

Private correspondence to the Editor: NOT FOR PUBLIC PEER REVIEW FILE

In my response to reviews (for the eventual public peer review) I have ignored the tone of comments that are within the appended pdf and the review in general of Reviewer 5 since they are not constructive and generally quite abrasive. For example,

Pg. 2 the present structure is extremely indigestible, and squanders any goodwill that the reader might still have after reading that pompous introduction;

Pg. 10 for heaven's sake, everyone knows this.;

Pg. 21 Without treatment yes, but what idiot does it that way? I don't know a single climate scientist dumb enough to compare models and data literally;

Pg. 26 dear lord, again?;

Pg. 27 one more for the road; etc. etc.

Pg. 2 I am surprised that experienced scientists like Didier Roche or Claire Waelbroeck let the paper be submitted once, let alone four times, with such flaws.

I bring them up here, though, because they also aren't professional, and I'm surprised that the reviewer has felt that this is appropriate. Especially, the following:

Pg. 2 I am surprised that experienced scientists like Didier Roche or Claire Waelbroeck let the paper be submitted once, let alone four times, with such flaws

which crosses the line into what can be considered an ad hominem attack. I am a quite an experienced scientist (started PhD in 2009, defended in 2013), and yet I found that these comments somewhat discouraged even me. However, I am encouraged by my co-authors to press on. I do worry, however, about how such a review would be received by a more easily discouraged / less experienced author, perhaps making their first submission, who may think that comments like these are typical of either the publication process, Climate of the Past, Copernicus or the EGU. Frustration at the author is not an excuse, we all read papers, theses, student submissions, et cetera where we would like the author to have done things differently, but it is our job as reviewers to tone our criticism in a way that is constructive and treats the work sent to us with a modicum of respect (even if we disagree about the science or the conclusions). Whilst I greatly value the open-discussion aspect of the journal, , can I suggest that future reviews for Copernicus from this reviewer be checked?

Editor comments

Comments to the Author: I have obtained two further reviews of your paper, both from referees who were not involved in the first round of reviewing. You will be relieved to know that both reviewers, in contrast to the previous rounds, do see merit in your approach - this means we have passed one barrier to publication. Referee 4 does not go into great detail after that, considering that your paper has already received a lot of comments. However referee 5 is very unhappy with the style of the paper and its readability, to the extent that they recommend rejection.

In view of the fact that we have votes in favour of the principle of the paper, and that I generally agree that this does seem like a useful approach, I have decided to let you proceed with a further revision. However I also agree with referee 5 that the paper is extremely tough going. As well as the review, referee 5 provided an annotated pdf with further comments and I am hoping Copernicus will load this for you to see soon. However the paper definitely needs yet another rewrite to make it easier to follow.

*The figures are called in a bizarre order (as far as I could see it went Fig 1, then 8 and 9, then in practice 5 and 7).

* The intro and methods are not too bad, but once we reach section 3, the paper needs to be structured in a logical and linear fashion so that the reader is guided from figure to figure and told what the conclusions from each one are. this is simply not the case at present and I could barely follow what you were doing.

*Additionally the density and number of figures is daunting. I strongly recommend that you strip down to a smaller number of essential panels and figures and put the rest into the supplement.

I realise this paper has already been through numerous iterations, but I'm afraid it is essential to make further improvements if it is to be accepted. I hope you will be able to achieve this.

Reply to Editor

Thank you for considering the value of our manuscript and sending it out for further review.

As suggested by yourself and Anonymous Reviewer 5 we have altered the layout / structure, in doing so we have also reduced the amount of text. We have reduced the number of panels in the figures because we have split the paper into five 'experiments' or 'mini papers', but the number of figures is still 9 (but with less panels per figure).

Reply to Reviewer 4: M. Kucera

Considering the original manuscript, the reviews and the response to the comments, I believe much of the criticism of the manuscript has been misdirected. The reviewers seem to wish that the manuscript would have dealt with a technique to reconstruct ENSO from fossil foraminifera. Anything else they consider not useful. I fear that this view is too simple and I concur with the statement that the authors have made in defense of their study:

“We seek to understand whether or not foraminifera populations in the water are intrinsically capable of recording ENSO dynamics, and at which locations. We consider this to be the most fundamental consideration for foraminifera-based ENSO studies..... Assessment of [sedimentary records] essentially becomes a moot point if the events to be reconstructed are not recorded by foraminifera in the water [in the first place].”

I cannot see anything in the statement that should not be true and contrary to the referees, I see much merit in the concept of preceding complex paleoceanographical interpretations by feasibility assessments of this kind. For this reason, I believe the revised manuscript can be considered for publication.

We thank the reviewer for the time in reviewing our manuscript.

Realizing that this is a contentious issue, I offer some additional suggestions:

- The title still does not reflect the essence of the approach. Foraminifera do not record ENSO, they record seawater properties that change in response to ENSO. I recommend amending the title towards something like: „Proxy modelling approach to assess the potential of extracting past ENSO signal from planktonic foraminifera proxies ...” or “...that ENSO-related patterns are recorded...”

We thank the reviewer for suggesting alternative titles and have gone with their first suggestion.

- Neither the abstract nor the conclusion acknowledge sufficiently the key assumption of the approach: that the FAME model is able to robustly predict the seasonality and the vertical dimension of foraminifera habitat. The author mention this issue clearly (e.g., in section 4.2), but what is missing is a qualifier in the statement of the result in the abstract and conclusion “provided the assumptions of the model are correct, our results indicate that....”.

We have added into the text qualifiers that the reviewer seeks. We agree that the model relies upon a set of assumptions, which do not necessarily reduce the usefulness of the model.

- This is not to say that I believe that the question of the validity of the model discredits the papers. I do have my doubts about the robustness of a niche model that is based exclusively on temperature, but I also note that is not really critical for the value of this study. If formulated correctly then all it does is showing how, assuming it is valid, an explicit forward model can be used a-priori to test if it is plausible that the given recording system can record an oceanographic signal to allow robust reconstructions. This, I believe, is the way forward in paleoceanography and it should be emphasized more in the manuscript.

Line 19 (commented text) should read vaterite, not verite

Section removed from paper.

Reply to Reviewer 5.

Structure: Rearranged format / Length

We thank the reviewer for their suggestion regarding the format and have taken this onboard, modifying the text accordingly. Each section is now presented as a self-contained mini-study. We also agree with the reviewer that by combining multiple questions our text has become unwieldy, something we ourselves alluded to in the first round of review. We appreciate the reviewer for giving our paper the consideration to read it afresh (*“As a prelude, I note that the paper has received prior reviews, which I refrained from reading to avoid biasing my judgment.”*). As they themselves may have guessed (*“I apologize in advance if the following comments are redundant, or if they contradict the recommendations of previous referees – I know firsthand that one cannot please every referee”*) a number of their comments were in fact referring to the previous reviewers suggestions and modifications which lengthened the manuscript. We apologize for the over emphasis of certain points which may have become tedious to the reader (*“though the tediousness of the exposition leads the authors to belabor obvious points at the expense of critical explanations”*; *“Pg. 14 Needlessly tedious. Cut to the chase!”*; *“Pg. 9 So why mention it at all?”*)

Grammar and Figures

We have altered the text and figures accordingly.

Choice of distribution fit

Pg.11 *If your results depend on the choice of kernel, you are in deep, deep trouble ;*

Pg. 11 *While the choice of kernel is typically unimportant, the fact that it is so awkwardly justified raises a red flag.”*

A statistical test that includes a fitted distribution will of course add an assumption to any results regarding the fit of the distribution. There are multiple kernel distributions, whilst it is unimportant which kernel (as we stated to the previous reviewers), in hindsight we felt it was wise to include the full title so that other researchers can (if they wish) replicate our work. Therefore, we are not justifying an Epanechnikov kernel over other kernels (*‘fail to see how this justifies an Epanechnikov kernel, as opposed to any other kernel’*) but over other types of distributions (beta, gamma, gaussian, etc.). Our justification for seeking a generic fit was added at the behest of the previous round of reviews, the aim of the paper is not to test which distribution best fits ocean parameters and therefore a generic fit was sought.

Matlab’s fit distribution *“fitdist uses a normal kernel smoothing function and chooses an optimal bandwidth for estimating normal densities, unless you specify otherwise”* (Matlab website).

QQ-plots

“Finally, and though I would be the last reader to request that the paper get any lengthier, I am surprised that the authors did not focus on the most obvious Achille’s heel of IFAs, as practiced, for instance, by White et al. [2018]. Contrary to the author’s claim, there is nothing inherently wrong with using quantile-quantile plots to compare distributions – indeed it can be a fine idea. The one issue with QQ plots as applied in studies like White et al. [2018] is that it is a handful of extreme values that determine the slope, making the results extremely brittle to outliers. To my mind, this is the most urgent statistical issue to address about the way IFAs are currently presented.”

Our study seeks to test the presence of water conditions that would be favourable for recording of ENSO by foraminifera populations in the water. We also briefly investigate if such parts of the ocean water coincide with areas of high sediment accumulation rate deep-sea archives or not. As the reviewer points out, our manuscript already covers

a lot of ground. We are of course aware of a number of other challenges associated with sediment based IFA reconstructions, such as the interpretive limitations of QQ plots when combined with limited sample sizes. Whilst we (generally) agree with the reviewer, after the previous rounds of review we felt that the manuscript was already long enough.

We are actually working on a manuscript that focuses on the recording of ENSO in the sediment domain on millennial timescales (i.e., regarding QQ plots and other aspects), but it would be much too much to fit into this current manuscript, which already covers substantial ground (and focuses on decadal timescales)

Rounds of review: Fourth submission?

Finally, the reviewer suggests that this is the fourth submission, as opposed to the second round of review, this wouldn't matter if it wasn't in part their rationale for rejection:

*“I am surprised that experienced scientists like Didier Roche or Claire Waelbroeck let the paper be submitted once, **let alone four times**, with such flaws”*

*“After what looks **like 4 trials**, it is unclear that another round of revisions could fix these fatal flaws.”*

*“In summary, this paper is not appropriate for publication in present form, and it is unclear if, **after this many trials**, it can ever be brought up to that standard.”*

References suggested by the reviewer:

Whilst the reviewer suggested the following references, we have altered the text, so that some do not fit into the revised manuscript. Or they are redundant.

D. Khider L. D. Stott, J. Emile-Geay, R. Thunell, D. E. Hammond, 2011. Assessing El Niño Southern Oscillation variability during the past millennium. *Palaeoceanography and Palaeoclimatology*, <https://doi.org/10.1029/2011PA002139>

D. Khider G. Huerta C. Jackson L. D. Stott J. Emile-Geay, 2015. A Bayesian, multivariate calibration for Globigerinoides ruber Mg/Ca. G^3 , <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015GC005844>

S. Dee J. Emile-Geay M. N. Evans A. Allam E. J. Steig D.M. Thompson, 2015. PRYSM: An open-source framework for PROXY System Modeling, with applications to oxygen-isotope systems. *Journal in Advances in Modelling Earth Systems*, <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015MS000447>

M. Comboul and J. Emile-Geay, 2015. Paleoclimate Sampling as a Sensor Placement Problem. AMS, <https://doi.org/10.1175/JCLI-D-14-00802.1>

Evans, M. N., S. E. Tolwinski- Ward, D. M. Thompson, and K. J. Anchukaitis (2013), Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Science Reviews doi:10.1016/j.quascirev.2013.05.024

Comments from Reviewer's annotated PDF file:

NOTE: Some comments by the anonymous reviewer appear to be 'off the cuff' remarks and therefore we have left them without reply.

Reviewer comments are in red, the text they are referring to in italic and our reply in bold.

Pg. 2 Proustian, and neither complete nor readable

This section was removed to streamline the text.

Pg. 2 What do you mean?

Sentence referred to: *Yet, the simulation of past ENSO using climate models has been fraught with difficulties due to the associated feedbacks of ENSO upon model boundary conditions (e.g., SST, pCO₂) (Ford et al., 2015).* – **ENSO as a source of the largest climate variability is an intrinsic component of the climate system, makes it a complex feature to model.**

Pg. 3 Surely these two studies do not have a monopoly on model-data comparison? Why single them out?

This section was removed to streamline the text.

Pg. 3: Also, Khider et al 2011: <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2011PA002139>

This section was removed to streamline the text.

Pg. 3 Period

This section was removed to streamline the text.

Pg. 3: "Bijective" is mathematically pedantic. "One-to-one mapping" will be just as accurate, and much easier to understand.

The reviewer could have just said simplify to one-to-one mapping. This section was removed to streamline the text.

Pg. 3 Add Khider et al. 2015 <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2015GC005844>

This section was removed to streamline the text.

Pg. 4: This is actually called a proxy system model, and the canonical reference is Evans et al 2013: Evans, M. N., S. E. Tolwinski- Ward, D. M. Thompson, and K. J. Anchukaitis (2013), Applications of proxy system modeling in high resolution paleoclimatology, Quaternary Science Reviews, 76 (0), 16i½28, doi:10.1016/j.quascirev.2013.05.024.

