084095638 | 0 | 0.010562523 | 0 | 0.182337407 | 0.024217783 | 0.002920639 | 0.002327332 | 0 | 0 | 21110770.18 |
| $F_{Upper}$ | 67 | 0.135654447 | 0.058680442 | 0.014357537 | 0.46677151 | 0 | 0.071371904 | 0 | 0.010812808 | 0 | 0.205276519 | 0.033256944 | 0.001976973 | 0.001840915 | 0 | 0 | 114989979.1 |
| $F_{Middle}$ | 66 | 0.251988743 | 0.034539801 | 0.010776254 | 0.472299892 | 0 | 0.02503696 | 0 | 0.007206227 | 0 | 0.177166556 | 0.020985567 | 0 | 0 | 0 | 0 | 116704727.8 |
| $F_{Middle}$ | 65 | 0.332223597 | 0.025759982 | 0.009007455 | 0.399968339 | 0 | 0.02848336 | 0 | 0.008479704 | 0 | 0.196077563 | 0 | 0 | 0 | 0 | 0 | 300582712.8 |
| $F_{Middle}$ | 64 | 0.290937382 | 0.027799995 | 0.008123123 | 0.48687222 | 0 | 0.018617498 | 0 | 0.007636598 | 0 | 0.158531956 | 0 | 0.000682082 | 0.000799146 | 0 | 0 | 62517011.62 |
| $F_{Middle}$ | 63 | 0.283184166 | 0.030704472 | 0.009523242 | 0.447923541 | 0 | 0.020965935 | 0 | 0.009415006 | 0 | 0.196144133 | 0 | 0.002139506 | 0 | 0 | 0 | 154536148.8 |
| $F_{Lower}$ | 62 | 0.566180517 | 0.077318333 | 0.016600134 | 0.193487423 | 0 | 0.015193538 | 0 | 0.001785307 | 0 | 0.123360545 | 0 | 0.006074203 | 0 | 0 | 0 | 35374351.49 |
| $F_{Lower}$ | 61 | 0.509197967 | 0.064083153 | 0.010357851 | 0.266514984 | 0 | 0.014594603 | 0 | 0.002328713 | 0 | 0.103982425 | 0.024742714 | 0.004197589 | 0 | 0 | 0 | 44262234.74 |
| $F_{Lower}$ | 60 | 0.182333942 | 0.059215869 | 0.013251196 | 0.388143269 | 0 | 0.041101296 | 0 | 0.005121668 | 0 | 0.273941343 | 0.03210656 | 0.003475865 | 0.001308992 | 0 | 0 | 50448336.04 |
| $F_{Lower}$ | 59 | 0.113850708 | 0.047414068 | 0.010640399 | 0.297776039 | 0 | 0.026725089 | 0 | 0.002670087 | 0 | 0.465599315 | 0.031174786 | 0.00300078 | 0.001148728 | 0 | 0 | 23107831.71 |
| $F_{Lower}$ | 58 | 0.286941443 | 0.034228904 | 0.008552593 | 0.460062131 | 0 | 0.027987933 | 0 | 0.00572862 | 0 | 0.140451849 | 0.036046527 | 0 | 0 | 0 | 0 | 56601504.71 |
| $F_{Lower}$ | 57 | 0.238707158 | 0.04398717 | 0.012545007 | 0.300764322 | 0.070281211 | 0.029933329 | 0.005633262 | 0.006352414 | 0 | 0.238987771 | 0.050589084 | 0.002219271 | 0 | 0 | 0 | 40057517.16 |
| $F_{Lower}$ | 56 | 0.373883969 | 0.037569673 | 0.009279493 | 0.42010548 | 0 | 0.023837543 | 0 | 0.006528458 | 0 | 0.104379605 | 0.022200336 | 0.002215441 | 0 | 0 | 0 | 53654147.77 |
| $F_{Lower}$ | 55 | 0.327282663 | 0.034843262 | 0.005652672 | 0.414007314 | 0 | 0.029257156 | 0 | 0.004566727 | 0 | 0.161479697 | 0.022910509 | 0 | 0 | 0 | 0 | 625891.081 |
| $F_{Lower}$ | 54 | 0.430956625 | 0.025258684 | 0.00575876 | 0.356838952 | 0 | 0.013513657 | 0 | 0.010286275 | 0 | 0.132233555 | 0.025153491 | 0 | 0 | 0 | 0 | 9367475.417 |
| $F_{Lower}$ | 53 | 0.289261876 | 0.059049969 | 0.017505102 | 0.451424927 | 0 | 0.038078517 | 0 | 0.012491093 | 0 | 0.119289593 | 0.012898923 | 0 | 0 | 0 | 0 | 90034734.48 |
| $F_{Lower}$ | 52 | 0.218204398 | 0.029829709 | 0.00749647 | 0.42557657 | 0 | 0.064596382 | 0 | 0.00804341 | 0 | 0.173039308 | 0.062534025 | 0.006066439 | 0.004613289 | 0 | 0 | 11885511.27 |
| $F_{Lower}$ | 51 | 0.453331774 | 0.029768904 | 0.007051592 | 0.324491349 | 0 | 0.024075928 | 0 | 0.005231543 | 0 | 0.122240454 | 0.033808457 | 0 | 0 | 0 | 0 | 18963930.3 |
| D | 50 | 0.506260704 | 0 | 0 | 0.084604517 | 0.237557591 | 0.025661906 | 0.004816755 | 0 | 0 | 0.024068182 | 0.083135684 | 0.009716685 | 0.013292991 | 0.004958954 | 0.005926031 | 14923641.94 |
| D | 49 | 0.165798882 | 0.049828977 | 0.013285791 | 0.407408317 | 0 | 0.043097001 | 0 | 0.007905812 | 0 | 0.311046865 | 0 | 0.001628356 | 0 | 0 | 0 | 9438631.101 |
| D | 48 | 0.683970806 | 0.009739322 | 0.00953636 | 0 | 0 | 0 | 0 | 0 | 0 | 0.080035219 | 0.172450993 | 0.044267301 | 0 | 0 | 0 | 9543887.057 |
| D | 47 | 0.745505795 | 0.009319523 | 0.005804498 | 0 | 0 | 0 | 0 | 0 | 0 | 0.061570435 | 0.177799748 | 0 | 0 | 0 | 0 | 10549335.59 |
| D | 46 | 0.170572036 | 0.046020414 | 0.011650369 | 0.409637639 | 0 | 0.040669871 | 0 | 0.008411152 | 0 | 0.313038519 | 0 | 0 | 0 | 0 | 0 | 10015754.18 |
| $C_{Upper}$ | 45 | 0.208731714 | 0.02784716 | 0.013579474 | 0.483059085 | 0 | 0.008059627 | 0.005459258 | 0 | 0 | 0.253263681 | 0 | 0 | 0 | 0 | 0 | 2618063.09 |
| $C_{Upper}$ | 44 | 0.232664879 | 0.024118837 | 0.010547066 | 0.421879108 | 0 | 0.019980949 | 0 | 0.009725673 | 0 | 0.24913095 | 0.031952537 | 0 | 0 | 0 | 0 | 1793050.428 |
| $C_{Upper}$ | 43 | 0.245245064 | 0.028862299 | 0 | 0.417379837 | 0 | 0.025152529 | 0.006441548 | 0.018608623 | 0 | 0.236767294 | 0.021542805 | 0 | 0 | 0 | 0 | 1694312.3 |
| $C_{Upper}$ | 42 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $C_{Middle}$ | 41 | 0.132965403 | 0.023096452 | 0.005300593 | 0.462016658 | 0 | 0.030270227 | 0.007769592 | 0.004010104 | 0 | 0.294616912 | 0.03995406 | 0 | 0 | 0 | 0 | 8635626.028 |
| $C_{Middle}$ | 40 | 0.280145414 | 0.058634935 | 0.02674809 | 0.288228549 | 0 | 0.096711022 | 0 | 0.013180449 | 0 | 0.126809673 | 0.097456692 | 0.012085176 | 0 | 0 | 0 | 1830624.012 |
| $C_{Middle}$ | 39 | 0.46160705 | 0.081678017 | 0.015056482 | 0.184970594 | 0.055405922 | 0.044146803 | 0 | 0.004954445 | 0.002505326 | 0.087695111 | 0.06198025 | 0 | 0 | 0 | 0 | 464406.564 |
