Peer Review File

Manuscript Title: Future increase in extreme El Niño supported by past glacial changes

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referee #1 (Remarks to the Author):

This is a very nice study that combines paleoclimate simulations and proxy data to investigate the evolution of ENSO variability during the last deglaciation. Analysis is focused on the Last Glacial Maximum and the comparison with the preindustrial. Model and proxy data consistently suggest a weaker ENSO variability and a less frequent extreme events during the LGM. Through a mechanistic understanding and additional simulation with 2xCO2, the authors suggest that the extreme El Nino events could be more frequent in response to future warming.

This is a solid study with in-depth analysis. However, I struggled to find sufficient relevance and novelty for a flagship journal such as Nature. Please see my reasoning below.

1. The modeling results are not novel. First, the increase of ENSO extremes with warming has been consistently simulated by the new generation of climate models. For example, Cai et al. (2021; doi:10.1038/s43017-021-00199-z) reported that “100% of 23 CMIP6 models indicate enhanced EP ENSO variance in the future, while 65% generate an increase in CP ENSO variance” and “The enhanced variability in EP and CP ENSO is associated with increased occurrence of extreme EP El Niño and extreme La Niña events. Indeed, extreme El Niño and La Niña events are both projected to increase from 5.6 events per century in the present day to 8.9 and 8.3 events per century in the future climate, respectively.” The improved consensus among models has been linked to the modest improvements in simulated Pacific mean state, and ENSO patterns and diversity. Second, the overall evolution of ENSO variation during the last deglaciation is similar to Liu et al. (2014; doi:10.1038/nature13963). Although CESM1 simulates a much more realistic ENSO variability than CCSM3, as was pointed out by the authors, both models show a weaker ENSO at the LGM that first weakens during the deglaciation and then strengthens towards the Late Holocene. In addition, the weaker ENSO in the CESM1 LGM simulation has been reported in Zhu et al. (2017; doi:10.1002/2017GL073406). The authors seem to focus their discussion more on the extreme ENSO events, but the extreme events are closely correlated with the overall ENSO variability (their Figure 1C).

2. Similarly to 1, results from the individual foraminiferal analyses (IFA) in this study are from published studies. The authors’ interpretation of the data is basically the same as the original authors, i.e., proxy data reflects a weaker ENSO variability during the LGM.

3. On multiple occasions, the authors seem to be overconfident with their results and the simulated ENSO dynamics in CESM1. ENSO events and their responses to external forcings are determined by a delicate balance of feedback processes and interactions with the mean state climate and the seasonal cycle. It is not clear whether the current GCMs, even when multiple models do produce consistent results, are able to simulate the ENSO dynamics in sufficient accuracy. (1) The current GCMs simulate an excessive eastern Pacific cold tongue and a double ITCZ, which will impact the delicate balance of ENSO feedbacks. (2) Current GCMs simulate a reduced east-west temperature gradient in the equatorial Pacific during the historical period while the instrumental record suggests otherwise (see Seager et al., 2019 for a recent discussion; doi: 10.1038/s41558-019-0505-x). The biases in the simulated mean state climate and trend in SST gradient to warming may potentially reflect a major deficiency of GCMs to simulate the coupled tropical dynamics. (3) Several recent studies suggest that models with a much higher horizontal resolution may be needed in order to partly resolve these biases and simulate the relevant ENSO dynamics. Wang et al. (2022; doi:10.1038/s41561-021-00890-2) use CESM1 at a higher horizontal resolution (~0.25deg atmosphere and ~0.1deg ocean) and find the equatorial eddies generate an anomalous dampening effect on ENSO events, which is missing in the majority of state-of-the-art climate models at a non-eddying resolution. Using CESM1 in a similar high-resolution configuration, Wengel et al. (2021; doi:10.1038/s41558-021-01132-4) show that the simulated ENSO variability weakens in response to an increase of the atmospheric CO_2, a result that is different from the coarser resolution simulations in the same model. In Wengel et al. (2021), the ENSO weakening is caused mainly by stronger latent heat flux damping and weaker advective feedbacks at the eddying resolution. Taken together, these high resolution CESM1 simulations demonstrate that the lower-resolution simulations with ~1–2deg horizontal resolution do not adequately simulate important ENSO-relevant ocean mesoscale processes and may provide incorrect ENSO responses to external forcings.

Referee #2 (Remarks to the Author):

DiNezio and Thirumalai produce a series of simulations using CESM at intervals of 3 ka over the deglaciation, forced by changes in greenhouse gasses, ice sheet extent, and insolation. Changes in meltwater input and associated changes in the Atlantic Meridional Overturning Circulation are not considered. They argue that ENSO is simulated very well in this particular model, and find that ENSO variability is linked to the mean state of the Walker Circulation through changes in the mixed layer depth and extent of the warm pool. They validate the results through the comparison of model results with individual foraminifera analysis from the LGM.

While others have examined the behavior of ENSO during the LGM, the significance here is the analysis of the factors leading to the weaker Bjerknes feedback, and the performance of the simulations in a model that does a good job with modern day ENSO. If they have correctly identified the major controls on ENSO variability, this helps us to better understand how ENSO will change in the future. The paper is clearly written and the illustrations appealing, although I think moving a few details into the main text from the supplement (as noted below) and providing more detail in the figure captions will help the reader better follow the paper. Overall, this is a strong and well-presented paper.

My expertise is in the paleo proxy data, so will focus on the model/data comparison in my review and leave others to comment on the mechanisms.

Evaluation of the CESM LGM mean state:

The logical layout of the paper is that the authors will check if the mechanism they identify operating in the model applies to the real world by validating using the LGM individual foraminifera analyses. It would then seem logical to first make the case that the controlling factors (mixed layer depth, trade-wind strength, warm pool edge) are well simulated by this particular model for the LGM. DiNezio (2011) showed that while the tropical overturning strengthens in all models as expected, the Pacific Walker circulation strength is highly variable in climate models and depends on how the maritime continent is handled. Maybe this is done in another paper and I missed the reference?

The authors mention a few observational studies suggesting a deeper thermocline, but a more thorough comparison of the model state with the data seems warranted. While there may not be direct proxies for mixed layer depth, trade-wind strength or the latitude of the warm pool edge, it would seem sensible to check the related quantities for which proxies do exist (thermocline depth, SST patterns, precipitation linked records in the tropics, etc.) Below are some recent LGM data compilations relevant to the tropical Pacific you might consider.

Tierney et al. (2020) Nature - Mg/Ca and alkenone data compilations
Monteagudo et al. (2021) GRL – Mg/Ca SST patterns and d18Osw (salinity)
Seltzer et al (2020) Nature – land temperature from noble gasses
McGee (2020) Annual Reviews – Precipitation sensitive proxies
DiNezio and Tierney (2013) Nature Geoscience- Precipitation sensitive proxies
Leech et al (2013) EPSL – thermocline structure

Relationship between ENSO variability and Climate Forcing over all runs at 15 ka minimum:

While the authors argue that trade wind strength, mixed layer depth, and warm pool edge can explain the patterns in variability, they do not explain how these factors are linked to the forcing mechanisms. How big a role is CO2 playing? Local Insolation? Ice Sheet size? Why do they find minimum variability at 15 ka? There are a number of tropical Pacific records with extremes around 15ka (e.g. Borneo Speleothems, EEP T and productivity, etc. --see refs below). Given the timing coincident with Heinrich Stadial 1, the extremes in the records have been interpreted as a consequence of AMOC change and the attendant atmospheric reorganization, but this model does not know about meltwater and AMOC, so the authors may have uncovered another mechanism to explain this pattern. So I am very curious about what is causing the minimum at 15ka the model. Can this be discussed, perhaps with the major forcings be added as a panel in Figure 3 (CO2, insolation during a critical month)?

