Response to Editor comments.

Dear Dr Abrook et al.,

Thanks for submitting your interesting manuscript to Biogeosciences and for your responses to the reviewer reports. Their feedback is very positive and I agree with their comments, which also comprise some questions that I also had during my personal review.

The only additional question that I have is about how many varves does one biomarker sample typically contain? Not sure if I missed it but how much time comprises the averaged GDGT signal of a sample? It would be good to add this to the manuscript as well. Because I understand that you cannot separate a sub-annual GDGT signature from the varved records, correct?

Other than that, I get a good understanding of how you will address the reviewer suggestions and comments based on your detailed reponses. Therefore, I invite you to submit a revised version of your manuscript which incorporates the requested changes and your responses.

Thank you and I look forward to your revised manuscript.

We would like to thank the editor for the positive comments around the submitted manuscript and for inviting us to revise our manuscript following the review process. We have addressed all reviewers comments in this document (in red) and state where the changes appear in the clean version of the manuscript.

In terms of your comment around the number of varve years per sample we address this here and below in a comment from Reviewer 1:

The 0.5 cm resolution contains multiple varves which change throughout the record and between site locations. All of the lake sites have sub millimetre scale varves throughout, which average 8 – 10 varve years per 0.5 cm sample (over the more heavily sampled area of the core). Unfortunately, we cannot separate a sub-annual GDGT signature from this record as we do not have any varve layers which approach 0.5 in thickness. This type of analysis would be interesting at those sites where thicker varves are preserved.

We have added this distinction to lines 136 and 461.

Response to Reviewer 1.

Dear Sebastian Naeher and Ashley Abrook,

This manuscript presents measurements of iso- and br-GDGTs in three Holocene varved sediment sequences in northern Europe, as well as at higher resolution for the last 300 years for two of the sites, allowing comparison with instrumental records. An impressive amount of data is presented, and the manuscript is well written with good structure. By presenting three new Holocene records, this manuscript is a valuable contribution to paleoclimate knowledge in northern Europe, and the investigation of GDGTs in varved lake sediments is of clear relevance to the proxy community. However, some interpretations, particularly for sparsely sampled late Holocene intervals, are not robustly supported by the data. A deeper discussion of discrepancies between GDGT based reconstructions and instrumental records would also improve the manuscript. Provided that these points and the more detailed comments below are adequately addressed, I see no issues with publication of this manuscript.

We would like to thank the reviewer for their positive review and for noting the 'impressive amount of data' that is presented, and that our manuscript is of 'clear relevance to the proxy community'. We also thank the reviewer for their constructive comments, which will undoubtedly improve the manuscript. Below we provide answers to each of the comments after they appear in red and state how we have addressed them in the clean version of the manuscript.

General comments:

- Why are the last 300-year data not included in the Holocene reconstructions presented in figs 5, 6, 7 and 9? It would be very useful to extend the long records towards the present if possible. If suitable, I would suggest adding it, or clarify in the manuscript why it is not suitable.

  This is a good and very valid point. We left our the last 300-year data from the long-term view as these systems are different in modern settings compared to the Holocene. For example, at Diss Mere the system is not varved from 2ka BP and is presently eutrophic. At Nautajarvi whilst the lake is still accumulating varves in the present day, human impact in the Finnish region of the lake began roughly 2- 2.5 ka BP. So we thought it best to separate out the disturbed / undisturbed parts of the record. Additionally, as you have pointed out, there are some caveats with the Diss Mere record in relation to the instrumental data (Nautajarvi follows instrumental data well).

  Nonetheless we have added the last-300 year data into Figures 5 and 6. We do not have this data for Figure 7 owing to oxygenation treatments in the last disrupting

modern sedimentation. We have not added this to Figure 9 as the goal here was to compare broad patterns in Holocene data.
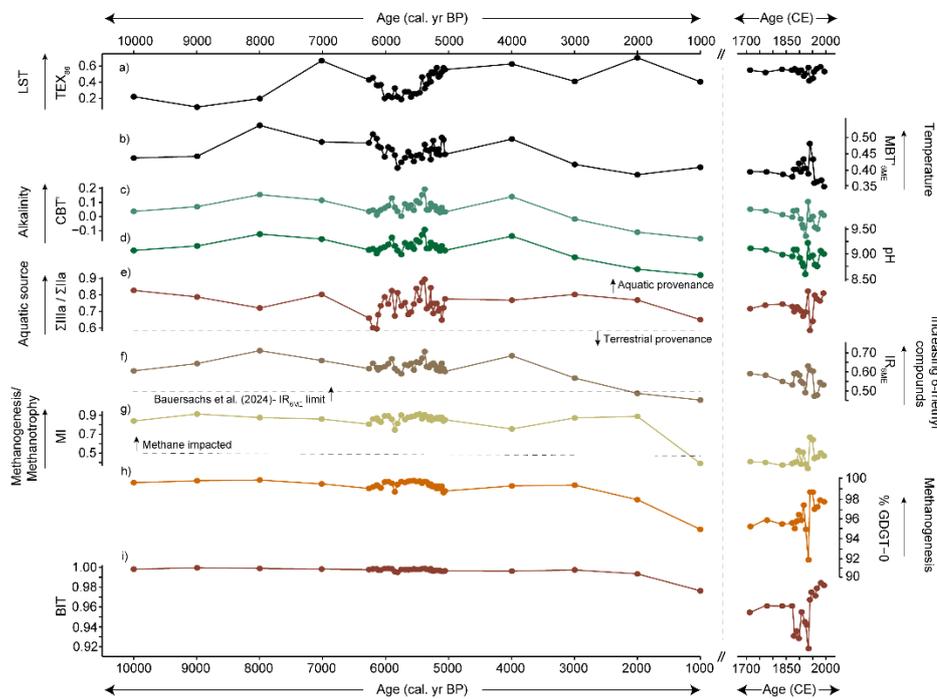
Figures 5 and 6 are now as follows:

Figure 5: Main GDGT reconstructions from Diss Mere. Shown are a) TEX$_{86}$; b) MBT'$_{5ME}$ (De Jonge et al., 2014); c) CBT' (De Jonge et al., 2014); d) reconstructed pH (Russell et al., 2018); e) ΣIIIa / ΣIIa ratios with cut offs from Xiao et al. (2016); f) IR$_{6ME}$ (De Jonge et al., 2014; Bauersachs et al., 2024); g) the methane index (Zhang et al., 2011); h) % GDGT-0 and i) the branched vs isoprenoid tetraether index (BIT).
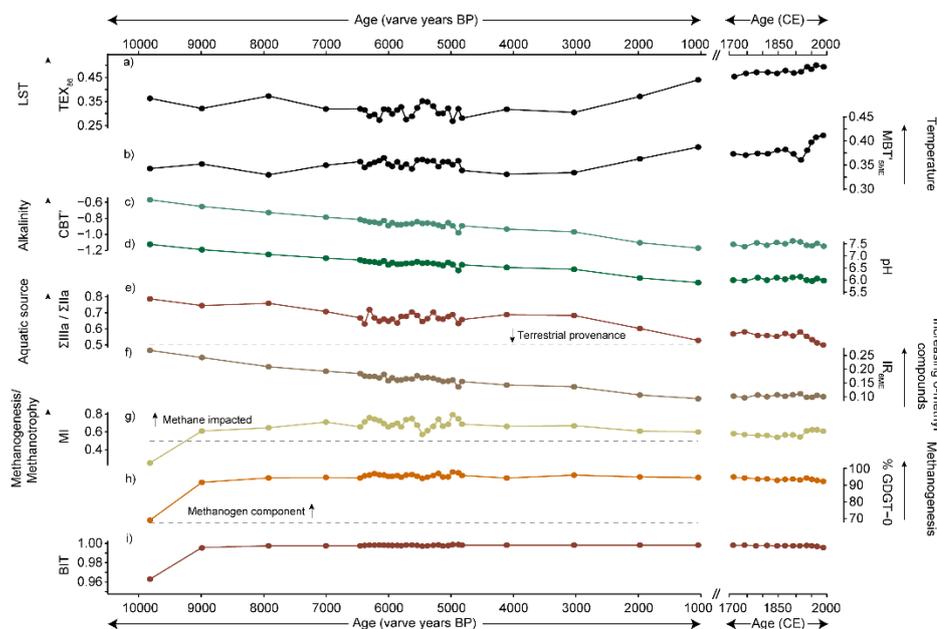

Figure 6: Main GDGT reconstructions from Nautajärvi. Shown are a) TEX$_{86}$; b) MBT'$_{5ME}$ (De Jonge et al., 2014); c) CBT' (De Jonge et al., 2014); d) reconstructed pH (Russell et al., 2018); e) ΣIIIa / ΣIIa ratios with a cut off from Xiao et al. (2016); f) IR$_{6ME}$ (De Jonge et al., 2014); g) the methane index (Zhang et al., 2011); h) % GDGT-0 and i) the branched vs isoprenoid tetraether index (BIT).