Sentence referred to: *Recent attempts at circumnavigating proxy related problems have employed isotope-enabled models (Caley et al., 2014; Roche et al., 2014; Zhu et al., 2017a), proxy models (Dolman and Laepple, 2018; Jonkers and Kučera, 2017; Roche et al., 2018) or uncertainty analysis (Thirumalai et al., 2013; Fraass and Lowery, 2017; Dolman and Laepple, 2018) to predict both the potential $\delta^{18}O_c$ values in foraminifera and/or the probability of detection of a climatic event.*

Altered

Pg. 4: There's already a section 1.2

Altered

Pg. 4: Do you need to cite 5 papers for "sedimentation rate"? That seems to be a basic notion that does not need much defining.

Altered. The interaction between sedimentation rate and bioturbation and their influence upon the temporal resolution of sediment archives is an often misunderstood principle.

Pg. 5: Too many methods for too many questions. I suggest organizing by question, so readers don't have to constantly flip between 2 and 3.

Altered

Pg. 5: More precisely?

Sentence referred to: *The difference between the constant of Hut (1987) and the dynamic value (Brand et al., 2014) is minor.*

Altered to streamline the text

Pg. 5: Ecological, no? Otherwise the sentence makes no sense.

Sentence referred to: *Foraminifera as modelled entities has been developed as a tool for translating, a climatic input (typically a reanalysis dataset or climate model output) into a (simulated-) **climatic** signal, a signal that aims to approximate the depth integrated growth of foraminifera (e.g., Pracht et al., 2019; Wilke et al., 2006; Steindhardt et al., 2015).*

Altered

Pg. 5: Cite Dee et al 2015

Sentence referred to: *Data-model comparison studies suffer from an inability to directly compare like with like so that there are differences in (i) the units used i.e., most proxies reconstructing temperature do not give values of temperature in degrees °C or K but in their own proxy units (e.g., per mil ‰; mmol/mol; species abundance or ratio) necessitating a conversion; and (ii) there is a reduction in scales, i.e., models give a wealth of information (multiple*

depth layers and high resolution time slices) in the time-depth domain. - **This is basic information so does not need a citation defining it.**

Pg. 6: Does a poor job of explaining how FAME differs from competitors. Rephrase to better motivate this model.

Sentence referred to: *A number of models and modelling studies exist to determine the foraminiferal responses to present (Fraile et al., 2008, 2009; Kageyama et al., 2013; Kretschmer et al., 2017; Lombard et al., 2009, 2011; Roy et al., 2015; Waterson et al., 2016; Žarić et al., 2005, 2006) , past (Fraile et al., 2009; Kretschmer et al., 2016) and future (Roy et al., 2015) climate scenarios, FAME uses the associated temperature and $\delta^{18}O_{eq}$ at each grid cell to compute a time averaged $\delta^{18}O_c$ and T_c for a given species. FAME was produced as an attempt to reduce the error associated with data-model comparisons by (i) generating simulated-proxy time-series from model runs that can be compared with age-depth values down core; and (ii) to reduce the model information for a given timeslice into a manageable and relevant value using an integration that would make sense on a biological point of view (Roche et al., 2018).*

We have added text throughout to better motivate this model

Pg. 6. From
Altered

Pg. 6. Weight
Altered

Pg. 6. Comma
Section altered

Pg. 6: So why mention it at all?

Sentence referred to: *The MARGO database does not include *N. dutertrei*, meaning that we concentrate our efforts mainly on *G. ruber* and *G. sacculifer*. – **Because it is a species that has been used for this type of study, therefore whilst we focus on the other two FAME has the option for more species than the MARGO database.***

Pg. 7: this seems to contradict what was just said above. Please clarify

Sentence referred to: *This was repeated four times, during which the lower depth limit of the growth rate computation was set to 60; 100; 200 and 400 m. – **Redundant sentence removed***

Pg. 7: for heaven's sake, everyone knows this.

Sentence referred to: *The tropical Pacific Ocean is divided into four Niño regions based on historical ship tracks, from east to west: Niño 1 and 2 (0° to $-10^\circ S$, $90^\circ W$ to $80^\circ W$), Niño 3 ($5^\circ N$ to $-5^\circ S$, $150^\circ W$ to $90^\circ W$), Niño 3.4 ($5^\circ N$ to $-5^\circ S$, $170^\circ W$ to $120^\circ W$) and Niño 4 ($5^\circ N$ to $-5^\circ S$, $160^\circ E$ to $150^\circ W$). – **Removed***

Pg. 8: If your results depend on the choice of kernel, you are in deep, deep trouble.

-see comment at start of reply

Pg. 8 Prodigiously unclear

Sentence referred to: Here, all values, i.e., the population, associated with a climatological state are compared with the other populations representing the different climatological state, the results plotted here are Neutral climate state vs. El Niño climate state. – **Altered to ‘For each test, comparison is made between all the values of one climatological state and all the values of another climatological state.’**

Pg. 8 You're not testing input data here, you're comparing to two other studies, one of which uses an isotope enabled GCM. but one would never guess from the text.

The reviewer is referring to the following section:

Sentence referred to: 2.5 Test of input data (Temperature and calculated $\delta^{18}O_{eq}$)

Foraminifera as modelled entities produces a modulated response that seeks to replicate how foraminifera modify the climate signal, several studies have approximated the foraminiferal signal in a different way (e.g., Thirumalai et al., 2013; Zhu et al., 2017a). In order to understand how FAME has altered the signal, and the degree to which the conclusions drawn depend upon the modelled growth rates, the input datasets of the sea water properties (Temperature and calculated $\delta^{18}O_{eq}$), underwent a similar statistical test (Figure 4). Unlike FAME, which integrates over several depth levels using the computed growth rate, the test of the input datasets was with fixed depths without any growth rate weighting. These fixed depths are 5, 149 and 235 m, giving a Eulerian view (Zhu et al., 2017a) in which to observe the implications of FAME’s dynamic depth habitat. As per the FAME output, each timestep value was placed into its climate state and an Anderson-Darling test performed to compare the (dis)similarity of on the resultant distributions.

This section is about testing our input data and not comparing to two other studies (which we do later on in the paper). As stated in the section header we are testing the input dataset used for FAME (i.e., In order to understand how FAME has altered the signal, and the degree to which the conclusions drawn depend upon the modelled growth rates, the input datasets of the sea water properties (Temperature and calculated $\delta^{18}O_{eq}$), underwent a similar statistical test (Figure 4)) using the same statistical tests we use later in the paper for FAME.

Pg. 8: Linear or not? the software is unimportant here - it's the method that counts

Sentence referred to: monthly anomalies and a **linear** trend removed (using the detrend function of MatLab 2019a) – the resultant data was left unfiltered (i.e., Zhu et al., 2017a used a 1-2-1 filter) – **Yes linear, as already stated. We are referring to the detrend function of Matlab, so it is important to refer to which software we are using.**

Pg. 9: most cryptic

Sentence referred to underlined, however a larger section is required for our reply: *Four ‘picking’ experiments were performed, as FAME computes the average value for a given time step and given the single foraminiferal isotope variance for an equivalent time step (e.g., weeks: Steinhardt et al., 2015) it is more than likely that this computation reduces the real spread in values. Therefore, rather than use the terminology specimen we prefer to use months.*

Pg. 9 Incoherent grammar

*Although we note that parameterisation of misidentification would be difficult, as it requires understanding of the variability in both standard deviation and absolute values for species co-occurring downcore (Feldmeijer et al., 2015; Metcalfe et al., 2015; 2019). - **Altered***

Pg. 9 Incoherent grammar

Sentence referred to: *Therefore, for each picked month between -0.40 and 0.40 ‰ is added to the picked month value (in intervals of 0.02 ‰), this is approximately $\pm 2^{\circ}C$ (i.e., $\sim 4^{\circ}C$). - **Altered***

Pg. 10 Enormous!

Sentence referred to: *Therefore, for each picked month between -0.40 and 0.40 ‰ is added to the picked month value (in intervals of 0.02 ‰), this is approximately $\pm 2^{\circ} C$ (i.e., $\sim 4^{\circ} C$). - As we outline in our rationale this is equivalent to the range found in trap data, this could reflect the variation in absolute depth habitat experienced by individual foraminifera, slight variation in growth or other aspects of the individual (size, food, metabolism, etc.).*

Pg. 10 Incoherent

Sentence referred to: *Each picked month has their own randomly selected error for both of these errors, i.e., each value is the sum of the month picked and their own error. - Altered*

Pg. 10 But you're already using KDE, which is basically a interpolation over quantiles! So why interpolate on top of that?; This not a large computation compared to what is described above

Section Removed

Pg. 10 Focus on what you have, not the rest [strikethrough]

This is in contrast to other proxies such as corals (Cole and Tudhope, 2017), speleotherms (Chen et al., 2016) and molluscs (Butler et al., 2013; Milano et al., 2017), where distinct time-specific banding is present (true 'time-series' proxies). – **As suggested section removed**

Pg. 10 Incoherent Grammer

Sentence referred to underlined: *This is especially apparent in $\delta 18O_c$ where there is a difference temporally of $\delta 18O_{sw}$ (e.g., the ice volume effect in glacial and interglacial cycles ~ 1.25 ‰) meaning that the same temperature can have radically different $\delta 18O_c$ values, a consequence of this is that a series of high magnitude, but low frequency El Niño events could be disturbed in a discrete-depth record.*

Pg. 11 Needlessly tedious. Cut to the chase!

Section referred to: *Whilst our intention here is a generalised view to be used as an approximate guide, it is important to note that the Pacific Ocean has the largest proportion globally of >1 km tall seamounts that are smaller than <100 km (Wessel, 1997). Which may have important, relatively shallow-water sedimentary sequences, which may also be of sufficient sediment accumulation rate, therefore we supplement the GEBCO bathymetric data with the locations of seamounts. However, whilst there are an estimated 50,000 seamounts in the Pacific that are taller than a km (Menard, 1964; Wessel and Lyons, 1997), only 12,000 have been documented on charts (Batiza, 1982), and approximately 291 have been dated (Koppers et al., 2003; Clouard and Bonneville, 2005; Hillier, 2007). It is these 291, $<1\%$ of the estimated seamounts, we have overlain onto the bathymetric data (Figure 8b), although this number is further reduced as we only plot between $20^{\circ}S$ and $20^{\circ}N$.*

Previous reviewers requested the addition of seamounts, yet there is no easily accessible database of such a feature. We have condensed this section (leaving the references behind) as we agree.

Pg. 11 Can you detail differences with TURBO2, with which I am more familiar?

The reviewer/reader is invited to read a recently published paper (Lougheed, 2020) in GMD detailing SEAMUS, where the similarities and differences with TURBO2 are outlined in detail.

Pg. 12 Incoherent grammer

Sentence referred to: *To investigate how much temporal signal is integrated into discrete-depth intervals for typical tropical Pacific SAR, we, therefore, utilised the single foraminifera sediment accumulation simulator (SEAMUS, Lougheed, 2019) to bioturbate, as the input climate signal (Figure's 9 to 11)...*

Pg. 12 Strange choice. Why not use simulated data, like SST from TraCE-21k? I understand the idea of using NGRIP as prototypical of a high-res paleoclimate record, but putting it in the middle of the Pacific is rather incongruous.

We wanted to use an oxygen isotope record, as there are global and regional changes which we discussed in the original manuscript. In the original MS and comments made in reply to the previous (first) round of reviewers we discussed that bioturbating a foraminifera between time intervals with different global $\delta^{18}\text{O}_{\text{sw}}$ may alter the interpretation. As we state in this section this is the only record that is long enough and whose data is available.

Regarding the other point(s) as we had expected the reader might think this, which is why we had added a section. We refer the reviewer to this section already in the text: *The use of the NGRIP timeseries here is purely as an input parameter to investigate the effect of bioturbation upon a given climate signal - it is important to stress that by using NGRIP as an input signal for SEAMUS we are neither implying that tropical Pacific cores should have signal similar to NGRIP or inferring some kind of causal relationship.* **We are not putting an ice core in the middle of the Pacific. We are using it, as the reviewer themselves points out, as a high-resolution palaeoclimate record that can be bioturbated using SAR and BD values that are similar to the Pacific Ocean.**

Pg. 13 Incoherent grammar
Altered

Pg. 14 Yes, could be
i.e. 10 different 'months' picked between grid-points may exacerbate or accentuate differences.

Pg. 14 At what level?