Partin et al., 2007, Nature 449, 452-455.
Kienast et al., 2013, Paleoceanography 28, 663-674
Kienast et al., 2006, Nature, 443, 846-849

Consideration of calcification depth of subsurface species:

The authors compare the IFA from both N. dutertrei and G. tumida to variability at 70 m model depth, based on studies of modern living depth (tow data) which reflects the standing stock in the water column including all life stages. However, calcification depth (the average depth at which foraminifera on the sea floor calcified) is more relevant to the IFA, since many species will add calcite at a deeper depth at the end of their life cycle. G. tumida has a deeper calcification depth than N. dutertrei, and the calcification depth of N. dutertrei partially follows thermocline depth (shallower where the thermocline is shallower, deeper where it is deeper. The 70 m depth is probably a good choice for N. dutertrei in the Eastern Pacific today, but not necessarily for the Western Pacific or for the LGM when the thermocline is deeper. G. tumida calcification depth is less sensitive to thermocline depth today, but is likely around 100 m deeper than N. dutertrei. I'm not sure how big a difference it would make, but would the model produce different LH-LGM variability differences for N. dutertrei if the calcification habitat were adjusted for the deeper thermocline during the LGM? Would the G. tumida data from site 849 be in better agreement with the model if a deeper calcification depth were used?

There doesn't seem to be much change at 150-200 m at Site 849 (Extended Data Figure 2), which would help explain the G. tumida IFA data at this site. But overall, I agree with the sense that it is more difficult to directly interpret the subsurface data and compare to the model given the calcification depth uncertainties and potential model biases in thermocline structure.

Some recent papers which will discuss many of the relevant studies on calcification depth for subsurface species, as well as look at the link between N. dutertrei calcification depth and thermocline depth:

Ford et al. (2018) GRL (see supplement)
Lakhani et al (2022) Marine Micropaleontology

Missing study:
The IFA results from the central tropical Pacific (Rustic et al., 2020) would seem quite relevant. It looks like their glacial data might not be close enough to 21 ka to include in this study, but do seem to contradict it (higher LGM variability) so should be mentioned. I wonder why they didn't analyze 21 ka? They also make a connection between ENSO variability and mixed layer depth, as well as ITCZ position.

Author Rebuttals to Initial Comments:

Reply to Reviewer 1

We are immensely thankful to reviewer 1 for their insightful comments that helped us distill the novelty of our study and clarify its key implications. We believe the manuscript has improved dramatically after addressing their critique. Here we provide responses to their comments (points 1 to 3) which we reproduce in italics together with our responses.

1. The modeling results are not novel. First, the increase of ENSO extremes with warming has been consistently simulated by the new generation of climate models. For example, Cai et al. (2021; doi:10.1038/s43017-021-00199-z) reported that "100% of 23 CMIP6 models indicate enhanced EP ENSO variance in the future, while 65% generate an increase in CP ENSO variance" and "The enhanced variability in EP and CP ENSO is associated with increased occurrence of extreme EP El Niño and extreme La Niña events. Indeed, extreme El Niño and La Niña events are both projected to increase from 5.6 events per century in the present day to 8.9 and 8.3 events per century in the future climate, respectively." The improved consensus among models has been linked to the modest improvements in simulated Pacific mean state, and ENSO patterns and diversity. Second, the overall evolution of ENSO variation during the last deglaciation is similar to Liu et al. (2014; doi:10.1038/nature13963). Although CESM1 simulates a much more realistic ENSO variability than CCSM3, as was pointed out by the authors, both models show a weaker ENSO at the LGM that first weakens during the deglaciation and then strengthens towards the Late Holocene. In addition, the weaker ENSO in the CESM1 LGM simulation has been reported in Zhu et al. (2017; doi:10.1002/2017GL073406). The authors seem to focus their discussion more on the extreme ENSO events, but the extreme events are closely correlated with the overall ENSO variability (their Figure 1C).

Reviewer 1 raises novelty concerns regarding studies focusing on multi-model predictions. We agree that models are starting to converge in their predictions of a future increase in ENSO variability as seen in our model. The study mentioned by the reviewer (Cai et al. 2021; doi:10.1038/s43017-021-00199-z) was cited in the introduction of our manuscript to highlight this emerging consensus from models (line 8 of the revised manuscript).

We recognize the power of multi-model studies, e.g., employed by Cai et al. (2021), but we also propose that validation against past changes is equally important to constrain model projections. As reviewer 1 points out, models exhibit pervasive, systematic errors, and therefore, inter-model consensus might not necessarily translate into better projections, particularly without validation against "real world" observations, historical or paleo. In addition to validating responses against past changes, we propose a novel mechanism that operates under past and future conditions, allowing the use of past (paleo) changes as targets for validation. This common mechanism represents a major advance for using paleodata to validate future changes predicted by models.

Ours is the first study combining past and future simulations with paleodata to produce validated projections of ENSO changes in a warming climate. Our mechanism is novel as it proposes that changes in the frequency of extreme El Nino governs the changes in ENSO. This mechanism is consistent with an emerging paradigm that ENSO is not an oscillation and instead a series of events initiated every time an El Nino is triggered.

In sum, our study makes the following novel advances: 1) proposes a mechanistic framework to connect past and future changes in ENSO, 2) validates our model with new and existing (paleo) observations capturing surface and subsurface variability associated with ENSO, 3) identifies the importance of mechanical coupling—a physical process previously overlooked—as the main coupled processes governing past and future ENSO changes, and 4) explains how changes in extreme El Nino events control ENSO's overall response to external forcings. Point #3 is essential since it considers asymmetries in ENSO dynamics – a new area of ENSO research initiated by studies by Kessler, Okumura, Deser, and DiNezio.

More importantly, we hope that Reviewer 1 will agree that validating models against past ENSO changes is critical to increase confidence in model-based predictions. Changes during the instrumental period cannot be used as target for validation because ENSO is highly variable making the detection of externally forced changes very challenging – a consistent feature of model simulations since the pioneering work of Wittenberg (2009).

Our study overcomes this issue using paleodata from the best sampled past climatic interval with changes in global mean temperatures, the Last Glacial Maximum, for which several reconstructions of past ENSO variability, show a consistent and rather large reduction in strength, allowing us to establish that ENSO is sensitive to externally forced climate changes. Such extensive and rigorous validation considering internal variability and changes in seasonality has not been performed before. We note that Zhu et al. (2017; doi:10.1002/2017GL073406) performed a limited validation focusing on just one site in the eastern Pacific, concluding that this site is dominated by changes in seasonality – in agreement with our analysis but without validating using data from ENSO sensitive sites. Our analysis spans a total of seven Pacific Ocean sites including the newly generated IFA data from the center of action of ENSO in the central Pacific (see Figs. 2-3). Data from all sites support the simulated changes and the two sites in the central and eastern equatorial Pacific show a large weakening of ENSO variability that is very likely to be forced (98% probability). These independent datasets allowed us to perform a more systematic and quantitative validation of the model than what could be performed with a single site as done by Zhu et al. (2017). Our quantitative model-IFA comparison quantifies the probability that internal variability and changes in seasonality explain the glacial changes. **No previous study, including those mentioned by reviewer 1, has attempted this, nor has provided a common mechanism with future responses.**

Furthermore, Liu et al. (2014) did not use the wealth of IFA data that we use here and thus could not arrive at robust conclusions regarding glacial ENSO changes. Furthermore, Liu et al. (2014) did not connect past with future via a common mechanism as we do here. Last, the CCSM3 model used by Liu et al. (2014) was excessively oscillatory as we show in our supplemental discussion, thus did not simulate the mechanism presented here. In fact, the CCSM3 predicted weaker ENSO under warming—in contrast to the consensus of CMIP models including our model. This model error could explain the rather muted sensitivity of ENSO to glacial forcings reported by Liu et al.

