RC1: Anonymous Referee #1

Referee comment in black
Author answer in blue

The manuscript applies individual foraminifera analysis (IFA) δ¹⁸O across four oceanographically distinct sites in the tropical Indian Ocean to evaluate variability resulting from seasonality and extreme climate events. While IFA is a well-established concept, the authors extend its application to the tropical Indian Ocean by combining new measurements with previously published records. The forward modelling approach uses ORAS5 reanalysis to simulate δ¹⁸O of calcite (δ¹⁸Oc) at surface (0.5m) and thermocline (97m) depths. The forward-modelled distributions are compared with core-top IFA δ¹⁸Oc of G .ruber and N. dutertrei, except for the lack of thermocline species data from the eastern tropical Indian Ocean site. Climate sensitivity is assessed by introducing ±75% perturbations in ORAS5 temperature and salinity, modifying only seasonal amplitude at the surface and both seasonal amplitude and temporal frequency at the thermocline

The scientific premise is relevant and the manuscript falls within the scope of CP. The core-top IFA δ¹⁸Oc data are sufficient to support the broad interpretations and conclusions. However, the forward modelling framework rests on several simplifying assumptions that limit the realism of the outputs. These assumptions need to be either improved or more rigorously justified. Significant revisions to the methodology are therefore required before the manuscript can be considered for publication.

We thank Reviewer 1 for their comments, which helped greatly in improving and clarifying the manuscript. We reply below point by point to their remarks.

General comments:

1) *The authors used ORAS5 reanalysis data for forward modelling δ¹⁸Oc at surface (0.5 m) and thermocline (97 m) depths. Since ORAS5 forms the foundation of the results and discussion, the authors should first evaluate how well ORAS5 reproduces observed conditions near the proxy sites. Comparisons with instrumental datasets (e.g., satellite-derived SST and SSS, buoy-based temperature and salinity observations) need to be presented (in supplementary). Additionally, a comparison with other reanalysis products such as HadISST (surface), GODAS (surface and thermocline), and SODA3 would strengthen confidence in the reanalysis-based approach and help assess the sensitivity of results to the choice of reanalysis product.*

We thank the reviewer for this suggestion. We used ORAS5 reanalysis because it was widely used in previous palaeoceanographic studies, including for analogous IFA sensitivity analysis already performed in the tropical Pacific Ocean (see e.g., *Glaubke et al., 2021*). We have now added to the Supplementary Material a validation of ORAS5 against three out of the four other sources suggested by the referee, including two reanalysis products (HadISST1.1, see *Rayner et al., 2003*) for surface temperature, GODAS (*Behringer & Xue, 2004*) for both surface and thermocline temperature and salinity. We also compare ORAS-5 with the instrumental dataset EN4 (*Good et al., 2013*) for both surface and thermocline temperature and salinity for the 1980–2018 period. We attempted to download the SODA reanalysis product, but all links towards each iteration of the SODA3 and their subsequent SODA4 products are down from their *University of Maryland* website.

Validation against independent datasets confirms that ORAS5 reliably captures both the seasonal cycle and interannual variability of surface temperature at the four proxy sites (see new Supplementary Figure 1 that will be added to the revised manuscript). Depending on the site and the depth we are tracking, there are small differences in the mean and/or amplitude recorded by other reanalysis datasets or instrumental data.

For the seasonal SST component that pilots at first order the computed calcite δ¹⁸Oc we target, we observe two sites where all products show near-perfect agreement (sites MD96-2060 and U1467), and two other sites that match with very high confidence, albeit with some features (such as the magnitude of the cooling occurring during summer monsoon upwelling at site NIOP905, and the upwelling season at SO189-39KL) not perfectly captured in one or more products. Yet all the seasonal variability, including at sites where two SST minima are recorded, are well captured both in the structure of the seasonal cycles and the absolute temperature values. For salinity, which only affects surface δ¹⁸Oc calculation at second order, a very good match is also observed, with the EN4 product having difficulties capturing the salinity maximum one would expect during the upwelling seasons at NIOP905 and SO189-39KL sites. It is unclear whether the reanalysis products are biased by the Ocean Model boundary effects and/or the gridding, or the instrumental lagrangian profiles used in the EN4 dataset were unable to capture the full structure of the mentioned sites, as we note that the dynamic Somalian and Indonesian margins are undersampled in comparison with more offshore locations (see https://www.metoffice.gov.uk/hadobs/en4).

For the interannual SST component, all products show very good agreement for temperature with r^2 ranging from 0.49–0.68 at MD96-2060, 0.54–0.57 at U1467, and 0.29–0.43 at SO189-39KL ($p < 0.001$ in all cases). NIOP 905 shows lower but still significant interannual correlations ($r^2 = 0.11$–0.26, $p < 0.001$).

For the thermocline depth that, as suggested later by the reviewer, will mainly drive the IFA interannual changes, we also find p-values>0.999 at each site for each tested product.

For surface salinity, interannual correlations are generally low at all sites ($r^2 < 0.08$, p values ranging from non-significant at MD96-2060 to < 0.001 at other sites), which is partly consistent with known difficulties that reanalysis products have in simulating salinity variability in the tropical Indian Ocean, especially for the subsurface. For subsurface salinity, interannual correlations are also generally low at most sites ($r^2 < 0.14$, p values ranging from non-significant to < 0.001). A notable exception is SO189-39KL, where interannual subsurface salinity correlations are significant across all products ($r^2 = 0.31$–0.55, $p < 0.001$), likely reflecting the strong influence of the Indonesian Throughflow on thermocline salinity variability at this site. However, the difference in amplitude between the different reanalysis data are less than 0.2 psu, so given that subsurface salinity contributes only marginally to the variance of forward-modelled δ¹⁸Oc as compared to temperature (See QQ-plot Fig. 7), computing a r^2 does not make much sense in this case.

Overall, this validation demonstrates that ORAS-5 faithfully reproduces the large-scale temperature and salinity conditions at the four proxy sites, and is therefore appropriate in our forward modelling framework. The supplementary figures and statistics are provided in the revised Supplementary Material (Figure S1).

Figure S1 — ORAS5 Validation: Monthly time series (anomaly) (1980–2018)

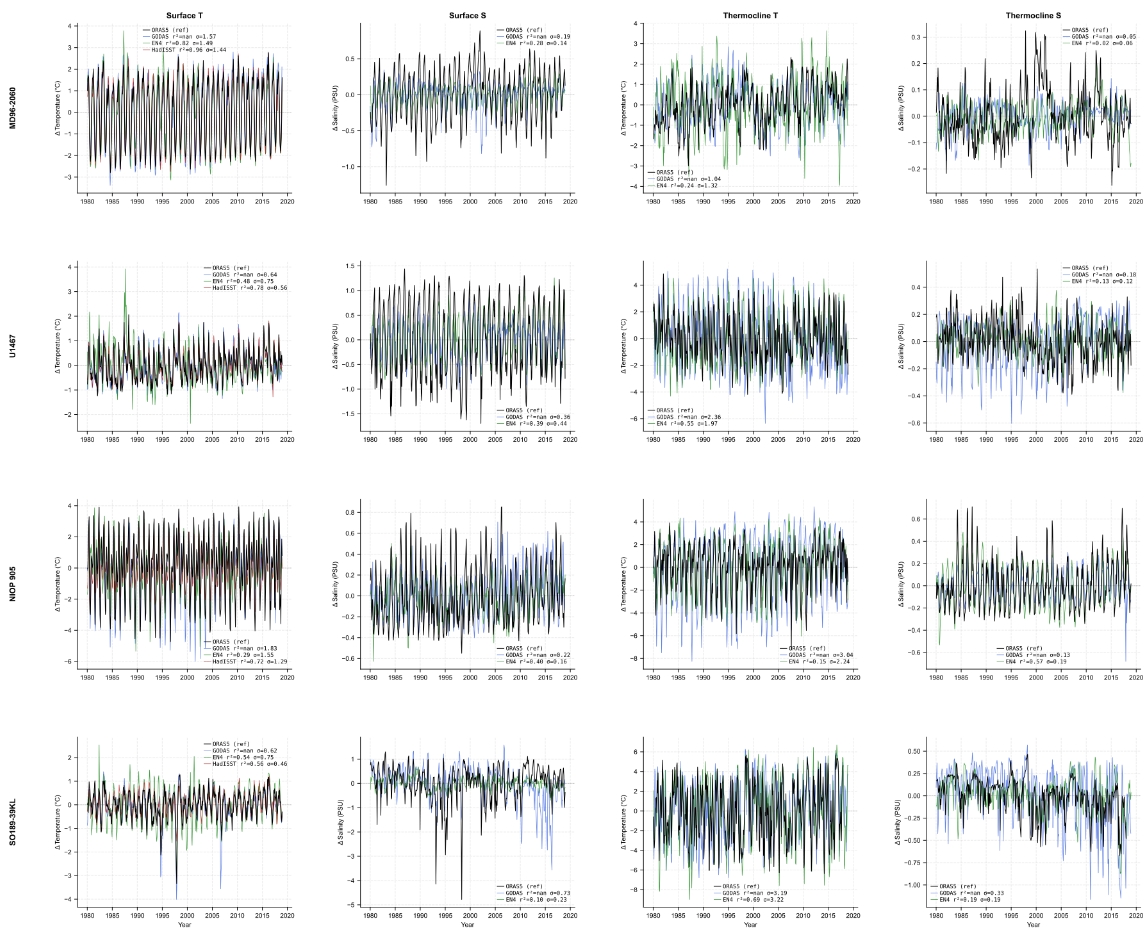

Figure S1b — ORAS5 Validation: Annual means (anomaly) (1980–2018)