- The high resolution recent (last 300 year) data that is compared to instrumental data is very interesting. Since one of the cores follow the instrumental data, and the other does not, I think this section deserves more attention and some further discussion would be useful, as mentioned below. Can the findings from the core-tops be further used to indicate which records are more robust down-core?

  We thank the reviewer for this comment and agree that this deserves more discussion. We demonstrated that brGDGT temperature reconstructions align closely with modern instrumental data at Nautajarvi, and whilst brGDGT reconstructions approximate instrumental data at Diss Mere, there are periods where this alignment is less strong (Section 4.4). We do want to note though that although the alignment is less strong the instrumental data are within calibration error of the reconstruction. At Diss Mere, we consider that the alignment may be less strong due to a number of factors, including sedimentation rate increases, a shift in diatom and blue-green algal communities and changes in catchment vegetation leading to eutrophication from 2 ka BP onwards, especially with intensified human activity over the last 1 kyr (see Boyall et al., 2023; 2024). These changes are driven by an increase in detrital input following continued human occupation (Boyall et al., 2023; 2024).

  We suggest this may have altered the microbial community and could yield slightly different brGDGT temperature estimates (as observed elsewhere; e.g. Russell et al., 2018). We have expanded Section 4.4.1, lines 524-536, to show the different relationships between brGDGT-derived and instrumental temperature data at Nautajarvi and Diss Mere.

- Section 4.5.1 discussed the temperature trends over the Holocene based on the three new records. Conclusions are made based on the changes in trends over the later Holocene, but the records only contain 2-4 data points over this period, leading to very high weight being given to single datapoints. Given that there can be large scatter in GDGT data, this is not robust, and the section should be reworked. Potentially this can be improved by adding the last 300 year data as mentioned above.

  We thank the reviewer for this comment - we acknowledge the sparse nature of some parts of the record. However, we have published independent climate data from these locations (using the same sample) that provide support for our long-term Holocene interpretations. For example, the Holocene calcium curve from Diss Mere (interpreted as reflecting a summer signal in Boyall et al., 2024) evolves in a similar fashion to our brGDGT temperature reconstruction. Similarly, the trends in palynological growing degree day reconstructions from Nautajarvi (Ojala et al., 2008; Lincoln et al., 2025) follows the brGDGT reconstructions over the early- and mid-Holocene. However, there is a clear mismatch in the late-Holocene
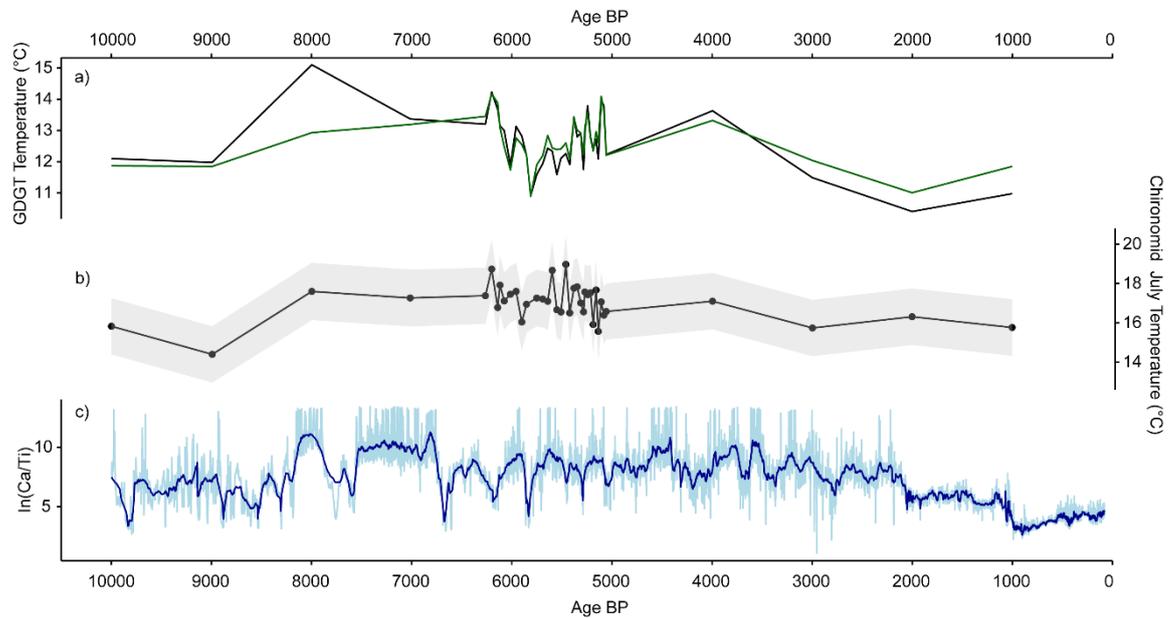
Figure S10. Additional site data from Diss Mere. Shown are GDGT reconstructions from the main body of the manuscript (black, Martinez-Sosa et al., (2021); green, Raberg et al. (2021), chironomid inferred July temperatures (Abrook et al., 2025); and the In(Ca/Ti) ratio representing variability in summer precipitated calcite and detrital inwash (Boyall et al., 2024).
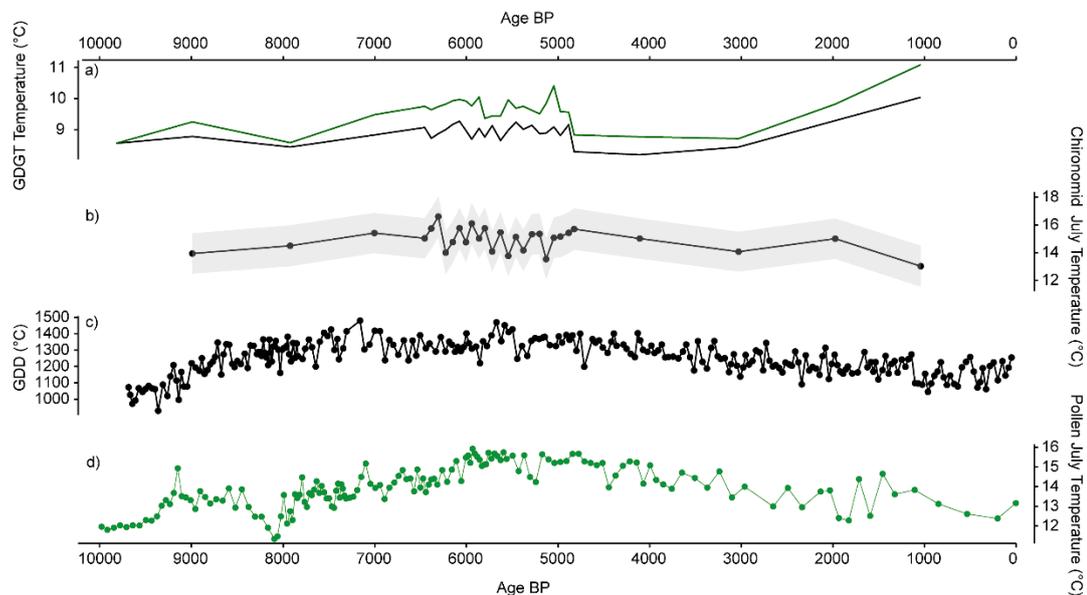
Figure S12. Additional site data from Nautajärvi. Shown are a) GDGT reconstructions from the main body of the manuscript (black, Martinez-Sosa et al., (2021); green, Raberg et al. (2021); b) chironomid inferred July temperatures (Abrook et al., 2025); c) a pollen-based growing degree day (GDD >5 °C) reconstruction (Ojala et al., 2008); and d) a pollen-based reconstruction of July temperatures using a six method ensemble from Lake Kuutsjärvi, Finland (Salonen et al., 2024).

- Is TOC data available from the studied cores? Comparing the GDGT indexes with TOC may be very useful, since organic contents may be an important variable that covaries with GDGT distribution changes. Sudden shifts in TOC contents are strong indicators of environmental and limnological shifts, and are therefore likely to influence also GDGT distributions, see for example Hällberg et al., 2023 (https://doi.org/10.1016/j.orggeochem.2023.104702).

  Unfortunately highly resolved TOC data is not available from these sediments. We only have spot loss-on-ignition data from Nautajarvi. Whilst TOC is absolutely an important additional variable to consider in this type of work with large swings in depositional environment (e.g., peat-to-lake, or soil-to-peat; e.g., see Inglis et al., 2019), our lake sediments are continuously varved across most of the Holocene (until 2ka at Diss Mere). As the style of deposition is rhythmic and stable throughout the Holocene, this implies largely stable TOC values. Whilst the thicknesses of individual varves do change (e.g., Martin-Puertas et al., 2025), this is unlikely to be driven by large shifts in TOC and is more reflective of more features inherent to the climatic season (length/persistence; see Martin-Puertas et al., 2025). Where varves are not preserved (last 2 ka at Diss Mere) TOC is likely to increase, however only two samples in our data cover this period.