Pg. 14 Again, do this here, not the methods

Don't understand what these comments are referring to

Pg. 14 Oxymoronic

Sentence referred to underlined, however a larger section is required for our reply: *Therefore, understanding the biological variability on shorter timescale (e.g., Steinhardt et al., 2015; Mikis et al., 2019) which, maybe here over exaggerated, may be crucial for understanding discrepancies between cores.* – **Not an example of an oxymoron, if the reviewer is referring to over exaggerated.**

Pg. 15 Better to report the p-values and effect sizes, rather than whether $p < \text{or} > 0.05$
We don't refer to $p < \text{or} >$ here but to the test metric (not the p value)

Pg. 15 Can you summarize what this shows?
Section altered

Pg. 16 Statistically?

We agree that using the wording significant implies statistical correlation and have altered this section

Pg. 16 What you are really testing here are the effects of dissolution and bioturbation
Section 7 in the revised version now refers to this as 'Approximation of sedimentary archives'

Pg. 17 And what? ("both" implies a second clause)
Altered

Pg. 17 That's what needs to be done, but figures are so messy i can't get to it.
We have altered the figures

Pg. 18 And that is where people would sample

Sentence referred to: *However, at certain locations, near islands or seamounts, the SAR and water depth may be high enough to allow for a signal to be preserved (Figure 8B) that may not be represented here. – We agree, but close to seamounts there may also be enhanced bioturbation as these underwater obstacles alter the circulation leading to resuspension or upwelling of nutrients. Likewise, the stability of the sediment on the seamount and potential for resuspension of older material may alter the preserved stratigraphy (through slumping, or winnowing).*

Pg. 18: Without treatment yes, but what idiot does it that way? I don't know a single climate scientist dumb enough to compare models and data literally.

Sentence referred to underlined, however a larger section is required for our reply: *Whilst we are principally interested in understanding whether living foraminifera can theoretically reconstruct ENSO, comparison with data requires an additional analysis. This is because data-model comparisons are subjective, nominally supposing that the data is the value to be achieved by the model. However, if the foraminifera modulate the original climate signal, then preservation selectively filters which specimens are conserved whereas bioturbation acts to reorder, transposing the order in which they are recovered from the depth domain. Once the sediment is recovered, the researcher acts as a final filter, which is in essence a random picking. – We are elaborating on all the factors that make the data potentially biased or erroneous, one can use a proxy system model to generate pseudo proxy time series so that a comparison between models and data can be made. The attempt to fit pseudo-data produced from model output to ‘garbage data’ is our point, rather than literally comparing data and models. In other words, we agree with the comment of the reviewer.*

Pg. 18 This review is off topic

We have altered the text

Pg. 19 Assuming you are right

Sentence referred to: *ENSO studies using palaeoceanography have exposed shifts, one can, therefore, question what is being reconstructed in such studies.*

Pg. 19 of what?

Text altered

Pg. 20 A point made originally by Thirumalai et al 2013

Sentence referred to: *our own analysis using the ratio of total to interannual variance also suggests that much of the variance in the simulated foraminiferal signal is dominated by interannual variance.*

Pg. 20 off topic

4.2.2 *The use of models in reconstructions*

The previous reviewers wanted justification for such a model, we have altered this section in light of this round of review

Pg. 20 Carbonate preservation

This sentence is referring to the carbonate ion effect and not carbonate preservation

Pg. 21 Agreed, but this is FAME, not this paper

Sentence referred to: *A dynamic depth habitat in which the environmental signal becomes a weighted average of the water column can further confound the original signal (Wilke et al., 2006).* – **I am not entirely sure what point the reviewer is making here. Wilke et al. show that the depth habitat of foraminifera can be approximated by a weighted average of the various calcifying depths rather than a specific water depth (the signal most researchers want to reproduce).**

Pg. 21 Gibberish. Makes zero sense

Sentence referred to: *The synthesis of pseudo-timeseries to discern the potential distribution for different scenarios, whilst a necessary approximation, is nonetheless one that is free of cause and causality. Modulating a timeseries for events with enhanced or weakened amplitude or fewer or greater number of events assumes in essence that there is limited feedback both regionally (between two sites) and internally within the timeseries (i.e., a process that operates on a higher level).* – **We have altered this section of text**

Pg. 21 Those won't have feedback btwn sites. That is not what climate models do

Sentence referred to: *or multi-model ensembles with prescribed boundary conditions can be used for the generation of timeseries in which the physics of atmospheric and oceanic circulation are constrained and feedbacks between sites can occur.*

We have altered this section of text

Pg. 21 Need to back up with refs. This is a serious claim

Sentence referred to: *The perceived failure of several climate models to resolve ENSO adequately, resulting in variable ENSO frequency and amplitude between models, could therefore be used to determine the proxy signal from model derived timeseries at different frequencies and intensities of ENSO.*

We have altered this section of text

Pg. 22 How can something be "somewhat binary"? Are we in fuzzy logic territory?

Sentence referred to: *This gives a somewhat binary view, the feature either occurs or does not occur, and if it occurs then it has either enhanced or weakened.* – **Colloquialism that does not refer to fuzzy logic. Binary thinking of researchers to either something occurs or does not occur. Have altered**

Pg. 22: Unclear and verbose

Sentence referred to: *Yet this can (though not always) preclude a scenario in which the feature has shifted.*

Pg. 22 You are discussing optimal sampling design. This has been done before, though not for forams: <https://journals.ametsoc.org/doi/10.1175/JCLI-D-14-00802.1>

Sentence referred to: *Climate models could therefore also be used to determine applicable core locations for comparison of proxy values with 'like with like' oceanographic features (similar to the analysis of Evans et al. (1998) for predicting coral sites), without necessarily the cost of a time-slice project (e.g., CLIMAP, MARGO).* – **I assume the point of this comment is that we should refer to Comboul and Emile-Geay, 2014? As we know it has not been done before hence why we refer to a paper on corals.**

Pg. 22 Number already taken. Get your numbering straight!

Referring to: *4.2 Limitations of the methods applied and assessment of model uncertainties*

Altered

Pg. 22 Isn't LeGrande and Schmidt only for surface values, not subsurface?

Sentence referred to: *The spatial variability in salinity, particularly within regions underlying the intertropical convergence zone (ITCZ) and the moisture transport from the Caribbean into the eastern Pacific along the topographic low that represents Panama Isthmus, the resultant conversion of salinity to $\delta^{18}O_{sw}$ and then*

$\delta^{18}O_{eq}$ may contain further error.

There are some parameters for subsurface in LeGrande and Schmidt.

Pg. 23: dear lord, again?

Referring to: **Whilst** the change in Mg/Ca with temperature has been validated (*e.g.*, Elderfield and Ganssen, 2000), the computation of a pseudo-proxy value for and from model parameters remains enigmatic.

Pg. 23: That is one enigmatic sentence

Sentence referred to: Whilst the change in Mg/Ca with temperature has been validated (*e.g.*, Elderfield and Ganssen, 2000), the computation of a pseudo-proxy value for and from model parameters remains enigmatic.

Pg. 24: The conclusions agree with my intuition, but the paper is too messy to back them up.

Review of
**Using a foraminiferal ecology model to test if tropical
Pacific planktonic foraminifera are suitable recorders of
ENSO, version 4**
by Metcalfe et al.

Recommendation: *Reject*

Summary

The manuscript uses a number of forward models to assess the possibility of recovering distinct ENSO states from individual foraminiferal analyses (IFAs). While the goal is excellent and the general approach sound, the presentation is so abstruse that it removes all credibility from the paper's claims. After what looks like 4 trials, it is unclear that another round of revisions could fix these fatal flaws.

1 General Comments

As a prelude, I note that the paper has received prior reviews, which I refrained from reading to avoid biasing my judgment. I apologize in advance if the following comments are redundant, or if they contradict the recommendations of previous referees – I know firsthand that one cannot please every referee.

1.1 Statistics

As said above, the general approach is sound, though the tediousness of the exposition leads the authors to belabor obvious points at the expense of critical explanations.

My biggest question mark is on the kernel density estimation (section 2.4). The authors use an Epanechnikov kernel with the following justification “*The use of an Epanechnikov kernel distribution to fit the data, as opposed to other types of distribution, represents a trade-off between keeping as many parameters constant whilst mimicking the underlying dataset for a large number of grid points.*”. I fail to see how this justifies an Epanechnikov kernel, as opposed to any other kernel. While the choice of kernel is typically unimportant, the fact that it is so awkwardly justified raises a red flag. If your results depend sensitively on the choice of kernel, you are in deep, deep trouble. It would be important to include (as a supplement) an analysis with a different kernel choice (e.g. Gaussian).

The authors also mention a variable bandwidth, which is fine, but do not explain how it is chosen (e.g. Silvermann criterion). Given that the entire premise of the paper is to compare distributions, this is a crucial detail that needs to be better explained, possibly with a sensitivity analysis.

On the broader point of reporting the results of the Anderson-Darling test, the authors rely

exclusively on whether the p-values are above or below 5%. As emphasized by *Wasserstein and Lazar* [2016], the American Statistical Association explicitly warns against relying exclusively on p-values, and recommends additional metrics like effect sizes and confidence intervals. I think mapping effect sizes, possibly stippled to indicate whether the p-values are above or below 5%, would be better practice.

Finally, and though I would be the last reader to request that the paper get any lengthier, I am surprised that the authors did not focus on the most obvious Achille's heel of IFAs, as practiced, for instance, by *White et al.* [2018]. Contrary to the author's claim, there is nothing inherently wrong with using quantile-quantile plots to compare distributions – indeed it can be a fine idea. The one issue with QQ plots as applied in studies like *White et al.* [2018] is that it is a handful of extreme values that determine the slope, making the results extremely brittle to outliers. To my mind, this is the most urgent statistical issue to address about the way IFAs are currently presented.

1.2 Structure

The paper follows the classic structure of Introduction/Methods/Results/Discussion. The only issue here is that, because they consider 3 distinct questions, the methods are varied and lengthy, and by the time the reader gets to the end of Section 3, they have largely forgotten the relevant methods. It would seem more natural to me to structure one section per question, with relevant methods introduced where needed. This would look like: 1) introduction 2) distinguishing variance statistics 3) distinguishing distributions, 4) sensitivity to input parameters, 5) impacts of dissolution and bioturbation 6) discussion. One thing is for sure: the present structure is extremely indigestible, and squanders any goodwill that the reader might still have after reading that pompous introduction.

1.3 Grammar

I have reviewed dozens of papers over my career, but this one takes the prize for the most abstruse writing coming from native English speakers. A few times I had to look up whether some of the quirks might be differences between British and American English, but I could find no justification in any grammar book for spelling figures "figure's" (P14L30), for starting sentences by "Whilst" followed by a comma, for writing Proustian run-on sentences, or for being generally so incoherent that, after reviewing the paper on my iPad, the number one suggestion from my autocomplete is "incoherent grammar" (see annotated manuscript). I am surprised that experienced scientists like Didier Roche or Claire Waelbroeck let the paper be submitted once, let alone four times, with such flaws.

1.4 Figures

The figures are a piece of work. First, this is the first time I have seen figures so large that they make the text pages of the PDF look like microfilm. To add insult to injury, they are all of different sizes, making the document's navigation extremely tedious. The substance is no better than the style, unfortunately, as (apart from Fig 10), they are all so poorly designed that I would tell my students to redo them. It seems like the authors cannot decide what point to make, so they bombard the reader with lots of similar, overloaded figures. It is imperative to focus the design around the key points, and put the other figures in an appendix/supplement.

2 Line by line Comments

see annotated manuscript.

In summary, this paper is not appropriate for publication in present form, and it is unclear if, after this many trials, it can ever be brought up to that standard.

References

Wasserstein, R. L., and N. A. Lazar (2016), The ASA Statement on p-Values: Context, Process, and Purpose, *The American Statistician*, *70*(2), 129–133, doi:10.1080/00031305.2016.1154108.

White, S. M., A. C. Ravelo, and P. J. Polissar (2018), Dampened El Niño in the Early and Mid-Holocene Due To Insolation-Forced Warming/Deepening of the Thermocline, *Geophysical Research Letters*, *45*(1), 316–326, doi:10.1002/2017GL075433.