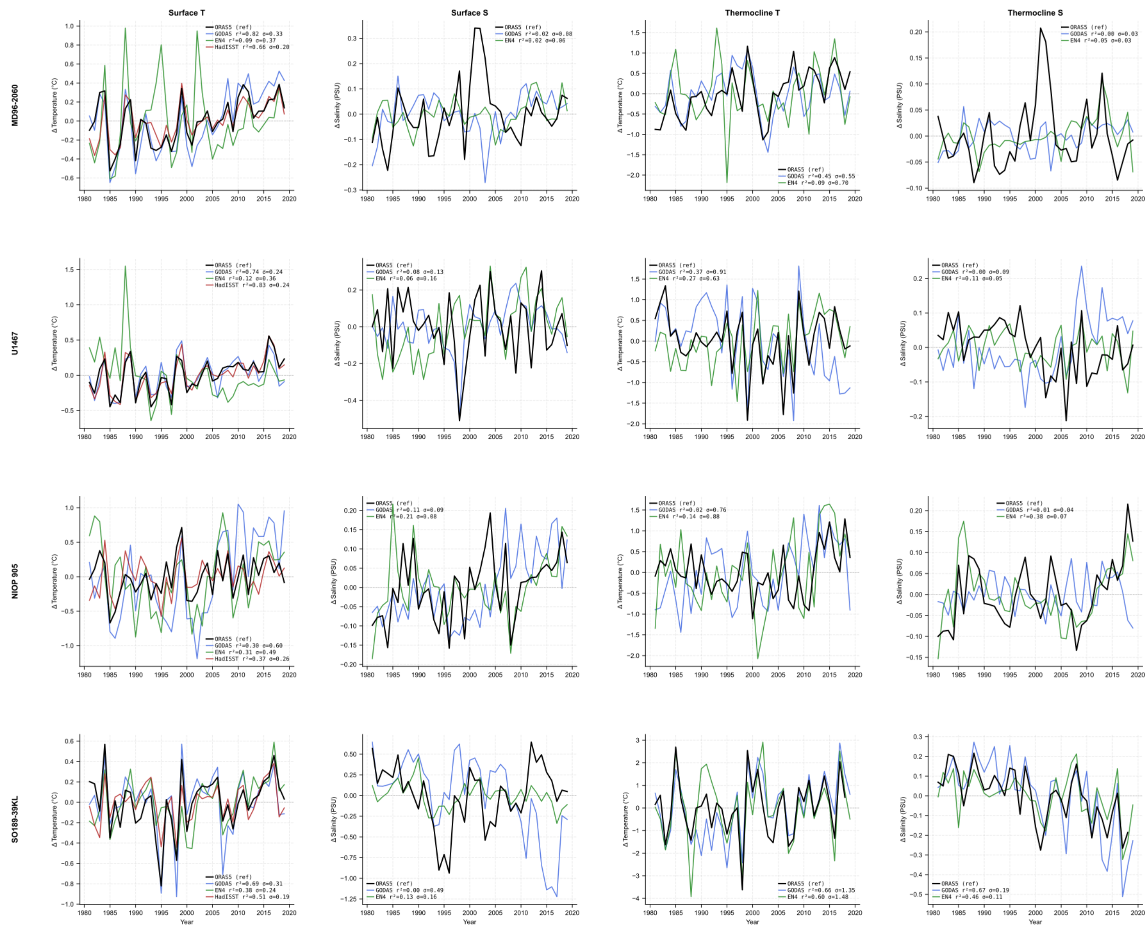

Figure S1c — ORAS5 Validation: Seasonal cycle (1980–2018)

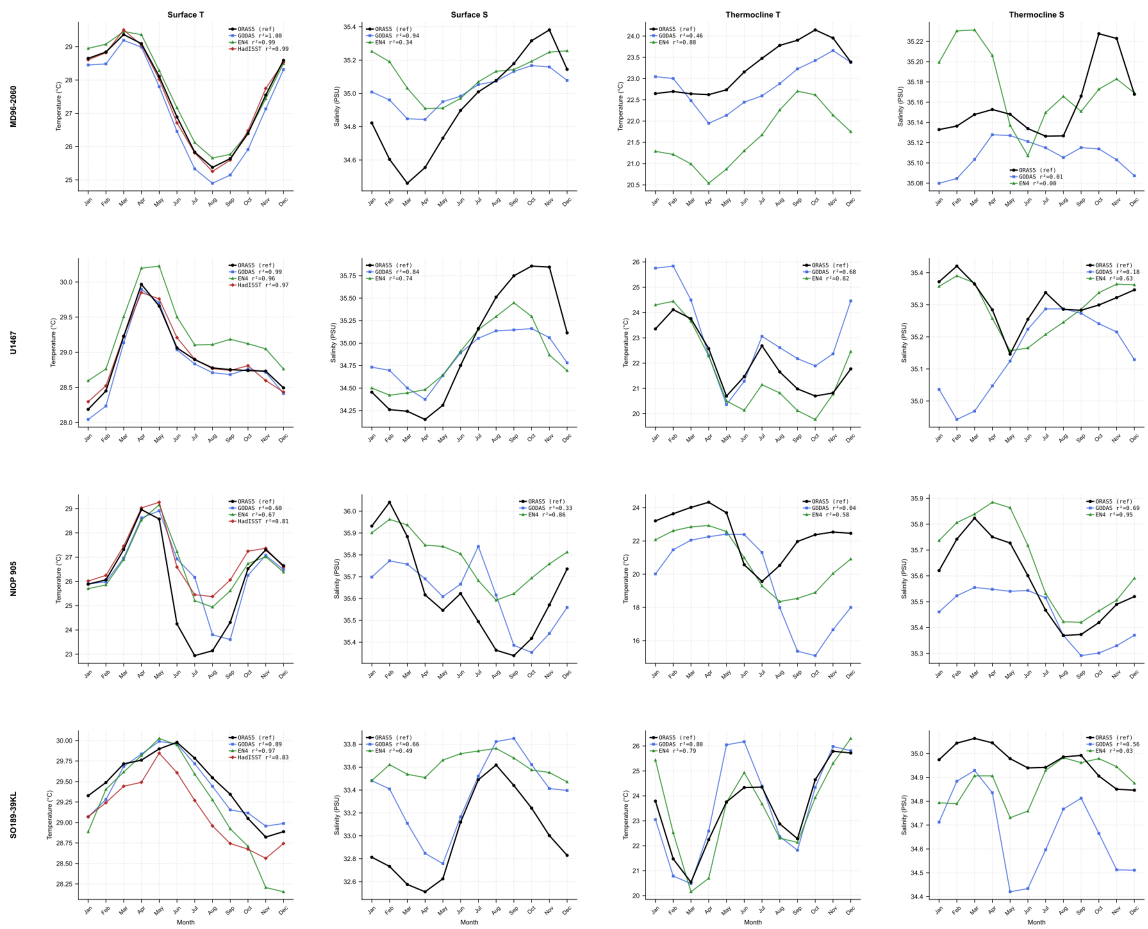

Figure S1. Validation of ORAS5 reanalysis against independent datasets at the four proxy sites (1980–2018). For each site (MD96-2060, NIOP 905, U1467, SO189-39KL), comparisons are shown for surface temperature, surface salinity, thermocline temperature, and thermocline salinity across three temporal modes: (i) mean monthly climatology, (ii) interannual anomalies (annual means), and (iii) seasonal cycle. ORAS5 (black) is used as the reference dataset and compared against GODAS (blue; Behringer & Xue, 2004), EN4 (green; Good et al., 2013), and HadISST (red; Rayner et al., 2003) for surface temperature only. Pearson r^2 values are reported in each panel. Surface and thermocline extractions correspond to 0.5 m and 97 m respectively. HadISST is not available for salinity or thermocline variables.

