Comments Anonymous Referee #2:

In their manuscript Maier and colleagues present a new alkenone SST record from the Arabian Sea. Their record shows substantial millennial scale SST variability of 4C during both the last glacial and the Holocene, with the exception a quiescent period during the LGM. However, the glacial/interglacial offset is small, only around 1C. This result is in stark contrast to other Arabian Sea SST records which show large glacial/interglacial changes, and small millennial scale variability. This intriguing result potentially offers new insights into either the spatial variability of SSTs in the Arabian Sea and their sensitivities to elucidate different components of the climate system, or as a sensitive recorder of the Oman Upwelling Zone or more broadly the east-west SST gradient in the Arabian Sea.

While the paper provides good detail on specific changes in the record and potential mechanisms for these changes, it misses the big picture. The knowledge gap and research question are not stated and it's not clear what the jump is in our understanding of the regional climate system provided by this paper. It is hard to assess the validity of changes through time when the timing of SST changes in the presented record are rarely stated or given uncertainties. The lack of glacial/interglacial offset is not explored in detail. The impact of changes in bathymetry due to deglacial sea-level rise is only lightly touched upon. While individual Heinrich events are discussed, a common response is not, and no distinction is made between the response to Heinrich events versus Greenland stadials. I have further misgivings over the event chronology used as a comparison, which is based on Mediterranean cosmogenic exposure dates with external error of 4kyr, rather than using the INTIMATE chronology to focus on Greenland stadials and interstadials, or a U/Th dated speleothem which might be able to distinguish Greenland stadials and Heinrich events. The spectral analysis section is reasonable, but the focus is on what frequencies are found, when perhaps the most interesting result is a 5kyr window of no periodicity surrounding the LGM. Overall, this is an exciting record that lacks a quantified robust discussion. In its current form it is not ready for publication.

Response: First, we would like to thank the anonymous referee for her/his helpful and insightful comments that have contributed to improving the manuscript.
Our central motivation of this study is that there is a lack of SST data, particularly in the Gulf of Oman, which our study seeks to fill by focusing on the core location. Moreover, there are still few records with high-resolution SST reconstructions from the Arabian Sea. The core site in the Gulf of Oman is situated in an atmospheric (influenced by the SW and NE monsoon and NW winds) and oceanic (eddies, partly affected by the Oman Upwelling as well as inflow of Persian Gulf Water and strong SST gradients) dynamic region. Compared to other parts of the Arabian Sea, we show that northern hemispheric climate dynamics might be damped or strengthened by more regional processes in this complex dynamic region. Notably, we find that there is minimal cooling during the LGM compared to other locations in the Arabian Sea, and our site is more sensitive to temperature changes at millennial-scales than others. We can only speculate what are the dynamics behind the SST pattern at our core site and it asks for further paleoclimatic research in this part of the Arabian Sea. We will clarify the knowledge gap and research question, particularly in the abstract and conclusion.
It is correct that we do not pay much attention to the impact of changing water depth due to sea level rise. The core site is located at a water depth of about 770 m. Red Sea sea