  We have added a short discussion on this to Section 4.3.2, lines 485 – 490.

- 7 methyl and GMGT data from these sites would be interesting to see from these sites, if that data is available. It may provide clues to the provenance of the GDGTs as well as potentially providing an additional indication of temperature variability (Baxter et al., 2019, https://doi.org/10.1016/j.gca.2019.05.039). Of course, this manuscript is already long, and 7-methyl brGDGTs and GMGTs should only be included if the authors deem this to have large explanatory power.

  We agree that 7-methyl GDGTs and GMGTs can provide additional insights into the bacterial community, but this is beyond the scope of this manuscript. The primary aim here is to test whether brGDGTs are robust temperature proxies in varved lakes. However, we acknowledge that GMGTs could be evaluated in future studies from these locations. 7-methyl GDGTs are likely to indicate changes in salinity which given the temperate locations of these lakes we do not consider as major contributors.

Specific comments:

- Fig 1:

  o in the B inset, lake depth isobars are not numbered as in a) and c) insets. This would be good to add.

  This has been added to Figure 1.

  o The brown symbols in depositional model for a) is not indicated in the legend what they are.

  Together with the green brush stroke these are organic remains. This has been added to Figure 1.

  o It is also not clear in the figure text why Diss Mere has three panels for the depositional model, so please add that explanation there. What does the light brown/green shading indicate?

  Diss mere has three panels as these are the differences in sediment facies for each of the lamination types. There are three varve types at Diss Mere represented by differences in sub-layer. Light brown / green also represents differences in sub-layer (calcite vs organic). The varve models for Nautajarvi and Meerfelder Maar, are, in contrast, very simple and stable. This explanation has been added to the Figure caption.

  o Letters in the lake insets for a) and b) are not explained in the figure text, and what are yellow markers in b)? If they are important in the manuscript, please explain them in the figure text. If not important, perhaps remove from figure?

  The letters relate to the different coring locations which we have added to the text. The yellow markers are other survey points and are not too important for this manuscript. We have removed the yellow markers. Figure 1 now reads:

  Figure 1: Location of each of the sites in Europe: a) Diss Mere, UK, b) Nautajärvi, Finland; c) Meerfelder Maar, Germany. Site boxes reveal lake bathymetry with depths and inflows/outflows. Varve depositional models are added to each box which demonstrate the main varve facies from each location, with three varve microfacies at Diss Mere and one each at Nautajärvi and Meerfelder Maar. Letters denote different cores from a) and b). Diss Mere adapted from Boyall et al. (2023); Nautajärvi adapted from Lincoln et al. (2025).

- Site description sections. Please provide references and explanation for the meromictic conditions at the lakes already here.

  We have added descriptors to the table that demonstrate mixing regime for each site and provided a reference for each from the lakes.

Table 1: Lake descriptors from each site. * This is the assumed mixing regime for periods of varve formation across each site, at Diss Mere and Meerfelder Maar this is likely different to today: [1] Martin-Puertas et al. (2021); [2] Lincoln et al. (2025); [3] Nürnberg et al. (2007). ** The origin of Diss Mere is debated, with possible origins dating to the Anglian stage (Bailey, 2005).

Lake	Climatology	Lake origin	Mixing regime*	Varve composition
Diss Mere	Maritime	Thermokarst**	Meromictic[1] (no lake water overturn)	Calcite – organic
Nautajärvi	Sub-Arctic	Glacial lake	Dimictic[2] (two lake water overturn periods annually)	Clastic – organic
Meerfelder Maar	Maritime (with continental influences)	Volcanic (maar lake)	Meromictic[3] (no lake water overturn)	Diatomaceous – detrital

- How does the 0.5 cm resolution compare with the varve thickness? It would be interesting to know if it is technically possible to reach sub-annual resolution, and if some samples represent that in this study.

  EDITOR COMMENT- How many varves does one biomarker sample typically contain? How much time comprises the averaged GDGT signal of a sample? It would be good to add this to the manuscript as well. Because I understand that you cannot separate a sub-annual GDGT signature from the varved records, correct?

  The 0.5 cm resolution contains multiple varves which change throughout the record and between site locations. All of the lake sites have sub millimetre scale varves throughout, which average 8 – 10 varve years per 0.5 cm sample (over the more heavily sampled area of the core). Unfortunately, we cannot separate a sub-annual GDGT signature from this record as we do not have any varve layers which approach 0.5 in thickness. This type of analysis would be interesting at those sites where thicker varves are preserved.

  We have added this distinction to lines 136 and 461.

- The manuscript refers to GDGT-0 vs crenarchaeol as %GDGT-0. I would expect that to refer to the fractional abundance of GDGT-0 (relative to all isoGDGTs). Instead, why not refer to it simply as GDGT-0/cren, as frequently done previously, for instance by Baxter et al., 2021 https://doi.org/10.1016/j.quascirev.2021.107263 ? This would make it much clearer what is meant.

  Thank you for this. We use %GDGT-0 as originally defined in Sinninghe Damsté et al. (2012) as this enables simpler inter site comparisons.

  The Baxter et al. (2021) approach is similar but the ratio is theoretically 'limitless' which makes comparisons between locations difficult. For the GDGT-0/cren ratio

a value >2 is indicative of methanogen contribution. For %GDGT-0, a value > 67% is indicative of a methanogen contribution. Regardless of the metric we use, the key conclusions are identical (i.e. each lake is dominated by methanogenic archaea). We have added a statement to this formula stating that is related to GDGT-0/cren, line 184.

- Refer to Hopmans et al., 2004 https://doi.org/10.1016/j.epsl.2004.05.012. Since that original publication is used in marine settings only, I would suggest also referring to a study showing that it can also be used as an aquatic/land signal for terrestrial sites.

  We have added Hopmans et al., 2004 in as a reference for BIT, line 190. We have also added Baxter et al. (2021) to show that it can be used in terrestrial settings, line 190.

- Is this 'wrong' classification correlated with an increased TOC from that core? Since TOC has previously been found to strongly correlate with GDGT distributions elsewhere (see for example fig 3 in Hällberg et al., 2023), I think it would be valuable to show this data, as mentioned earlier, if available.

  Unpublished loss-on-ignition spot samples from Nautajarvi suggest there are no rapid changes in TOC during the Holocene. Crucially, the varve structures also remain constant suggesting no significant changes in TOC content across the Holocene. So whilst TOC is important we do not consider it a driver here. We have added this statement to the manuscript, Section 4.3.2, lines 396 and 485 – 490.

- "*Across all three sites the BIT index is very high and is > 0.92*". A BIT value of 0.92 is a significant change from >0.99 which is normally found in soils/terrestrial sites, and I would therefore suggest that a value of 0.92 likely represents a significantly different GDGT distribution, likely resulting from some environmental change. However, looking at figures 5-7, I do not see any such low values, so perhaps this is a typo?

  Thank you for this observation. All of our 'Holocene' samples (i.e. those from 1ka to 10ka) have high BIT (Diss Mere = >0.98; Nautajarvi = >0.99 (the first sample in the record Is 0.96); Meerfelder Maar = >0.99). It is when we include the 'modern' samples from Diss Mere where this statement comes from. The BIT in the modern samples range from 0.92 to 0.98. We have added the modern sample range to Figures 5 and 6 and have added a statement to line 356 – 359 to clarify this.

  See Figure 5 above.

- Fig 4. What are the "modern" samples in a) and b)? Surface samples?

  The modern samples represent the samples covering the last 300 years. We have clarified that in the figure caption:

Figure 4: Principal components analysis (PCA) left, and redundancy analysis (RDA) right. a) and b) detail the plots from Diss Mere; c) and d) from Nautajärvi; and e) and f) from Meerfelder Maar. a) and c) show modern (last 300-years) and Holocene sample distributions within the PCA plots. Explanatory variables in the RDA plots are major elements from μ-XRF data (from Martin-Puertas et al. (2012); Boyal et al. (2024); Lincoln et al. (2025)) with numerical points showing the median age of samples.

Fig 4. What are the plot labels in the RDA plots? Ages? In that case, I would suggest adding a fill color gradient based on that value, to increase readability.

Yes the plot labels are the sample ages. We attempted to add a colour fill of the individual points to the sample ages but this did not help with readability. For ease we added the following line to the figure caption:

Explanatory variables in the RDA plots are major elements from μ-XRF data (from Martin-Puertas et al. (2012); Boyal et al. (2024); Lincoln et al. (2025)) with numerical points showing the median age of samples.