2) *Assigning a single fixed depth for the surface (0.5m) and thermocline (97m) uniformly across all sites and through time is a significant simplification. Thermocline depth is neither spatially constant (Fig. 2A) nor temporally stable, it varies with monsoon phase, extreme climate events, and regional oceanographic dynamics. The authors should consider integrating temperature and salinity data over the mixed layer and thermocline depth ranges, accounting for their spatial and temporal variability, to produce a more physically realistic forward-modelled $\delta^{18}Oc$ distribution.*

We agree with the reviewer, and this choice to allow one fixed depth for the depth model grid at 0.5 and 97 m was oversimplified. We have tested the sensitivity of our forward modelling time series δ^{18}Oc distribution by comparing outputs obtained at fixed depth with an average temperature and salinity defined over a depth range.

For the surface, we have compared the 0.5 m fixed depth with the integrated 0–50 m depth range to compute the forward-modelled δ^{18}Oc timeseries. In three out of the four sites we see no difference in the computed *G. ruber* δ^{18}Oc forward modelled timeseries. At site SO189-39KL, the integration of the upper 50 m removes part of the variability, in particular at the interannual time scale. This is quantitatively confirmed by the raw δ^{18}Oc time series: the standard deviation of the difference between the two extraction methods is 3–4 times larger at SO189-39KL (σ = 0.12–0.20‰) than at the other three sites (σ = 0.03–0.07 ‰; Fig. S2), consistent with the sharper vertical T/S gradients associated with this upwelling-influenced site. This is because the site lies within an upwelling zone, where interannual anomalies are expressed as short, intense signals concentrated very close to the surface; averaging over the upper 50 m dilutes these signals with more stable subsurface waters. Preserving this surface-concentrated variability is essential, as it is precisely the signal used in Thirumalai et al. (2019) to infer changes in interannual variability, and the sampling strategy at this site was specifically designed around this upwelling signature. We therefore retain *G. ruber* as the interannual variability recorder at this site, consistent with the original publication by Thirumalai et al. (2019), even though this species is known to live within the upper part of the mixed layer. We hence think that integrating the first 50 meters of the water column in that case would contribute to alter our model-forward exercise as it would remove the extreme interannual events we track in our study. We further note that this feature fits with the real sedimentary core-tops and its forward-modelled IFA we present in the second part of the article as being the most variable site despite being derived from surface-dwelling foraminifera.

For the thermocline, we have compared the 97 m fixed depth with the integrated 70-120 m water depth range to compute the forward-modelled δ^{18}Oc timeseries derived, omitting that of the SO189-39KL because we only use *G. ruber* at this site. Both water depth assignments methods in the forward-modelled δ^{18}Oc of subsurface-dwelling foraminifera (*N. dutertrei*) lead to nearly identical timeseries, with the standard deviation of the difference between the two methods remaining below 0.045‰ at all three sites (MD96-2060, NIOP 905, U1467), demonstrating that defining the modelled *N. dutertrei* depth habitat at a fixed or an integrated water column does not change the modelled δ^{18}Oc timeseries. This is because in the thermocline, the temperature gradient with depth is so high that even changing an assigned living habitat depth by a few tens of meters would not alter the full thermocline variability as modelled in IFA for the three sites where *N. dutertrei* were used (Fig. S2).

We further checked by quantifying this water depth assignment method its impact on the simulated S.D. on random 60 IFA for MD96-2060, NIOP 905 and U1467 sites. The results confirm that the choice of a depth range vs. a fixed depth has a negligible effect (ΔSD remaining below 0.05‰ for the simulated IFA SD with raw ORAS5 data, and below 0.065‰ after seasonal and/or interannual amplification, representing less than 15% of the total simulated IFA SD at these sites.

The sensitivity test figures for fixed depth versus depth ranges are provided in the revised Supplementary Material (Fig. S2). For readability and to ease comparison between sites to evaluate their sensitivity, we kept the simulations as initially produced and we discuss in the revised manuscript both case with fixed or range depth.

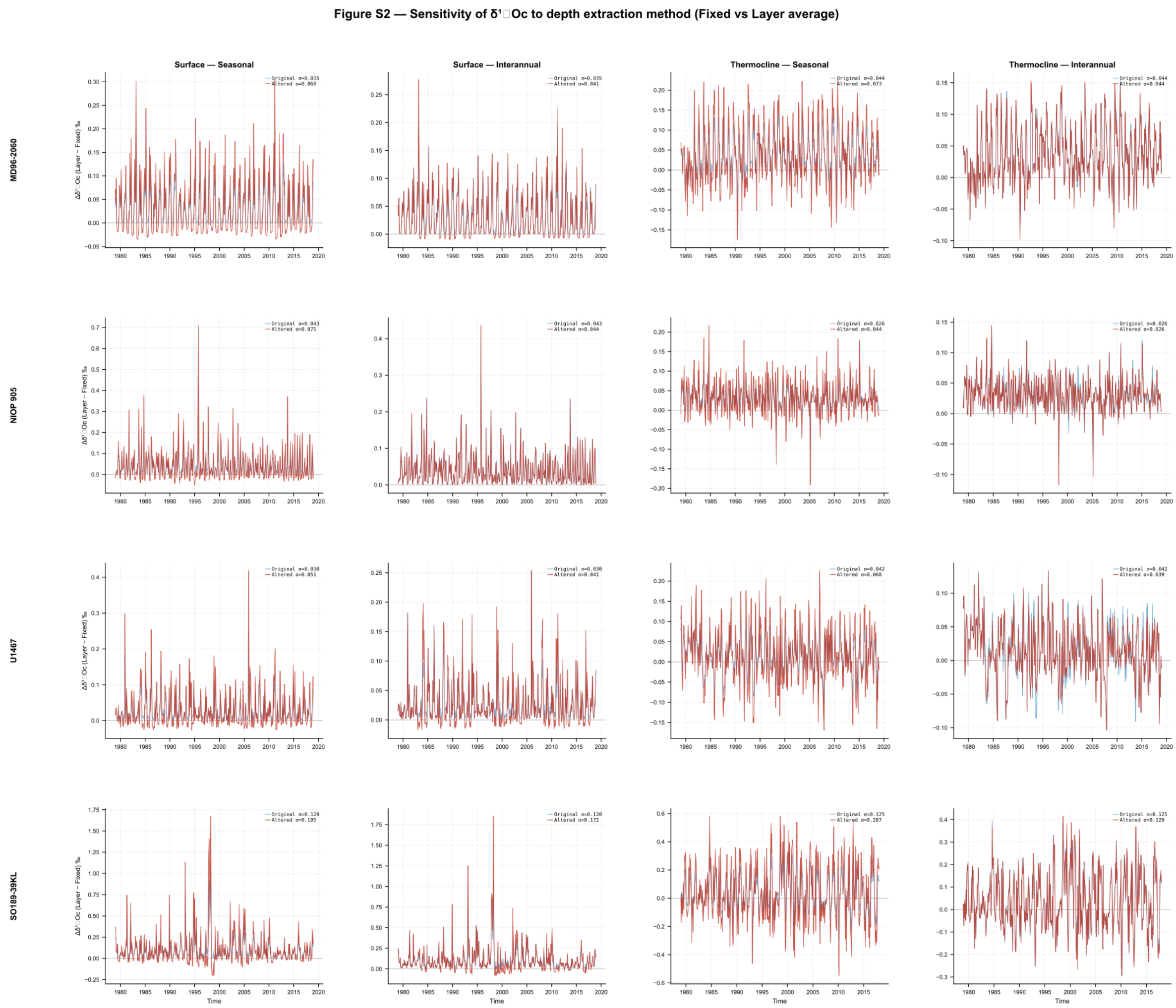

Figure S2. Sensitivity of forward-modelled δ^{18}Oc time series to the choice of depth extraction method (fixed depth vs. layer average). Each panel shows the time series of the difference in modelled δ^{18}Oc between a layer-averaged extraction (0–50 m and 70–120 m) and a fixed extraction depth (0.5 m and 97 m), for the four sites and for seasonal and interannual variability. The grey horizontal line marks zero difference between the two methods, serving as a visual reference.