level reconstructions by Siddal et al. (2003) and Rohling et al. (2008) show a sea level rise of about 80 to 100 m from the last glacial to the Holocene, which led to a connection between the Persian Gulf and the Arabian Sea. We do not expect a strong impact of sea level/water depth fluctuations on SST during the Pleistocene. However, the connection to the Persian Gulf since the postglacial sea level rise may indeed have influenced SST variations at our core site. The core site is located in the southern part of the Gulf of Oman and, therefore, in the area of PGW outflow. We will add further explanation to the discussion.
For the INTIMATE event chronology, Rasmussen et al. (2014) stated that the last about 44 ka are based on NGRIP and GRIP data sets: "*In the most recent extension of the INTIMATE event stratigraphy scheme, all events down to and including GI-12 were defined using NGRIP and GRIP δ18O data (Blockley et al., 2012). While most of the existing definitions are well supported by the [Ca2+] record, with δ18O shifts typically aligning with the early or middle parts of the corresponding [Ca2+] shifts, the [Ca2+] records support the choice of older onset points of stadials GS-9 and GS-10. However, yet again we retain current definitions to maintain compatibility with the existing event stratigraphy scheme.*" The reviewer is correct that the uncertainties of the Mediterranean cosmogenic exposure dates are comparably higher (about 4 ka) than those of the INTIMATE record (about 1 ka). However, firstly the timing and duration of most Heinrich events in our record are similar to those of Greenland Ice Core records. An exception is Heinrich 3 event, which is dated between 32.7 and 31.3 ka by Allard et al. (2021) and around 30.6 ka (Greenland Stadial (GS) 5.1 within H3 (Pedro et al. 2022)) and the GS 5.2 around 32.0 ka by the INTIMATE chronology (Rasmussen et al. 2014), with the latter falling in the timing of H3 by Allard et al. (2021). Secondly, the H3 event is followed by D-O 4. Fleitmann et al. (2009) have shown an age offset of the D-O 4 by about 590 years between the Sofular Cave in Turkey (earlier) and the NGRIP GICC05 chronology. Further, our $\delta^{15}N$ record (Burdanowitz et al. 2024) matches the D-O timings of the more nearby Sofular Cave by Fleitmann et al. (2009) very well. For instance, the pronounced strong OMZ around 30 ka is associated with D-O 4, which would fall within H3 after the NGRIP records. Further, from the 21 AMS ^{14}C dates, two are around the timing of H3 with ages of 30250 +- 340 yr BP and 32590 +- 540 yr BP, respectively. Lastly, as discussed in Burdanowitz et al. (2024), some of the existing records in the Arabian Sea are tuned to the NGRIP ice cores, and the peaks then of course fit in the timing of the Heinrich and D-O events. Altogether, we are convinced that our age model is robust and independent, and the observed age offsets between the NGRIP cores and the more regional cores are not due to age uncertainties. Therefore, we will stick to the Allard et al. (2021) timings of the events. Further we will stick with the nomenclature of Dansgaard-Oeschger events to be consistent with our earlier findings and as they are used according to the Greenland Interstadials (e.g. Rasmussen et al. 2014)
We will provide specific responses to each of the line-by-line comments in your specific comments section.

Specific Comments

Line 20 states that the dynamics of monsoon periods are influenced by northern hemisphere solar radiation, but the paper is largely about Heinrich events and Dansgaard-Oeschger interstadials and other abrupt millennial scale events, not solar radiation.

Response: Thank you for your insightful comment regarding the influence of northern hemisphere solar radiation on monsoon dynamics. We recognize that the primary focus of our paper is on Heinrich events, Dansgaard-Oeschger interstadials, and other abrupt millennial-scale events as they overprint the general influence of the solar radiation. However, we included the discussion of northern hemisphere solar radiation to provide a broader context for understanding the relevant climatic mechanisms during these periods.

Knowledge Gap/Research Question: It's not clear what the purpose of this study is. What is it trying to find out? At the paragraph break line 51-52 we need to understand what is the unknown which this study intends to investigate. A one or two sentence summary of Gaye et al. 2018 of key results might be useful, as this might outline what the knowledge gap is.

Response: We will clarify the purpose of our study giving a brief summary of the key results from Gaye et al. (2018) covering the past about 25 ka. They found that the glacial SST were about 4°C lower than SST's during the Holocene in the Arabian Sea. Further they postulate that the general glacial SST gradient within the Arabian Sea had a stronger N-S insolation driven component than during the Holocene with a more pronounced NW-SW circulation driven component. We will also ensure that the knowledge gap and the purpose of our study are explicitly highlighted in both the abstract and conclusion sections to reinforce their significance.

Age model uncertainty: Dating uncertainty and age model uncertainty need to be stated in section 3: i.e. a key summary of the Burdanowitz et al., 2024 age model.

Response: We agree with the referee (and the other two referees) that we need to discuss our age model in more detail. Therefore, we will add a paragraph and the figure of the age-depth model by Burdanowitz et al. (2024) in the material and methods section. We will also add the measured ages including their uncertainties as diamonds to figure 4.