- Fig 5e. The arrow with "marine provenance" needs to be relabeled, since i assume you mean aquatic rather than marine?

Thank you for this observation. We have changed to aquatic for all figures. See figure 5 above.

- Is the value 0.9 correct? It looks like it's higher based on the figures.

We assume this is in reference to the BIT. This value includes the modern samples as suggested in the comment above. We have changed this to reflect Holocene and modern (>300 years) samples lines 356 – 359:

'Across all three lakes, the BIT index is high for all samples (Holocene and modern) (> 0.9), which implies consistently high input of soil or peat organic matter into the lake'.

- 375 section, on source attribution. Additional evidence for soil input can be derived from degree of cyclization, IR6-methyl and 7 methyl GDGTs as done by Martin et al., 2019. GMGTs and their isomers may also provide further clues and indicate bacterial community shifts, see Hällberg et al., 2023.

Many thanks for these comments on source attribution. We agree that 7-methyl GDGTs and GMGTs can provide additional insights into the bacterial community, but this is beyond the scope of this manuscript and were not routinely identified. We do observe GMGTs in some samples but this will form the basis of a subsequent manuscript. Our primary aim here was to test whether brGDGTs are robust temperature proxies in varved lakes.

We do calculate IR6me values, however we do not have terrestrial soil values as comparators. As we already use three / four approaches to disentangle provenance, which are mostly in agreement with each other, we do not feel IR6me will add too much to these observations.

- Please clarify how it is reflected in fig 4: "*Nautajärvi are classified as 'peat-type' (this is also reflected in Fig. 4)*"

  Thank you for this observation. What we meant by this was that the samples that have different GDGT distributions are grouped together separately from the rest of the Holocene. For ease we have removed 'this is also reflected in Fig. 4'.

- Reference needed.

  We are unsure where in text this refers to.

- It is probably useful to also mention the results of Baxter et al., 2021 (https://doi.org/10.1016/j.quascirev.2021.107263) from meromictic Lake Challa here.

  We have added the Baxter et al., 2024 reference: 'In permanently stratified (meromictic) lakes, pentamethylated and hexamethylated brGDGTs increase in abundance with depth and are associated with anoxic conditions (Weber et al., 2018; Yao et al., 2020; van Bree et al., 2020; Baxter et al., 2024).'

  We aren't 100% sure where you would like this discussed. We added the Baxter et al. 2024 reference as this details brGDGTs and the 2021 reference details only the isoGDGTs. We mention Baxter et al. 2021 in Section 4.1.

- Section 4.3.2, and in particular lines 468-475. CBT' calculation includes 6-methyl brGDGTs, and is therefore a mix of cyclization and isomerization, despite the (perhaps misleading) name *cyclization of branched tetraethers*. It would therefore be better to compare degree of cyclization (Sinninghe Damsté et al., 2009 https://doi.org/10.1016/j.gca.2009.04.022) and IR6me instead of CBT' and IR6me, to disentangle these compound influences.

  Thank you for the comment. In section 4.3.2 we employ CBT' and IR6me to evaluate how the GDGT and XRF data respond to changes in pH. Both metrics have a global, linear relationship with soil pH and thus are deemed most suitable here. Although the degree of cyclisation offers an alternative approach to assess pH (see Table 1; Baxter et al., 2024), it also contains 5- and 6-methyl brGDGTs in its formulation and does not overcome the caveats raised by the reviewer.

  We have altered the text in lines 453 – 456 to better clarify what we were attempting to do with this section.

- Fig 8. The figure text is quite messy.

  - The panels should be a, b, c, d, since it's four panels.

    We have altered the labelling in the panels of Figure 8 to a), b), c), etc as requested. See new figure below:
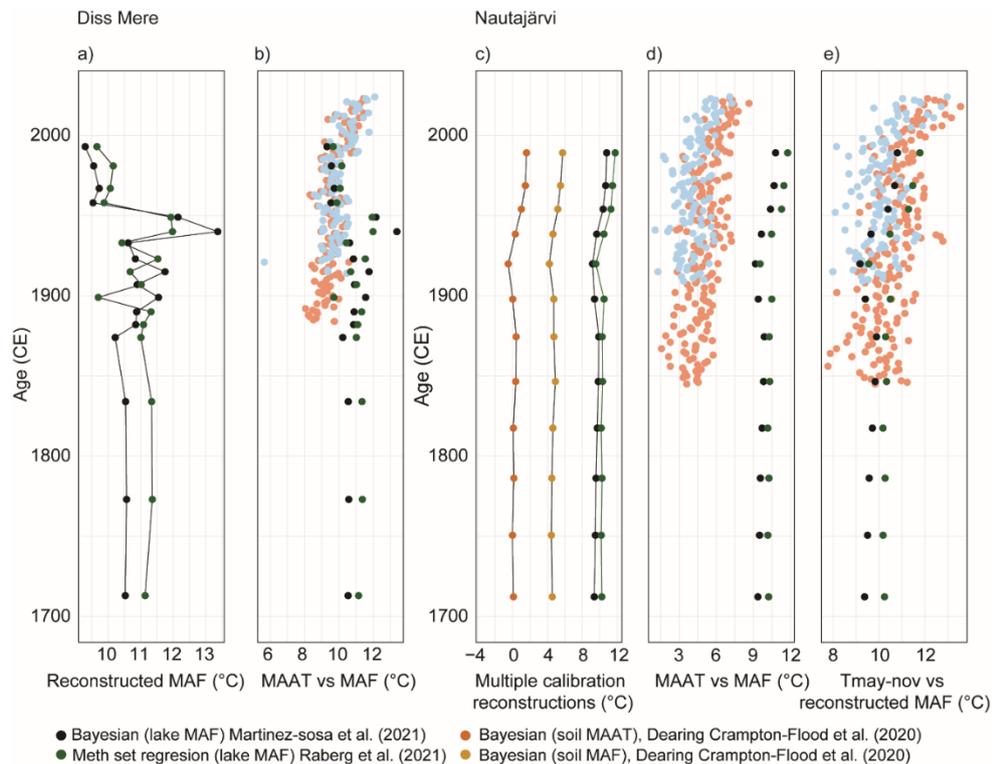
Figure 8: Modern (>300-year) temperature reconstructions from Diss Mere and Nautajärvi. Shown are a) GDGT-based MAF reconstructions at Diss Mere; b) this data alongside modern MAAT instrumental data from both the Lowestoft meteorological station (light blue) and from 5 km gridded temperatures using the HadUK dataset (light orange; Hollis et al., 2019); c) GDGT-based MAAT and MAF reconstructions at Nautajärvi; d) these GDGT-based MAF reconstructions alongside modern MAAT instrumental data from the Heinola Asemantaus (light blue) and Helsinki Kaisaniemi (light orange) meteorological stations; and e) Nautajärvi GDGT reconstructions against re-sampled May-November meteorological station air temperature data.

- o Currently, panel b) is not specified in the text, only a).

  Thank you for this observation. We have edited the figure caption to reflect each panel see above caption.

- o I would propose labelling the panels for easier readability, such as "MAAT", "Tmay-Nov" or similar.

  Thank you for this observation, we have altered the figures labels to reflect the reconstructed and instrumental temperatures.

- o Better to use common era (CE) instead of BC/AD?

  We agree this is better and have changed accordingly.

- It would be interesting with a deeper discussion of your results from the short core presented in fig 8. Nautajärvi has a very good agreement with the instrumental record, but Diss Mere shows the opposite trend. Please elaborate on this, and potential causes for it. The sudden offsets in temperature around 1940 in Diss Mere may be useful in investigating this. What happens in the GDGT distributions

(or other data) to cause this offset? After the offset, the reconstructed temperature is lower than before, contrary to instrumental data which show warming. Any clues to why?

We thank the reviewer and agree that this section requires some deeper thought and discussion. As outlined above this offset could be due to various mechanisms. Firstly, these sediments are not varved and as a result may host a different bacterial community compared to the laminated sediments of the Holocene- the upper 50-years of sediment display different GDGT distributions compared to much of the Holocene (e.g., less GDGT-1a). Secondly, around 1940, we observe increases in IR6me and changes in the ΣIIIa /ΣIIa ratio which may equally reveal a slightly different GDGT community. Thirdly, there is evidence that the lake is eutrophic at this time (owing to human influences). These additional environmental factors likely explain why the reconstruction does not fit the instrumental data as well as at Nautajarvi.

Of note, this feature of core-top cooling has previously been identified in lakes from the northern hemisphere (e.g., Miller et al., 2018; Zhao et al., 2021). The Miller et al. (2018) record also shows a mid-20th century temperature peak before declining, with greater methylation observed after the peak, which is similar to our observations. These authors also suggest shifts in the bacterial community may overprint the GDGT – temperature relationship.