3) *The authors artificially introduced ±75% variability in temperature and salinity to simulate extreme climate events but do not adequately justify this choice beyond citing Thirumalai and Clemens (2020), whose study focuses on the northern Bay of Bengal, a region highly sensitive to salinity changes due to large freshwater fluxes and therefore oceanographically distinct from the four sites used here. The authors first need to quantify the realistic upper and lower bounds of ENSO- and IOD-related variability at each of the four oceanographically distinct sites before applying a uniform ±75% perturbation. For example, the annual SST range at site MD96-2060 (western tropical Indian Ocean) is approximately 3.5°C (Fig. 1a). Increasing this variability by 75% would yield a range of ~6.1°C. It is therefore necessary to demonstrate whether such variability is actually observed during ENSO or IOD events at this site. Without such evidence, the interpretation appears overstated*

We understand that the ±75% alteration applied invariably at all sites could be confusing given the flowchart of the analysis we describe in the article. As the flowchart targets two distinct objectives with different methodologies to validate the IFA method at the oceanic scale with only few sites, we will now introduce a flowchart at the end of the introduction with more detailed description of each step (flowchart conceptualized using A.I. for the form and colors).

Indeed, adopting Thirumalai & Clemens (2020) or any other amount of alteration would not much change the results given the objectives of the analysis: describing which site and which species could be valuable to detect an increase or a decrease of seasonality or interannual variability. We never claim to estimate by how much the variability of another period in the past had changed because, given the uncertainties, we think that having an amount of e.g.+/- half a seasonal/interannual shift in variability would just be unreachable. However, our analysis is framed to identify regions and species that will likely be most sensitive to track past changes. By essence, the 75% alteration has to be invariable to allow comparison of the sites and depths sensitivities to distinguish among sites and species those able to reconstruct changes in larger-scale structures of interest such as the Indian monsoon or the IOD.

This choice is further supported by the analysis of peak ENSO and IOD events in the ORAS5 reanalysis (1958–2018), which shows that thermocline temperature anomalies during the strongest events range from 2.69°C to 6.58°C across the four sites, representing approximately 50–75% of the mean seasonal amplitude at each site. The ±75% perturbation therefore represents a physically meaningful upper bound for altered variability states, rather than an arbitrary scaling factor. For example, the MD96-2060 site could well have behaved completely differently during a given glacial period, so we must consider such a possibility to envision whether IFA could be helpful to resolve it in later paleo studies. We hope that the new flowchart will help illustrate the main aim of the analysis we performed: first perform an idealized simulation of a virtual different ocean with increased or decreased variabilities, and second provide realistic model-data comparison using regional specific parameters.

4) *"The authors state (Lines 207–211 and 330–332) that only seasonal amplitude is altered at the surface, whereas both amplitude and frequency of interannual events are modified at thermocline depths. What is the rationale for not introducing interannual frequency changes at the surface as well? This methodological choice requires clarification and justification."*

We thank the reviewer for this methodological point. The choice to modify only seasonal variability at the surface was for two reasons. The first is that in the western tropical Indian Ocean, temperature variability is dominated by the seasonal cycle rather than interannual variability. The second reason is more methodological as outlined in our response to general comment 3, this idealized design a seasonal perturbation at the surface versus a combined amplitude and frequency perturbation at the thermocline was deliberately chosen as part of the simplification exercise inherent to our forward-modelling approach, to test how sensitive the IFA distribution mean and SD are to distinct modes of variability change under an altered simulated climate. However, ORAS5 time series show that our four sites have mean surface temperature anomalies during peak ENSO/IOD events (July–November) that range from ±0.1 to ±0.4°C, compared to climatological seasonal ranges of 1.2–6.0°C depending on the site. Interannual variability thus represents a minor fraction of total surface variability.

To directly address this concern, we additionally ran the forward model under three configurations for *a* surface range depth (0-40 m) at all four sites. We compared seasonal change only, interannual change and both combined. The results reveal distinct patterns depending on the site. At MD96-2060 and NIOP 905, interannual changes produce nearly identical or lower IFA SD than seasonal (0.57 vs 0.58 ‰ and 0.57 vs 0.69 ‰ respectively), confirming that seasonal variability dominates surface IFA dispersion at these western sites. At U1467 and SO189-39KL, we observe a dominance of interannual variability rather than seasonality (0.54 vs 0.43 ‰ and 0.48 vs 0.30 ‰) reflecting the stronger ENSO/IOD surface signature at these central and eastern Indian ocean sites. We acknowledge that a combined seasonal and interannual perturbation at the surface increases IFA SD by 0.17 to 0.25 ‰ across sites, and note that this sensitivity test will be added and discussed in the revised manuscript to the to allow the reader to assess the impact of this methodological choice.

5) *The use of site names throughout the discussion is difficult to follow. It would improve readability if the authors consistently referred to the sites using regional identifiers such as western Arabian Sea, western tropical Indian Ocean, central tropical Indian Ocean, and eastern tropical Indian Ocean. Otherwise, readers must repeatedly refer back to Figure 1 to identify the site locations.*

We have done this in the revised manuscript.

6)

The authors need to report nearby seasonal flux, and habitat-depth information for G.ruber and N.dutertrei from the FORCIS database (https://doi.org/10.1038/s41597-023-02264-2) in the Supplementary Material. Since the study interprets the SD of d18O G.ruber and N.dutertrei as reflecting climate variability, forward modelling using realistic habitat depth and seasonal flux data would be more appropriate than assuming constant carbonate flux at fixed depths (0.5 m and 97 m). Incorporating these parameters would substantially improve the paleoclimate relevance of the modelling framework.

First, regarding the habitat depth, constraining the calcification depth of planktonic foraminifera from the FORCIS plankton net observational database in the Indian Ocean is inherently difficult. Also, the data present two biases identified as a lack of data in some regions of the Indian Ocean, and the fact that distributions are complexified by the fact that non-stratified nets integrate over large depth intervals, potentially capturing specimens well above their preferred habitat depth. This means that a surface-dwelling species like *G. ruber* may appear at 100 m simply because the net was deployed between 0 and 200 m. Still, we have compiled available plankton net data from the FORCIS database from locations close to our sites. Three out of four study sites (NIOP 905, U1467, SO189-39KL) nonetheless confirm the expected qualitative pattern, with *G. ruber* consistently captured shallower in the water column than *N. dutertrei*, even though the absolute depth estimates themselves remain unreliable for the reasons outlined above. No plankton net data are available near MD96-2060, which reflects a data gap. Given these uncertainties, we cannot use these observations with confidence to constrain calcification depths in the tropical Indian Ocean.

The modelling depth of 0.5 m is therefore retained as representative of the surface mixed layer signal recorded by *G. ruber*. consistent with its well-documented surface habitat (Fairbanks et al., 1982). For *N. dutertrei*, the FORCIS-based average occurrence depth in the Indian Ocean is 41.31 ± 3.59 m (range 10–150 m; Chaabane et al., 2026), reflecting its broad vertical range spanning the base of the mixed layer to the sub-thermocline. We note however that this estimate is likely biased towards shallower depths, as non-stratified net tows integrate over large depth intervals and may capture specimens well above their preferred habitat depth. Analysis of ORAS5 reanalysis data at our four sites indicates that the thermocline depth averages 83–107 m (±19–37 m) across sites, placing the 97 m reference depth consistently within one standard deviation of the mean thermocline depth at all sites. Also, as we argue in our manuscript, this species is mainly related to the Deep Chlorophyll Maximum and not directly to the thermocline. In response to the previous comment, the forward model has been updated to extract temperature and salinity averaged over a depth layer of 70–120 m rather than at a fixed point, better capturing the habitat variability of this species depending also on the site. This modification does not substantially alter the IFA standard deviations reported in the study (Fig. S2).