Greenland stadials vs Heinrich events: This paper uses the nomenclature of Dansgaard-Oeschger interstadials and Heinrich events. It is not clear why Heinrich events in particular are different from Greenland stadials in this record. I suggest switching to a Greenland stadial nomenclature throughout the manuscript. This would also allow for more precise timings from the INTIMATE chronology to be used in preference to the Mediterranean cosmogenic dataset currently used.

Response: As stated earlier, the INTIMATE chronology is based on NGRIP and GRIP data sets. Further, some of the Greenland stadials are Heinrich events (e.g., GS 5.1 is related to H3 (Pedro et al. 2023)). As mentioned above it is correct that the uncertainties of the Mediterranean cosmogenic exposure dates are comparably higher (about 4 ka) than those of the INTIMATE record (about 1 ka). However, firstly most of the timing and duration of the Heinrich events are similar to those of Greenland Ice Core records, except H3. Secondly, the

H3 event is followed by D-O 4. Fleitmann et al. (2009) have shown an age offset of the D-O 4 by about 590 years between the Sofular Cave in Turkey (earlier) and the NGRIP GICC05 chronology. Further, our $\delta^{15}N$ record (Burdanowitz et al. 2024) matches the D-O timings of the more nearby Sofular Cave by Fleitmann et al. (2009) very well. For instance, the pronounced strong OMZ around 30 ka is associated with D-O 4, which would fall within H3 after the NGRIP/INTIMATE chronology. Further, from the 21 AMS ^{14}C dates, two are around the timing of H3 with ages of 30250 +- 340 yr BP and 32590 +- 540 yr BP, respectively. Lastly, as discussed in Burdanowitz et al. (2024), some of the existing records in the Arabian Sea are tuned to the NGRIP ice cores, which peaks then of course fit in the timing of the Heinrich and D-O events. Altogether, we are convinced that our age model is robust and the observed age offsets between the NGRIP cores and the more regional cores are not due to age uncertainties. Therefore, we will stick with the Allard et al. timings of the events. Further we will stick with the nomenclature of Dansgaard-Oeschger events to be consistent with our earlier findings and as they are used equivalent to the Greenland Interstadials (e.g. Rasmussen et al. 2014).

The timings of the shaded bars in figures 4 and 5 are inaccurate. The 4.2 ka event is not 5-4 kyr BP. The 8.2 ka event is not 9-8 kyr BP. The YD is not 13-12 kyr BP. H1 is not 18-15.5. H2 is tricky to tie down, but I think it is more likely to be where DO-2 is currently labelled (this SST drop would help the case being made that Heinrich events are associated with SST decreases). I believe H3 is more likely to be GS5.1 than GS5.2. I don't know how much of this is a plotting issue and how much is an issue with the choice of external event chronology.

Response: Thank you for your detailed observations regarding the timings of the shaded bars in Figures 4 and 5. We intentionally designed the bars to be larger for visibility, as smaller bars would be difficult to discern. Additionally, we acknowledge that events can have varying onset times, and thus we have allowed for some flexibility in their representation. The designations for glacial and interglacial periods have been carefully chosen and align with those established by Burdanowitz et al. (2024), ensuring accuracy in our representation. Further, we disagree that a decrease of SST is strictly associated with Heinrich Events as the study is highly complex which needs to be taken into account. Only the H4 event led to a significantly drop of SST's at the core site. As stated above, we are convinced that our independent age model is robust and we will not tune it to get a "better fit" to the NGRIP/INTIMATE chronology. We clearly have no signal of GS 5.1 in our SST record and our d15N record (Burdanowitz et al. (2024)) shows at the timing of GS 5.1 (after the INTIMATE chronology) indications of the D-O 4, which matches with the cave record of Fleitmann et al. (2009).

Glacial/Interglacial change: The lack of a large change in SST from glacial to interglacial (about 24.5 to 25.5 C by my eye) compared to nearby records (2-4 C) is one of the remarkable results of this paper. Yet it is not discussed. The role of changing bathymetry is similarly only lightly discussed.