Using a foraminiferal ecology model to test if tropical Pacific planktonic foraminifera are suitable recorders of ENSO

Brett Metcalfe^{1,2}, Bryan C. Lougheed^{1,3}, Claire Waelbroeck¹ and Didier M. Roche^{1,2}

¹Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

²Earth and Climate Cluster, Department of Earth Sciences, Faculty of Sciences, VU University Amsterdam, de Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands

³Department of Earth Sciences, Uppsala University, Villavägen 16, 75236 Uppsala, Sweden

Correspondence to: Brett Metcalfe (b.metcalfe@vu.nl)

A complete understanding of past El Niño-Southern Oscillation (ENSO) fluctuations is important for the future predictions of regional climate using climate models. One approach to reconstructing past ENSO dynamics uses planktonic foraminifera as recorders of past climate to assess past spatiotemporal changes in upper ocean conditions. In this paper we utilise a model of planktonic foraminifera populations, *Foraminifera as Modelled Entities (FAME)*, to forward model the potential monthly average $\delta^{18}\text{O}_c$ and temperature signal proxy values for *Globigerinoides ruber*, *Globigerinoides sacculifer* and *Neogloboquadrina dutertrei* from input variables covering the period of the instrumental record. We tested whether the modelled foraminifera population $\delta^{18}\text{O}_c$ and T_c associated with El Niño events statistically differ from the values associated with other climate states. For these foraminiferal species our results suggest that the values of El Niño events can be differentiated from other climate states. Our model computes the proxy values of foraminifera in the water, suggesting that, in theory, water locations for a large proportion of the Tropical Pacific should be suitable for differentiating El Niño events from other climate states. However, in practice it may not be possible to differentiate climate states in the sediment record. Specifically, comparison of our model results with the sedimentological features of the Pacific Ocean shows that a large proportion of the hydrographically/ecologically suitable regions, coincide with low sediment accumulation rate at the sea floor and/or regions that lie below the critical water depths for calcite preservation (lysocline and CCD).

1. Introduction

1.1 El Niño-Southern Oscillation (ENSO)

Predictions of short-term, abrupt changes in regional climate are imperative for improving spatiotemporal precision and accuracy when forecasting future climate. Coupled ocean-atmosphere interactions (wind circulation and sea surface temperature) in the tropical Pacific, collectively known as the El Niño-Southern Oscillation (ENSO) on interannual timescales and the Pacific Decadal Oscillation on decadal timescales, represent global climate's largest source (Wang et al.,

2017) of inter-annual climate variability (Figure 1). Due to ENSO's major socio-economic impacts upon pan-Pacific nations, which, depending on the location, can include flooding, drought and fire risk, it is imperative to have an accurate understanding of both past and future behaviour of ENSO (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). The instrumental record of the past century provides important information (that can be translated into the Southern Oscillation Index; SOI), however, detailed oceanographic observations of the components of ENSO (both the El Niño and Southern Oscillation), such as the Tropical Oceans Global Atmosphere (TOGA; 1985-1994) experiment only provide information from the latter half of the twentieth century (Wang et al., 2017). To acquire longer records, researchers must turn to the geological record using various archives that are available from the (pan-)Pacific region, including: corals (Cole and Tudhope, 2017); foraminifera (Ford et al., 2015; Garidel-Thoron et al., 2007; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; White et al., 2018); stalagmite (Asmerom et al., 2007; Zhu et al., 2017b); ice detritus (Patterson et al., 2004; Skrivaneck and Hendy, 2015); lake (Anderson et al., 2005; Bennett et al., 2001; Benson et al., 2002; Conroy et al., 2008; Enzel and Wells, 1997; Higley et al., 2018; Loubere et al., 2013; Zhang et al., 2014); terrestrial (Barron et al., 2003; Barron and Anderson, 2011; Caramanica et al., 2018; Hendy et al., 2015; Staines-Urías et al., 2015); and sedimentological parameters (Moy et al., 2002) including varves (Du et al., 2018; Nederbragt and Thurow, 2001, 2006) to reconstruct long-term variations in proxies, linked to climate, that may provide clues to ENSO and its impact upon both regional and global climate. An integrated approach combining palaeoclimate proxies (Ford et al., 2015; Garidel-Thoron et al., 2007; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; White et al., 2018) and computer models (Zhu et al., 2017a) can help shed light on the triggers of past ENSO events, their magnitude and their spatiotemporal distribution. Yet, the simulation of past ENSO using climate models has been fraught with difficulties due to the associated feedbacks of ENSO upon model boundary conditions (e.g., SST, $p\text{CO}_2$) (Ford et al., 2015). One way to deduce the relative impact and importance of various feedbacks and, in turn, reduce model-dependent noise in our predictions, is to compare model output with proxy data (Roche et al., 2017; Zhu et al., 2017a).

1.2 Foraminiferal Proxies

Such an approach, however, requires an abundance of reliable spatiotemporal proxy data from the entire Pacific Ocean. Moreover, such proxy reconstructions are subject to several unknowns, uncertainties and biases. For the specific case of foraminifera populations in the water, it particularly arises from the species-specific ecological niche. The mapping of proxy value to climate value can therefore be skewed, a major factor governing the spatiotemporal distribution of a given planktonic foraminiferal species is the presence of an ideal water temperature. Proxies of past ENSO and Pacific SST (Ford et al., 2015; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; Sadekov et al., 2013; White et al., 2018) are based upon the biomineralisation of the calcite, or a polymorph such as vaterite (Jacob et al., 2017), shells of foraminifera (Emiliani, 1955; Evans et al., 2018; Zeebe and Wolf-Gladrow, 2001). In general, there are three major types of foraminifera-based palaeoceanographic proxies:

(1) those associated with the faunal composition and their abundance within deep-sea sediments that utilises either a qualitative approach (Phleger et al., 1953; Schott, 1952); a weighted average (Berger and Gardner, 1975; Jones, 1964; Lynts and Judd, 1971); a selected species approach (e.g. coiling direction, or warm-water species presence; Ericson et al., 1964; Ericson and Wollin, 1968; Hutson, 1980b; Parker, 1958; Peeters et al., 2004; Ruddiman, 1971; Schott, 1966); a regression analysis (Hecht, 1973; Imbrie and Kipp, 1971; Williams and Johnson, 1975); or, a transfer function (CLIMAP Project Members, 1976; McIntyre et al., 1976; Williams, 1976; Williams and Johnson, 1975) that compares the down-core records with a dataset of 'modern' values and their associated water column parameters (Hutson, 1977, 1978);

(2) those associated with the stable oxygen isotope composition of a whole shell analysed either individually (Ganssen et al., 2011; Koutavas et al., 2006; Koutavas and Joanides, 2012; Leduc et al., 2009) or pooled (Garidel-Thoron et al., 2007; Koutavas et al., 2002; Stott et al., 2002, 2004), herein $\delta^{18}\text{O}_c$ (c = calcite), which can be used to reconstruct SST and past oxygen isotope values in seawater $\delta^{18}\text{O}_{\text{sw}}$ (sw = seawater) when paired with a proxy that can either reconstruct temperature or salinity;

(3) those associated with trace metal geochemistry (e.g., Ford et al., 2015; Sadekov et al., 2013; Stott et al., 2002, 2004; White et al., 2018), more specifically the natural logarithm of the relative concentration of Mg and Ca ($\ln(\text{Mg}/\text{Ca})$), of the shell, based upon the temperature dependent (Elderfield and Ganssen, 2000; Nürnberg et al., 1996) incorporation and substitution of a Mg cation into the calcite lattice (Branson et al., 2013, 2016).

The interpretation of these proxies, however, is not straightforward, for example, calibration of foraminiferal assemblage based transfer functions with surface temperatures as opposed to a deeper temperature signal may in fact skew the reconstructed temperature (Telford et al., 2013); $\delta^{18}\text{O}_c$ can be affected by species-specific size effects (Feldmeijer et al., 2015; Metcalfe et al., 2015; Pracht et al., 2018); disequilibria or vital effects, which clouds the accurate reconstruction of past SST and $\delta^{18}\text{O}_{\text{sw}}$. There is also no simple bijective function between $\delta^{18}\text{O}_c$ and the oceanic variables $\delta^{18}\text{O}_{\text{sw}}$ and temperature used in its calculation, with variability in $\delta^{18}\text{O}_{\text{sw}}$ limiting the use of $\delta^{18}\text{O}_c$ as a pure temperature proxy. Likewise, researchers have not been able to discount the impacts of the ambient salinity (Allen et al., 2016; Gray et al., 2018; Groeneveld et al., 2008; Kısakürek et al., 2008) and carbonate ion concentration (Allen et al., 2016; Evans et al., 2018; Zeebe and Sanyal, 2002) on the Mg/Ca content of foraminifera, nor biological effects such as growth banding (Eggins et al., 2003; Hori et al., 2018; Sadekov et al., 2008, 2009; Vetter et al., 2013). Foraminifera are also not passive recorders of environmental conditions such as SST, in that the very ambient environment that researchers wish to reconstruct also modifies the foraminiferal population as well (Mix, 1987; Mulitza et al., 1998). Sensitivity to the variable being

reconstructed may increase or decrease the relative contribution of individual events through modulation of the flux to the seafloor, increasing or decreasing the chance of sampling such occurrences (Mix, 1987; Mulitza et al., 1998). Culture experiments have identified temperature (Lombard et al., 2009, 2011), light (Bé et al., 1982; Bé and Spero, 1981; Lombard et al., 2010; Rink et al., 1998; Spero, 1987; Wolf-Gladrow et al., 1999), carbonate ion concentration [CO_3^{2-}] (Bijma et al., 2002; Lombard et al., 2010) and ontogenetic changes (Hamilton et al., 2008; Wycech et al., 2018) as variables that drive, alter or induce changes in foraminiferal growth.

Computation of the influence of biological and vital effects upon physiochemical proxies, such as those based on foraminifera should be a fundamental consideration for any accurate data-model comparison. Recent attempts at circumnavigating proxy related problems have employed isotope-enabled models (Caley et al., 2014; Roche et al., 2014; Zhu et al., 2017a), proxy models (Dolman and Laepple, 2018; Jonkers and Kučera, 2017; Roche et al., 2018) or uncertainty analysis (Thirumalai et al., 2013; Fraass and Lowery, 2017; Dolman and Laepple, 2018) to predict both the potential $\delta^{18}\text{O}_c$ values in foraminifera and/or the probability of detection of a climatic event. The use of ecophysiological models (Kageyama et al., 2013; Lombard et al., 2009, 2011) can help circumvent some of the problems associated with a purely mathematical approximation (e.g., Caley et al., 2014) of the translation of an ambient signal into a palaeoclimate proxy. They are not limited to foraminifera and can provide an important way to test whether proxies used for palaeoclimate reconstructions are suitable for the given research question. Several studies have investigated the response of planktonic foraminifera from core material or computed pseudo foraminiferal distributions, their proxy values, and the resultant (likely) distribution of these proxy values with respect to ENSO (e.g., Leduc et al., 2009; Thirumalai et al., 2013; Ford et al., 2015; Zhu et al., 2017).

1.2 Aims and Objectives

Here, we investigate whether living planktonic foraminifera can be theoretically used in ENSO reconstructions, differing from previous research by using a foraminiferal growth model, *Foraminifera as modelled entities* (FAME; Roche et al., 2018), to tackle the dynamic seasonal and depth habitat of planktonic foraminifera (Wilke et al., 2006; Steinhardt et al., 2015; Mix, 1987; Mulitza et al., 1998). To be a valid proxy for the reconstruction of ENSO, the proxy values of populations of planktonic foraminifera associated to different climatic states (*i.e.*, El Nino, Neutral, La Nina) should be significantly different from one another. In order to test our research question, ‘*are the distributions of proxy values associated with El Niño months statistically different from distributions of proxy values associated with neutral or La Niña months?*’, our methodology follows a forward modelling approach in which the computed values of the temperature recorded by calcite (T_c - a pseudo temperature aimed at mimicking Mg/Ca albeit one uninfluenced by secondary factors) and $\delta^{18}\text{O}_c$ are assigned to one of these climatological states. This forward modelling approach does not pre-suppose foraminifera can record ENSO variability (‘Can we detect?’) *i.e.*, what is done when inverting the core top pooled $\delta^{18}\text{O}$ or individual foraminiferal $\delta^{18}\text{O}$ distributions and infer changes in ENSO (‘How could we detect?’). A secondary objective is to compare the output of this approach with secondary factors that further modulate the climatic signal through post-mortem processes. We identify regions in the Pacific Ocean where the sedimentation rate (Berger, 1970a, 1971; Boltovskoy, 1994; Lougheed et al., 2018;

Olson et al., 2016) may be too low or the water depth (Berger, 1967, 1970b; Boltovskoy, 1966; Loughheed et al., 2018) too deep (causing dissolution of carbonate sediments) thus preventing the capture and preservation of the foraminiferal signal.