For the seasonal flux, we extracted sediment trap data from the FORCIS database (Chaabane et al., 2023) within a ±10° box around each core site to characterise the seasonal flux of *G. ruber* and *N. dutertrei.* The availability of the data and the proximity of the traps can vary considerably and lead to uncertainties. NIOP 905 benefits from a near trap (AS8, <1 km), providing a robust local flux record. For the remaining three sites, the nearest available traps are located 1312 km (U1467), 1223 km (SO189-39KL), and 946 km (MD96-2060) from the core sites respectively, reflecting the chronic undersampling of sediment trap records in the tropical Indian Ocean. These distances introduce uncertainty in the seasonal flux weights applied at these sites, as the trap records may not fully capture the local flux seasonality. Nevertheless, the seasonal patterns derived from these traps are coherent with known regional oceanographic forcing by the Indian SW monsoon at U1467 and NIOP 905 and the Australian SE monsoon upwelling at SO189-39KL, suggesting that the broad seasonal signals are regionally representative despite the spatial offset. The MD96-2060 record additionally comprises only 39 sediment cup samples spanning 2.3 years, further limiting its temporal resolution.

In response to the reviewer's comment, we assessed a sensitivity test to the forward model which has been updated to weight the monthly sampling of δ^{18}Oc values by the normalized monthly flux weights derived from the FORCIS trap data (Fig. S3). This modification introduces a seasonally realistic bias in the simulated IFA population, favouring months of high productivity. Incorporating this flux weighting modifies the IFA SD by +0.0003‰ to −0.082‰ across sites and species (Wilcoxon test, non-significant at MD96-2060 *N. dutertrei* and U1467 *G. ruber*; $p<0.001$ elsewhere). The largest absolute difference is recorded at SO189-39KL for N. dutertrei (−0.082‰, −9.1%). At NIOP 905, the flux-weighted IFA SD for *N. dutertrei* increases significantly (+9.5%) rather than decreasing, reflecting the site-specific complexity and influence of upwelling dynamics on subsurface flux seasonality, consistent with sampling occurring predominantly during the SW monsoon at this site.

We therefore show that the global dynamic and the comparison of site by total variability under both configurations and with both species remains the same. The highest surface SDs are still in the western site with MD96-2060 and NIOP 905 and subsurface in the east at site SO189-39KL. We then updated the forward-modelled script with this suggestion for future comparison. As we look at the SD and the range in IFA distribution, we demonstrated here that the main conclusion of the study is still robust regardless of the method used.

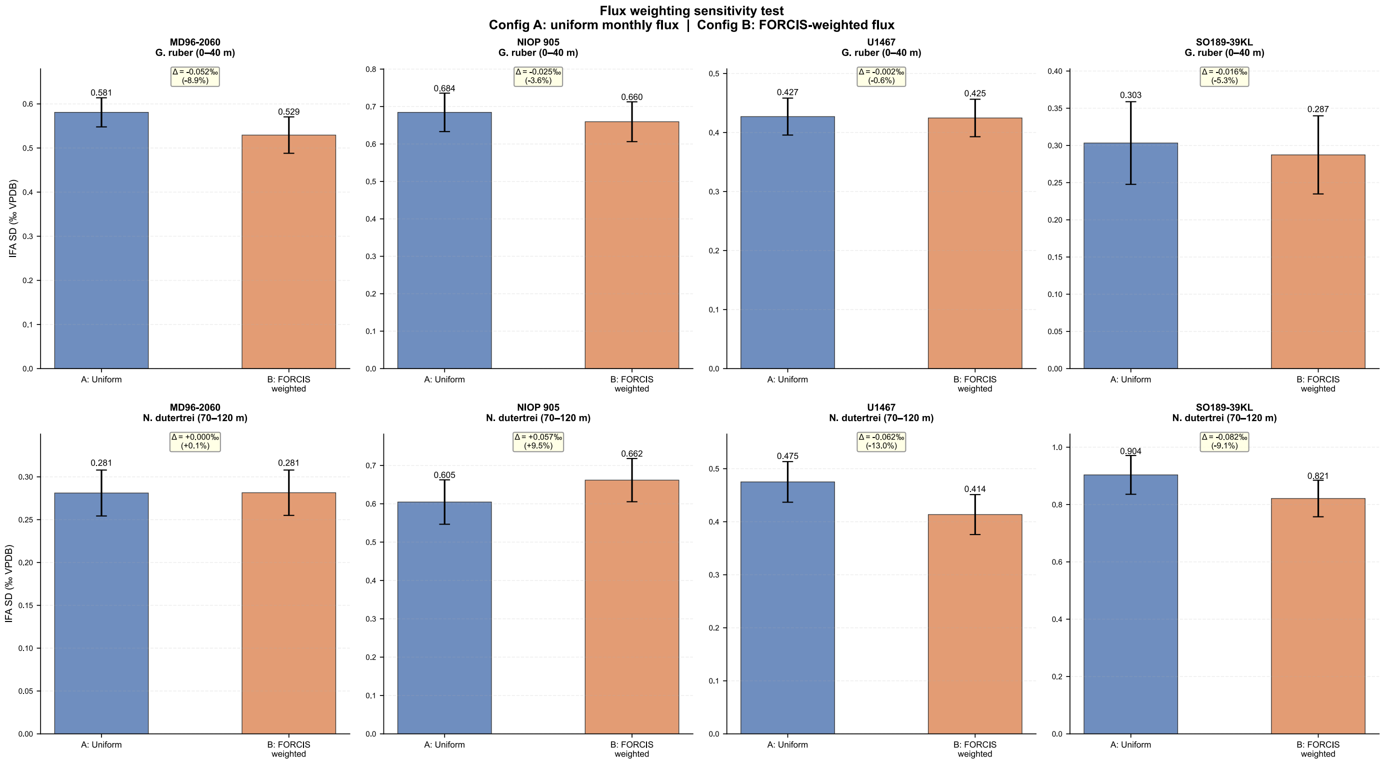

Figure S3. Sensitivity of forward-modelled IFA δ¹⁸Oc SD to seasonal flux weighting**.** IFA SD (mean ± 1σ across 1000 Monte Carlo simulations, N = 60 foraminifera) is shown for *G. ruber* (surface, 0–40 m; top row) and *N. dutertrei* (thermocline, 70–120 m; bottom row) at the four proxy sites, under seasonal ±75% amplification.

7) *"At present, the manuscript relies on forward modelling mean and variability of δ18Oc at fixed depths (0.5 m and 97 m), assuming constant monthly carbonate flux, a single δ18O paleotemperature equation across species, and constant δ18Ow–salinity relationships across all sites. These assumptions move the model outputs away from realistic oceanographic conditions. To make the forward-modelling framework more robust and useful for the paleoclimate community, the following revisions are recommended: a) Define the habitat depth and its variability for G. ruber and N. dutertrei at each site using nearby plankton-net or sediment-trap observations (e.g., FORCIS database). b) Incorporate seasonal flux information from sediment-trap records near the proxy sites (e.g., FORCIS database). c) Use regional and depth-specific δ18Ow–salinity relationships for both surface and thermocline waters, given that the four sites are oceanographically distinct. d) Apply species-specific δ18O paleotemperature equations instead of using the G .bulloides equation of Bemis et al. (1998) for all species. e) Propagate uncertainties through each modelling step and include them in the final outputs. f) Introduce realistic climate perturbations at both surface and thermocline depths based on observed ENSO/IOD variability near the study sites. Such modifications would make the forward-model outputs substantially more realistic and scientifically valuable for the IFA and paleoclimate communities."*

Each of the recommended modifications has been addressed individually in our responses above. Specifically: habitat depth (point a) and seasonal flux (point b) are addressed in our response to General Comment 6; regional δ^{18}Ow–S relationships (point c) are addressed in our response to Specific Comments 5–6; species-specific paleotemperature equations (point d) are addressed in our response to Specific Comment 4; realistic climate perturbations at surface and thermocline depths (point f) are addressed in our response to General Comment 3. Regarding uncertainty propagation (point e), we address this below and include this recommendation in our forward-modelling steps.

Specific comments:

1) *Line 47: The authors state "drier northeast monsoon during boreal winter." However, the northeast monsoon delivers substantial rainfall over southern India, Sri Lanka, and parts of Southeast Asia.*

Done, we have revised the sentence accordingly to better reflect the regional pattern of northeast monsoon precipitation.

2) *Figure 1: The authors used annual range for surface waters and SD for thermocline waters. Please specify the depth range over which the averages were calculated. Since the study objective is related to validating IFA δ18O SD, why not also present SD for surface waters in the figure?*

We thank the reviewer for this point. We have updated the caption to specify the depth used for each panel. As discussed before, we use a fixed depth that is representative of two distinct vertical depth zones: surface and subsurface. The goal of this figure is to introduce the global Indian Ocean variability with the studied site but also the difference between the surface and subsurface variability. We prefer to use surface-layer annual range instead of SD in this figure, because it illustrates where in the Indian Ocean the maxima and minima of seasonality occur.