Response: The most significant impact of sea-level change occurred at the end of the Pleistocene leading to the flooding of the Persian Gulf, which we discussed in detail. For smaller sea-level changes, the influence of water depth changes becomes less critical, as these variations do not significantly affect SST unless they result in substantial alterations of the ocean circulation (Rohling et al., 2008; Siddall et al., 2003). The core site is located in an area with several atmospheric and oceanic influences. Firstly, the LGM has not a strong impact on SST at the core site. Secondly, during the Holocene, the SST variations are strong but lying between SST's in the northern/northeastern Arabian Sea and the Oman Upwelling region. This shows that the core site is affected by both the NE and the SW monsoon. Further, the connection to the Persian Gulf became important during the Holocene. Lastly, the strong SST gradient, which is also visible in modern SST records, within the Gulf of Oman have a strong impact on SST reconstruction as a small shift can have a large impact (see figure 2). However, we will point out the lack of the glacial/interglacial changes more clearly in the discussion.

Mid-Holocene cooling: What is the timing of the SST decrease at 5ka relative to the end of the Green Sahara? And does this have any mechanistic implications?

Response: Thank you for your inquiry regarding the timing of the SST decrease at 5 ka in relation to the end of the Green Sahara. We agree that the SST decrease around 5 ka coincides closely with the transition out of the Green Sahara period. This timing suggests a potential link between the cessation of the Green Sahara and regional climate dynamics, as shifts in monsoon intensity, changes in atmospheric/oceanic circulation patterns may have contributed to observed SST cooling in our record. We will discuss this in more detail in the discussion.

2 ka event: I am skeptical here without a detailed description of timing and uncertainty. A climate anomaly between 5 and 4 kyr BP is no longer sufficient to be labelled as the 4.2 ka event. It needs to be coincident with the Carolin et al., (2016) window of 4.26 to 3.97 kyr BP. I think the discussion of the event starting at line 312 is reasonable. But line 316 goes against the previous discussion. I would remove references to 4.2 from elsewhere in the paper (such as the results and conclusions) but keep the brief discussion around 312 as this part is nuanced and reasonable. Unless of course, specific data can be provided to support the assertions made.

Response: Thank you for your comments regarding the 4.2 ka events and its timing. As previously mentioned, we used the broader time frame of 4–5 ka for clarity in our figure, as a more specific delineation may not be easily discernible. While we acknowledge the time frame from Carolin et al. (2016), we believe that uncertainties and potential shifts in the timing of events can lead to variations, making it plausible for the 4.2 ka event to occur slightly earlier or later. Therefore, we would like to retain references to both the 8.2 ka and 4.2 ka events in the manuscript.
We appreciate your observation regarding the sentence in line 316, which appears contradictory to the preceding discussion. We will revise this sentence to ensure consistency and clarity throughout the text. Thank you again for your constructive feedback, which will help to improve the manuscript.

Spectral Analysis: There is a significant reduction in cyclicity between 22 and 17kyr BP (ish). This quiescent LGM might warrant further investigation.

Response: Thank you for your insightful observation regarding the significant reduction in cyclicity between approximately 22 and 17 ka BP during the Last Glacial Maximum (LGM). We totally agree and acknowledge that this stable period represents an intriguing aspect of our data that may indeed warrant further investigation.

The 7200 year cycle is a potential subharmonic of precession, modulated by obliquity.

Might the high number of potential cyclicity peaks around 500 years be one periodicity on top of an uncertain chronology?

Response: We agree that the 7200 year cycle might be a subharmonic of the precession. We will add this potential mechanism to the discussion part. Regarding the high number of potential cyclicity peaks around 500 years, we recognize that these could be indicative of a true periodicity. However, we also acknowledge that they may reflect the uncertainties inherent in the chronology of the data. The uncertainties of our age model are between 170 to 240 years for the Holocene and up to 770 years for the oldest part of the record.

Technical Corrections

Line 29: insert modern: "of the total modern annual precipitation"

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 41: "Northern Hemisphere glacial ice-sheets"

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 48: no need for a comma between 'both' and 'the'.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 49: "in the past" is redundant and could be deleted.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 65: The 'Study Area' section is more of a 'Modern Climate Dynamics' section.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 93: New Paragraph?