A weaker ENSO under LGM conditions was reported in Zhu et al. (2017) as pointed out by reviewer 1—however that study had a much more limited scope than ours despite using the same model. First, Zhu et al. did not validate the changes against reconstructions of past ENSO variability. It only focused on demonstrating that one site in the far eastern Pacific is not sensitive to changes in ENSO and instead is sensitive to changes in seasonality. Second, Zhu et al. did not explore other intervals rather than the LGM as potential targets for validation. Third, Zhu et al. (2017) did not consider the asymmetry of ENSO in the analysis, potentially conflating mechanisms. Last, Zhu et al. (2017) did not connect changes in ENSO to the changes in mean climate, such as the depth of the mixed layer, and particularly, they did not connect these changes to global cooling and global warming via changes in the strength of the Walker Circulation as we do in our study. This connection is essential to link past and future changes, one of the novel features of our study. For these reasons our analysis represents a major and novel advance relative to Zhu et al. (2017).

It is important to emphasize that Zhu et al. (2017) did not consider the asymmetric dynamics of El Nino and La Nina events to quantify feedbacks (Kessler 2002; DiNezio and Deser 2014). This feature of the ENSO phenomenon is now well accepted (Timmermann et al. 2018, https://doi.org/10.1038/s41586-018-0252-6). Not considering asymmetry is a key limitation of both the Liu et al. (2014) and Zhu et al. (2017) studies which used the BJ index technique to quantify the changes in the strength of feedbacks underlying ENSO dynamics. This technique groups El Niño and La Niña events together and therefore conflates their sensitivity to changes in climate. This technique was useful for previous generations of models which erroneously simulated ENSO as a symmetric oscillation (*e.g.*, Liu et al. 2014). But models have improved in their simulation of ENSO dynamics. Among them, CESM1 simulates these asymmetries with striking realism (*e.g.*, DiNezio et al. 2017, doi:10.1007/s00382-017-3575-3). Therefore, an approach considering ENSO asymmetry is needed to estimate mechanisms driving ENSO changes, otherwise using the BJ index could lead to erroneous conclusions.

For instance, Zhu et al. (2017) conclude that under LGM conditions, the Bjerknes feedback and ENSO weaken due to a weaker thermocline and upwelling feedbacks in contrast to our finding of a dominant role by zonal advection feedback. The thermocline and upwelling feedbacks are more important in the dynamics of La Niña, involving vertical displacements in the thermocline (e.g. Extended Data Figure 4). Therefore, according to Zhu et al. (2017), the reduction in the BJ index seems dominated by the changes in La Nina and not by the changes in El Niño.

Our finding that extreme El Niño events are the key mechanism whereby the ENSO phenomenon mainly responds to climate changes is another breakthrough of our study that is entirely novel. The connection between extreme El Niño and overall ENSO variability is not trivial. This link is a direct result of the fact that the ENSO phenomenon is not an oscillation and instead a series of events energized every time a new El Nino event is triggered (Kessler 2002, DiNezio and Deser 2014). Our study is the first to show that extreme El Niño events are particularly sensitive to changes in climate via changes in the responsiveness of zonal currents involved in the Bjerknes feedback. Explicitly accounting for asymmetric mechanisms driving El Niño and La Niña is key to quantifying the strength of feedback, something that the technique used by Liu et al. (2014) and Zhu et al. (2017) is not capable of doing. Furthermore, we consider the seasonality of these feedbacks across past climates, something that no previous study has done and is essential to study ENSO since this phenomenon is "phase locked" to the annual cycle – hence its name.

To better convey these nuances, we rewrote the paragraph discussing the implication of our results (lines 252 – 267). This paragraph emphasizes:

1. The importance of validation against multiple independent datasets,
2. The importance of identifying a common mechanism operating under past and future conditions,
3. The importance of extreme El Nino as the most sensitive aspect of ENSO to external forcing. La Nina is determined by El Nino. Therefore, extreme El Nino govern the changes in ENSO variability,
4. The novelty of mechanical coupling as the critical mechanism governing the influence of mean state on ENSO (via extreme El Nino). This mechanism, which plays a role in the strength of the Bjerknes feedback, is the critical "knob" for the frequency at which El Nino events reach peak magnitude.

2. Similarly to 1, results from the individual foraminiferal analyses (IFA) in this study are from published studies. The authors' interpretation of the data is basically the same as the original authors, i.e., proxy data reflects a weaker ENSO variability during the LGM.

In the revised version of our manuscript, we include a new dataset from the central Pacific bolstering the robustness of the existing data. We would like to emphasize that our study includes novel quantitative model-based analysis to determine if the changes seen in the paleodata reflect externally forced changes in ENSO. This is the first time that existing IFA datasets are used to evaluate models using these quantitative methods that account for habitat, seasonality, and unforced variability.

In addition, previous studies presenting IFA data did not compare those data against models considering changes in seasonality, internal variability, or calcification depth. Our new IFA dataset of stable isotopes in individual shells of *Globigerinoides ruber* (surface-dwelling taxa) from central equatorial Pacific sediments—the center of action of ENSO—robustly supports our main conclusion of a weaker ENSO during this interval. Note that variability in this site is unambiguously influenced by changes in ENSO and not by changes in the annual cycle of the Pacific cold tongue. The new data are presented in lines 83 – 115 of the revised manuscript (also in the new Figures 2 and 3).

Furthermore, we show that each dataset, by virtue of its location, captures different aspects of the signature of the ENSO changes. This is particularly important for subsurface records which may be sensitive to reduced ENSO conditions or amplified seasonal variability. This is what we meant in our statement in our manuscript "a systematic, model-based attribution of these changes has not been carried out". Our validation, in addition to calcification depths, also includes collective statistical tests of changes in seasonality as a driver for overall changes in variability–something that no previous study has done.

We note that neither Liu et al. (2014) nor Zhu et al. (2017) perform a systematic comparison with proxy data as we do here–thus, they are unable to arrive to a systematic conclusion of the probability that the ENSO changes are forced. This is important because ENSO is highly variable, therefore IFA data showing weaker ENSO could be the result of sampling a period of naturally occurring lower ENSO variability. By taking advantage of multiple sites with ENSO-sensitive changes, we can quantify the probabilities that ENSO changes were forced during the LGM. In particular, the eastern equatorial site of Ford et al. (2015) and our new central equatorial Pacific site showes very large reductions in upper-ocean temperature variability that are best explained by an externally forced reduction (Figure 3 and Extended Data Figure 1). In the revised version of the manuscript, we focus on these two datasets to attribute ENSO changes (Figs. 2 & 3).

3. On multiple occasions, the authors seem to be overconfident with their results and the simulated ENSO dynamics in CESM1. ENSO events and their responses to external forcings are determined by a delicate balance of feedback processes and interactions with the mean state climate and the seasonal cycle. It is not clear whether the current GCMs, even when multiple models do produce consistent results, are able to simulate the ENSO dynamics in sufficient accuracy. (1) The current GCMs simulate an excessive eastern Pacific cold tongue and a double ITCZ, which will impact the delicate balance of ENSO feedbacks. (2) Current GCMs simulate a reduced east-west temperature gradient in the equatorial Pacific during the historical period while the instrumental record suggests otherwise (see Seager et al., 2019 for a recent discussion; doi: 10.1038/s41558-019-0505-x). The biases in the simulated mean state climate and trend in SST gradient to warming may potentially reflect a major deficiency of GCMs to simulate the coupled tropical dynamics.

We recognize that models suffer from pervasive errors in the simulation of the mean climate, yet their simulation of ENSO variability has improved dramatically in recent years. CESM1 simulates ENSO with high realism, as discussed in the supplemental, but also in a recent intercomparison of models by Planton et al. (2021, https://doi.org/10.1175/BAMS-D-19-0337.1). We also acknowledge and discuss the issue of the observed cooling trends in the Pacific (lines 272-277). We note that many of studies cited by reviewer 1 (e.g. Cai et al. 2021*)* also suffer from this uncertainty and focus on a subset of models identified using similar metrics of ENSO realism. Regarding observed trends, we note that the paleoclimatic evidence of an El Nino-like pattern during the Pliocene suggests that as CO2 concentrations and climate equilibrates, the Pacific will start warming according to models.