| $C_{Middle}$ | 38 | 0.177715911 | 0.033898247 | 0.013703993 | 0.362431637 | 0.041586023 | 0.032032885 | 0.009454528 | 0.002268339 | 0.002706413 | 0.272910112 | 0.051291912 | 0 | 0 | 0 | 0 | 2062418.14 |
| $C_{Lower}$ | 37 | 0.28263696 | 0.045532494 | 0.037598339 | 0.300701609 | 0 | 0.080040194 | 0 | 0 | 0 | 0.20155999 | 0.051930414 | 0 | 0 | 0 | 0 | 836094.94 |
| $C_{Lower}$ | 36 | 0.278470683 | 0.130201833 | 0.093060436 | 0.303016005 | 0 | 0 | 0 | 0 | 0 | 0.195251043 | 0 | 0 | 0 | 0 | 0 | 121464.98 |
| $C_{Lower}$ | 35 | 0.207584527 | 0.049396987 | 0.027153542 | 0.306388525 | 0.075843551 | 0.052942446 | 0 | 0 | 0 | 0.22896951 | 0.051720912 | 0 | 0 | 0 | 0 | 998027.983 |
| $C_{Lower}$ | 34 | 0.160613333 | 0.108062687 | 0.105898647 | 0.201629509 | 0 | 0.190702065 | 0 | 0 | 0 | 0.233093759 | 0 | 0 | 0 | 0 | 0 | 426899.091 |
| $C_{Lower}$ | 33 | 0.251412399 | 0.120189498 | 0 | 0.266612113 | 0 | 0.031386306 | 0 | 0 | 0 | 0.220070687 | 0.110328998 | 0 | 0 | 0 | 0 | 300801.056 |
| $C_{Lower}$ | 32 | 0.525267632 | 0.017976175 | 0 | 0.107372413 | 0.15059942 | 0.023264826 | 0.001415723 | 0 | 0 | 0.126001327 | 0.048102485 | 0 | 0 | 0 | 0 | 1061314.514 |
| $C_{Lower}$ | 31 | 0.470384333 | 0.014225843 | 0 | 0.192523496 | 0.089787175 | 0 | 0 | 0 | 0 | 0.126210467 | 0.106868687 | 0 | 0 | 0 | 0 | 122841.23 |
| $C_{Lower}$ | 30 | 0.373487173 | 0.023221341 | 0.003576195 | 0.283510518 | 0 | 0.021765084 | 0.00627782 | 0.002655938 | 0.000440548 | 0.219446608 | 0.065618776 | 0 | 0 | 0 | 0 | 842180.308 |
| $C_{Lower}$ | 29 | 0.30706048 | 0.027289422 | 0.008581465 | 0.14684824 | 0.112510032 | 0.048348867 | 0 | 0.005339922 | 0.001099612 | 0.264106478 | 0.06338775 | 0.007291958 | 0.008135776 | 0 | 0 | 1483597.644 |

...

*Table B.2 continued on next page*

*Continuation of Table B.2*

...

| Sample Information | | brGDGT [%] | | | | | | | | | | | | | | | Total brGDGT |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bed | ID | Ia | Ib | Ic | IIa | IIa' | IIb | IIb' | IIc | IIc' | IIIa | IIIa' | IIIb | IIIb' | IIIc | IIIc' | |
| $B_{Upper}$ | 28 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $B_{Upper}$ | 27 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $B_{Upper}$ | 26 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $B_{Upper}$ | 25 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $B_{Upper}$ | 24 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| $B_{Lower}$ | 23 | 0.163668499 | 0 | 0 | 0.409338255 | 0 | 0 | 0 | 0 | 0 | 0.426993246 | 0 | 0 | 0 | 0 | 0 | 458133.157 |
| $B_{Lower}$ | 22 | 0.171199697 | 0 | 0 | 0.409670234 | 0 | 0 | 0 | 0 | 0 | 0.419130069 | 0 | 0 | 0 | 0 | 0 | 743438.665 |
| $B_{Lower}$ | 21 | 0.152805274 | 0 | 0 | 0.380440718 | 0 | 0 | 0 | 0 | 0 | 0.466754008 | 0 | 0 | 0 | 0 | 0 | 1197096.447 |
| $B_{Lower}$ | 20 | 0.186595844 | 0 | 0 | 0.419329148 | 0 | 0 | 0 | 0 | 0 | 0.394075008 | 0 | 0 | 0 | 0 | 0 | 1516570.029 |
| $B_{Lower}$ | 19 | 0.162516977 | 0 | 0 | 0.419352614 | 0 | 0 | 0 | 0 | 0 | 0.418130408 | 0 | 0 | 0 | 0 | 0 | 1191744.715 |
| d/e | 18 | 0.331040363 | 0 | 0 | 0.494634744 | 0 | 0 | 0 | 0 | 0 | 0.174324893 | 0 | 0 | 0 | 0 | 0 | 118858.636 |
| d/e | 17 | 0.239701127 | 0 | 0 | 0.760298873 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 6824.849 |
| d/e | 16 | 0.296554628 | 0 | 0 | 0.703445372 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 295885.684 |
| d/e | 15 | 0.293318662 | 0.023092879 | 0 | 0.474229761 | 0 | 0 | 0 | 0 | 0 | 0.209358698 | 0 | 0 | 0 | 0 | 0 | 318038.652 |
| d/e | 14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| d/e | 13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| d/e | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| d/e | 11 | 0.46958913 | 0 | 0 | 0.53041087 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 22339.94 |
| d/e | 10 | 0.381785328 | 0 | 0 | 0.381785328 | 0 | 0.236429343 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 11135.441 |
| A | 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A | 8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A | 7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A | 6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A | 5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 5231.107 |
| A | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.761190597 | 0.238809403 | 0 | 0 | 0 | 0 | 15878.085 |
| A | 3 | 0.099383801 | 0.054973652 | 0 | 0.121024448 | 0.129264186 | 0.094162109 | 0 | 0 | 0 | 0.235395213 | 0.165146008 | 0.075832583 | 0.024818 | 0 | 0 | 334052.991 |
| A | 2 | 0.084146937 | 0 | 0 | 0.074383219 | 0 | 0.084130751 | 0 | 0.069868688 | 0 | 0.275702984 | 0.150595866 | 0.124078126 | 0.047026685 | 0.090066743 | 0 | 325771.167 |
| A | 1 | 0.05774595 | 0.035663082 | 0 | 0.073881501 | 0.074148132 | 0.047740788 | 0.018806902 | 0.051749969 | 0 | 0.268210002 | 0.184655604 | 0.077724468 | 0 | 0.064462372 | 0.04521123 | 325501.564 |

Table B.2: brGDGT relative abundance data of Barmur Group lignite samples. Total sum of all brGDGT (absolute abundance) also provided.