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 119: This sentence is a bit confusing. I suggest 'Alkenones were measured at 2cm for the upper 162cm and 4cm resolution below 162cm by combining consecutive subsamples due to lower organic content'.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 120: suggest '3 to 18g of sediment'.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 122: space needed between sentences.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 123: This sentence is out of place. Methods should be written in chronological order.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 144: This is precision not accuracy

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 164: Figure reference needed here.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 187: 'mitigated' is not correct here. Is 'reduced' sufficient?

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 214: 'westerlies'

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 267: suggest 'minor decrease in Central Arabia'

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 308: New paragraph.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 320: period missing at the end of the sentence.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 338: 'Implies' is incorrect here. Suggests or is linked to might be better.

Response: We agree with the reviewer's comment and have revised the point accordingly.

Line 80: This sentence might need a rewrite to make it clearer. At present (4-5C) as a quantification of several hundred kilometers makes no sense.

Response: We will revise this sentence.

Line 65-85: Better disambiguation between spatial and temporal SST gradients is needed.

Response: We will correct the sentence as followed: A spatial SST gradient of 4-5 °C develops over several hundred kilometers with the onset of the SW monsoon, creating a temperature low near the coast of Oman and a high of approximately 29 °C in the western Gulf of Oman (Figure 2c).
We are referring specifically to the spatial SST gradient generated by the SW monsoon onset, with cooler temperatures near the coast of Oman and warmer temperatures in the western Gulf of Oman. We did not refer to the temporal SST gradient here. We understand that this may have caused confusion due to the title, which includes both spatial and temporal SST variability.

Line 101: Need a linking sentence or signpost sentence to help the read transfer into this new paragraph.

Response: We concur with the reviewer´s observation and updated the sentence: Building on complex interaction within the Arabian Sea, mesoscale eddies are cyclonic and anticyclonic rotating water masses, contrary to the surrounding main currents and emerge as key players in the regulation of surface ocean circulation (Fischer et al., 2002; de Marez et al., 2019; Al Saafani et al., 2007; Trott et al., 2019).

Line 151: The resolution of the evenly spaced dataset needs to be stated alongside the resolution of the actual data, so that this choice can be evaluated.

Response: We will add the missing information to the material & methods part. As mentioned in the manuscript we are using the package ncdf4.helpers v.0.3-6 (Bronough, 2021) and the approx. function. We used the highest resolution of the record by using the function "get.f.step.size()" resulting into a minimum step size of 40 years between two measurements. With that we were able to define the amount of time steps (in total 890) to generate an even spaced dataset. We are aware that this may lead to uncertainties, especially for the parts of the core with a lower resolution. The resolution of the original data sets vary between 40 and 800 years (mean: 181 +- 124 years, median: 168 years), with lowest resolution during the LGM.

Line 250: 17ka is not earlier than 19ka

Response: In this case, "earlier" response not to 19 ka at the beginning of the paragraph but to the B-A around 14 ka in the sentence before the usage of "earlier". Following the usage of "earlier around 17 ka" is correct.

Line 258: its not clear what is meant by the NW winds moving in the opposite direction to the SW monsoon.

Response: We will rewrite the sentence to make it more clear: NW winds, peaking between 15 and 13 ka (Sirocko et al., 2000), may have contributed to the earlier warming, while SW monsoon conditions weakened (Leuschner and Sirocko, 2000; Sirocko et al., 2000).

Line 299: Lead with the evidence, then the interpretation, not the other way round.

Response: We will correct the sentence as followed: With the beginning of the mid-Holocene, strengthening of NE monsoon conditions likely led to a temporarily interrupted transport of upwelled water masses to the core location.

Line 337: D-O interstadial?

Response: The intention of our sentence was to illustrate the contrasting effects associated with different climatic events, such as a weakening (or strengthening) of the AMOC during the 8.2 ka cold event compared to D-O interstadials. We aimed to highlight the opposing shifts of the ITCZ and the corresponding impacts on the ISM. We will ensure this contrast is more explicitly stated in the text for clarity.

Thank you again for your time and effort. Your comments and suggestions helped us to improve our manuscript.

On behalf of all co-authors,

Jan Maier