Also, we feel our study using CESM1 is still relevant because the most updated version, CESM2, has a degraded ENSO relative to CESM1 (Capotondi et al. 2020) as well as an unrealistic glacial simulation (Zhu et al. 2021).

(3) Several recent studies suggest that models with a much higher horizontal resolution may be needed in order to partly resolve these biases and simulate the relevant ENSO dynamics. Wang et al. (2022; doi:10.1038/s41561-021-00890-2) use CESM1 at a higher horizontal resolution (~0.25deg atmosphere and ~0.1deg ocean) and find the equatorial eddies generate an anomalous dampening effect on ENSO events, which is missing in the majority of state-of-the-art climate models at a non-eddying resolution. Using CESM1 in a similar high-resolution configuration, Wengel et al. (2021; doi:10.1038/s41558-021-01132-4) show that the simulated ENSO variability weakens in response to an increase of the atmospheric CO2, a result that is different from the coarser resolution simulations in the same model. In Wengel et al. (2021), the ENSO weakening is caused mainly by stronger latent heat flux damping and weaker advective feedbacks at the eddying resolution. Taken together, these high resolution CESM1 simulations demonstrate that the lower-resolution simulations with ~1–2deg horizontal resolution do not adequately simulate important ENSO-relevant ocean mesoscale processes and may provide incorrect ENSO responses to external forcings.

We recognize that these mechanisms involving mesoscale features are not fully active in our model or any IPCC-class model. However, the inclusion of this mechanism would act in unison with our mechanism based on the zonal advection feedback since both would act to strengthen ENSO variability under warming.

More importantly, reviewer 1 also mentions that in these high-resolution simulations, the changes in ENSO are dominated by a stronger atmospheric damping under 4XCO2 levels. We agree that it is important to consider higher levels of CO2 forcing in our mechanistic explanation of ENSO changes. To address this issue, we include an equilibrated 4XCO2 simulation in our revised manuscript and analysis. We find that, indeed, equilibration to 4XCO2 forcing produces warming in the eastern equatorial Pacific that limits the amplitude of El Nino via a much stronger atmospheric damping. However, after careful analysis of our simulations we find that this mechanism is not dominant under 2XCO2 forcing because the center of action of El Nino events shifts to the east, and thus, events can maintain the same amplitude as in the PI climate. We include results from this analysis in lines 157 – 198 of the revised manuscript and as an Extended Data Figure 2. This analysis also compares the climatological and El Niño sea-surface temperatures against observations to show the realism of our model in terms of mean climate (Extended Data Figure 2E, blue vs. black lines).

Our analysis of the new 4xCO2 simulation shows that our proposed mechanism breaks down under this high CO2 forcing, consistent with Wengel et al. (2021). This level of CO2 forcing is consistent with high emission scenarios that seemed plausible a few years ago. However, such high CO2 levels are becoming less likely due to progress in the transition to renewable energy (https://www.nature.com/articles/d41586-020-00177-3). Therefore, our mechanism under 2xCO2 levels will likely be valid for the rest of this century and thanks to our validation against past changes represents a robust constrain on future ENSO changes.

Wengel et al. (2021) show that ENSO remains unchanged under 2xCO2 forcing. However, their simulations are too short to isolate forced changes given the high unforced variability (Wittenberg 2009). For instance, we subsampled ENSO variability in our simulations for comparison with analysis period used by Wengel et al. to estimate ENSO variability (see Figure R1.1 below in this response). We find many 80-year intervals in which ENSO variability is comparable under PI and 2xCO2 conditions. This suggests that the high-resolution simulations analyzed by Wengel et al. might be too short to isolate forced changes.

In sum, our results are not different from Wengel et al. (2021), suggesting that ENSO simulation in the high and the standard resolution configurations of CESM1 are not structurally different -- other than the effect of mesoscale features discussed above. Our mechanism remains valid for the range of CO2 forcing expected for this century, and therefore, our analysis provides a robust assessment of the risk of increased extreme El Nino as our climate warms.

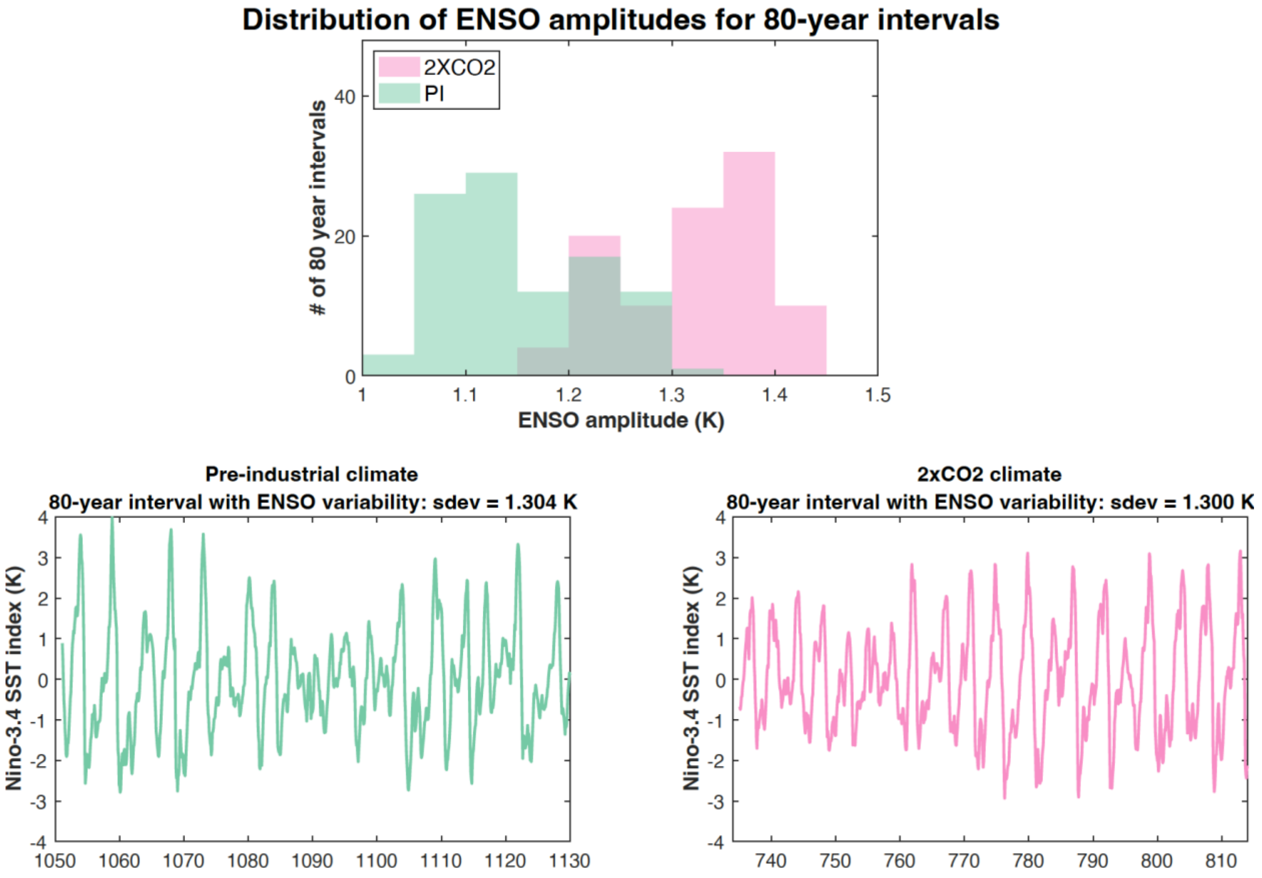

Figure R1.1 Distribution of the standard deviation of the Nino-3.4 SST index over overlapping 80-year intervals simulated by the standard resolution CESM1 (our model) under pre-industrial (aqua) and CO2 doubling (pink) conditions. The 80-year interval is selected for comparison with analysis period used by Wengel et al. (2021) to estimate ENSO variability. Two 80-year intervals with comparable ENSO variability, measured by the standard deviation of the Nino-3.4 SST index, are also shown.