Nevertheless, we have updated the main figure to also show standard deviation for the surface and we will use this new version of this figure in the revised manuscript. In general, it shows the same pattern of variability for each variable, with the northwestern area most sensitive to changes in seasonal variability. Moreover, the standard deviation of δ^{18}Oc is spatially the same as the temperature pattern.

3) Lines 160–161: Please report the mean thermocline depth and its standard deviation for the four sites based on ORAS5 reanalysis.

Done, we have added in this paragraph the mean thermocline depth and the associated standard deviation for all sites extracted from ORAS5 based on the depth of maximum vertical temperature gradient over 1958–2018: 83.2 ± 30.7 m at MD96-2060, 103.3 ± 37.9 m at NIOP 905, 85.1 ± 17.6 m at U1467, and 90.2 ± 27.2 m at SO189-39KL. We note that the selected thermocline depth (97 m) falls within the standard deviation of the mean thermocline at all sites.

4) Lines 163–164: Why are species-specific equations not being used? Instead, the study applies the G. bulloides and O. universa equations from Bemis et al. (1998) to simulate δ18Oc for G. ruber and N. dutertrei.

Done, the forward model now applies the *G. ruber*-specific equation of Farmer et al. (2007) (a = 15.4, b = 4.78) for surface simulations, and the *N. dutertrei*-specific equation of Farmer et al. (2007) (a = 14.6, b = 5.09) for thermocline simulations. A choice of equations will be available in the new revised version of the python script. We also performed a sensitivity analysis comparing all equations. For *G. ruber*, the two equations (Bemis et al., 1998 and Farmer et al., 2007) yield a systematic mean offset of ~+0.09‰ ($p < 0.001$), reflecting the difference in the intercept coefficient, while the IFA δ^{18}Oc SD differs by less than 0.005‰ and is statistically indistinguishable at three of four sites (Wilcoxon test, $p > 0.05$). For *N. dutertrei*, the mean offset is ~+0.03–0.04‰ ($p < 0.001$), and the IFA SD is reduced by 0.014–0.020‰ with the Farmer et al. (2007) equation, a difference that is statistically significant ($p < 0.001$) but smaller than the analytical error of d18O measurements (0.08‰) and well below the between-site differences in IFA SD. We therefore conclude that the adoption of species-specific equations does not significantly affect the conclusions of this study, although we agree with the reviewer that using species-specific equations is logical from an ecophysiological point of view.

5) Lines 166–167: The authors apply the Singh et al. (2010) δ18Ow–salinity regression for surface waters. However, this regression mainly reflects northern Indian Ocean conditions influenced by strong riverine freshwater input (https://doi.org/10.1016/j.dsr.2010.08.002). Please justify its applicability to the study locations. Please also report the R^2 value and associated uncertainty.

Our forward-modelling methodology is built around different steps and objectives. As discussed in the paper, in order to forward-model, we need to implement with ORAS5 reanalysis a salinity-$\delta^{18}O_{sw}$ equation for the surface and subsurface. The idea is to mimic the different planktonic foraminiferal species which record salinity and temperature signals.
Our framework is divided into two main parts:

1) The first part is to test the sensitivity of each regional site changes in seasonal and interannual variability. In order to have a basis of comparison we need to affect each site by the same coefficient of variability and the same method and equation to calculate the δ^{18}Oc of synthetic IFA.
   In the first step, a uniform δ^{18}Ow–S relationship is deliberately applied across all sites. This ensures that any differences in the simulated δ^{18}Oc response reflect contrasting local oceanographic conditions alone, rather than equation-dependent offsets.
   As a consequence, we chose to apply the same paleotemperature and δ^{18}Osw-S equations for each site to be sure that we test the response in an altered climate under the same parameters and sensitivity.
   In fact, the Singh et al. (2010) equation likely represents an approximation of the true δ^{18}Ow–S relationship at each site, which may affect how realistically the extrema of variability are reproduced in the synthetic IFA. For this reason, we differentiate this first step with a second one, in which regional equations are used instead. Unlike the first step, this second step is not designed to compare the sensitivity of each site to imposed variability changes; rather, its goal is to best fit the synthetic IFA data to real core-top data, by minimizing the systematic bias introduced by a single generic δ^{18}Ow–S relationship.

2) The second step in our forward-modelling approach aims to show how and at what point the modern data from reanalysis are able or not to reproduce the SD and mean of real IFA of Late Holocene core-tops. In order as realistic and precise as possible, site-specific regional δ^{18}Ow–S equations are used to minimize systematic biases when comparing modelled and measured δ^{18}Oc distributions. As a robustness check, we reran the variability sensitivity scenarios (seasonal and interannual, ±75%) using these regional equations at all four sites (Fig. S4). At two sites where the regional slope differs from the uniform value (U1467: 0.28 vs. 0.26; SO189-39KL: 0.16 vs. 0.26), the absolute IFA SD changes significantly (Wilcoxon $p<0.001$; +4.3% and −12.4% respectively, under original conditions), reflecting the direct scaling effect of the δ^{18}Ow–S slope on the salinity-driven component of δ^{18}Oc variance. However, the relative amplitude of the response to imposed climate variability (modified to the original SD ratio) is preserved at both sites confirming that our site ranking by sensitivity to climate variability the main conclusion of this study is robust to the choice of δ^{18}Ow–S equation. In this second step, we do not use the Singh et al., 2010 equation to compared synthetic and real IFA. Instead, we use regional specific equations.

To illustrate the effect of using regional equations even in step one of our frameworks, we used the site-specific equations reported in Section 5.3.3 (Stainbank et al., 2021 for U1467; Thirumalai et al., 2019 for SO189-39KL; LeGrande & Schmidt, 2006 for MD96-2060; Delaygue et al., 2001 for NIOP 905) in place of the uniform Singh et al. (2010) equation, and reran the sensitivity scenarios described above (Fig. S4). The Delaygue et al. (2001) equation used at NIOP 905 shares an identical slope (0.26) with the uniform equation and therefore leaves the IFA SD unchanged at this site, only shifting the mean modelled δ^{18}Oc by the difference in intercept (−8.9 vs. −8). We confirm that the mean modelled δ^{18}Oc shifts consistently with the change in intercept and slope at each site, and that the absolute IFA SD is also affected where the regional slope differs from the uniform one. Critically, the site ranking by sensitivity to seasonal and interannual variability the primary comparison underpinning our conclusions remains unchanged between the two configurations.

Taken together, these results show that while the choice of δ^{18}Ow–S equation has a measurable effect on the absolute IFA SD at some sites, it does not affect the site ranking or the relative amplitude of the sensitivity to climate variability, which is the central comparison of step one. We therefore retain the uniform Singh et al. (2010) derived equation ($\delta^{18}Ow = 0.26 \times S - 8$) for this first step, both as part of the broader simplification strategy underlying our forward-modelling framework, and to ensure direct comparability with the forward-modelling approach of Thirumalai & Clemens (2020) in the Bay of Bengal, who applied this same coastal equation (derived from Singh et al., 2010) in their own forward model. We have clarified this two-step framework explicitly in the revised manuscript with the new framework figure and added the R^2 value and associated uncertainties.