We have added this discussion to the manuscript lines 524 – 536.

- Fig 9.
    - I would propose to show all records with the same y axis spacing, so that it is possible to see which sites show very little change versus larger change.

      This is a very good point and will help with latter sections of the manuscript. We have changed the y-axis spacing so that all match. See figure 9 below.

    - The arrows indicating trends appear quite arbitrary and require some better explanation.

      The goal here was to assist in the visual representation of trends for what we discuss in text. We have removed these arrows from the new figure version. See below.
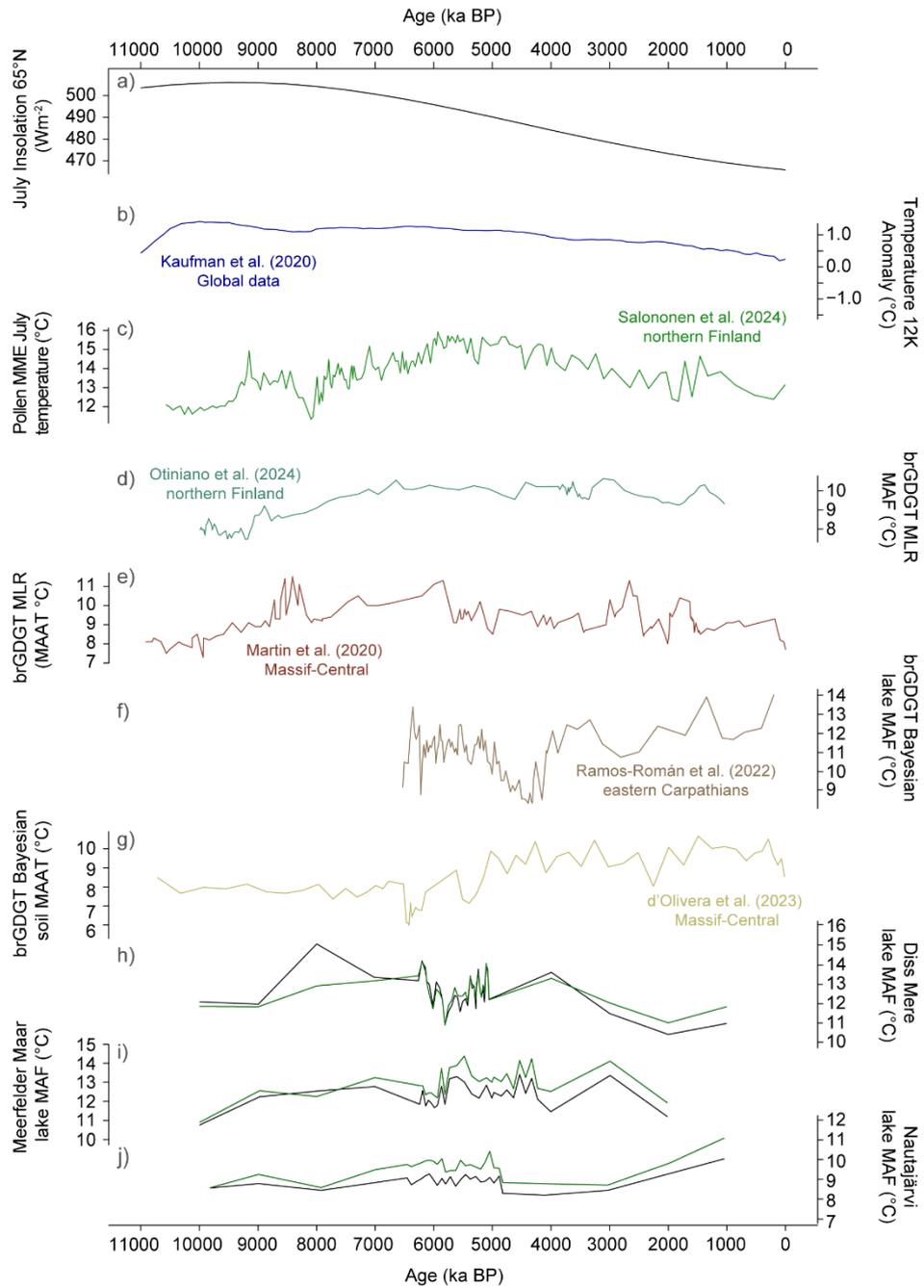
Figure 9: Comparative Holocene climate records. a) summer insolation at 65°N; b) Holocene global surface temperature anomalies (Kaufman et al., 2020); c) a pollen-based multi-method ensemble of Holocene July temperatures (Salonen et al., 2024); d) brGDGT multiple linear regression MAF temperatures, northern Finland (Otiniano et al., 2024); e) brGDGT multiple linear regression MAAT temperatures, Massif-Central, France (Martin et al., 2020); f) brGDGT Bayesian MAF temperatures, eastern Carpathians, Romania (Ramos-Román et al., 2022); g) brGDGT Bayesian MAAT temperatures, Massif-Central, France (d'Olivera et al., 2023); and h, I, j) the Bayesian lake MAF (black; Martinez-Sosa et al., 2021) and methylation set multivariate `regression MAF reconstructions (green; Raberg et al., 2021) from sites in this study.

- Specify if you mean the *trend* here rather than amplitude.

  I believe this comment is in relation to Section 4.5.1. Here we are referring to trends so have added this into the section heading of 4.5.1.

- 541: "*Peak warmth occurs at Diss Mere at ~ 8 ka BP*" this seems to be based on only a single datapoint, and is only true for the MBT based calibration, but at odds with the Raberg calibration. The way I read that graph (9h), the Diss Mere reconstruction shows slightly lower temperatures at 10-9 ka BP and after ~3 ka BP, but this is based on very few datapoints. The period 8-4 ka BP has very slightly elevated but variable temperatures. The temperature at ~5.9 ka BP is for example the coolest of the full record. I therefore find the discussion of the Diss Mere temperature evolution to be lacking in robustness, and higher sampling resolution would likely be needed to draw these conclusions. That said, like mentioned earlier, if the near surface data are added to the full reconstruction, the trends for the Late Holocene may be clearer.

We thank the reviewer for this comment and agree that the sparseness of the sampling interval in the early and late Holocene does not necessarily always help with observations of temperature trends. Whilst we agree with the reviewer that temperatures between 10 and 9 ka and after 4 ka are based on few data points and care does indeed need to be taken to avoid over interpretation we can provide evidence here from other proxy data from these sites (which we will include as supplementary figures) that dominant climate evolution broadly matches published climate data from these sites (see figures above- added to the supplement).

The ca. 3°C temperature rise from 10 – 8 ka and the 3°C temperature decline at the end of the Holocene is also comparable in evolution to the observations from other GDGT and traditional proxy reconstructions in Europe and shown on figure 9 (e.g. Salonen et al., 2024; Otiniano et al., 2024; Martin et al., 2020), which we believe adds credence to our observations (combined with the additional site evidence that we show above and in the supplement).

We have however modified the paragraph to reflect the observations more precisely lines 564 – 575: 'Diss Mere and Meerfelder Maar exhibit ~ 2°C - 3°C warming between 10 and 8 / 7 ka BP, respectively. From the available data, peak warmth occurs at Diss Mere between 8 and 4 ka BP, with the warmest temperature at 8 ka BP reconstructed from a single data point in a sparsely sampled section. Peak warmth at Meerfelder Maar appears between 5.6 and 3 ka BP with the observation at 3 ka again from a single datapoint. At Diss Mere temperatures generally follow a decreasing pattern from 8 ka BP (depending on calibration). In contrast, Nautajärvi reveals only a small temperature increase from the early to mid-Holocene from 9.8 ka BP, mid-Holocene peak warmth between 6 and 4.8 ka BP and rising temperatures across the late-Holocene (Fig. 9). Whilst caution is required when interpreting sparsely sampled datapoints, our reconstructions are broadly comparable with climate and physical proxy data

obtained from these records (with the exception of warm temperatures during the late Holocene from Nautajärvi and Meerfelder Maar; S10, S11, S12), and comparative reconstructions from across Europe (Fig. 9) adding greater credence to our interpretations.

- The statements about peak temperatures around 5.6-4.3ka BP at Meerfelder Maar also doesn't appear robust, with at least the datapoint at around 3k showing comparable temperature, with only one datapoint after that showing a slight cooling.

  We have altered the text in this paragraph to reflect the data more precisely. See above paragraph.

- Specify what is meant by "this" at the start of the paragraph.

  We have removed 'This' and reformed the first sentence to 'Spatial differences in reconstructed temperatures also appear in the mid-Holocene.' Line 610.

- But it is highly uncertain how much of this 2-3 degree temperature variability stems from methodological uncertainty/scatter versus actual climate shifts. This needs to be mentioned in the text.