| Sample Information | | | brGDGT Abundance | | | | | | Total Organic Matter and Carbon | |
|---|---|---|---|---|---|---|---|---|---|---|
| Bed | ID | Stratigraphic Height [m] | Sample Mass [g] | TLE [mg] | TLE Analysed [mg] | Raw HPLC-MS Signal | Adjusted HPLC-MS Signal | Normalised HPLC-MS Signal | TOM [%] | TOC [%] |
| $J_{Upper}$ | 92 | 430 | 5.07 | 20.1 | 10.05 | 606297456.7 | 60328105.14 | 0.319200848 | 16 | - |
| $J_{Upper}$ | 91 | 430 | 5.01 | 32.95 | 16.475 | 833720710.0 | 50605202.43 | 0.267756189 | 17 | - |
| $J_{Upper}$ | 90 | 430 | 5.53 | 19.55 | 9.775 | 563914779.8 | 57689491.54 | 0.305239731 | 17 | - |
| $J_{Upper}$ | 89 | 430 | 5.29 | 9.4 | 4.7 | 656521390.3 | 139685402.2 | 0.739086677 | 16 | - |
| $J_{Upper}$ | 88 | 430 | 5.15 | 10.6 | 5.3 | 587395709.2 | 110829379.1 | 0.586407143 | 17 | - |
| $J_{Upper}$ | 87 | 430 | 5.13 | 6.6 | 3.3 | 488375829.9 | 147992675.7 | 0.783041129 | 14 | - |
| $J_{Upper}$ | 86 | 430 | 5.05 | 7.2 | 3.6 | 680390356.5 | 188997321.3 | 1.0 | 19 | - |
| $J_{Upper}$ | 85 | 426 | 5.69 | 0.13 | 0.065 | 553214.393 | 8510990.662 | 0.045032335 | - | 16 |
| $J_{Upper}$ | 84 | 426 | 4.79 | 0.27 | 0.135 | 403348.262 | 2987764.904 | 0.015808504 | - | 17 |
| $J_{Upper}$ | 83 | 426 | 5.55 | 0.3 | 0.15 | 138347.242 | 922314.9467 | 0.004880042 | - | 17 |
| $J_{Upper}$ | 82 | 426 | 5.58 | 0.33 | 0.165 | 1242800.388 | 7532123.564 | 0.03985307 | - | - |
| 10 | 81 | 251 | 5.35 | 8.1 | 4.05 | 25562885.69 | 6311823.628 | 0.033396366 | 19 | - |
| 10 | 80 | 251 | 5.17 | 10.3 | 5.15 | 43956031.16 | 8535151.682 | 0.045160173 | 24 | - |
| 10 | 79 | 251 | 5.74 | 8.1 | 4.05 | 24522332.24 | 6054896.85 | 0.032036945 | 15 | - |
| 10 | 78 | 251 | 5.24 | 9.9 | 4.95 | 33450148.29 | 6757605.715 | 0.035755034 | 21 | - |
| 10 | 77 | 251 | 5.48 | 8.6 | 4.3 | 34478110.89 | 8018165.323 | 0.042424756 | 28 | - |
| 10 | 76 | 251 | 5.28 | 8.9 | 4.45 | 18172654.76 | 4083742.643 | 0.02160741 | 27 | - |
| 10 | 75 | 251 | 5.3 | 48.3 | 24.15 | 42084023.69 | 1742609.676 | 0.009220288 | 26 | - |
| $F_{Upper}$ | 74 | 200 | 4.79 | 8.97 | 4.485 | 395434.831 | 88168.30123 | 0.000466506 | 36 | 57 |
| $F_{Upper}$ | 73 | 200 | 4.67 | 15.23 | 7.615 | 1235944.969 | 162304.0012 | 0.000858764 | 49 | 47 |
| $F_{Upper}$ | 72 | 200 | 4.76 | 4.2 | 2.1 | 80567.791 | 38365.61476 | 0.000202996 | 6 | 68 |
| $F_{Upper}$ | 71 | 200 | 4.95 | 8.57 | 4.285 | 284536.617 | 66402.94446 | 0.000351343 | 39 | - |
| $F_{Upper}$ | 70 | 200 | 5.29 | 9.5 | 4.75 | 10767680.09 | 2266880.019 | 0.011994244 | 25 | 29 |
| $F_{Upper}$ | 69 | 200 | 5.23 | 7.8 | 3.9 | 56894657.19 | 14588373.64 | 0.077188256 | 39 | 29 |
| $F_{Upper}$ | 68 | 200 | 5.01 | 4.12 | 2.06 | 21110770.18 | 10247946.69 | 0.054222709 | 29 | 37 |
| $F_{Upper}$ | 67 | 200 | 5.34 | 7.8 | 3.9 | 114989979.1 | 29484610.02 | 0.156005439 | 29 | - |
| $F_{Middle}$ | 66 | 199 | 5.5 | 13.6 | 6.8 | 116704727.8 | 17162459.97 | 0.090807954 | 37 | 39 |
| $F_{Middle}$ | 65 | 199 | 4.97 | 15.3 | 7.65 | 300582712.8 | 39291857.88 | 0.207896375 | 16 | 25 |
| $F_{Middle}$ | 64 | 199 | 5.49 | 18.15 | 9.075 | 62517011.62 | 6888926.901 | 0.036449865 | 64 | 39 |
| $F_{Middle}$ | 63 | 199 | 4.86 | 9.63 | 4.815 | 154536148.8 | 32094734.95 | 0.169815819 | 19 | - |
| $F_{Lower}$ | 62 | 198 | 5.58 | 7.32 | 3.66 | 35374351.49 | 9665123.358 | 0.051138944 | 57 | 16 |
| $F_{Lower}$ | 61 | 198 | 5.12 | 10.76 | 5.38 | 44262234.74 | 8227181.179 | 0.043530676 | 47 | 64 |
| $F_{Lower}$ | 60 | 198 | 5.54 | 14.16 | 7.08 | 50448336.04 | 7125471.192 | 0.03770144 | 68 | 19 |
| $F_{Lower}$ | 59 | 198 | 5.24 | 8.6 | 4.3 | 23107831.71 | 5373914.352 | 0.028433812 | 32 | - |
| $F_{Lower}$ | 58 | 198 | 5.4 | 4.52 | 2.26 | 56601504.71 | 25044913.59 | 0.132514648 | 40 | - |
| $F_{Lower}$ | 57 | 198 | 5.3 | 19.28 | 9.64 | 40057517.16 | 4155344.104 | 0.021986259 | 38 | - |
| $F_{Lower}$ | 56 | 198 | 5.41 | 8.1 | 4.05 | 53654147.77 | 13247937.72 | 0.070095902 | 27 | - |
| $F_{Lower}$ | 55 | 198 | 5.66 | 7.63 | 3.815 | 625891.081 | 164060.5717 | 0.000868058 | 49 | - |
| $F_{Lower}$ | 54 | 198 | 5.94 | 15.9 | 7.95 | 9367475.417 | 1178298.795 | 0.006234474 | - | |