2. Methods

2.1 Input variables (Temperature; Salinity and $\delta^{18}\text{O}_{\text{sw}}$)

5 For input variables, temperature and salinity of the ocean reanalysis data product (Universiteit Hamburg, DE) ORA S4 (Balmaseda et al., 2013) were extracted at one-degree resolution for the tropical Pacific (-20°S to 20°N and 120°E to -70°W), with each single grid cell comprised of data for 42 depth intervals (5 – 5300 m water depth) and 696 months (January 1958 – December 2015). For computation of the oxygen isotope of seawater ($\delta^{18}\text{O}_{\text{sw}}$), a global 1-degree grid was generated, and each grid cell was classified as belonging to one of 27 distinct ocean regions, as defined by either societal and scientific agencies, for identifying regional $\delta^{18}\text{O}_{\text{sw}}$ – salinity relationships (LeGrande and Schmidt, 2006). Using the $\delta^{18}\text{O}_{\text{sw}}$ database of LeGrande and Schmidt (2006) a regional $\delta^{18}\text{O}_{\text{sw}}$ – salinity relationship was defined, of which the salinity is the salinity measured directly at the isotope sample collection point (included within the database). Two matrices were computed; one giving values of the slope (m) and the other of intercept (c) of the resultant linear regression equations, these were used as look-up tables to define the monthly $\delta^{18}\text{O}_{\text{sw}}$ from the monthly salinity Ocean reanalysis product ORAS S4 (Balmaseda et al., 2013), which was used for the calculation of $\delta^{18}\text{O}_{\text{eq}}$, *i.e.* the expected $\delta^{18}\text{O}$ for foraminiferal calcite formed at a certain temperature (Kim and O’Neil, 1997). The $\delta^{18}\text{O}_{\text{eq}}$ is calculated from a rearranged form of the following temperature equation:

$$T = T_0 - b \cdot (\delta^{18}\text{O}_c - \delta^{18}\text{O}_{\text{sw}}) + a \cdot (\delta^{18}\text{O}_c - \delta^{18}\text{O}_{\text{sw}})^2, (1)$$

Specifically, we used the quadratic approximation (Bemis et al., 1998) of Kim and O’Neil (1997), where $T_0 = 16.1$, $a = 0.09$, $b = -4.64$ and converted from V-SMOW to V-PDB using a constant of -0.27‰ (Hut, 1987; Roche et al., 2017):

$$\Delta = b^2 - 4a \cdot (T_0 - T_{\text{sw}}), (2)$$

$$\delta^{18}\text{O}_{c,\text{eq}} = \frac{-b - \sqrt{\Delta}}{2a} + \delta^{18}\text{O}_{\text{sw}} - 0.27, (3)$$

The difference between the constant of Hut (1987) and the dynamic value (Brand et al., 2014) is **minor**.

2.2 Foraminifera as modelled entities (FAME)

25 Foraminifera as modelled entities has been developed as a tool for translating, a climatic input (typically a reanalysis dataset or climate model output) into a (simulated-) climatic signal, a signal that aims to approximate the depth integrated growth of foraminifera (e.g., Pracht et al., 2019; Wilke et al., 2006; Steindhardt et al., 2015). Data-model comparison studies suffer from an inability to directly compare like with like so that there are differences in (i) the units used *i.e.*, most proxies reconstructing temperature do not give values of temperature in degrees °C or K but in their own proxy units (e.g., per mil ‰; mmol/mol; species abundance or ratio) necessitating a conversion; and (ii) there is a reduction in scales, *i.e.*, models give a



wealth of information (multiple depth layers and high resolution time slices) in the time-depth domain. A number of models and modelling studies exist to determine the foraminiferal responses to present (Fraile et al., 2008, 2009; Kageyama et al., 2013; Kretschmer et al., 2017; Lombard et al., 2009, 2011; Roy et al., 2015; Waterson et al., 2016; Žarić et al., 2005, 2005), past (Fraile et al., 2009; Kretschmer et al., 2016) and future (Roy et al., 2015) climate scenarios, FAME uses the associated temperature and $\delta^{18}\text{O}_{\text{eq}}$ at each grid cell to compute a time averaged $\delta^{18}\text{O}_{\text{c}}$ and T_{c} for a given species. FAME was produced as an attempt to reduce the error associated with data-model comparisons by (i) generating simulated-proxy time-series from model runs that can be compared with age-depth values down core; and (ii) to reduce the model information for a given time-slice into a manageable and relevant value using an integration that would make sense ⁶⁸ a biological point of view (Roche et al., 2018).

10 The FAME model utilises the temperature-growth rate equations of Lombard *et al.* (2009) to simulate temperature-derived growth rate (Kageyama et al., 2013; Lombard et al., 2009, 2011), this growth rate is then used as a ⁷⁰ weighing to produce a growth rate-weighted proxy value (Roche et al., 2018). The original Lombard et al. (2009, 2011) equations are based upon a synthesis of culture studies, pooled together irrespective of experimental design or rationale, therefore they can be considered to conceptually represent the fundamental niche of a given foraminiferal species, *i.e.* the range in environment that the species can survive. The basic structure of FAME is based upon temperature based Michaelis-Menton kinetics to predict growth rate, described in Lombard et al. (2009), without using the parameters (e.g., light, respiration, food) associated with FORAMCLIM (Lombard et al., 2011). The absence of known values or proxy values for the full set of parameters associated with FORAMCLIM has led us to an Occam's Razor favoured approach in model parameterisation for FAME (Roche et al., 2018). Although other processes may also impact species such as mixed layer depth and nutrients, these variables for now can be set aside, as temperature provides the dominant signal, it is worth noting that in all probability some variance will arise from these processes and deviation between observed and expected values should consider this.

Using the MARGO core top $\delta^{18}\text{O}_{\text{c}}$ database (MARGO Project Members*, 2009), Roche *et al.* (2018) validated and computed the optimum depth habitat (the depth habitat that exhibits the strongest correlation when comparing FAME $\delta^{18}\text{O}_{\text{c}}$ and MARGO $\delta^{18}\text{O}_{\text{c}}$) for each species in the MARGO database (MARGO Project Members*, 2009). Whilst, both models, FAME and FORAMCLIM, can compute the growth rate of eight foraminiferal species from culture studies (Kageyama et al., 2013; Lombard et al., 2009, 2011; Roche et al., 2017), the limited number of species available for a global core top comparison led to a reduction in the number of species modelled (Roche et al., 2018). Here the output of FAME is further restricted to three species that have been the main focus of foraminifera-based studies that have been used to infer ENSO variability, namely the upper ocean dwelling *Globigerinoides sacculifer* and *Globigerinoides ruber*, as well as the thermocline dwelling *Neogloboquadrina dutertrei* (Ford et al., 2015; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; Sadekov et al., 2013). The MARGO database ⁷² does not include *N. dutertrei*, meaning that we concentrate our efforts mainly on *G. ruber* and *G. sacculifer*.

In this study, ORA S4 temperature was used as the input variable, with the growth rate computations artificially constrained to arbitrary values of the upper 60; 100 and 200 m to reflect the presence of photosymbiotic algae in the various

foraminiferal species. By identifying the optimum depth habitat, Roche *et al.* (2018) established the realised niche, *i.e.* the range in environment that the species can be found, for these species for the late Holocene. Unlike some foraminiferal models, FAME does not include limiting factors such as competition, respiration or predation variables as no reliable proxy exists for such parameterisation in the geological record, therefore aspects such as interspecific competition that may limit the niche width of a species are not computed. As these depth constraints (<60 m; <100 m; and <200 m) may induce some variability we opted to include an extreme value of <400 m that grossly exaggerates the potential depth window. It is important to note however that as the computation of FAME is based on growth occurring within a temperature window it does not necessarily mean that for a given grid point modelled foraminifera will grow down to 400 m (or whichever cut-off value is used), only that the model in theory can do so (depending if optimal temperature conditions are met) to capture the total theoretical niche width. As the optimised depths computed from the MARGO dataset of Roche *et al.* (2018) are shallower, and upper ocean water is more prone to temperature variability, our approach likely dampens both the modelled $\delta^{18}\text{O}_c$ and T_c . The modelled growth rate was used to compute the monthly depth-weighted oxygen isotope distribution for each species, using the aforementioned computed $\delta^{18}\text{O}_{eq}$ for a given ~~latitudinal and longitudinal~~ grid point (Figures 2 and 3). This was repeated four times, during which the lower depth limit of the growth rate computation was set to 60; 100; 200 and 400 m. No correction for species specific disequilibria, such as vital effect, was applied to the data.

2.4 Statistical analysis

~~The tropical Pacific Ocean is divided into four Niño regions based on historical ship tracks, from east to west: Niño 1 and 2 (0° to 10°S, 90°W to 80°W), Niño 3 (5°N to 5°S, 150°W to 90°W), Niño 3.4 (5°N to 5°S, 170°W to 120°W) and Niño 4 (5°N to 5°S, 160°E to 150°W).~~ Pan-Pacific meteorological agencies differ in their definition (An and Bong, 2016, 2018) of an El Niño, with each country's definition reflecting socio-economic factors. Therefore, for simplicity we use the Oceanic Niño Index (ONI), based upon the Niño 3.4 region (because of the region's importance for interactions between ocean and atmosphere) which is a 3-month running mean of SST anomalies in ERSST.v5 (Huang *et al.*, 2017). We utilise a threshold of $\chi \geq +0.5^\circ\text{C}$ (where χ is the value of ONI) as a proxy for El Niño, $-0.5^\circ\text{C} \leq \chi \leq +0.5^\circ\text{C}$ for neutral climate conditions and $\chi \leq -0.5^\circ\text{C}$ for a La Niña in the Oceanic Niño Index. Many meteorological agencies consider that five consecutive months of $\chi \geq +0.5^\circ\text{C}$ must occur for the classification of an El Niño event. However, here the only difference is that we consider that any single month falling within our threshold values as representative of El Niño, neutral or La Niña conditions (grey bars in Figure 1). This simplification reflects the lifecycle of planktonic foraminifera (~4 weeks) seeing that the population at time step t does not record what happened at $t-1$ or what will happen at $t+1$. As we are producing the mean population growth weighted $\delta^{18}\text{O}$ values, 'almost' El Niño or 'almost' La Niña would be indistinguishable from the build-up and subsequent climb-down of actual El Niño and La Niña events. Therefore, these 'almost' El Niño or 'almost' La Niña are placed within their respective climatological pools as El Niño or La Niña.

Each time-step for the entirety of the Pacific was classified as one of three climate states (El Niño; Neutral; and La Niña), where after the resultant $\delta^{18}\text{O}_c$ and T_c at each timestep produced by FAME for each grid-point were binned into their

respective categories. An Epanechnikov-kernel distribution was first fitted to the binned monthly output of a single climate state, the bandwidth varies between grid-points to provide for an optimal kernel distribution. The use of an Epanechnikov-kernel distribution to fit the data, as opposed to other types of distribution, represents a trade-off between keeping as many parameters constant whilst mimicking the underlying dataset for a large number of grid points. The conversion of the data from dataset to distribution may induce some small error induced by: rounding to whole integers; the use of a $\delta^{18}\text{O}_{\text{mid-point}}$ which gives an error associated with the bin size ($\pm 0.05\%$) that is symmetrical close to the distributions measures of central tendency but asymmetrical at the sides; and finally, the associated rounding error at the bin edges within a histogram ($\pm 0.005\%$). Subsequently any two desired distributions can be compared for (dis)similarity using an Anderson-Darling test (1954). Here, all values, i.e., the population, associated with a climatological state are compared with the other populations representing the different climatological state, the results plotted here are Neutral climate state vs. El Niño climate state.

2.5 Test of input data (Temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$)

Foraminifera as modelled entities produces a modulated response that seeks to replicate how foraminifera modify the climate signal, several studies have approximated the foraminiferal signal in a different way (e.g., Thirumalai et al., 2013; Zhu et al., 2017a). In order to understand how FAME has altered the signal, and the degree to which the conclusions drawn depend upon the modelled growth rates, the input datasets of the sea water properties (Temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$), underwent a similar statistical test (Figure 4). Unlike FAME, which integrates over several depth levels using the computed growth rate, the test of the input datasets was with fixed depths without any growth rate weighting. These fixed depths are 5, 149 and 235 m, giving a Eulerian view (Zhu et al., 2017a) in which to observe the implications of FAME's dynamic depth habitat. As per the FAME output, each timestep value was placed into its climate state and an Anderson-Darling test performed to compare the (dis)similarity of on the resultant distributions.