Last, previous studies have overlooked major deficiencies in the simulation of ENSO in the HR CESM1 that could affect its sensitivity to external forcings. ENSO in the HR CESM does not show consistent El Nino to La Nina transitions as observed. This is seen in a lagged autocorrelation function of the Nino-3.4 SST index that does not show negative correlations at 1-to-2-year lags as observed (Figure R1.2, left). Furthermore, the levels of variability simulated by the HR CESM1 are much lower than observed or than simulated by the standard resolution configuration (Figure R1.2, right).

These are a major weakness of the HR CESM because it indicates that simulated La Niña do not follow El Niño as observed. In contrast our configuration of the model and the standard (Figure R1.2 and Extended Data Figure 3). Therefore, we feel it is premature to consider the responses in the HR CESM as robust.

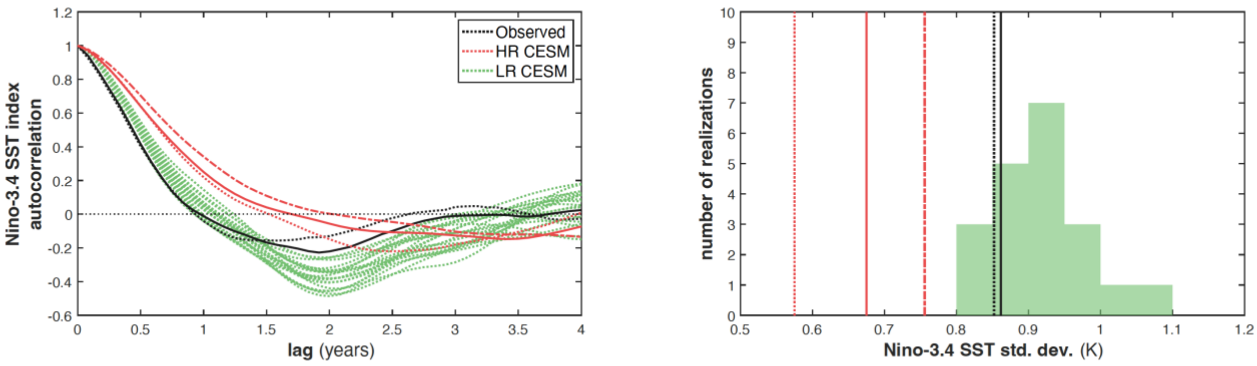

Figure R1.2 Observed and simulated lagged autocorrelation (left) and standard deviation (right) of the Nino-3.4 SST index. Simulated data are from preindustrial simulations performed with the high resolution (HR) and standard resolution (LR) configurations of CESM1. Observed data are from the ERSST5 gridded dataset. Both metrics are computed over contiguous 90-year periods for all three datasets. The HR and LR pre-industrial simulations are 270 years and 1800 years long respectively.

References

Cai, W., Santoso, A., Collins, M. *et al.* Changing El Niño–Southern Oscillation in a warming climate. *Nat Rev Earth Environ* **2**, 628–644 (2021). https://doi.org/10.1038/s43017-021-00199-z.

Capotondi, A., Deser, C., Phillips, A. S., Okumura, Y., & Larson, S. M. (2020). ENSO and Pacific Decadal Variability in the Community Earth System Model Version 2. *Journal of Advances in Modeling Earth Systems*, 12, e2019MS002022. https://doi.org/10.1029/2019MS002022

Chang, P., *et al.* 2020: An unprecedented set of high-resolution earth system simulations for understanding multiscale interactions in climate variability and change. *J. Adv. Model. Earth Syst.*, **12**, e2020MS002298. https://doi.org/10.1029/2020MS002298.

DiNezio, P. N., and C. Deser, Nonlinear Controls on the Persistence of La Niña. *J. Climate*, **27**, 7335–7355, (2014), https://doi.org/10.1175/JCLI-D-14-00033.1.

Kessler, W. S., Is ENSO a cycle or a series of events? *Geophys. Res. Lett.*, 29(23), 2125, (2002), https://doi.org/10.1029/2002GL015924.

Timmermann, A., An, SI., Kug, JS. *et al.* El Niño–Southern Oscillation complexity. *Nature* **559**, 535–545 (2018). https://doi.org/10.1038/s41586-018-0252-6.

Wang, S., Jing, Z., Wu, L. *et al.* El Niño/Southern Oscillation inhibited by submesoscale ocean eddies. *Nat. Geosci.* **15**, 112–117 (2022). https://doi.org/10.1038/s41561-021-00890-2.

Wengel, C., Lee, SS., Stuecker, M.F. *et al.* Future high-resolution El Niño/Southern Oscillation dynamics. *Nat. Clim. Chang.* **11**, 758–765 (2021). https://doi-org.colorado.idm.oclc.org/10.1038/s41558-021-01132-4.

Wittenberg, A. T. Are historical records sufficient to constrain enso simulations? Geophys- ical Research Letters 36 (2009). https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009GL038710.

Zhu, J., et al., Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model, *Geophys. Res. Lett.*, 44, 6984–6992 (2017), https://doi.org/10.1002/2017GL073406.

Zhu, J., Otto-Bliesner, B. L., Brady, E. C., Poulsen, C. J., Tierney, J. E., Lofverstrom, M., & DiNezio, P. (2021). Assessment of equilibrium climate sensitivity of the Community Earth System Model version 2 through simulation of the Last Glacial Maximum. *Geophysical Research Letters*, 48, e2020GL091220. https://doi.org/10.1029/2020GL091220

Response to Reviewer #2:

Reviewer Comment: *DiNezio and Thirumalai produce a series of simulations using CESM at intervals of 3 ka over the deglaciation, forced by changes in greenhouse gasses, ice sheet extent, and insolation. Changes in meltwater input and associated changes in the Atlantic Meridional Overturning Circulation are not considered. They argue that ENSO is simulated very well in this particular model, and find that ENSO variability is linked to the mean state of the Walker Circulation through changes in the mixed layer depth and extent of the warm pool. They validate the results through the comparison of model results with individual foraminifera analysis from the LGM.*

While others have examined the behavior of ENSO during the LGM, the significance here is the analysis of the factors leading to the weaker Bjerknes feedback, and the performance of the simulations in a model that does a good job with modern day ENSO. If they have correctly identified the major controls on ENSO variability, this helps us to better understand how ENSO will change in the future. The paper is clearly written and the illustrations appealing, although I think moving a few details into the main text from the supplement (as noted below) and providing more detail in the figure captions will help the reader better follow the paper. Overall, this is a strong and well-presented paper.

Response: We are very grateful to Dr. Jean Lynch-Stieglitz for the strongly supportive review of our manuscript. Based on their suggestions, we have made significant changes to our manuscript that affirm our original work and interpretations. Aside from including the analysis of a new simulation (4xCO2), we now include new *G. ruber*-based IFA measurements generated from central equatorial Pacific core site ML1208-MC21/BB20 at late Holocene and ~21 ka timeslices (Costa et al. 2016; Lynch-Stieglitz et al. 2015). At this site and at ODP Site 849, where the model indicates that the signal of change in variability is largest, we find striking agreement between simulated and reconstructed temperature variations (Fig. 3) that support externaly forced reductions in ENSO variability during the LGM. In our model-data comparison we now explicitly account for seasonal vs. interannual variability at these sites, including the impact of depth-stratified calcification of various species of planktic foraminifera for our revised model-data comparison (see Fig. 2), based on the Reviewers' inputs. By including a new site in the center of action of ENSO we were able to dramatically increase our confidence that the IFA-inferred changes are highly likely externally forced (98% probability).