| $F_{Lower}$ | 53 | 198 | 5.32 | 15.07 | 7.535 | 90034734.48 | 11948869.87 | 0.06322243 | - | 49 |
| $F_{Lower}$ | 52 | 198 | 5.12 | 17.68 | 8.84 | 11885511.27 | 1344514.849 | 0.007113936 | - | 6 |
| $F_{Lower}$ | 51 | 198 | 5.69 | 6.3 | 3.15 | 18963930.3 | 6020295.334 | 0.031853866 | - | - |
| D | 50 | 145 | 5.07 | 12.6 | 6.3 | 14923641.94 | 2368832.054 | 0.012533681 | 85 | - |
| D | 49 | 145 | 5.17 | 15.2 | 7.6 | 9438631.101 | 1241925.145 | 0.006571126 | 37 | - |
| D | 48 | 145 | 5.05 | 14.1 | 7.05 | 9543887.057 | 1353742.845 | 0.007162762 | 56 | - |
| D | 47 | 145 | 5.02 | 12.8 | 6.4 | 10549335.59 | 1648333.686 | 0.008721466 | 93 | - |
| D | 46 | 145 | 5.04 | 9.2 | 4.6 | 10015754.18 | 2177337.865 | 0.011520469 | 37 | - |
| $C_{Upper}$ | 45 | 118 | 5.72 | 8.15 | 4.075 | 2618063.09 | 642469.4699 | 0.003399358 | - | 11 |
| $C_{Upper}$ | 44 | 118 | 5.76 | 6.2 | 3.1 | 1793050.428 | 578403.3639 | 0.003060379 | - | 12 |
| $C_{Upper}$ | 43 | 118 | 5.71 | 6.1 | 3.05 | 1694312.3 | 555512.2295 | 0.00293926 | - | 24 |
| $C_{Upper}$ | 42 | 118 | 5.34 | 12.55 | 6.275 | 0.0 | 0.0 | 0.0 | 21 | - |
| $C_{Middle}$ | 41 | 114 | 5.25 | 7.44 | 3.72 | 8635626.028 | 2321404.846 | 0.012282739 | 14 | 28 |
| $C_{Middle}$ | 40 | 114 | 5.54 | 5.03 | 2.515 | 1830624.012 | 727882.3109 | 0.003851284 | 11 | 21 |
| $C_{Middle}$ | 39 | 114 | 4.93 | 8.03 | 4.015 | 464406.564 | 115667.8864 | 0.000612008 | 12 | 14 |
| $C_{Middle}$ | 38 | 114 | 5.31 | 7.0 | 3.5 | 2062418.14 | 589262.3257 | 0.003117834 | 24 | - |
| $C_{Lower}$ | 37 | 109 | 5.15 | 20.1 | 10.05 | 836094.94 | 83193.52637 | 0.000440184 | 57 | - |
| $C_{Lower}$ | 36 | 109 | 5.04 | 19.4 | 9.7 | 121464.98 | 12522.16289 | 6.63e-05 | 56 | - |
| $C_{Lower}$ | 35 | 109 | 5.04 | 18.2 | 9.1 | 998027.983 | 109673.4047 | 0.000580291 | 61 | - |
| $C_{Lower}$ | 34 | 109 | 5.02 | 16.3 | 8.15 | 426899.091 | 52380.25656 | 0.000277148 | 49 | - |
| $C_{Lower}$ | 33 | 109 | 5.01 | 17.7 | 8.85 | 300801.056 | 33988.81989 | 0.000179838 | 49 | - |
| $C_{Lower}$ | 32 | 109 | 4.7 | 6.67 | 3.335 | 1061314.514 | 318235.2366 | 0.001683808 | 47 | 47 |
| $C_{Lower}$ | 31 | 109 | 4.94 | 8.68 | 4.34 | 122841.23 | 28304.43088 | 0.000149761 | 51 | 51 |
| $C_{Lower}$ | 30 | 109 | 5.77 | 4.93 | 2.465 | 842180.308 | 341655.2974 | 0.001807726 | 54 | 54 |
| $C_{Lower}$ | 29 | 109 | 5.09 | 4.28 | 2.14 | 1483597.644 | 693269.9271 | 0.003668147 | 28 | - |
| $B_{Upper}$) | 28 | 63 | 5.04 | 27.3 | 13.65 | 0.0 | 0.0 | 0.0 | - | - |
| $B_{Upper}$) | 27 | 63 | 5.04 | 24.8 | 12.4 | 0.0 | 0.0 | 0.0 | - | - |
| $B_{Upper}$) | 26 | 63 | 5.78 | 30.8 | 15.4 | 0.0 | 0.0 | 0.0 | - | - |
| $B_{Upper}$) | 25 | 63 | 5.13 | 24.6 | 12.3 | 0.0 | 0.0 | 0.0 | - | - |
| $B_{Upper}$) | 24 | 63 | 5.57 | 30.2 | 15.1 | 0.0 | 0.0 | 0.0 | - | - |
| $B_{Lower}$ | 23 | 62 | 5.31 | 20.3 | 10.15 | 458133.157 | 45136.27163 | 0.00023882 | 23 | - |
| $B_{Lower}$ | 22 | 62 | 5.16 | 14.6 | 7.3 | 743438.665 | 101840.913 | 0.000538848 | 29 | - |
| $B_{Lower}$ | 21 | 62 | 5.06 | 15.9 | 7.95 | 1197096.447 | 150578.1694 | 0.000796721 | 29 | - |
| $B_{Lower}$ | 20 | 62 | 5.01 | 18.2 | 9.1 | 1516570.029 | 166656.0471 | 0.000881791 | 27 | - |
| $B_{Lower}$ | 19 | 62 | 5.14 | 13.4 | 6.7 | 1191744.715 | 177872.3455 | 0.000941137 | 27 | - |
| d/e | 18 | 14 | 4.92 | 3.7 | 1.85 | 118858.636 | 64247.91135 | 0.000339941 | - | 19 |
| d/e | 17 | 14 | 4.71 | 4.13 | 2.065 | 6824.849 | 3305.011622 | 1.75e-05 | - | 14 |
| d/e | 16 | 14 | 5.83 | 2.93 | 1.465 | 295885.684 | 201969.7502 | 0.001068638 | - | 15 |
| d/e | 15 | 14 | 5.2 | 4.48 | 2.24 | 318038.652 | 141981.5411 | 0.000751236 | - | - |
| d/e | 14 | 14 | 5.05 | 19.5 | 9.75 | 0.0 | 0.0 | 0.0 | 15 | - |
| d/e | 13 | 14 | 5.01 | 18.3 | 9.15 | 0.0 | 0.0 | 0.0 | 29 | - |
| d/e | 12 | 14 | 5.19 | 17.1 | 8.55 | 0.0 | 0.0 | 0.0 | 29 | - |
| d/e | 11 | 14 | 5.0 | 10.3 | 5.15 | 22339.94 | 4337.852427 | 2.3e-05 | 13 | - |
| d/e | 10 | 14 | 5.12 | 11.4 | 5.7 | 11135.441 | 1953.58614 | 1.03e-05 | 27 | - |
| A | 9 | 0 | 5.56 | 9.55 | 4.775 | 0.0 | 0.0 | 0.0 | - | 43 |
| A | 8 | 0 | 5.91 | 9.05 | 4.525 | 0.0 | 0.0 | 0.0 | - | 37 |
| A | 7 | 0 | 5.06 | 16.95 | 8.475 | 0.0 | 0.0 | 0.0 | - | 50 |
| A | 6 | 0 | 5.22 | 14.52 | 7.26 | 0.0 | 0.0 | 0.0 | - | - |