2.6 Alternative statistical tests

In order to compare our results with previously published studies using planktonic foraminifera we employed a series of simple statistical tests, mimicking those applied to sediment archives by the palaeoclimate community. A chief parameter that has been employed in previous ENSO proxy work using foraminiferal analysis (more specifically, individual foraminiferal analysis; IFA) is the measure of individual foraminifera downcore standard deviation ($\sigma(\delta^{18}\text{O}_c)$). Increased $\sigma(\delta^{18}\text{O}_c)$ is considered to correlate to increased variation in SST and, in turn, increased ENSO incidence and/or magnitude (Leduc et al., 2009; Zhu et al., 2017a) or increased interannual variance (Thirumalai et al., 2013). The variance ($\sigma^2(\delta^{18}\text{O}_c)$) of the timeseries were computed both as the total variance and as the interannual variance, the latter is computed as outlined in Zhu et al. (2017a). For the interannual variance, the mean monthly climatology is subtracted from the dataset, producing monthly anomalies and a linear trend removed (using the detrend function of MatLab 2019a) – the resultant data was left unfiltered (i.e., Zhu et al., 2017a used a 1-2-1 filter). Four 'picking' experiments were performed, as FAME computes the average value for a given time step and given the single foraminiferal isotope variance for an equivalent time step (e.g.,

weeks: Steinhardt et al., 2015) it is more than likely that this computation reduces the real spread in values. Therefore, rather than use the terminology specimen we prefer to use months. Given the complexity in reconstructions of trace metal geochemistry (Elderfield and Ganssen, 2000; Nürnberg et al., 1996): the potential error associated with determining which carbonate phase is first used when foraminifera biomineralise (Jacob et al., 2017); growth-band integration; secondary factors (e.g., salinity, carbonate ion) the focus of the picking here has been on the $\delta^{18}\text{O}_c$. Irrespective of which experiment, 60 months were drawn, with replacement, and the number of Monte Carlo iterations is set at 10,000. We assume that the 'picker' is taxonomically well-trained and/or has a procedure in which species can be checked taxonomically post-analysis (e.g. photographing all specimens prior to analysis, Pracht et al. (2019)) and therefore do not include an error that deals with incorrect identification. Although we note that parameterisation of misidentification would be difficult, as it requires understanding of the variability in both standard deviation and absolute values for species co-occurring downcore (Feldmeijer et al., 2015; Metcalfe et al., 2015; 2019). For each run of experiment's (i) to (iii) the drawn months were saved in order to perform (iv):

(i) In Picking Experiment-I (Figure 3D), the months drawn for each iteration of the Monte Carlo were selected and each grid-point was sampled (i.e., there are 10000·60 selected months). This assumes that the same months are selected at grid point A as point B.

(ii) In Picking Experiment-II (Figure 3E), at each grid-point a Monte Carlo was run (i.e., there are 170·40·10000·60 selected months). This assumes that different months could be selected between grid point A and point B.

(iii) In Picking Experiment-III (Figure 3F), at each grid-point a Monte Carlo was run using the growth rate weighting for each month (i.e., there are 170·40·10000·60 selected months), this assumes that in periods of higher growth there will be a higher flux of the species and therefore a greater chance of selecting that month. The rationale being that not only are different months selected between grid point A and point B, but if A and B differ climatologically there may be an over subscription of ecologically beneficial habitats in one core location compared to the other.

(iv) In Picking Experiment-IV (Figures 3G to 3I), the experiment of (ii) was re-run but with the addition of two sources of error: The first error is based upon FAME producing the average value for a given time slice, therefore short-term variability in temperature and/or the spread in the population (i.e., variance in depth of an individual; variance in chamber growth per individual), as evidenced by single foraminiferal analysis of sediment trap samples (e.g., Steinhardt et al., 2015), is potentially lost. Therefore, for each picked month between -0.40 and 0.40 ‰ is added to the picked month value (in intervals of 0.02 ‰), this is approximately



±2° C (*i.e.*, ~4° C). The second error is the analytical error that an individual measurement will have. Machine measurement error is assumed to lie between -0.12 and 0.12 ‰ (in intervals of 0.005 ‰ – the 3rd decimal place is an exaggeration of machine capabilities although it will have repercussions for rounding) the 1σ of within run (as opposed to long-term average) of international stable isotope standards. The intervals of both errors (0.02 ‰ and 0.005 ‰) were chosen to give a similar number (n = 41 and 49) of potential randomly selected error for each picked month. Each picked month has their own randomly selected error for both of these errors, *i.e.*, each value is the sum of the month picked and their own error. The values for within month variability (Figures 3G) and machine error (Figure 3H) are calculated separately and then combined (Figure 3I), as they may have a corresponding or conflicting signs, either ‘cancelling’ out each other or amplifying the difference.

An associated statistical methodology is the graphical summary (as opposed to a numerical summary via a test value) of plotting the quantiles of two probability or the quantiles of sample probability distribution against a theoretical distribution distributions also referred to as a Quantile-Quantile, or Q-Q plot (*e.g.*, Ford et al., 2015; White et al., 2018). A complimentary (*i.e.*, used in association with, not as replacement, Filliben 1975) test metric, the Probability plot correlation coefficient (Filliben, 1975) can be used as a numerical summation of this approach, which bases its rationale on near linearity between the two tested distributions. This graphical technique is not used here for the following reasons, (i) the climatic categories (*i.e.*, El Nino, Neutral, La Nina) imposed upon the data give uneven sized sample distributions requiring an interpolated quantile estimate; and (ii) the large graphical computation required (170-40).

2.7 An approximation of sedimentary archives: Water depth & Sedimentation Rate

Discrete sediment intervals retrieved from systematically bioturbated deep-sea sediment cores contain foraminifera with ages spanning many centuries (Lougheed et al., 2018; Peng et al., 1979). This is in contrast to other proxies such as corals (Cole and Tudhope, 2017), speleotherms (Chen et al., 2016) and molluscs (Butler et al., 2013; Milano et al., 2017), where distinct time-specific banding is present (true ‘time-series’ proxies). The ambient signal following translation into a foraminiferal signal within the water is therefore further modulated by several post mortem processes, which include: the latitudinal-longitudinal shift in position of sinking foraminifera - the so-called ‘funnel affect’ (van Sebille et al., 2015; Deuser et al., 1981); dissolution of calcium carbonate either in the water (Schiebel et al., 2007), at the seafloor, or due to pore fluids; and bioturbation. As mentioned, mixing by bioturbation, results in an apparent smoothing of the downcore, discrete-depth multi-specimen signal (Hutson, 1980a; Löwemark, 2007; Löwemark et al., 2005, 2008; Löwemark and Grootes, 2004; Cole and Tudhope, 2017; Mix, 1987), thus leading to the possibility of interpreting single outlying foraminifera values within a specific depth as representing an ‘extreme’ climate, when they may in fact represent climate from a different time or epoch. This is especially apparent in δ¹⁸O_c where there is a difference temporally of δ¹⁸O_{sw} (*e.g.*, the ice volume effect in glacial and interglacial cycles ~1.25 ‰) meaning that the same temperature can have radically different δ¹⁸O_c values, a consequence of this is that a series of high magnitude, but low frequency El Niño events could be disturbed in a discrete-depth record.

Therefore, in order to reliably extract short-term environmental information from foraminiferal-based proxies, the signal that one is testing or aiming to recover must have a large enough magnitude, be largely unaffected by dissolution (*i.e.*, above the lysocline) so as not to adversely affect the population and the sedimentation rate must be high enough to give sufficient temporal coverage and rule out upwards bioturbation of single foraminifera from significantly different climate periods.

5 In our first step in consideration of post-mortem signal alteration we focus on dissolution. The lysocline, the depth at which dissolution first becomes apparent (Berger, 1968; 1970), and the Calcite (or Calcium Carbonate) Compensation Depth (CCD; Bramlette, 1961) vary between the different ocean basins; the Atlantic Ocean in which deep water forms has a relatively deep CCD as a by-product of 'young' well ventilated bottom waters whereas the Pacific Ocean the final section of the thermohaline circulation conveyor belt, has a shallower CCD. In order to highlight the potential for dissolution, the bathymetry of the Pacific was extracted from the General Bathymetric Chart of the Oceans GEBCO 2014 30 arc-second grid (version 20150318, www.gebco.net) between -20°S to 20°N and 120°E to -70°W (Figure 8). Depths of 3500 m below sea-level (bsl), 4000 m bsl and 4500 m bsl  used here as cut-off values, these depths represent multiple possible depths under which there is the **potential for noticeable dissolution (*i.e.*, lysocline) or be dissolved** (*i.e.*, CCD). Whilst our intention here is a generalised view to be used as an approximate guide, it is important to note that the Pacific Ocean has the largest proportion globally of >1 km tall seamounts that are smaller than <100 km (Wessel, 1997). Which may have important, relatively shallow-water sedimentary sequences, which may also be of sufficient sediment accumulation **rate**  therefore we supplement the GEBCO bathymetric data with the locations of seamounts. However, whilst there are an estimated 50,000 seamounts in the Pacific that are taller than a km (Menard, 1964; Wessel and Lyons, 1997), only 12,000 have been documented on charts (Batiza, 1982), and approximately 291 have been dated (Koppers et al., 2003; Clouard and Bonneville, 2005; Hillier, 2007). It is these 291, <1% of the estimated seamounts, we have overlain onto the bathymetric data (Figure 8b), although this number is further reduced as we only plot between 20°S and 20°N.

The second step when considering post-mortem signal alteration is the sediment accumulation rate (SAR). We first plot the time-averaged deep-sea SAR (Olson et al., 2016), adapted by Lougheed *et al.* (2018) for the Tropical Pacific (Figure 9). New geochronological tools, such as dual ¹⁴C-^δ¹⁸O measurements on single foraminifera (Lougheed et al., 2018), show that low sedimentation rate cores can have large variances in age between individual foraminifera present within a discrete 1 cm depth interval (Berger and Heath, 1968; Lougheed et al., 2018). In order to model bioturbation, a number of papers have used a **diffusion** style approach that reduces the parameters down to sediment mixing intensity and sediment mixing depth (herein referred to as bioturbation depth, BD), although this may be an artificial division purely driven by mathematical need rather than biological constraints (Boudreau, 1998). The bioturbation depth has been shown to have a global average of 9.8 cm (1 σ : \pm 4.5 cm) that is independent of both water depth and sedimentation rate (Boudreau, 1998), though likely controlled as a result of the energy efficiency of foraging, *e.g.* deeper burrows may cost more energy to produce than can be offset in extracted food resources, and potential decay in labile food resources with sediment depth. It is not possible to carry out a transient bioturbation model upon the temperature and salinity ocean reanalysis data that we used for FAME, as it only covers half a century of data, whereas thousands of years of input data are required to force a transient bioturbation model.

To investigate how much temporal signal is integrated into discrete-depth intervals for typical tropical Pacific SAR, we, therefore, utilised the single foraminifera sediment accumulation simulator (SEAMUS, Lougheed, 2019) to bioturbate, as the input climate signal (Figure's 9 to 11), 0-40,000 year $\delta^{18}\text{O}_w$ of NGRIP (North Greenland Ice Core Project Members, 2004; Rasmussen et al., 2014; Seierstad et al., 2014). The ice core time series is an ideal input for a bioturbation simulator, as it represents a highly temporally resolved climate input signal. SEAMUS simulates foraminifera in 10-year timesteps. The use of the NGRIP timeseries here is purely as an input parameter to investigate the effect of bioturbation upon a given climate signal - it is important to stress that by using NGRIP as an input signal for SEAMUS we are neither implying that tropical Pacific cores should have signal similar to NGRIP or inferring some kind of causal relationship. As we seek to investigate the effect of bioturbation, no attempt has been made to modulate the input signal's absolute values to mimic expected $\delta^{18}\text{O}_c$ values and this is why each plot of the synthetic down core time series retains the use of V-SMOW, despite carbonates being required to be V-PDB (Coplen 1995). Keeping all things constant, and varying a single parameter between experiments with SEAMUS, the sediment accumulation rate (SAR) was varied to fixed values of either 1, 2, 5 or 10 cm kyr^{-1} (representative of typical Pacific SAR) and a bioturbation depth (BD) of either 5, 10 or 15cm based upon the global estimate and its error bounds (Boudreau, 1998). For each experiment, the selected values of SAR and BD were kept constant for the entire SEAMIS model run (i.e., the intensity and magnitude of bioturbation was not varied). In reality, SAR and BD may vary temporally depending on local conditions (e.g., food, oxygen). Finally, the FAME results for the three species are overlaid with a water depth mask that highlights whether grid points are above or below 3500 m below sea-level (mbsl), to also show seafloor areas under the CCD depth, where carbonate material is not preserved (Berger, 1967, 1970b). A comparison between water depth and time-averaged deep-sea SAR (Olson et al., 2016), adapted by Lougheed *et al.* (2018) is shown in Figure's 7 and 9.

3. Results

The results of the forward model (Figure 2 and 3) are compared with the input values (Figure 4) in order to identify regions in which the values are statistically distinct for different climate states (Figure's 5-7). These results are then shown against the water depth (Figure's 7 to 9) and the SAR (Figure's 9-11) for the region. The results utilise Foraminifera as Modelled Entities (FAME; Figures 2, 3, 5, 6, and 7); the original Ocean Reanalysis data with computed $\delta^{18}\text{O}_{eq}$ (Figure 4); the General Bathymetric Chart of the Oceans (GEBCO; Figures 7 to 9); and the single foraminifera sediment accumulation simulator (SEAMUS; Figures 9 to 11).