We have additionally expanded our text and added a new supplementary section discussing mean state changes across the surface and subsurface Pacific Ocean during the LGM.

Reviewer Comment: *Evaluation of the CESM LGM mean state:*
The logical layout of the paper is that the authors will check if the mechanism they identify operating in the model applies to the real world by validating using the LGM individual foraminifera analyses. It would then seem logical to first make the case that the controlling factors (mixed layer depth, trade-wind strength, warm pool edge) are well simulated by this particular model for the LGM. DiNezio (2011) showed that while the tropical overturning strengthens in all models as expected, the Pacific Walker circulation strength is highly variable in climate models and depends on how the maritime continent is handled. Maybe this is done in another paper and I missed the reference? The authors mention a few observational studies suggesting a deeper thermocline, but a more thorough comparison

of the model state with the data seems warranted. While there may not be direct proxies for mixed layer depth, trade-wind strength or the latitude of the warm pool edge, it would seem sensible to check the related quantities for which proxies do exist (thermocline depth, SST patterns, precipitation linked records in the tropics, etc.)

Response: To address this comment, we have expanded our discussion of the LGM mean state in the main manuscript (see Revised manuscript lines 234-251) and have added a supplemental section in our Supplementary Information file detailing these points below.

We now reference DiNezio et al. (2016), DiNezio et al. (2018), and DiNezio et al. (2020) in the supplemental information section, all of which focus on CESM1.2 under LGM boundary conditions, so readers may obtain full details. DiNezio et al. (2018) also focused on a model-data comparison of the Indo-Pacific mean state during the LGM. However, we contend that assessments of surface-ocean cooling across the tropical Pacific during the LGM are much more challenging due to conflicting proxy records, sharp seasonal and vertical temperature gradients potentially confounding mean state SST changes, and due to the larger size of the basin relative to the coverage of available records (e.g., Tierney et al., 2020, Monteagudo et al., 2021). Nevertheless, paleoclimate assimilation techniques applied to the LGM explicitly consider these sources of (multi-)proxy uncertainty (either via inverse- or forward-modeling approaches; see, e.g., Cooper et al., 2024) and coverage limitation to provide globally complete temperature reconstructions. A recent assessment of the multi-model mean of four independent data assimilation reconstructions of the LGM show that surface-ocean cooling in the eastern equatorial Pacific was larger in magnitude compared to the western Pacific (Cooper et al. 2024)—in agreement with our CESM1.2 simulations. This feature is associated with a deeper thermocline in accordance with a stronger Pacific Walker Circulation (e.g., Lynch-Stieglitz et al. 2015) during the LGM. We suggest that recent assessments of subsurface proxies in the eastern (Ford et al., 2018) and western (Hollstein et al. 2018) support this inference as well. Our supplemental section is reproduced below for ease of reading:

Recent applications of paleoclimate data assimilation techniques have made significant advances in circumventing the aforementioned biases by explicitly incorporating proxy system models to blend model output and paleoproxy information (Tierney et al. 2020). Additionally, data assimilation techniques provide globally complete reconstructions by leveraging local and remote information arising from available proxies. A recent assessment of four independent paleoclimate assimilation reconstructions of the LGM (using both inverse and forward modeling of proxies) indicates that SSTs in the eastern tropical Pacific cooled more than the western Pacific (Cooper et al., 2024) after accounting for proxy and spatial biases. This inference (see multimodel mean in Fig. 1e in Cooper et al., 2024) is in striking agreement with the simulated pattern and magnitude of surface-ocean cooling in our simulations. In CESM1.2, the eastern equatorial Pacific cooling response is associated with a deeper thermocline during the LGM (DiNezio et al. 2016; Fig. 4 in main text). Whereas available estimates of changes in LGM thermocline depths have also been largely equivocal (Andreasen and Ravelo,1997, Sagawa et al., 2012, Leech et al., 2013), recent reappraisals of proxy data in the eastern (Ford et al., 2018) and western (Hollstein et al., 2018) tropical Pacific strongly support a deeper thermocline. Moreover, CESM1.2 shows thermocline deepening in the southern tropical western Pacific (Hollstein et al., 2018) and shoaling in the off-equatorial northern tropical Pacific and in the southwest Pacific (see Fig. 8a in DiNezio et al., 2016), both features that are observed in proxy reconstructions (Andreasen and Ravelo,1997, Sagawa et al., 2012, Leech et al., 2013, Hollstein et al., 2018). Accordingly, we suggest that these lines of evidence point to a deeper mixed layer and a thicker mixed layer across the Pacific during the LGM pointing towards a stronger Walker Circulation (Lynch-Stieglitz et al., 2015), as in the CESM1.2. simulations.

Reviewer Comment: *Relationship between ENSO variability and Climate Forcing over all runs at 15 ka minimum: While the authors argue that trade wind strength, mixed layer depth, and warm pool edge can explain the patterns in variability, they do not explain how these factors are linked to the forcing mechanisms. How big a role is CO2 playing? Local Insolation? Ice Sheet do not explain how these factors are linked to the forcing mechanisms. How big a role is CO2 playing? Local Insolation? Ice Sheet size? Why do they find minimum variability at 15 ka? There are a number of tropical Pacific records with extremes around 15ka (e.g. Borneo Speleothems, EEP T and productivity, etc. --see refs below). Given the timing coincident with Heinrich Stadial 1, the extremes in the records have been interpreted as a consequence of AMOC change and the attendant atmospheric reorganization, but this model does not know about meltwater and AMOC, so the authors may have uncovered another mechanism to explain this pattern. So I am very curious about what is causing the minimum at 15ka the model. Can this be discussed, perhaps with the major forcings be added as a panel in Figure 3 (CO2, insolation during a critical month)?*

Response: In our revised manuscript, we state that the simulated changes arise due to a combination of glacial (CO_2, ice-sheet, sea-level) and orbital forcings. Unfortunately, we do not yet have single-forcing experiments available for these timeslices (particularly the deglacial timeslices) from which we would be able to parse out the spatiotemporal footprint of different climate forcings (see DiNezio et al., 2018). Despite not having analyzed single-forcing runs, we are confident that the bulk of the changes in ENSO and mean state are governed by changes in the Walker circulation via changes in global mean temperature, i.g. global cooling and global warming. As such, both ice sheets and greenhouse gases are likely to play a dominant role under glacial conditions. We clarify this in lines 280-282 of the revised manuscript.

Given the availability of multiple IFA reconstructions for the LGM, we choose to focus on the 21 ka simulation—which shows detectable changes relative to modern (0 ka) conditions—as our validation target. We suggest that our simulations may be used as a baseline of comparison with future studies that incorporate freshwater forcing under deglacial conditions using CESM1.2. We feel that such an approach is better suited to comment on the forcing of ENSO variability at 15 ka (although very few ENSO-resolving records exist to validate such changes).

Reviewer Comment: *The authors compare the IFA from both N. dutertrei and G. tumida to variability at 70 m model depth, based on studies of modern living depth (tow data) which reflects the standing stock in the water column including all life stages. However, calcification depth (the average depth at which foraminifera on the sea floor calcified) is more relevant to the IFA, since many species will add calcite at a deeper depth at the end of their life cycle. G. tumida has a deeper calcification depth than N. dutertrei, and the calcification depth of N. dutertrei partially follows thermocline depth (shallower where the thermocline is shallower, deeper where it is deeper. The 70 m depth is probably a good* *choice for N. dutertrei in the Eastern Pacific today, but not necessarily for the Western Pacific or for the LGM when the thermocline is deeper. G. tumida calcification depth is less sensitive to thermocline depth today, but is likely around 100 m deeper than N. dutertrei. I'm not sure how big a difference it would make, but would the model produce different LH-LGM variability differences for N. dutertrei if the calcification habitat were adjusted for the deeper thermocline during the LGM? Would the G. tumida data from site 849 be in better agreement with the model if a deeper calcification depth were used?*

There doesn't seem to be much change at 150-200 m at Site 849 (Extended Data Figure 2), which would help explain the G. tumida IFA data at this site. But overall, I agree with the sense that it is more difficult to directly interpret the subsurface data and compare to the model given the calcification depth uncertainties and potential model biases in thermocline structure.