| A | 5 | 0 | 5.02 | 9.1 | 4.55 | 5231.107 | 1149.693846 | 6.08e-06 | 34 | - |
| A | 4 | 0 | 5.08 | 7.4 | 3.7 | 15878.085 | 4291.374324 | 2.27e-05 | 43 | - |
| A | 3 | 0 | 5.03 | 8.6 | 4.3 | 334052.991 | 77686.74209 | 0.000411047 | 38 | - |
| A | 2 | 0 | 4.97 | 10.2 | 5.1 | 325771.167 | 63876.69941 | 0.000337977 | 57 | - |
| A | 1 | 0 | 5.24 | 8.3 | 4.15 | 325501.564 | 78434.11181 | 0.000415001 | 37 | - |

Table B.3: Barmur Group lignite sample information. Sample mass analysed (g) and Total Lipid Extract (TLE - mg). TLE analysed represents the mass of TLE used for silica gel chromatography after splitting. Raw HPLC-MS signal represents the sum of all brGDGT fractions, which is then adjusted by the analysed TLE amounts (signal/TLE). The adjusted HPLC-MS signal is then normalised to the largest value. Also shown, Total Organic Matter (TOM) and Total Organic Carbon (TOC)

| Sample Information | | n-Alkane [%] | | | | | | | | | | | | | | | | | Total n-Alkane | CPI |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Bed | ID | $C_{19}$ | $C_{20}$ | $C_{21}$ | $C_{22}$ | $C_{23}$ | $C_{24}$ | $C_{25}$ | $C_{26}$ | $C_{27}$ | $C_{28}$ | $C_{29}$ | $C_{30}$ | $C_{31}$ | $C_{32}$ | $C_{33}$ | $C_{34}$ | $C_{35}$ | | |
| $J_{Upper}$ | 92 | 0.062 631 6 | 5.107 310 0 | 0.004 704 3 | 0.054 736 2 | 0.040 725 8 | 0.007 781 1 | 0.014 825 3 | 0.064 824 3 | 0.345 277 0 | 0.018 522 3 | 0.215 994 0 | 0.028 826 8 | 0.044 512 7 | 0.004 627 4 | 0.075 693 8 | 0.001 440 4 | 0.014 825 9 | 1934739998 | 5.738 976 4 |
| $J_{Upper}$ | 91 | 0.003 414 9 | 0.017 072 2 | 0.017 638 5 | 0.039 330 5 | 0.101 925 2 | 0.021 523 8 | 0.023 625 7 | 0.036 595 5 | 0.328 192 7 | 0.028 581 8 | 0.143 837 9 | 0.025 325 5 | 0.159 777 1 | 0.002 049 5 | 0.040 107 9 | 0.002 802 6 | 0.008 198 7 | 1981864382 | 6.695 702 4 |
| $J_{Upper}$ | 90 | 0.019 064 7 | 0.012 335 4 | 0.002 429 3 | 0.079 986 4 | 0.007 876 4 | 0.081 629 5 | 0.204 634 1 | 0.012 497 3 | 0.219 443 4 | 0.030 977 1 | 0.173 577 3 | 0.019 951 7 | 0.002 434 5 | 0.013 497 4 | 0.100 958 3 | 0.003 999 2 | 0.014 708 1 | 1987490325 | 6.542 352 9 |
| $J_{Upper}$ | 89 | 0.044 514 0 | 0.032 607 3 | 0.109 857 0 | 0.026 275 9 | 0.010 658 4 | 0.073 324 0 | 0.078 044 0 | 0.003 771 6 | 0.212 583 1 | 0.041 216 4 | 0.115 110 8 | 0.002 432 9 | 0.167 492 1 | 0.016 446 1 | 0.055 329 4 | 0.002 841 0 | 0.007 495 9 | 2286299352 | 7.002 084 3 |
| 10 | 80 | 0.011 745 3 | 0.003 068 1 | 0.050 606 8 | 0.003 742 9 | 0.129 717 0 | 0.027 184 2 | 0.211 832 2 | 0.035 979 6 | 0.098 751 3 | 0.027 349 8 | 0.296 088 1 | 0.000 440 4 | 0.046 847 8 | 0.004 657 3 | 0.046 846 6 | 0.002 664 9 | 0.002 477 7 | 2846957785 | 8.588 336 8 |
| 10 | 79 | 0.009 474 1 | 0.015 554 4 | 0.050 098 7 | 0.049 752 3 | 0.127 228 9 | 0.015 888 0 | 0.273 723 0 | 0.018 245 3 | 0.187 786 6 | 0.031 917 2 | 0.083 522 3 | 0.024 367 0 | 0.088 847 3 | 0.004 456 6 | 0.009 330 5 | 0.003 313 8 | 0.006 494 2 | 2552599072 | 7.297 538 6 |
| 10 | 78 | 0.045 160 8 | 0.006 921 2 | 0.063 254 2 | 0.080 257 1 | 0.071 115 0 | 0.078 121 8 | 0.189 185 1 | 0.001 554 7 | 0.181 599 3 | 0.035 577 5 | 0.150 563 1 | 0.005 319 6 | 0.027 627 1 | 0.000 301 0 | 0.057 328 5 | 0.002 089 4 | 0.004 024 5 | 1879631440 | 9.268 382 0 |
| $F_{Upper}$ | 74 | 0.008 074 4 | 0.000 520 0 | 0.048 890 4 | 0.025 619 5 | 0.195 762 7 | 0.068 998 9 | 0.007 270 6 | 0.036 777 5 | 0.035 077 0 | 0.003 266 6 | 0.351 697 7 | 0.015 613 7 | 0.166 577 4 | 0.004 416 2 | 0.025 000 2 | 0.001 887 3 | 0.004 549 8 | 2135098782 | 6.994 276 4 |
| $F_{Upper}$ | 73 | 0.006 783 0 | 0.030 024 3 | 0.050 918 4 | 0.001 033 9 | 0.015 720 7 | 0.019 581 6 | 0.190 183 7 | 0.060 523 2 | 0.203 860 5 | 0.011 612 2 | 0.295 363 4 | 0.007 730 4 | 0.039 276 1 | 0.002 123 4 | 0.059 712 8 | 0.000 062 8 | 0.005 489 7 | 2524385013 | 8.685 274 3 |
| $F_{Upper}$ | 72 | 0.031 843 0 | 0.021 624 1 | 0.010 631 5 | 0.013 515 0 | 0.073 137 5 | 0.024 088 2 | 0.308 363 7 | 0.038 456 7 | 0.021 518 2 | 0.022 116 0 | 0.275 624 6 | 0.015 111 1 | 0.126 436 5 | 0.007 203 6 | 0.001 142 3 | 0.002 540 9 | 0.006 647 1 | 2287459369 | 7.717 055 9 |