3.1 FAME Output: Variance

We compute growth-weighted $\delta^{18}\text{O}_c$ (Figure 5 and 7) and temperature (Figure 6 and 7) distributions for each grid cell in the fifty-eight year simulation using FAME (Roche et al., 2018), constraining the calculation to the Tropical Pacific Ocean (between -20°S to 20°N and 120°E to -70°W). Our model produces 696 individual monthly maps for all three species

(Figure 2). While two of the three species (*G. ruber* and *G. sacculifer*) have similar ecologies, they show differences in their resultant $\delta^{18}\text{O}_c$ for the same ocean conditions (Figure 2). A comparison of our computed variance with measured data (Supplementary Table 1) is made, we compare both the value of the nearest grid-cell and because of the size of the grid and drift of foraminifera (van Sebille et al., 2015) an average of a 3 by 3 grid in which the nearest grid-cell to the core location is in the center. A comparison is made with both the iCESM model output and the core's that match this output (Zhu et al., 2017a). For the Late Holocene sample (~1.5 ka) MD02-2529 (08°12.33'N 84°07.32'W; 1619 m) in which *N. dutertrei* individual foraminifera were analyzed from >250 μm (Leduc et al., 2009) giving a calculated standard deviation of measured foraminifera of 0.38 ‰. Whereas, the full ~60 year time series (n = 696) of FAME presented here, gives a standard deviation for all species, at depth cut off 60 m between 0.26 and 0.32 ‰; at depth cut off 100 m between 0.20 and 0.29 ‰; at depth cut off 200 m between 0.20 and 0.25 ‰; and at depth cut off 400 m between 0.20 and 0.24 ‰ (see Table 1). Although these values vary if the average of the surrounding grid cells is used (see Table 1). In comparison the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.4 ‰; Eulerian 100 m: 0.6 ‰; and Lagrangian value of 0.49 ‰. There are three samples (Koutavas and Joanides, 2012; Sadekov et al., 2013) located south of core site MD02-2529, these are the Late Holocene (~1.6 ka) samples of V21-30 (01°13'S 89°41'W; 617 m) and (~1.1 ka) V21-29 (01°03'S 89°21'W; 712 m) in which *G. ruber* was measured individually (Sadekov et al., 2013). For these two sites the measured standard deviation is 0.507 ‰ and 0.510 ‰ for V21-30 and V21-29 respectively (Koutavas and Joanides, 2012). The third core site at a similar location is (~1.6ka) CD38-17P (01°36'04 S 90°25'32W; 2580 m) was not analysed individually, instead replicates of pooled samples of 2 or 3 shells of *N. dutertrei* (Sadekov et al., 2013) were made these measured values give a standard deviation of 0.28 ‰. The full ~60 year time series (n = 696) of FAME presented here, gives a standard deviation for all species, at depth cut off 60 m between 0.33 and 0.41 ‰; at depth cut off 100 m between 0.27 and 0.40 ‰; at depth cut off 200 m between 0.25 and 0.35 ‰; and at depth cut off 400 m between 0.25 and 0.34 ‰ (see Table 1). Although these values vary if the average of the surrounding grid cells is used (see Table 1). In comparison the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.53 ‰; Eulerian 100 m: 0.75 ‰; and Lagrangian value of 0.35 ‰.

25 The study of ENSO has focused on whether the variability is entirely in response to ENSO or whether it is dominated by interannual variability (Xie, 1994, 1995; Wang et al 1994, 2010), here the interannual (Figure 3C) and total variance (Figure 3A) was computed and a ratio between the two calculated (Figure 3B; see Supplementary Table 1). Like the same analysis of interannual and total variance computed for iCESM and SODA reanalysis (Carton et al., 2000), outlined in Zhu et al. (2017a), there is also high ratio of interannual to total variance in our computed FAME dataset (Figure 3B). Although there are regions in the Eastern Equatorial Pacific wherein this ratio reduces. Despite this reduction, the ratio between total and interannual variance is still above > 0.5.

The Monte-Carlo experiments (Figure 3D-I) highlight the variation in picking a subset of the months, here 60, from the full timeseries. The FAME- $\delta^{18}\text{O}_{eq}$ *G. sacculifer* with a depth cut-off of 60 m is plotted here, the values for each grid point is the range in standard deviation (*i.e.*, the maximum standard deviation minus the minimum standard deviation) between iterations

of the Monte-Carlo ($n = 10,000$). The range in standard deviations between iterations is plotted instead of the mean of the standard deviations; with increasing n the mean converges toward the sample mean, however as the point of the Monte-Carlo is to generate plausible ‘samples’ it is more important to take into account the range in possible values which would help to establish the potential variability of subsampling. For the most part, regions with high total variance (Figure 3A) also have a larger range in standard deviations between the iterations ‘picked’. It is interesting to note that by changing from the same months picked for each grid-point (Monte-Carlo I: Figure 3D) to varying the months picked between grid-points (Monte-Carlo II: Figure 3E or Monte-Carlo III: 3F) the range goes from ‘smooth’ to a more noisy dataset. Whilst the values plotted here are not the absolute values (as they are the range in standard deviation for a given grid point for the entire 10,000 iterations), it can be seen that some of the inter-core comparisons could in essence relate to differences in picking, *i.e.* different ‘months’ picked between grid-points may exacerbate or accentuate differences. Likewise, adding random variability, between -0.4 and 0.4 ‰ (Figure 3G and 3I), may also reduce the differences between areas of high Total variance and low Total variance. Though the values associated with machine error (-0.12 to 0.12 ‰) appear to do little to affect the range (Figure 3H and 3I). Whilst again the values plotted are not the absolute values, the variability added in an attempt to mimic biological variation of a given time slice increases the range of possible standard deviations in regions with low Total variance (Figure 3G and 3I). Therefore, understanding the biological variability on shorter timescale (e.g., Steinhardt et al., 2015; Mikis et al., 2019) which, maybe here over exaggerated, may be crucial for understanding discrepancies between cores.



3.2 FAME Output: Anderson-Darling test

Using a basin-wide statistical test, we examine whether the $\delta^{18}\text{O}_c$ values of a given El Niño foraminifera population (FP_{EN}) and a given non-El Niño (‘Neutral conditions’) foraminifera population (FP_{NEU}) can be expected to be significantly different at any given specific location. Where FP_{EN} and FP_{NEU} exhibit significantly different distributions, ENSO events can potentially be detected by paleoceanographers. In cases where FP_{EN} and FP_{NEU} do not exhibit significantly different values, then the chosen species and/or location represent a poor choice to study ENSO dynamics. Each simulation time step was placed into a climate states: identification of timesteps that represent El Niño (EN), Neutral (NEU), and La Niña climate conditions was done using the Oceanic Niño Index (ONI) derivative (Huang et al., 2017) (Figure 1). Comparison, for each species, FAME’s predicted growth-weighted $\delta^{18}\text{O}_c$ and T_c distributions associated with each climate event was done using an Anderson-Darling (AD) test. This statistical test can be used to determine whether or not two distributions can be said to come from the same population. The results of this test are presented in the following way, in which there are four criteria: areas where the population distributions of the two climate states are found to be statistically similar have black grid cells in all panels referring to the Anderson-Darling test results (Figure's 4-7); the results in which the areas where the populations distributions of two climate states are found to be statistically distinct are shown with two distinct colour schemes depending on whether a computable error can be included (Grey and Hashed) or not (White).

For FAME- $\delta^{18}\text{O}_c$ the results where the populations are dissimilar are either plot as grey and hashed for *G. ruber* and *G. sacculifer* or white for *N. dutertrei* (Figure 5). This is because for these two species (*G. ruber* and *G. sacculifer*) we have the possibility to determine how robust these results are. We use the 1σ values of the observed (FAME) minus expected (MARGO), as computed by Roche et al. (2018) with the MARGO core top $\delta^{18}\text{O}_c$ database, as the potential error associated with the FAME model. Regions in which the difference between the two populations are larger than the potential error are associated with grey, whereas those less than the potential error as hashed regions (Figure 5A and B), these errors should be seen as a guide rather than a rejection of a site. Because the MARGO database does not contain *N. dutertrei* we have given the panels concerning this species a separate colour scheme, black represents grid-cells for which the two populations cannot be said to be statistically different, white grid-cells are those in which the two populations can be said to be statistically different (Figure 5C). As we do not have a similar way to calculate the error for T_c , FAME- T_c results are shown (in Figure 6) with this same binary pattern (*i.e.*, white grid-cells are those in which the two populations can be said to be statistically different and black are those in which the two populations can be said to be similar). To reduce the complexity, the overlay of the species Anderson-Darling results (Figure 7) also uses the binary colour scheme (white or black).

Our results show that much of the Pacific Ocean can be considered to have statistically different population between FP_{EN} and FP_{NEU} for both $\delta^{18}\text{O}$ (Figure 5) and T_c (Figure 6). We consider that the likely cause for such a remarkable result is due to FAME computing a weighted average and, therefore, the lack of a signal found exclusively within the regions demarked in Figure 1 as El Niño regions could represent how the temperature signal is integrated via an extension of the growth rate; growing season and depth habitat of distinct foraminiferal populations. Taking into account the FAME- $\delta^{18}\text{O}_c$ error for *G. ruber* and *G. sacculifer*, we have computed regions in which the difference in oxygen isotopes between the two populations ($\Delta\delta^{18}\text{O}_c$) compared with the AD-test is smaller than the aforementioned error (Hatched in Figure 5), *i.e.* where the mean difference between FP_{EN} and FP_{NEU} is within the error. The hatched regions in Figure 5 considerably reduce the areal extent of significant difference between FP_{EN} and FP_{NEU} , with the remaining regions aligning with the El Niño 3.4 region (Figure 1). It is important to note that this error relates to the model and in reality, the difference between the climate states could be larger or smaller. No such test was performed on the *N. dutertrei* dataset, because of its absence from the MARGO dataset.

To further test the model-driven results and to assess if they are still consistent when the depth limitation is varied, the analysis was rerun with depths of 100, 200 and an extreme value of 400 m (Figure 5-7). Whilst it is possible to discern differences between the depths, it is important to note that a large percentage of the tropical Pacific remains accessible to palaeoclimate studies. A shallower depth limitation in the model increases the area for the ‘warm’ species, suggesting that the influence of a reduced variability in temperature or $\delta^{18}\text{O}_{eq}$ with a deeper depth limit causes the differences between FP_{EN} and FP_{NEU} to be reduced. Overlaying the results of the Anderson-Darling test for all three species (Figure 7) per depth for 60, 100 and 200 m highlights the areas where multi-species comparisons could be made. To account for potential differences in depth habitat we make a combination of shallower depth for *G. ruber* and deeper depths for *G. sacculifer* and *N. dutertrei* (Pracht et al., 2019) in the final panels (Figures 7D and 7H).

3.3 Test of input parameters (fixed depth: temperature and $\delta^{18}\text{O}_{\text{eq}}$)

The model-driven results were assessed with the underlying input dataset (temperature and $\delta^{18}\text{O}_{\text{eq}}$), these underwent the same statistical test (Figure 4), although with fixed depths of 5 m, 149 m and 235 m (see section 2.5). The results for each grid point are presented as either black, grey or hashed. Areas where the population distributions of the two climate states are found to be statistically similar have black grid cells. Regions in which the difference between the two populations are larger than the potential error are associated with grey, whereas those less than the potential error as hashed regions. The threshold error (i.e., the difference between the means of each distribution) is for temperature (Figure 4A-C) 0.5 °C and for $\delta^{18}\text{O}_{\text{eq}}$ (Figure 4D-F) 0.10 ‰, these errors should be seen as a guide rather than a rejection of a site. The results of this fixed depth, non-FAME, test show that the shallowest depths produce populations that are significantly different both in terms of their mean values and their distributions. In the upper panel of Figure 4, the canonical El Niño 3.4 region is clearly visible at 5 m depth. Whilst differences exist between Anderson-Darling results for the input data (Figure 4) and the FAME $\delta^{18}\text{O}$ (Figure 5) and T_c (Figure 6), for instance close to the Panama isthmus, there are significant similarities between the plots. These plots also show that our FAME data (Figure 5-7), in which we allow foraminiferal growth down deeper than the depths in Roche et al. (2018), are a conservative estimate and thus are on the low-end (Figure 4), to account for potential discrepancies with depth habitats. In the original paper on depth habitats based upon temperatures derived from $\delta^{18}\text{O}_c$, Emiliani (1954) cautioned that the depth habitats obtained would represent a weighted average of the total population, and while foraminiferal depth habitats are likely to vary spatiotemporally, the average depth habitat is skewed toward the dominant signal (Mix, 1987).