Some recent papers which will discuss many of the relevant studies on calcification depth for subsurface species, as well as look at the link between N. dutertrei calcification depth and thermocline depth: Ford et al. (2018) GRL (see supplement) Lakhani et al (2022) Marine Micropaleontology

Response: We are very grateful to the Reviewer for pointing out this issue. In our revised manuscript, alongside our new IFA measurements on *G. ruber*, we explicitly account for depth-stratified foraminiferal calcification habitats in our model-data comparison. To do so, we rely on the mean calcification depths (and confidence intervals) provided by Lakhani et al. (2022)—see our Extended Data Fig. 1 for the full depth profiles (0-200 m) at each site. We also separated depth-stratified changes in seasonal and interannual variability in the 21 ka simulation in order to compare with species-specific IFA at the various sites (Fig. 2): 0-30 m for *G. ruber*, 40-80 m for *T. sacculifer* and *N. dutertrei*, and 100-150 m for *G. tumida* IFA reconstructions. This exercise resulted in better matches between the data and the simulation, and the complicated nature of surface-to-subsurface variability in the eastern equatorial Pacific (see also disagreement in multiproxy reconstructions of surface-ocean variability documented by Zhang et al., 2022) explains the (seemingly-)equivocal IFA datasets. Finally, in our site-specific model-data comparison at ML1208 and ODP Site 849, we use 20 m averages for *G. ruber* and ~50 m water depth changes for *T. sacculifer*, based on mean calcification depths provided in Lakhani et al. (2022). We have detailed our approach towards the surface and subsurface model verification in our revised Methods section. We thank the Reviewer for their suggestion.

Response to Minor Comments:

Comment: *Missing study:*
The IFA results from the central tropical Pacific (Rustic et al., 2020) would seem quite relevant. It looks like their glacial data might not be close enough to 21 ka to include in this study, but do seem to contradict it (higher LGM variability) so should be mentioned. I wonder why they didn't analyze 21 ka? They also make a connection between ENSO variability and mixed layer depth, as well as ITCZ position.
Rustic et al., 2020, Nature Communications

Response: We now include this study in line 16 and 245-247 of our revised manuscript. Also, given the strong precessional forcing seen in our model is not unreasonable to expect stronger variability under the 25 ka BP interval analyzed by Rustic et al. Our mechanism based on zonal advective processes clarifies the link between levels of ENSO variability and mixed layer depth proposed by Rustic et al. We note that the upwelling feedback is a minor feedback governing ENSO and therefore we think that zonal advective processes governing warm pool expansion during the onset of extreme El Nino is a more plausible mechanism.

Comment: Line 109-111: It is not clear which sites are Mg/Ca which are d18O and which are G. tumida and which are N. dutertrei. Either point to Extended Data Table 3 or rewrite to clarify.
Response: We have now pointed the reader to Extended Data Table 3 in the revised manuscript.

Comment: Line 115-116: Ford et al. argue that since the thermocline deepens during the LGM, the G. tumida which are currently calcifiying in the relatively homogeneous layer below the thermocline, may have been calcifying within the thermocline during the LGM (so the water column migrating, not the forams). But that is no reason to dismiss this data from only this core but not the other subsurface cores which will be subject to the same effect. Your model should also capture this same process.
Response: Our new approach focusing on calcifying depth demonstrates consistency between the *G. tumida* IFA data and the simulation.

Comment: Line 170-189: At some point in this paragraph briefly explain how you are defining the warm pool for this study (based on rising motion not T threshold)
Response: We have made this change in the caption of figure 4.

Comment: Line 192-195: Please include your definition of mixed layer (based on 0.5 degree of T change), since this can be defined differently in different fields.
Reponse: We have made this change in the caption of figure 4.

Comment: Line 241-243: What is the length of each simulation that is used for the analysis?
Response: This has been clarified in Extended Data Table 2.

Comment: Line 293: what is "this site"?
Response: This has been removed.

Comment: Figure 2: I suggest not to use the triangle symbol to indicate sites where the data and model don't agree. At first, I thought they were indicating that it was ambiguous whether the sites were seasonal or ENSO sensitive. Just use a square to indicate the sites are ENSO sensitive, and the readers can see the lack of agreement for themselves. It might also be good to put the N. duterei and G. tumida data on different base maps if the change in variability differs between the calcification depths of these two species.
Response: We agree and appreciate the suggestion. This figure has been completely redesigned where the shapes refer to foraminiferal species. We now include all data on the comparison plot.

Comment: Figure 3c: Indicate how ENSO variability and box and whiskers are calculated (or refer reader to Figure 1).
Response: We now refer the reader to Figure 1.

Comment: Figure 3d: Mention specific metric used to determine mixed layer depth and warm pool edge. Are the winds averaged over 2.5S to 2.5N?
Response: We have made these changes in caption of figure 4.

Comment: *Figure 4: Need Latitude, Longitude and Depth markings and color scales on plots. Also, more information in caption. Presumably this is the model output. What model runs, averaged over what time period? Presumably A is a difference plot? Difference between what time periods?*
Response: We now explicitly state that the image is a schematic, which although based on model output, is intended to focus on outlining our mechanism. We now include additional details in the caption.

Comment: Extended Data Figure 2: Could you also include the depth range for G. tumida calcification as well as N. dutertrei, and indicate which core sites are tumida and which are dutertrei with different symbols. Alternatively you could break into 2 plots—one for tumida sites and one for dutertrei sites.
Response: This plot has been revised to include all the sites, including the depth ranges for all species prescribed by Lakhani et al. (2022).

Comment: Extended Data Figure 4 Caption "indeces" should be "indices"
Response: This type has been fixed in the revised version of the manuscript.

Comment: Extended Data Table 3: Please indicate Latitude and Longitude. Is ODP 849 not 846 meant? Please also include all the cores shown in Figure 2.
Response: This has been fixed based on the Reviewer's suggestion. Apologies for the confusion, thankful for identifying this typo.

References

Cooper, V. T. , Armour, K. C. , Hakim, G. J. , Tierney, J. E. , Osman, M. B. , Proistosescu, C. et al.. (2024). Last Glacial Maximum pattern effects reduce climate sensitivity estimates. Science Advances, 10(16), eadk9461.

Costa, K. M. , McManus, J. F. , Anderson, R. F. , Ren, H. , Sigman, D. M. , Winckler, G. et al.. (2016). No iron fertilization in the equatorial Pacific Ocean during the last ice age. Nature, 529(7587), 1-16.

DiNezio, P. N. , Timmermann, A. , Tierney, J. E. , Jin, F.-F. , Otto-Bliesner, B. L. , Rosenbloom, N. et al.. (2016). The climate response of the Indo-Pacific warm pool to glacial sea level. Paleoceanography, 31(6), 866-894.

DiNezio, P. N. , Tierney, J. E. , Otto-Bliesner, B. L. , Timmermann, A. , Bhattacharya, T. , Rosenbloom, N. et al.. (2018). Glacial changes in tropical climate amplified by the Indian Ocean. Science Advances, 4(12), eaat9658.

DiNezio, P. N. , Puy, M. , Thirumalai, K. , Jin, F.-F. , & Tierney, J. E. . (2020). Emergence of an equatorial mode of climate variability in the Indian Ocean. Science Advances, 6(19), eaay7684.

Ford, H. L. , McChesney, C. L. , Hertzberg, J. E. , & McManus, J. F. . (2018). A Deep Eastern Equatorial Pacific Thermocline During the Last Glacial Maximum. Geophysical Research Letters, 45(21), 11,806-11,816.