| $F_{Upper}$ | 71 | 0.022 066 3 | 0.026 486 7 | 0.060 566 8 | 0.013 257 8 | 0.027 916 7 | 0.021 579 5 | 0.156 014 1 | 0.035 150 7 | 0.402 798 0 | 0.027 982 5 | 0.026 866 1 | 0.007 643 5 | 0.087 498 8 | 0.000 369 3 | 0.071 356 8 | 0.003 066 2 | 0.009 380 1 | 3151965298 | 9.030 955 7 |
| $F_{Lower}$ | 62 | 0.001 205 5 | 0.034 867 7 | 0.033 671 7 | 0.068 513 9 | 0.083 042 0 | 0.010 309 1 | 0.110 928 1 | 0.036 201 8 | 0.171 215 5 | 0.022 576 8 | 0.115 993 3 | 0.009 056 9 | 0.180 929 7 | 0.003 423 2 | 0.099 665 5 | 0.004 716 0 | 0.013 683 2 | 2505173117 | 8.627 363 9 |
| $F_{Lower}$ | 61 | 0.037 143 2 | 0.001 604 0 | 0.022 619 8 | 0.026 605 9 | 0.106 464 6 | 0.003 890 1 | 0.118 091 6 | 0.052 202 9 | 0.270 922 3 | 0.003 682 7 | 0.121 505 2 | 0.017 300 0 | 0.175 926 5 | 0.008 101 7 | 0.028 462 5 | 0.002 579 7 | 0.002 897 4 | 2290881515 | 8.458 729 5 |
| $F_{Lower}$ | 60 | 0.014 428 0 | 0.004 769 0 | 0.022 306 4 | 0.035 722 5 | 0.039 980 1 | 0.021 513 7 | 0.204 595 4 | 0.006 302 7 | 0.235 186 9 | 0.054 159 8 | 0.134 752 0 | 0.007 772 3 | 0.178 415 2 | 0.004 259 9 | 0.034 772 4 | 0.000 028 7 | 0.001 035 1 | 3390246538 | 9.620 446 3 |
| $F_{Lower}$ | 59 | 0.001 592 8 | 0.048 924 4 | 0.065 429 7 | 0.020 426 8 | 0.001 275 9 | 0.069 390 6 | 0.111 702 1 | 0.026 415 4 | 0.232 382 7 | 0.017 570 1 | 0.212 555 4 | 0.009 998 5 | 0.111 802 9 | 0.002 670 0 | 0.057 937 2 | 0.001 370 7 | 0.008 554 7 | 2337168036 | 9.140 675 4 |
| $F_{Lower}$ | 58 | 0.008 009 7 | 0.023 662 0 | 0.066 387 6 | 0.012 593 0 | 0.020 768 1 | 0.021 231 6 | 0.055 021 1 | 0.051 880 7 | 0.006 137 5 | 0.021 292 4 | 0.384 634 3 | 0.001 789 7 | 0.193 183 0 | 0.005 312 7 | 0.118 203 3 | 0.002 264 3 | 0.007 629 2 | 2605341631 | 8.316 436 9 |
| $F_{Lower}$ | 57 | 0.040 348 6 | 0.006 248 4 | 0.024 825 9 | 0.036 990 0 | 0.062 050 2 | 0.015 433 7 | 0.112 944 7 | 0.040 855 8 | 0.228 964 7 | 0.015 943 6 | 0.222 395 6 | 0.020 503 6 | 0.095 370 5 | 0.000 503 2 | 0.071 377 8 | 0.001 779 3 | 0.003 464 3 | 3416806953 | 8.513 154 7 |
| D | 49 | 0.005 096 2 | 0.007 682 8 | 0.025 634 4 | 0.027 651 0 | 0.080 418 4 | 0.035 760 7 | 0.106 927 5 | 0.059 977 8 | 0.296 066 4 | 0.037 486 1 | 0.182 546 9 | 0.009 348 8 | 0.070 411 5 | 0.005 003 1 | 0.041 774 4 | 0.002 601 2 | 0.005 613 0 | 1022464579 | 5.413 006 0 |
| D | 46 | 0.004 524 9 | 0.005 954 1 | 0.018 121 6 | 0.022 092 3 | 0.073 800 3 | 0.031 541 4 | 0.110 129 3 | 0.053 170 1 | 0.277 391 1 | 0.035 504 4 | 0.188 382 0 | 0.008 220 1 | 0.098 260 7 | 0.005 188 6 | 0.058 791 0 | 0.002 298 6 | 0.006 629 5 | 1364371356 | 6.253 516 1 |
| $C_{Lower}$ | 37 | 0.006 934 0 | 0.008 274 9 | 0.018 403 2 | 0.014 844 1 | 0.045 710 3 | 0.021 126 8 | 0.113 772 6 | 0.033 113 4 | 0.219 816 8 | 0.035 139 7 | 0.192 056 2 | 0.017 250 8 | 0.155 397 3 | 0.012 002 9 | 0.095 703 1 | 0.003 210 7 | 0.007 243 2 | 3830688738 | 7.129 777 5 |
| $C_{Lower}$ | 36 | 0.006 692 7 | 0.008 382 1 | 0.013 127 6 | 0.013 070 0 | 0.036 416 2 | 0.020 562 7 | 0.096 154 6 | 0.034 106 3 | 0.219 063 3 | 0.041 200 4 | 0.226 816 5 | 0.019 322 7 | 0.158 422 7 | 0.011 713 4 | 0.085 035 1 | 0.003 677 5 | 0.006 236 5 | 1553634081 | 6.664 558 0 |
| $C_{Lower}$ | 35 | 0.007 914 9 | 0.011 506 5 | 0.024 142 4 | 0.020 351 1 | 0.051 295 5 | 0.025 186 8 | 0.176 396 8 | 0.033 585 6 | 0.262 033 9 | 0.030 673 6 | 0.177 071 9 | 0.012 230 8 | 0.097 229 4 | 0.007 204 0 | 0.055 735 4 | 0.002 589 2 | 0.004 852 4 | 1816565225 | 7.982 116 0 |
| $C_{Lower}$ | 34 | 0.009 136 7 | 0.011 917 8 | 0.019 908 5 | 0.018 261 0 | 0.056 101 9 | 0.022 683 1 | 0.095 658 7 | 0.032 774 1 | 0.184 491 0 | 0.038 851 2 | 0.213 908 4 | 0.019 140 6 | 0.166 619 4 | 0.012 215 0 | 0.087 506 3 | 0.003 401 8 | 0.007 424 5 | 2068725689 | 6.493 391 1 |
| $C_{Lower}$ | 33 | 0.010 723 1 | 0.012 827 2 | 0.019 449 4 | 0.017 720 8 | 0.045 762 2 | 0.023 816 3 | 0.105 750 6 | 0.037 295 2 | 0.221 093 5 | 0.038 441 9 | 0.216 052 0 | 0.016 833 0 | 0.142 526 7 | 0.010 086 0 | 0.072 636 4 | 0.003 131 1 | 0.005 854 5 | 1606267661 | 6.579 881 4 |
| $B_{Lower}$ | 23 | 0.037 482 8 | 0.026 496 4 | 0.022 665 1 | 0.013 588 4 | 0.001 187 8 | 0.010 399 9 | 0.029 892 5 | 0.062 799 0 | 0.229 453 4 | 0.048 544 0 | 0.303 542 6 | 0.012 562 4 | 0.090 758 5 | 0.004 057 7 | 0.100 867 7 | 0.000 169 0 | 0.005 532 7 | 8180371841 | 5.670 859 1 |