  Thank you for this comment, we have added this note into the manuscript, line 620 – 622: 'By exploiting varved lake sediments, we show that during the mid-Holocene, Diss Mere and Meerfelder Maar median reconstructions exhibit ~ 2°C - 3°C of temperature variability at multi-decadal scales. Whilst this 2-3°C temperature variability may be a product of calibration uncertainty, comparative climate reconstructions reveal similar patterns. Pollen and brGDGT-based reconstructions…'

Technical or minor corrections:

- The original reference for tex86 is Schouten et al., 2002 https://doi.org/10.1016/S0012-821X(02)00979-2

  We have added this to line 57

- This sentence comes a little out of the blue and requires a little more clarification. Is this threshold applied in this study?

  We are sorry but we do not know which sentence is being referred to here.

- Equation 9. Typo. As written now, it mathematically makes no sense. It should be GDGT0/(GDGT0+cren).

  Thank you for this observation, we have corrected this formula (equation 9).

- Why capitalization of crenarchaeol?

  This is an oversight on our behalf and we have corrected all mentions of crenarchaeol in text.

- Section 3.1. Please refer to figure numbers when presenting index results.

  We have added numbers into the manuscript when presenting results.

- Fig 4. To improve readability of this figure, I would suggest removing "brGDGT" in front of each compound. It is unambiguous that Ia, IIa etc. are brGDGTs. This can also be done for figure 2 axis labels.

  Thank you for this advice. We have tidied up Figure 4 labels by removing brGDGT from the PCA. However we have kept the labels in Figure 2 as they shift from iso to brGDGTs in each sub-figure. New Figure 4 as:
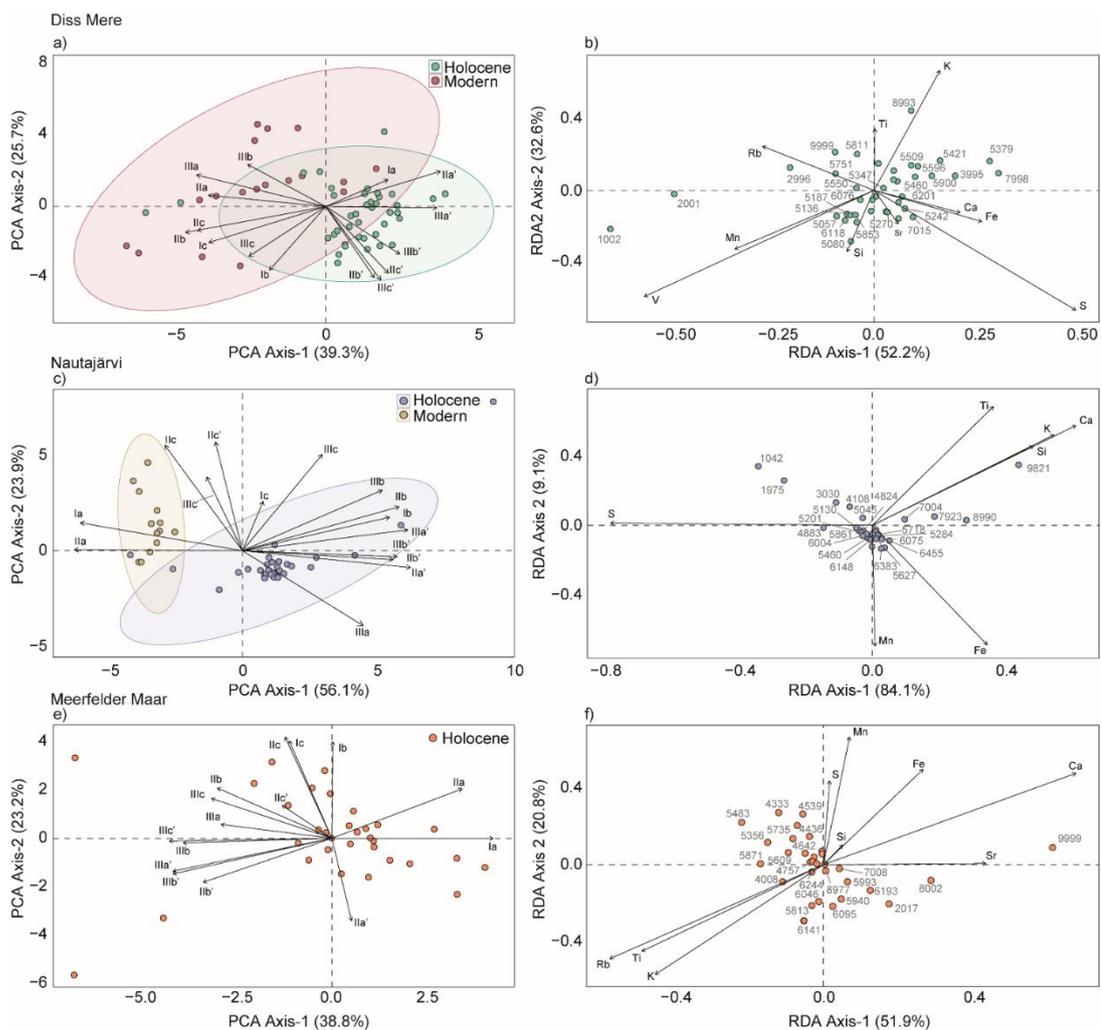

Figure 4: Principal components analysis (PCA) left, and redundancy analysis (RDA) right. a) and b) detail the plots from Diss Mere; c) and d) from Nautajärvi; and e) and f) from Meerfelder Maar. a) and c) show modern (last 300-years) and Holocene sample distributions within the PCA plots. Explanatory variables in the RDA plots are major elements from µ-XRF data (from Martin-Puertas et al. (2012); Boyal et al. (2024); Lincoln et al. (2025)) with numerical points showing the median age of samples.

- Fig s1. Add in text that it is based on TEX86.

  We have added into the text that this is a reconstruction from TEX:

  Figure S1. TEX86 reconstructed Lake Surface Temperature (LST) reconstructions from the three sites in this study a) Diss Mere, b) Meerfelder Maar, c) Nautajärvi. add into text that the reconstruction is based on TEX86

- Specify that you mean in *reconstructed* LST?

  As above:

  Figure S1. TEX86 reconstructed Lake Surface Temperature (LST) reconstructions from the three sites in this study a) Diss Mere, b) Meerfelder Maar, c) Nautajärvi. add into text that the reconstruction is based on TEX86

- Figures s4, s6, s8. Please make the order of the elements the same in all figures. Currently, the s6 figure has other order than s4 and s8, which reduces readability.

  We have changed the order of the elements in Figure S7 and S8 to match the other two figures. We have also moved two figure positions so that the correlation matrices appear in order as the reconstructions above.

- "vs" doesn't need to be in italics. But if you decide to still do that, be consistent throughout. For example, it is not in italics on line 362.

  We have changed all instances of vs.

- Reference should be in parentheses.

  We have changed this in text.

- Replace imperfect with moderate, or similar.

  We have changed this to 'moderate' – line 494.

Response to Reviewer 2.

In this manuscript Abrook et al. present a compilation of three european lake records from Holocene-aged varved sediments. These records are analized for their capability to preserve GDGT signals within the varves and whether the conditions that generate the varves may override the temperature signals of the GDGTs. Within this work the authors show that unique signals to the different lakes are captured, which can be related to specific conditions of the water chemistry, nevertheless, taken as a whole, the records suggest that the GDGT temperature signal is preserved, suggesting that varved sediments provide a useful tool to generate high resolution GDGT records.

I find this manuscript very interesting and it presents a very exciting dataset, additionally I consider that the main takeaways from the study are adequate and follow the data presented. The manuscript is also generally well written. I do however have some comments which I hope could help improve the manuscript and I will be happy to support the publication of his work once these have been addressed.

We would like to thank the reviewer for their constructive comments and suggestions on our manuscript. We also thank the reviewer for their positive comments on 'finding the manuscript very interesting and presenting a very exciting dataset'. Below we provide answers to each of the comments after they appear in red and state how we have addressed them in the clean version of the manuscript.

Main comments

I agree with Reviewer #1 that further comparison between the modern and Holocene records could be done for Nautajärvi and Diss Mere. In addition to the points made by Reviewer #1 I would suggest discussing further the separation observed for the Holocene and modern samples in Nautajärvi. Did you considered further PCs in the Diss Mere PCA that may show a similar spread (that may indicate a common feature explaining this)? Particularly since the first two PCs for Diss Mere account only for ~60% of the variance.