Hollstein, M. , Mohtadi, M. , Rosenthal, Y. , Moffa Sanchez, P. , Oppo, D. , Martínez Méndez, G. et al.. (2017). Stable oxygen isotopes and Mg/Ca in planktic foraminifera from modern surface sediments of the Western Pacific Warm Pool: Implications for thermocline reconstructions. Paleoceanography, 32(11), 1174-1194.

Hollstein, M. , Mohtadi, M. , Rosenthal, Y. , Prange, M. , Oppo, D. W. , Martínez Méndez, G. et al.. (2018). Variations in Western Pacific Warm Pool surface and thermocline conditions over the past 110,000 years: Forcing mechanisms and implications for the glacial Walker circulation. Quaternary Science Reviews, 201, 429-445.

Lakhani, K. Q. , Lynch-Stieglitz, J. , & Monteagudo, M. M. . (2022). Constraining calcification habitat using oxygen isotope measurements in tropical planktonic foraminiferal tests from surface sediments. Marine Micropaleontology, 170, 102074.

Lynch-Stieglitz, J. , Polissar, P. J. , Jacobel, A. W. , Hovan, S. A. , Pockalny, R. A. , Lyle, M. et al.. (2015). Glacial-interglacial changes in central tropical Pacific surface seawater property gradients. Paleoceanography, 30(5).

Monteagudo, M. M. , Lynch-Stieglitz, J. , Marchitto, T. M. , & Schmidt, M. W. . (2021). Central equatorial pacific cooling during the last glacial maximum. Geophysical Research Letters, 48(3), e2020GL088592.

Tierney, J. E. , Zhu, J. , King, J. , Malevich, S. B. , Hakim, G. J. , & Poulsen, C. J. . (2020). Glacial cooling and climate sensitivity revisited. Nature, 584(7822), 569-573.

Zhang, S. , Yu, Z. , Wang, Y. , Gong, X. , Holbourn, A. , Chang, F. et al.. (2022). Thermal coupling of the Indo-Pacific warm pool and Southern Ocean over the past 30,000 years.. Nat Commun, 13(1), 5457.

Reviewer Reports on the First Revision:

Referee #1 (Remarks to the Author):

In the revised manuscript, the authors have developed a new proxy record for the LGM ENSO variability and performed new simulations to clarify comparison with other simulations with the same model but at higher resolution. The authors' revision also highlighted the novelty of their work, including (1) developing and using paleo proxies to validate model simulations and (2) in-depth analyses of the mechanisms in a suite of paleo and future climate simulations with the consideration of new understanding of the ENSO dynamics in the recent literature. Overall, I now agree with the authors that the manuscript represents significant advancement over the published work.

I do have a few points that I hope the authors can address in the revision.
1. The authors use autocorrelation and skewness of NINO index as metrics to assess the modelled ENSO variability (Extended Data Figure 3) but CESM1 has substantial biases in simulating the other aspects of ENSO, such as the power density and amplitude being too large. What's the implication of these biases on the conclusion in the present study? Berner et al. (2020, https://doi.org/10.1175/JCLI-D-19-0545.1) noted that CESM1 biases in ENSO amplitude and damping time scale may compromise the ENSO response to global warming by overestimating the increase of ENSO variance.
2. I don't think the data assimilation products of the LGM (e.g., Figure 1e of Cooper et al, 2024) offer irrefutable support of an increased SST E-W gradient in the tropical Pacific. In fact, the changes in the LGM SST gradient in the data assimilation may result from the climate model priors, as the evidence from proxy data is inconclusive (–0.2±1℃; Tierney et al., 2020). In light of this, I would soften the tone when using the zonal SST gradient to support the model simulation, such as in Supplementary Text 2.
3. Is it possible to assess the reliability of IFA results from a small sample size of 63? This could be done by creating pseudo proxy records with the preindustrial simulation.

Minor comments
1. Check the author's name "Liu"
2. Figure caption for Extended Data Figure 3 C & D is missing.

Referee #2 (Remarks to the Author):

I thank the authors for their thorough response to my initial review. I'm also very excited to see new LGM IFA results from the Line Islands in this round. It's a very nice paper, and I hope to see it published soon.

Just some random things I caught on my read through:
Line 232—"Walker Circulation" meant?
There are issues with the citations in supplemental information.

Author Rebuttals to First Revision:

Response to Reviewer #1:

Comment: In the revised manuscript, the authors have developed a new proxy record for the LGM ENSO variability and performed new simulations to clarify comparison with other simulations with the same model but at higher resolution. The authors' revision also highlighted the novelty of their work, including (1) developing and using paleo proxies to validate model simulations and (2) in-depth analyses of the mechanisms in a suite of paleo and future climate simulations with the consideration of new understanding of the ENSO dynamics in the recent literature. Overall, I now agree with the authors that the manuscript represents significant advancement over the published work.

Response: We thank the reviewer for their endorsement of our work as a significant advancement over published work.

Comment: 1. The authors use autocorrelation and skewness of NINO index as metrics to assess the modelled ENSO variability (Extended Data Figure 3) but CESM1 has substantial biases in simulating the other aspects of ENSO, such as the power density and amplitude being too large. What's the implication of these biases on the conclusion in the present study? Berner et al. (2020, https://doi.org/10.1175/JCLI-D-19-0545.1) noted that CESM1 biases in ENSO amplitude and damping time scale may compromise the ENSO response to global warming by overestimating the increase of ENSO variance.

Response: Although ENSO simulation has dramatically improved in realism in CESM1 (relative to previous versions and other CMIP models), we agree with reviewer 1 that the remaining errors should be treated carefully, particularly the stronger amplitude of ENSO. We note that Berner et al. (2020) report that CESM1 simulates higher levels of variability than observed measured by the standard deviation of the Nino-3.4 SST index. This is just one way of assessing ENSO fidelity and perhaps does not reflect the fidelity regarding mechanisms. A recent intercomparison of models indicates that CESM1 is amongst the best models in terms of the simulation of mechanisms and feedbacks (Planton et al. 2021, https://doi.org/10.1175/BAMS-D-19-0337.1). Indeed, CESM1 is also one of the best models according to previous studies focusing on ENSO response to warming (Cai et al. 2018, https://doi.org/10.1038/s41586-018-0776-9). In particular, CESM1-CAM5., a very close version to our model, exhibits skewness of the E and C indices - additional metrics indicating that our model simulates extreme El Niño with realism.

Comment: 2. I don't think the data assimilation products of the LGM (e.g., Figure 1e of Cooper et al, 2024) offer irrefutable support of an increased SST E-W gradient in the tropical Pacific. In fact, the changes in the LGM SST gradient in the data assimilation may result from the climate model priors, as the evidence from proxy data is inconclusive (–0.2±1℃; Tierney et al., 2020). In light of this, I would soften the tone when using the zonal SST gradient to support the model simulation, such as in Supplementary Text 2.

Response: We have now softened the tone in our text when discussing the zonal SST gradient.

Comment: 3. Is it possible to assess the reliability of IFA results from a small sample size of 63? This could be done by creating pseudo proxy records with the preindustrial simulation.

Response: We have performed such simulations for our work (see Methods). If the reviewer is interested, we refer them to Thirumalai et al. (2013) and Glaubke et al. (2021), which demonstrate that ~60 individual foraminiferal measurements in the central Pacific can provide robust estimates of variability (standard deviation, percentiles).

Minor comments

1. Check the author's name, "Liu"

2. Figure caption for Extended Data Figure 3 C & D is missing.

Response: Fixed.

Response to Reviewer #2 (Dr. Jean Lynch-Stieglitz):

Comment: I thank the authors for their thorough response to my initial review. I'm also very excited to see new LGM IFA results from the Line Islands in this round. It's a very nice paper, and I hope to see it published soon.

Response: We greatly appreciate Dr. Lynch-Stieglitz's thorough reading of our work and suggestions to improve it!

Comment: Just some random things I caught on my read through:

Line 232—"Walker Circulation" meant?

There are issues with the citations in supplemental information.

Response: These issues have been fixed.