| $B_{Lower}$ | 22 | 0.146 615 9 | 0.007 275 8 | 0.088 282 1 | 0.021 885 5 | 0.079 647 2 | 0.050 100 9 | 0.014 645 3 | 0.022 728 0 | 0.338 256 9 | 0.036 914 3 | 0.054 789 4 | 0.011 149 1 | 0.112 517 3 | 0.001 984 8 | 0.005 519 6 | 0.001 822 0 | 0.005 866 0 | 3476665266 | 5.662 981 0 |
| $B_{Lower}$ | 21 | 0.028 517 4 | 0.043 389 6 | 0.027 408 2 | 0.147 270 8 | 0.053 578 5 | 0.054 176 5 | 0.191 533 8 | 0.019 079 6 | 0.021 783 5 | 0.042 589 3 | 0.206 079 9 | 0.004 312 8 | 0.051 220 8 | 0.014 043 0 | 0.070 843 3 | 0.003 221 0 | 0.020 952 1 | 1038622431 | 5.269 537 7 |
| $B_{Lower}$ | 20 | 0.002 021 1 | 0.006 632 4 | 0.015 420 5 | 0.113 351 9 | 0.105 245 5 | 0.037 057 2 | 0.177 694 6 | 0.057 995 2 | 0.163 490 5 | 0.014 460 6 | 0.118 812 2 | 0.009 772 8 | 0.095 404 7 | 0.010 816 9 | 0.069 083 5 | 0.000 282 7 | 0.002 457 9 | 5650242187 | 5.745 619 3 |
| $B_{Lower}$ | 19 | 0.007 499 2 | 0.038 826 9 | 0.036 984 4 | 0.021 051 0 | 0.009 069 1 | 0.102 575 1 | 0.173 341 6 | 0.028 347 3 | 0.233 427 1 | 0.057 039 9 | 0.130 185 3 | 0.014 247 6 | 0.092 142 8 | 0.003 472 6 | 0.041 060 6 | 0.002 383 8 | 0.008 345 8 | 4860999595 | 4.805 476 0 |
| d/e | 14 | 0.006 501 6 | 0.011 972 4 | 0.018 443 8 | 0.021 534 3 | 0.035 807 5 | 0.037 725 9 | 0.162 678 8 | 0.062 388 1 | 0.309 076 5 | 0.051 211 0 | 0.175 039 5 | 0.012 787 0 | 0.044 924 0 | 0.007 512 2 | 0.022 456 4 | 0.006 549 8 | 0.013 391 2 | 1095727170 | 4.623 125 0 |
| d/e | 13 | 0.010 419 9 | 0.014 693 1 | 0.020 844 0 | 0.024 795 0 | 0.042 524 7 | 0.043 888 4 | 0.173 734 1 | 0.068 295 1 | 0.298 063 6 | 0.049 722 1 | 0.138 088 0 | 0.011 560 1 | 0.046 854 7 | 0.007 861 7 | 0.025 161 6 | 0.007 567 2 | 0.015 926 7 | 1488948525 | 4.231 590 2 |
| d/e | 12 | 0.004 807 5 | 0.009 608 0 | 0.015 130 4 | 0.015 369 4 | 0.032 588 2 | 0.038 915 4 | 0.125 516 1 | 0.064 354 5 | 0.275 222 5 | 0.056 288 0 | 0.207 867 1 | 0.016 511 2 | 0.077 531 0 | 0.009 026 2 | 0.030 154 8 | 0.007 418 2 | 0.013 691 5 | 362690455 | 4.266 633 0 |
| d/e | 11 | 0.013 807 9 | 0.017 916 8 | 0.023 703 1 | 0.028 141 5 | 0.046 165 6 | 0.044 105 3 | 0.132 701 8 | 0.057 400 7 | 0.263 538 4 | 0.044 424 4 | 0.178 415 2 | 0.013 258 9 | 0.085 867 9 | 0.006 307 2 | 0.029 675 6 | 0.004 998 3 | 0.009 571 5 | 449104355 | 4.815 681 2 |
| d/e | 10 | 0.012 434 4 | 0.017 955 9 | 0.023 621 8 | 0.028 216 0 | 0.044 527 8 | 0.044 288 8 | 0.126 258 3 | 0.061 346 1 | 0.245 503 1 | 0.051 278 4 | 0.190 042 5 | 0.017 896 9 | 0.082 685 4 | 0.008 220 3 | 0.028 892 0 | 0.004 624 9 | 0.012 207 5 | 475075610 | 4.187 991 0 |
| A | 5 | 0.028 035 0 | 0.039 755 9 | 0.063 583 7 | 0.072 519 5 | 0.107 644 8 | 0.078 833 8 | 0.112 298 6 | 0.070 613 1 | 0.160 817 7 | 0.046 338 3 | 0.108 642 2 | 0.013 775 0 | 0.056 988 3 | 0.006 374 7 | 0.026 383 4 | 0.003 322 2 | 0.004 073 9 | 1510083446 | 2.733 186 4 |
| A | 4 | 0.026 063 3 | 0.030 944 2 | 0.045 473 4 | 0.053 784 2 | 0.078 321 3 | 0.064 151 9 | 0.099 181 0 | 0.068 649 1 | 0.204 781 2 | 0.056 327 5 | 0.139 186 1 | 0.023 008 6 | 0.072 489 2 | 0.008 604 2 | 0.022 375 6 | 0.003 596 9 | 0.003 062 4 | 3995343975 | 2.897 987 0 |
| A | 3 | 0.020 497 0 | 0.023 744 0 | 0.034 711 7 | 0.042 015 9 | 0.066 606 0 | 0.054 277 2 | 0.095 459 2 | 0.064 587 2 | 0.221 746 7 | 0.058 452 2 | 0.156 753 4 | 0.024 690 8 | 0.090 179 8 | 0.009 024 9 | 0.030 703 7 | 0.003 003 6 | 0.003 546 7 | 4105986675 | 3.271 058 8 |
| A | 2 | 0.031 896 6 | 0.036 007 0 | 0.053 850 4 | 0.062 027 0 | 0.082 946 6 | 0.073 336 9 | 0.107 451 7 | 0.078 298 8 | 0.177 797 6 | 0.061 562 8 | 0.114 281 8 | 0.021 609 6 | 0.056 121 7 | 0.008 307 3 | 0.025 422 8 | 0.004 015 6 | 0.005 065 9 | 6145298819 | 2.373 435 1 |
| A | 1 | 0.035 121 4 | 0.044 826 7 | 0.064 534 0 | 0.072 540 9 | 0.099 030 7 | 0.079 427 2 | 0.108 629 0 | 0.078 035 9 | 0.154 617 6 | 0.055 258 5 | 0.100 641 9 | 0.017 916 3 | 0.049 836 5 | 0.006 698 1 | 0.022 040 5 | 0.004 529 8 | 0.006 315 2 | 2853930435 | 2.259 359 5 |

Table B.4: *n*-Alkane $C_{19}$–$C_{35}$ relative abundance data of Barmur Group lignite samples. Total sum of *n*-Alkanes $C_{19}$–$C_{35}$ (absolute abundance) and Carbon Preference Index (CPI) also provided.