Many thanks for your comments, we agree that further comparisons between the modern and instrumental data are worthwhile at, for example, Diss Mere, where there are small offsets between the reconstructed and instrumental data (albeit the reconstructed temperature approximates the instrumental data). The reasons for this, we postulate, is that it is reflective of a slightly different environment. The last 2ka of the Diss Mere record coincides with a rapid sedimentation rate increase, a shift in diatom and blue-green algal communities and a shift in catchment vegetation leading to eutrophication (see Boyall et al., 2023; 2024). This is driven by an increase in detrital input following human occupation and the subsequent building of the Diss Mere town (Boyall et al., 2023; 2024). Critically this leads to eutrophication over the last 1kyr at Diss Mere, which may impact this part of the record which we did not explore.

We have added this discussion to the manuscript in Section 4.4.1 in lines 524 – 536.

In terms of the separation of 'modern' and 'Holocene' samples at Nautajarvi within the PCA, we consider this to be a component of the slightly different distribution of GDGTs between the 'Holocene' and 'modern' samples. Similar distributions are also observed within the late-Holocene, which plot in the 'modern' space of the PCA. The two PCs at Nautajarvi appear to be pulling this out (and also account for greater variation in the dataset). At Diss Mere, whilst the modern samples do plot in a slightly different area of the PCA, they also overlap with the Holocene data.

The features that we observe between the two sites (Diss Mere as disturbed / eutrophic in the modern and not in the Holocene vs Nautajarvi with continued varve preservation throughout the whole record) may indicate that the processes that control these features are different for each site.

We explored the different PCs at Diss Mere (PC-1 vs PC-2, PC1vs PC3, PC1 VS PC4 and combinations of the above) and did not observe similar separation as at Nautajarvi. We therefore consider that there is not a common process controlling the distributions.

Additionally, given the dataset presented here, I think this could be an excellent opportunity to compare these samples and the changes in mixing regimes with the proportion of cren/cren', as was proposed by Baxter et al., 2021 (10.1016/j.quascirev.2021.107263).

Thank you very much for this suggestion. We did consider the cren/cren' (f[CREN']) ratio as we thought that this might be useful in detecting changes in the position of the oxycline / shifts in mixing regime. Unfortunately, levels of the crenarchaeol regioisomer are very low (and even below detection at Diss Mere) so we did not pursue any further.

The low abundance of the crenarcaheol regioisomer indicates that all f[CREN'] values are <0.08 from both Meerfelder Maar and Nautajarvi (not available at Diss Mere). However, intriguingly we do observe an increase in (f[CREN']) in the mid-Holocene at Nautajarvi (which may suggest greater stratification as Baxter et al. (2021)). However, Lake Chala is meromictic, and therefore a permanently stratified and very deep lake (90m, much deeper than any lake here).

Data from Nautajarvi (Lincoln et al., 2025) suggests that the Finnish lake did undergo periods with both strengthened and weaker stratification (therefore relatively more hypolimnetic oxidation and Fe precipitation and colloid formation) in the mid-Holocene. Increased f[CREN'] in the mid-Hol from Nautajarvi, may suggest based on Baxter et al. (2021) increased stratification (more anoxia). The changes between higher and lower f[CREN'] values in the mid-Holocene on an inter-sample basis could account for this strengthened and weakened stratification as observed in the physical varve data.

We have added this discussion around f[CREN'] and lake water stratification (focussing on Nautajarvi to Section 4.3.1, lines 433 – 444. Given the number of panels already presented in Figures 5-7 we did not add further panels. We did however add a supplementary figure (S9) detailing f[CREN'] (see below):

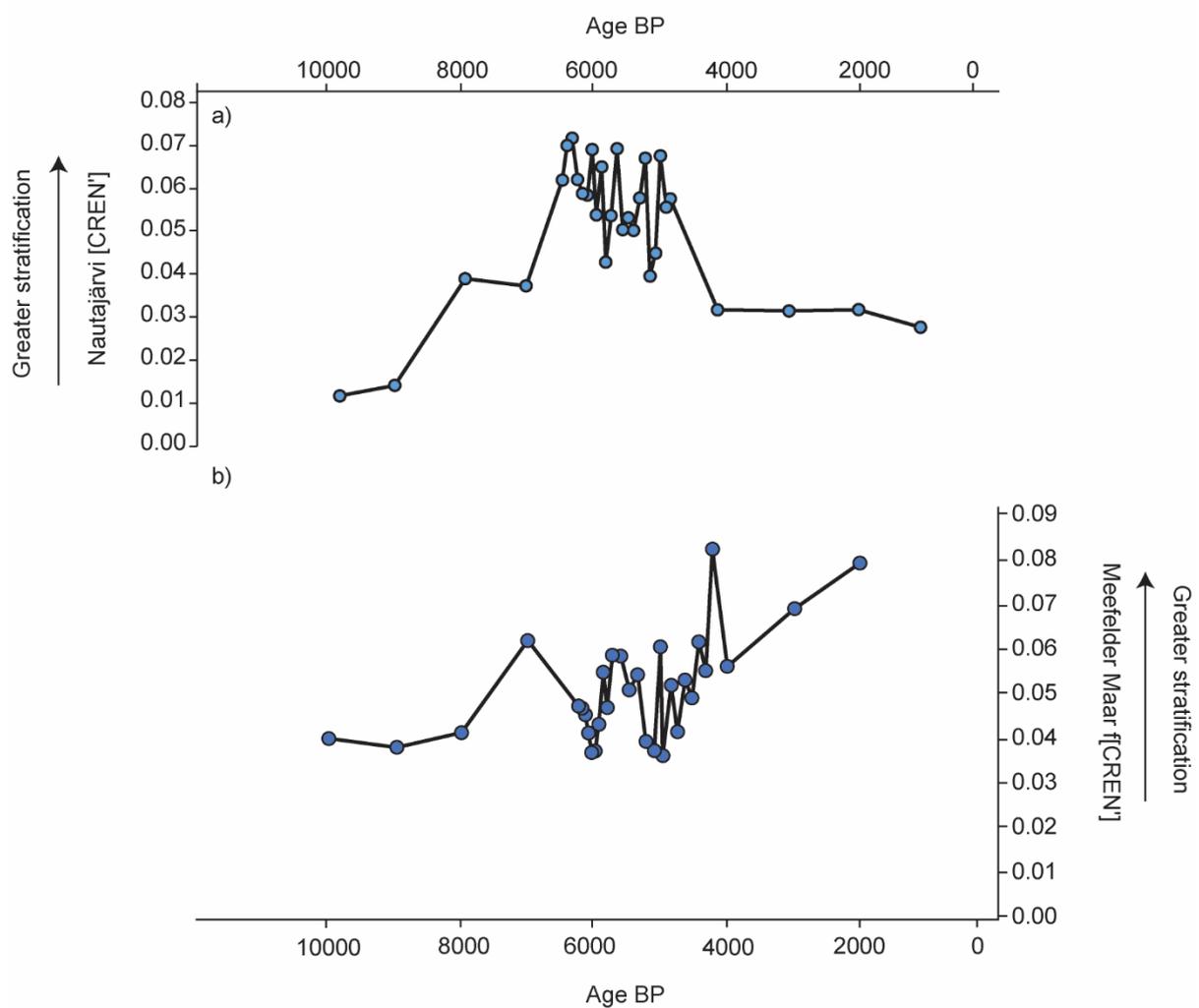**Figure S9: f[CREN'] from a) Nautajärvi and b) Meerfelder Maar.**

Finally, throughout the text, I find that the Figures and figure captions could be modified to be clearer and easier to read. See detailed comments below.

Many thanks for this comment. We have edited all figures and captions to make them easier to follow and read.

Specific comments

-Figure 1: The labelling of the different panels is confusing and starts over in the diagrams of he couplettes. Additionally the maps contain features that are not explained in the text.