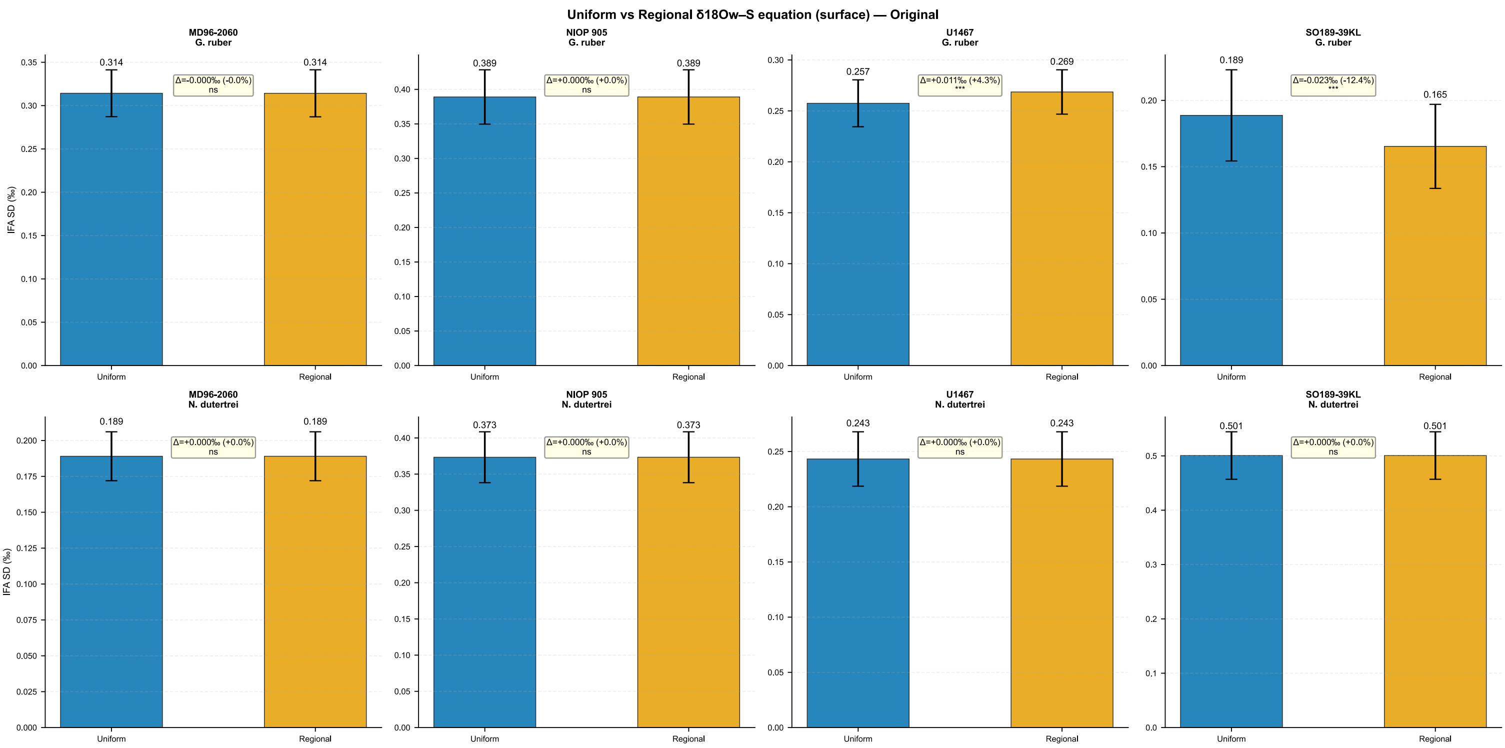

Figure S4. Sensitivity of IFA standard deviation to the choice of δ^{18}Ow–S equation (uniform vs. regional). Bars show mean IFA-derived standard deviation (± SD across 500 Monte Carlo simulations) for each site and species, using a uniform tropical Indian Ocean equation (blue) versus site-specific regional equations (orange). Δ and % values indicate the difference (regional − uniform); ns = not significant (Wilcoxon signed-rank test, paired).

6) *Lines 169–171: For subsurface δ18Ow–salinity relationships, the authors use a regression derived from a global dataset. Please show the locations of the water samples used from LeGrande and Schmidt (2006) along with proxy sites in a supplementary figure. Also provide the R^2 value and uncertainties in the slope and intercept.*

We have added a supplementary figure showing the location of the gridded data points from LeGrande & Schmidt (2006) used to derive the subsurface δ^{18}Ow–S regression. Also, we increase the precision of the regression by adding more data and the resulting equation is $\delta^{18}Ow = 0.449 (\pm 0.003)*S - 15.491 (\pm 0.089)$ ‰, with $R^2 = 0.931$. The R^2 value and uncertainties on slope and intercept have been added to the manuscript.

7) *Figure 2: The manuscript suddenly introduces World Ocean Atlas 2023 data, whereas ORAS5 is used throughout the forward modelling. Please use ORAS5 consistently for the depth-wise temperature and salinity plots at all four sites. The selection of 97 m as the thermocline depth appears to rely on WOA23 seasonality. However, the manuscript does not demonstrate how SD of salinity and temperature vary interannually with depth. The rationale for selecting 97 m is therefore not properly justified*

Done and text edited in the revised manuscript. The bottom panel shows the vertical structure of temperature SD at each site with a maximum for each site between 75 m and 125 m in the range for the 97 m value already discussed above. Also, the choice is justified by the mean thermocline depths computed from ORAS5 (83.2 ± 30.7 m at MD96-2060, 103.3 ± 37.9 m at NIOP 905, 85.1 ± 17.6 m at U1467, and 90.2 ± 27.2 m at SO189-39KL) (See Specific Comment 3).

8) Lines 198–200: This section is confusing. In Section 3.1.1, the authors use Singh et al. (2010) for surface waters and LeGrande and Schmidt (2006) for subsurface δ18Ow–salinity relationships. However, in Section 3.1.2, the manuscript refers to the δ18Ow–salinity relationship of Delaygue et al. (2001). Please clarify which relationship was actually used. In reality, δ18Ow–salinity relationships can vary seasonally. The authors should also evaluate how the SD of forward-modelled δ18Oc changes when different δ18Ow–salinity relationships are applied, and justify the use of a constant relationship.

The reference to Delaygue et al. (2001) in Section 3.1.2 concerns the second step of our forward-modelling framework (comparison with IFA core-tops), where region-specific δ^{18}Ow–S relationships are used, drawn from the original regional publications when available or from Delaygue et al. (2001) otherwise. This two-step structure is clarified further in our response to Specific Comment 5.
Regarding the seasonal variability of δ^{18}Ow–S relationships, we agree that these relationships can vary seasonally, particularly in regions strongly influenced by monsoonal freshwater input (Singh et al., 2010). As demonstrated in our sensitivity analysis (see response to Specific Comment 5 and Fig. S4), the choice of δ^{18}Ow–S equation has a measurable but modest effect on the simulated δ^{18}Oc SD, with absolute differences remaining on the order of 0.01–0.04‰ across sites even though this corresponds to a relative change of up to 12.4% at the most sensitive site (SO189-39KL). Critically, the ranking of sites by sensitivity to imposed climate variability, the central comparison of this study, is preserved regardless of the equation used. The use of a constant relationship for our framework is justified for the reasons detailed in our response to Specific Comment 5.

9) Lines 203–204: The authors state that "Each individual foraminifera is assumed to record the average monthly environmental conditions of its calcification period." Please provide justification based on the average lifespan of G. ruber and N. dutertrei.

The assumption that each individual foraminifera records the average monthly environmental conditions of its calcification period is justified by the known lifespan of both species. *G. ruber* and *N. dutertrei* have average lifespans of approximately 2–4 weeks (Spero, 1998; Bijma et al., 1990; Jonkers et al., 2022), which is comparable to or shorter than the monthly temporal resolution of the ORAS5 reanalysis used in this study. Under this assumption, each individual is considered to have calcified entirely within a single monthly interval, such that the monthly mean temperature and salinity provide a reasonable approximation of the environmental conditions experienced during calcification. We have added a brief justification to this effect in the revised manuscript (Lines 203–204).

10) Table 1: Please include uncertainties associated with the core-top ages.

Done, we have updated Table 1 by including the uncertainties. All uncertainties are based on 1σ ^{14}C AMS age. For U1467, no ^{14}C AMS age direct dating is available at the core-top level so the age model is based on orbital tuning using the Prob-stack of Ahn et al. (2017), with a sedimentation rate of 8.1 cm/kyr at the core-top.

11) Figure 3: Please include SD values in panels B and D, and discuss them in the interpretation section.

Done, we have added the SD for both panels B and D. At the surface (panel B), interannual SD values in $\delta^{18}Oc$ range from ~0.07 ‰ at MD96-2060 in summer to ~0.19 ‰ at U1467 in December, confirming that seasonal variability dominates over interannual variability at all sites. The highest SD are observed for temperature at site NIOP 905 during boreal summer and at site SO189-39KL for salinity throughout the year.
At the thermocline depth (panel D) the SD values are larger, particularly at SO189-39KL (~0.55 ‰) and MD96-2060 (~0.35 ‰), highlighting the greater sensitivity of these sites to interannual forcing. All the results have been discussed in the interpretation section.

12) Lines 318–320: The authors state that "At sites NIOP 905 and U1467, thermocline variability on interannual timescales is more muted, and seasonal variations in SST and SSS at thermocline depths are larger (Fig. 3D)." However, site NIOP 905 is influenced by the Somali Current. Does the absence of interannual variability in thermocline temperature and salinity imply that the interannual strength of the Somali Current is nearly constant? Is this interpretation supported by instrumental observations?