| Sample Information | | Hopane Isomers [%] | | | Total Hopane | Geohopanoid Index |
|---|---|---|---|---|---|---|
| Bed | ID | C30 $\beta\alpha$ | C30 $\alpha\beta$ | C30 $\beta\beta$ | | |
| $J_{Upper}$ | 92 | 0.027 996 6 | 0.078 303 6 | 0.893 699 8 | 77962173.64 | 0.893 699 8 |
| $J_{Upper}$ | 91 | 0.085 602 3 | 0.092 475 5 | 0.821 922 2 | 329496117.9 | 0.821 922 2 |
| $J_{Upper}$ | 90 | 0.032 056 5 | 0.123 795 2 | 0.844 148 3 | 283104272.8 | 0.844 148 3 |
| $J_{Upper}$ | 89 | 0.012 816 1 | 0.034 538 6 | 0.952 645 3 | 489534312.9 | 0.952 645 3 |
| 10 | 80 | 0.013 620 0 | 0.082 966 3 | 0.903 413 7 | 475552730.3 | 0.903 413 7 |
| 10 | 79 | 0.039 393 0 | 0.134 860 1 | 0.825 746 9 | 331562440.5 | 0.825 746 9 |
| 10 | 78 | 0.015 342 6 | 0.059 489 7 | 0.925 167 8 | 747484831 | 0.925 167 8 |
| $F_{Upper}$ | 74 | 0.015 510 1 | 0.192 140 2 | 0.792 349 8 | 341014435.5 | 0.792 349 8 |
| $F_{Upper}$ | 73 | 0.089 645 7 | 0.054 948 5 | 0.855 405 9 | 191513990.5 | 0.855 405 9 |
| $F_{Upper}$ | 72 | 0.018 105 0 | 0.138 954 8 | 0.842 940 2 | 351040747.5 | 0.842 940 2 |
| $F_{Upper}$ | 71 | 0.061 078 4 | 0.160 450 6 | 0.778 470 9 | 377077539.5 | 0.778 470 9 |
| $F_{Lower}$ | 62 | 0.059 139 5 | 0.011 186 0 | 0.929 674 6 | 84995110.01 | 0.929 674 6 |
| $F_{Lower}$ | 61 | 0.059 139 5 | 0.011 186 0 | 0.929 674 6 | 84995110.01 | 0.929 674 6 |
| $F_{Lower}$ | 60 | 0.146 190 7 | 0.078 037 5 | 0.775 771 8 | 285491088.9 | 0.775 771 8 |
| $F_{Lower}$ | 59 | 0.084 747 6 | 0.019 404 6 | 0.895 847 8 | 81208313.91 | 0.895 847 8 |
| $F_{Lower}$ | 58 | 0.018 514 5 | 0.146 296 6 | 0.835 188 9 | 663485850.9 | 0.835 188 9 |
| $F_{Lower}$ | 57 | 0.059 821 7 | 0.122 918 6 | 0.817 259 8 | 357288270.5 | 0.817 259 8 |
| D | 49 | 0.127 879 2 | 0.078 304 4 | 0.793 816 3 | 35665730.92 | 0.793 816 3 |
| D | 46 | 0.145 766 2 | 0.106 636 6 | 0.747 597 2 | 41901814.13 | 0.747 597 2 |
| $C_{Lower}$ | 37 | 0.093 087 7 | 0.117 712 3 | 0.789 200 1 | 103863033.4 | 0.789 200 1 |
| $C_{Lower}$ | 36 | 0.086 575 3 | 0.083 152 4 | 0.830 272 4 | 53577810.79 | 0.830 272 4 |
| $C_{Lower}$ | 35 | 0.094 198 5 | 0.108 886 6 | 0.796 914 9 | 71883939.31 | 0.796 914 9 |
| $C_{Lower}$ | 34 | 0.087 086 7 | 0.115 342 6 | 0.797 570 7 | 65774640.59 | 0.797 570 7 |
| $C_{Lower}$ | 33 | 0.117 126 9 | 0.099 971 6 | 0.782 901 6 | 63372181.41 | 0.782 901 6 |
| $B_{Lower}$ | 23 | 0.063 702 9 | 0.040 038 9 | 0.896 258 3 | 1139301085 | 0.896 258 3 |
| $B_{Lower}$ | 22 | 0.088 740 8 | 0.066 869 2 | 0.844 390 0 | 817945544.9 | 0.844 390 0 |
| $B_{Lower}$ | 21 | 0.048 964 7 | 0.041 475 1 | 0.909 560 2 | 464485608.8 | 0.909 560 2 |
| $B_{Lower}$ | 20 | 0.098 530 1 | 0.126 474 8 | 0.774 995 1 | 626983992.8 | 0.774 995 1 |
| $B_{Lower}$ | 19 | 0.050 908 3 | 0.149 194 1 | 0.799 897 5 | 382335636.1 | 0.799 897 5 |
| d/e | 14 | 0.141 123 6 | 0.134 044 0 | 0.724 832 4 | 61391640.03 | 0.724 832 4 |
| d/e | 13 | 0.142 689 1 | 0.126 861 5 | 0.730 449 4 | 111975958.6 | 0.730 449 4 |
| d/e | 12 | 0.134 008 7 | 0.146 714 8 | 0.719 276 4 | 24584901.95 | 0.719 276 4 |
| d/e | 11 | 0.150 802 4 | 0.162 732 9 | 0.686 464 7 | 31495224.54 | 0.686 464 7 |
| d/e | 10 | 0.153 056 7 | 0.183 611 1 | 0.663 332 2 | 32620077.63 | 0.663 332 2 |
| A | 5 | 0.356 603 4 | 0.320 560 3 | 0.322 836 3 | 84497928.55 | 0.322 836 3 |
| A | 4 | 0.350 998 5 | 0.277 644 2 | 0.371 357 4 | 233521791.2 | 0.371 357 4 |
| A | 3 | 0.341 099 5 | 0.273 148 8 | 0.385 751 7 | 169559675.3 | 0.385 751 7 |
| A | 2 | 0.334 700 5 | 0.295 083 0 | 0.370 216 5 | 400633497.4 | 0.370 216 5 |
| A | 1 | 0.345 118 0 | 0.330 671 5 | 0.324 210 5 | 170809052.1 | 0.324 210 5 |

Table B.5: Hopane C30 $\beta\alpha$, $\alpha\beta$, $\beta\beta$ relative abundance data of Barmur Group lignite samples. Total sum of all hopane stereoisomers (absolute abundance) and Geohopanoid Index also provided.


Stretch
Stretch
Thank you for submitting
your PhD thesis
Wanna play fetch?