Thank you for this comment. We have rearranged the figure and aligned the varve depositional models with each of the site locations so that these are more 'joined up'. As reviewer 1- we have removed the elements that were not explained in the figure caption. New Figure 1 as below:
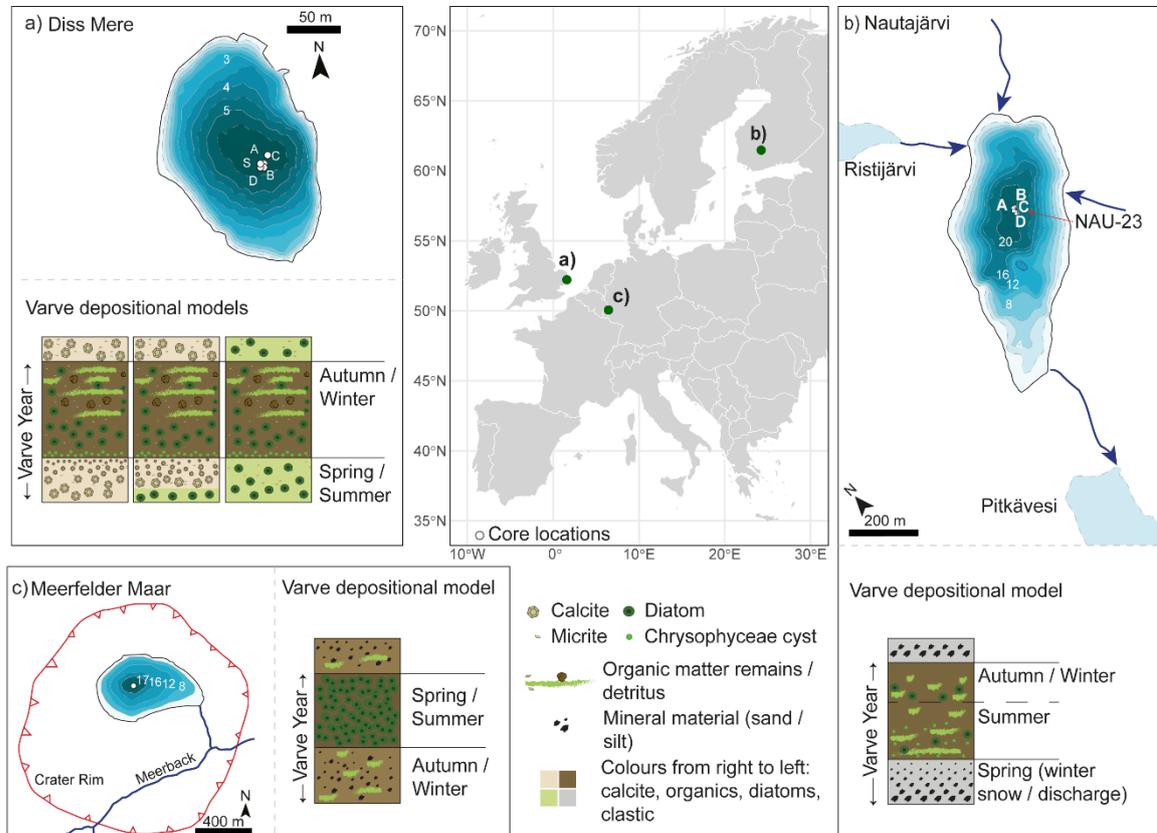

Figure 2: Location of each of the sites in Europe: a) Diss Mere, UK, b) Nautajärvi, Finland; c) Meerfelder Maar, Germany. Site boxes reveal lake bathymetry with depths and inflows/outflows. Varve depositional models are added to each box which demonstrate the main varve facies from each location, with three varve microfacies at Diss Mere and one each at Nautajärvi and Meerfelder Maar. Letters denote different cores from a) and b). Diss Mere adapted from Boyall et al. (2023); Nautajärvi adapted from Lincoln et al. (2025).

-Line 219 and throughout the text: capitalize brGDGT and isoGDGT when at the beginning of a sentence (i.e., BrGDGT and IsoGDGT).

We have edited this throughout the text when they appear.

-Figure 3. Is b) a zoom in on triplot on a). Please explain in legend.

Yes this is correct. We have added this to the figure caption, which now reads as:

Figure 3: Ternary plots showing a) the relative proportions of tetramethylated, pentamethylated and hexamethylated compounds from Diss Mere, Meerfelder Maar and Nautajärvi, respectively against the same data from a global brGDGT peat and soil dataset (grey; Dearing Crampton-Flood et al., 2020); a global lake dataset (black; Martinez-Sosa et al., 2021); Arctic lakes (green; Raberg et al., 2021) and Alpine lakes (red; Bauersachs et al., 2024); and b) an inset figure of a).

-Figure 4. Label each plot with a letter for clarity.

We had hoped to show that the colours of individual points follows in a left to right motion. However we have labelled each plot with a letter. The new figure as below:
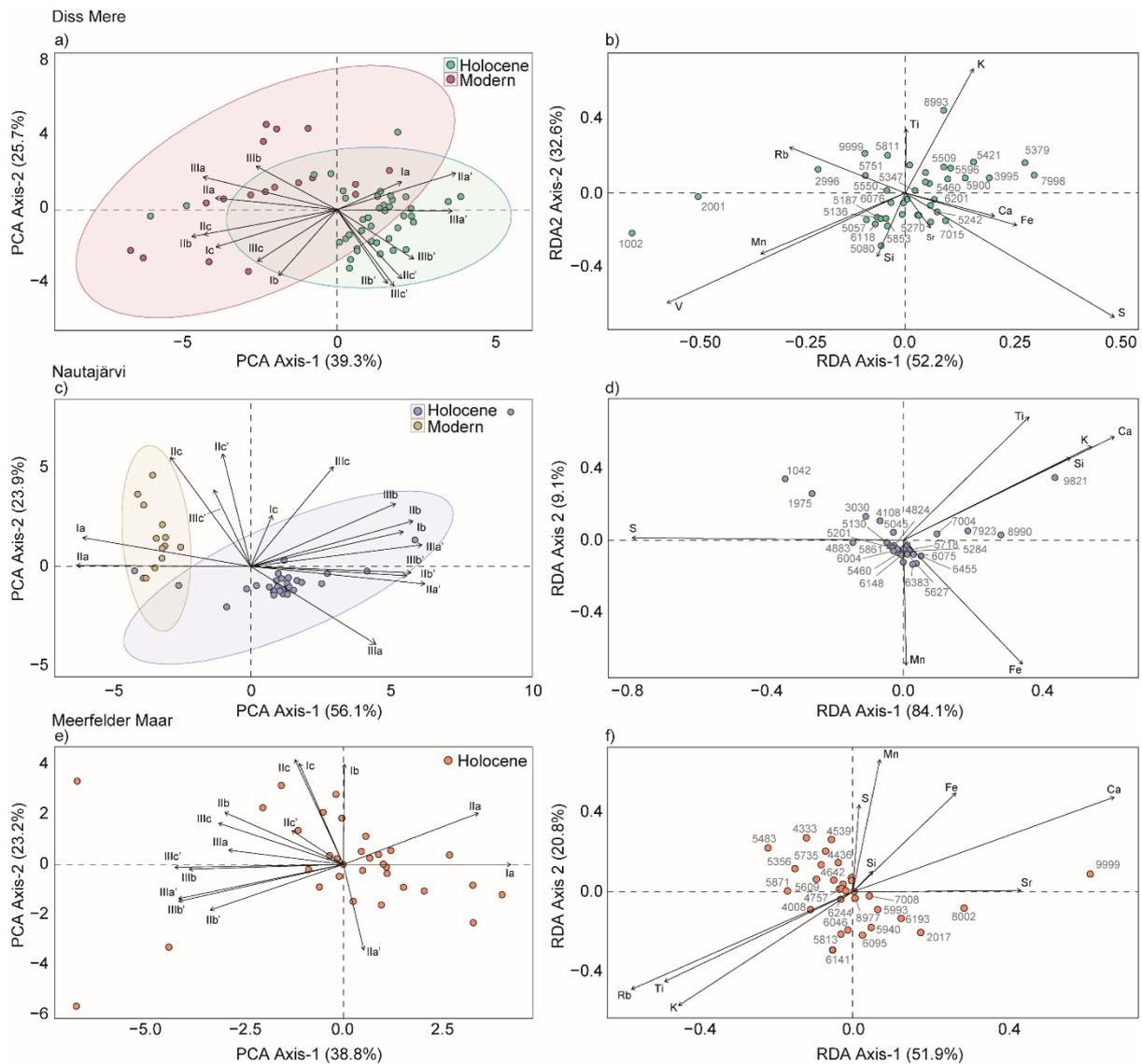
Figure 4: Principal components analysis (PCA) left, and redundancy analysis (RDA) right. a) and b) detail the plots from Diss Mere; c) and d) from Nautajärvi; and e) and f) from Meerfelder Maar. a) and c) show modern (last 300-years) and Holocene sample distributions within the PCA plots. Explanatory variables in the RDA plots are major elements from µ-XRF data (from Martin-Puertas et al. (2012); Boyal et al. (2024); Lincoln et al. (2025)) with numerical points showing the median age of samples.

-Figure 8. Legend is very hard to follow, should be rephrased.

We agree that the figure caption here is difficult to follow. We have rephrased to ensure that it adequately and simply describes the figure:

Figure 8: Modern (>300-year) temperature reconstructions from Diss Mere and Nautajärvi. Shown are a) GDGT-based MAF reconstructions at Diss Mere; b) this data alongside modern MAAT instrumental data from both the Lowestoft meteorological station (light blue) and from 5 km gridded temperatures using the HadUK dataset (light orange; Hollis et al., 2019); c) GDGT-based MAAT and MAF reconstructions at Nautajärvi; d) these GDGT-based MAF reconstructions alongside modern MAAT instrumental data from the Heinola Asemantaus (light blue) and Helsinki Kaisaniemi (light orange) meteorological stations; and e) Nautajärvi GDGT reconstructions against re-sampled May-November meteorological station air temperature data.