The phrasing of "more muted" was probably misleading as it could mean an absence of interannual thermocline variability at NIOP 905, which is not the case. Some studies describe the general behaviour of the Somali Current and its variability at large scale or seasonal scale (Beal and Donohue., 2013; Schott & McCreary., 2001) confirming that there is an interannual variability but which is not directly associated with IOD or/and ENSO events. At site NIOP 905, we have also checked in the reanalysis and instrumental dataset the interannual variability in this area. It reveals substantial interannual variability in thermocline temperatures, with a standard deviation of 1.25°C (ORAS5) and 1.47°C (EN4), and maximum anomalies reaching 5.4°C during strong IOD and ENSO events. The SD of interannual thermocline temperature is highest during boreal summer (JJA: 1.6°C, ORAS5), precisely the season of maximum Somali Current intensity and coastal upwelling along Somalia. It confirms in fact that the interannual variability strength of the Somalian current consistent with our observation.

The original statement referred to the fact that the interannual variability at this site is smaller and "muted" relative to the strong seasonal cycle (~11°C amplitude) and relative to the other thermocline variability observed at SO189-39KL and MD96-2060. We have clarified this in the revised manuscript.

13) Figure 4: It would be easier to follow if the panels were arranged longitudinally from west to east. This would reduce the need for readers to repeatedly identify site locations.

Figure 4 has been reorganized so that panels are arranged from west to east following the longitude of the study sites.

14) Line 446: Figure 1c represents the annual range of surface-water δ18Oc, not δ18O SD. The figure citation is therefore incorrect.

Done, corrected with the new Fig. 1C

15) Line 458: Incorrect citation of Figure 6B.

Done, corrected and figure was also range by longitude

16) Lines 527–529: This statement is misleading. Since interannual perturbations were not introduced at the surface, it is unclear how the authors conclude that interannual variability is observed in the surface simulations.

As clarified in the answer to general comment 4, we now update the forward-modelling with the possibility of modifying the interannual variability in the surface. We now present a sensitivity test comparing seasonal-only, interannual-only, and combined perturbations at the surface across all four sites, the results of which have been added to the Supplementary Material. We have revised the manuscript to clarify this distinction and refer to new Supplementary Material.

17) Lines 547–548: This statement appears overstated. I would describe the agreement between forward-modelled δ18O and core-top IFA δ18O as only partial, with approximately half of the sites showing agreement.

We have revised Lines 547–548 to more accurately reflect the partial nature of the agreement. At Sites MD96-2060, U1467, and SO189-39KL, the forward-modelled and measured IFA SD values are in reasonable agreement (differences < 0.03–0.05‰). However, at Site NIOP 905, the forward model systematically underestimates the measured IFA SD at both surface (synthetic σ = 0.377‰ vs. real σ = 0.565‰) and thermocline depths (synthetic σ = 0.431‰ vs. real σ = 0.526‰).

Technical comment:

1) Line 74: "Northeastern" appears incorrect; perhaps "northwestern" is intended.

Done

References:

Ahn, S., Khider, D., Lisiecki, L. E., and Lawrence, C. E. (2017): A probabilistic Pliocene–Pleistocene stack of benthic $\delta^{18}O$ using a profile hidden Markov model. Dynamics and Statistics of the Climate System, 2(1), dzx002. https://doi.org/10.1093/climsys/dzx002

Beal, L., & Donohue, K. (2013). The great whirl: Observations of its seasonal development and interannual variability. Journal of Geophysical Research: Oceans, 118, 1–13. https://doi.org/10.1029/2012JC008198

Behringer, D. W., and Xue, Y. (2004): Evaluation of the global ocean data assimilation system at NCEP: The Pacific Ocean. Eighth Symposium on Integrated Observing and Assimilation Systems for Atmosphere, Oceans, and Land Surface, AMS 84th Annual Meeting, Seattle, WA.

Bemis, B. E., Spero, H. J., Bijma, J., and Lea, D. W. (1998): Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. Paleoceanography, 13(2), 150–160. https://doi.org/10.1029/98PA00070

Bijma, J., Erez, J., and Hemleben, C. (1990): Lunar and semi-lunar reproductive cycles in some spinose planktonic foraminifers. Journal of Foraminiferal Research, 20(2), 117–127.

Chaabane, S., de Garidel-Thoron, T., Giraud, X., Schiebel, R., et al. (2023): The FORCIS database: A global census of planktonic Foraminifera from ocean waters. Scientific Data, 10, 354. https://doi.org/10.1038/s41597-023-02264-2

Chaabane, S., Schiebel, R., Meilland, J., Brummer, G.-J. A., Mortyn, P. G., Sulpis, O., Chalk, T. B., Giraud, X., Howa, H., Kuroyanagi, A., Beaugrand, G., and de Garidel-Thoron, T. (2026): Reassessment of the global distribution and diversity of modern planktonic foraminifera from the FORCIS database. Journal of Micropalaeontology, 45, 195–217. https://doi.org/10.5194/jm-45-195-2026

Delaygue, G., Bard, E., Rollion, C., Jouzel, J., Stiévenard, M., Duplessy, J.-C., and Ganssen, G. (2001): Oxygen isotope/salinity relationship in the northern Indian Ocean. Journal of Geophysical Research-Oceans, 106(C3), 4565–4574. https://doi.org/10.1029/1999JC000061

Fairbanks, R. G., Sverdlove, M., Free, R., Wiebe, P. H., and Bé, A. W. H. (1982): Vertical distribution and isotopic fractionation of living planktonic foraminifera from the Panama Basin. Nature, 298, 841–844.

Farmer, E. C., Kaplan, A., de Menocal, P. B., & Lynch-Stieglitz, J. (2007). Corroborating ecological depth preferences of planktonic foraminifera in the tropical Atlantic with the stable oxygen isotope ratios of core top specimens. Paleoceanography, 22, PA3205. https://doi.org/ 10.1029/2006PA001361

Glaubke, R. H., Thirumalai, K., Schmidt, M. W., and Hertzberg, J. E. (2021): Discerning changes in high-frequency climate variability using geochemical populations of individual foraminifera. Paleoceanography and Paleoclimatology, 36(2), e2020PA004065. https://doi.org/10.1029/2020PA004065

Good, S. A., Martin, M. J., and Rayner, N. A. (2013): EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates. Journal of Geophysical Research: Oceans, 118, 6704–6716. https://doi.org/10.1002/2013JC009067

Jonkers, L., Brummer, G.-J. A., Meilland, J., Groeneveld, J., and Kucera, M. (2022): Variability in Neogloboquadrina pachyderma stable isotope ratios from isothermal conditions: implications for individual foraminifera analysis. Climate of the Past, 18(1), 89–101. https://doi.org/10.5194/cp-18-89-2022

LeGrande, A. N., and Schmidt, G. A. (2006): Global gridded data set of the oxygen isotopic composition in seawater. Geophysical Research Letters, 33, L12604. https://doi.org/10.1029/2006GL026011

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A. (2003): Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. Journal of Geophysical Research, 108(D14), 4407. https://doi.org/10.1029/2002JD002670

Schott, F. A., and McCreary, J. P. (2001). The monsoon circulation of the Indian Ocean. Progress in Oceanography, 51, 1–123. https://doi.org/10.1016/S0079-6611(01)00083-0

Singh, A., Jani, R. A., and Ramesh, R. (2010): Spatiotemporal variations of the $\delta^{18}O$–salinity relation in the northern Indian Ocean. Deep Sea Research Part I, 57, 1422–1431.

Spero, H. J. (1998): Life history and stable isotope geochemistry of planktonic foraminifera. The Paleontological Society Papers, 4, 7–36.

Stainbank, S., Spezzaferri, S., Rüggeberg, A., Raddatz, J., de Leau, E. S., Yu, S. M., Zhang, M., and Kroon, D. (2021): Monsoon and tropical climate forcing on the physicochemical and thermocline characteristics of the Maldives Inner Sea: Insights from Marine Isotope Stages 1-2 and 10-13. Paleoceanography and Paleoclimatology. https://doi.org/10.1029/2020PA004105

Thirumalai, K., DiNezio, P. N., Tierney, J. E., Puy, M., and Mohtadi, M. (2019): An El Niño mode in the glacial Indian Ocean? Paleoceanography and Paleoclimatology, 34(8), 1316–1327. https://doi.org/10.1029/2019PA003669

Thirumalai, K., and Clemens, S. (2020): Monsoon reconstructions using bulk and individual foraminiferal analyses in marine sediments offshore India. Current Science, 119, 328–334.