3.4 Water depth and SAR

Our analysis uses reanalysis data for the time period 1958-2015, a hypothetical core that had a comparable resolution would essentially be analogous with a sediment core with a rapid sediment accumulation rate (SAR), representing 50 yr cm^{-1} (or 20 cm kyr^{-1}). Based on our analysis, such a hypothetical core could allow for the possible disentanglement of El Niño related signals from the climatic signal, but only in a best-case scenario involving minimal bioturbation, which is unlikely in the case of oxygenated waters. Extracting the oxygen saturation (SO_2) state, of the Pacific Ocean bottom waters from the Annual Climatology WOA13 give values that are predominantly >40 % (Figure 9B). Oxygen saturation is the concentration of Oxygen in a medium against the maximum that can be dissolved in the same medium. Whilst annual variability may exist, it is unlikely that bioturbation would be prevented by low oxygen. Therefore, using a cut off value that has been considered sufficiently high enough to outpace bioturbation (e.g., Koutavas and Lynch-Stieglitz, 2003) of 5 cm kyr^{-1} (Figure 9A) it can be demonstrated that much of the Pacific has an inferred lower sedimentation rate (< 5 cm kyr^{-1} ; Figure 9C) than this cut off value. To test the influence of bioturbation, the bioturbation simulator SEAMUS was run using the NGRIP time series. The results of SEAMUS highlight the potential single foraminifera depth displacement that low sedimentation rates can result in (Figure 9). Following the current available geochronological method (i.e., age-depth method) such specimens that are

displaced in depth are assigned the average age of the depth that they were displaced to, which could result in erroneous interpretations of climate variability when analysis such as IFA is applied (Lougheed et al., 2018). The results of SEAMUS are plotted both as time series of the bioturbated ‘NGRIP’ signal (Figure 10) and as histograms of the probability of finding a particularly pseudo-foraminifera with a given age in the bioturbation depth (Figure 11). As the bioturbation depth varies between 5, 10 and 15 cm for the different simulations of SAR, the histogram in each panel (in Figure 11) represent different thicknesses of sediment, i.e., for Figure 11 panels a, d, g, and j histograms represent data with a BD thickness of 5 cm. Likewise, the timeseries is plotted with the discrete 1 cm depth median age; the median age of the bioturbation depth (Figure 11) is the reason why each timeseries does not ‘start’ at 0 age (Keigwin and Guilderson, 2009).

The variance within a single depth in a core largely represents the integrated time signal for that depth (Figure 11), as opposed to the variance of a climatic signal for an inferred (or measured) average age for the depths in question. The proxy variance will be based both upon a non-uniform distribution in temporal frequency of specimens, i.e., older specimens are few compared to younger specimens. A large proportion of the specimens in the BD come from years that are ‘proximal’ (i.e., close to the youngest age) this may give undue confidence that the probability of picking a specimen from these years is higher, however the long-tail of the distribution means that there is an equally high chance of picking a specimen that has come from several thousand years earlier than the discrete-depth’s median age. If we consider for the moment this as picking specimens from a box, there is a high chance of picking from a single box that represents the age you want however there is an equally high chance of picking from numerous boxes with varying age. If the spread in the climatic variable is uniform throughout this time then it can be possible to reproduce a similar signal, although this would not by definition represent the actual spread in the actual climatic variable for a given time, however the spread in the climate variable is unlikely to be constant. With a varying spread in the climatic signal bioturbation can introduce the possibility of spurious interpretations, but it is of course more obvious where the measured distributions over-exaggerate the climate signal (e.g., Wit et al., 2013). Furthermore, if we consider that researchers do not pick as randomly as they profess, there is both a size and preservation bias to specimens selected, and size is not constant down-core (e.g., Metcalfe et al., 2015) we can further introduce bias within the dataset. The SEAMUS output that corresponds with our chosen SAR cut-off value of 5 cm kyr⁻¹ (Figure 10 and 11), the lower limit of our mask (Figure 9), is shown in panels in Figures 10H to 10J and Figures 11G to 11I. It is important to note however, that much of the region for which FAME is calculated upon has inferred sedimentation rates lower than this cut-off value (Figure 9C to 9H).

An additional factor in the post-mortem preservation of the oceanographic signal in foraminiferal shells is whether the shells can be preserved. The GEBCO bathymetry data is binned into 250 m wide bins, and the data normalised to 1.0. As the data contains both bathymetric and topographic (below and above 0 m), the grey area in each histogram represent > 0 m (Figure 8). Whilst, there are differences depending on the cut-off value (Figures 8C to 8E) much of the canonical El Niño 3.4 region (Wang et al., 2017) used in oceanography (Figure 1) is also excluded from these suitable areas. Overlaying the water depth and the SAR with the Anderson-Darling results (Figure 7) highlights that of the total area where FP_{EN} is significantly different from FP_{NEU} (i.e. those areas where planktonic foraminiferal flux is suitable for reconstructing past ENSO

dynamics), only a small proportion corresponds to areas where the sea floor is both above the CCD (< 3500 mbsl) and SAR is at least 5 cm/ka (Figure 9). However, at certain locations, near islands or seamounts, the SAR and water depth may be high enough to allow for a signal to be preserved (Figure 8B) that may not be represented here.

4. Discussion

5 4.1 From Life to Sedimentary Assemblages

Whilst we are principally interested in understanding whether living foraminifera can theoretically reconstruct ENSO, comparison with data requires additional analysis. This is because data-model comparisons are subjective, nominally supposing that the data is the value to be achieved by the model. However, if the foraminifera modulate the original climate signal, then preservation selectively filters which specimens are conserved whereas bioturbation acts to reorder, transposing the order in which they are recovered from the depth domain. Once the sediment is recovered, the researcher acts as a final filter, which is in essence a random picking. Although technically most researchers will pick whole shells so alongside size selectivity (e.g., Metcalfe et al., 2015) there is also preservation bias associated with picking of foraminifera (e.g. Koutavas and Lynch-Stieglitz, 2003). Whilst the presence of depths in the ocean whereupon calcite is absent from sediments was described in the earliest work (e.g., Murray and Renaud, 1891; Sverdrup, 1942), overlaying maps of measured surface sediment carbonate percentage with water depth in the Pacific Ocean led Bramlette (1961) to coin the term ‘compensation depth’ (Wise, 1978). This work highlighted the ‘narrow’ depths (4-5000 m) in the Central Pacific of the CCD. Conceptually Berger (1971) placed three levels in the Pacific ocean that were descriptive of the aspects (e.g., chemical, palaeontological and sedimentological) of the calcite budget; the saturation depth, demarking supersaturated from undersaturated; the lysocline, the depth at which dissolution becomes noticeable (Berger 1968, 1971); and compensation depth (Bramlette, 1961), in which supply is compensated through dissolution. The aspects of the lysocline was estimated by the faunal assemblages of Parker and Berger (1971, figure’s 14 and 15 of that publication), for much of the equatorial Pacific the lysocline is estimated at ~3800 m. As the lysocline is where dissolution becomes apparent, ergo it is a sample already visibly degraded, we therefore set the limit of the water depth mask shallower, at 3500 m bsl. In fact, in regions of high fertility, such as the Eastern Equatorial Pacific, the lysocline was estimated to be present at ~2800 m (Thunell et al., 1981) or ~3000 m (Berger, 1971; Parker and Berger, 1971). For instance, core V21-28 close to the Galapagos Islands (01°05’N, 87°17’W) has a shallower dissolution effect than either of these two values despite being collected from a water depth of 2714 m (Luz, 1973). A comparison between the hydrographic and sedimentary lysocline, using a mooring in the Panama Basin showed that the sedimentary lysocline is a product of where the hydrographic lysocline meets the seafloor (Thunell et al., 1981), therefore, this could lead to dissolution within the water of the settling flux (e.g., Schiebel et al., 2007). In the EEP region the shallower lysocline is accompanied by an equally shallower CCD (located at ~3600 m) for which the highly fertile is considered responsible for its shoaling, lowering the pH through increased CO₂ (Berger et al., 1976). The correspondence between lysocline depth and CCD depth does not hold true for the entirety of the Pacific, plotting a N-S cross-section from

50°N to 50°S Berger (1971) noted that in the Central Equatorial Pacific, the high fertility region generates a larger zone of dissolution resistant facies even with a shoaled lysocline. If we factor in the sedimentation rate of the Pacific, which has been estimated to be considerably lower than 1 cm (Blackman and Somayajulu, 1966; Berger, 1969; Menard, 1964), then dissolution may become further exacerbated. The longer a shell remains at the sediment-water interface the greater the prospects for it to be dissolved become, therefore low SAR increases the chance of dissolution (Bramlette, 1961). For instance, in 15 equatorial Pacific cores, below 4000 m, the average SAR was presented (Hays et al., 1969; here calculated) at 0.96 cm kyr⁻¹ (1σ ± 0.43 cm kyr⁻¹). Although there are regions and/or core locations in which the SAR is higher, for instance eight EEP cores shallower than the lysocline depth (Thunell et al., 1981) of ~2800 m were presented by Koutavas and Lynch-Stieglitz (2003) which have an average SAR, calculated at 7.20 cm kyr⁻¹ (1σ ± 2.82 cm kyr⁻¹). The average age for these same core's 0 cm core depth is 2184 years (1σ ± 1521 yrs), whilst it cannot be assumed that there has been no loss during recovery (i.e., core top is not sediment-water interface), a non-zero core top age is expected for both bioturbation (Keigwin and Guilderson, 2009) and dissolution. Alongside, the potential for dissolution there is the also the mixing of ocean sediments by the benthos (Bramlette and Bradley, 1942). For instance, Arrhenius (1961) noted that ash beds present in cores of the EEP (Worzel, 1959; Ewing et al., 1959) had a 2-3 cm layer above and below what should have originally been a sharp boundary in which they estimated that ~50% of the material originated from the other side of the boundary. If one assumes 1 cm kyr⁻¹ sedimentation rates, then the range in age of the obviously 6 cm mixed sediments is minimally ~6000 years per cm, comparison with an analogous SEAMUS simulation (bioturbation depth 5 cm; SAR 1cm) highlights the considerable spread in age, placing the 95.45% range between 110 and 18954 years (Figure 11). Much of this temporal variability (either through bioturbation or dissolution) will be hidden, especially when proxy values correspond with the expected values, and more obvious when the values are larger than expected (e.g., Wit et al., 2013). Owing to the lack of absolute variability during the Holocene the apparent confirmation of similarity between proxy values and modern distributions of the 'to be reconstructed' variable is not a confirmation of proxy reliability. Especially in the tropics wherein seasonal variability is limited. The effects of both bioturbation and dissolution are further amplified when combined with finite sampling strategies. Therefore, the results of the sedimentological features, presented here, imply that much of the Pacific Ocean is not suitable for preserving (Figures 7-9) the ENSO signal, despite the possibility of the species of foraminifera having unique values for different climate states (Figures 4-7). ENSO studies using palaeoceanography have exposed shifts, one can, therefore, question what is being reconstructed in such studies.

4.2 Palaeoceanographic Implications

4.2.1 Pacific climate reconstructions

30 One artefact of sampling (Dolman and Laepple, 2018) is the potential occurrence of aliasing (Pisias and Mix, 1988; Wunsch, 2000; Wunsch and Gunn, 2003): a fundamental problem with proxy records is that they can be confounded by local regional climate, and/or ENSO's teleconnections, that mimic ENSO changes albeit at a different temporal frequency. Our own

analysis using our FAME $\delta^{18}\text{O}_c$ and T_c output mimics foraminiferal sedimentary archives, pooling several decades worth of data in which the resolution is coarse enough to obscure and prevent individual El Niño events being visible but allowing for some kind of long-term mean state of ENSO activity to be reconstructed (Cole and Tudhope, 2017). The results of our Anderson-Darling test may be unduly influenced by the Pacific decadal variability (PDV), also referred to as the Pacific Decadal Oscillation (PDO) (Pena et al., 2008). In much of the tropical Pacific the ratio of decadal to interannual σSST suggests that they are comparable in magnitude, therefore fluctuations in SST are more obviously apparent outside of the purely canonical regions of ENSO (Wang et al., 2017). It could be that the areas outside of these canonical ENSO regions (Figure 1) reflect the PDO (Pena et al., 2008; Wang et al., 2017). The use of the variance $\sigma^2(\delta^{18}\text{O}_c)$, or standard deviation $\sigma(\delta^{18}\text{O}_c)$, as an indicator of ENSO is dependent on whether the original climate signal's variance was or was not dominated by interannual variance. Zhu *et al.* (2017) computed the total variance change with and without the annual cycle suggesting that, for some cores the increased assumed ENSO variability at the LGM as deduced by proxy records (Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003) may be purely a by-product of the annual cycle or dominated by it. Although the values of El Niño can be considered significantly different from other climate states, our own analysis using the ratio of total σ^2 interannual variance also suggests that much of the variance in the simulated foraminiferal signal is dominated by interannual variance. There are differences in the ratio of total to interannual variance between species and in different regions of the tropical Pacific, however, even with a dynamic depth habitat the ratio is still high (Figure 3; Supplementary Table 1).

4.2.2 The use of models in reconstructions

Reconstructions of the past climate of the Pacific have inferred a relatively weaker Walker circulation, a displaced ITCZ and equatorial cooling (Koutavas and Lynch-Stieglitz, 2003); both a reduction (Koutavas and Lynch-Stieglitz, 2003) and intensification (Dubois et al., 2009) in eastern equatorial Pacific upwelling; and both weakened (Leduc et al., 2009) and strengthened ENSO variability (Koutavas and Joanides, 2012; Sadekov et al., 2013). However, a number of the inferences are contentious, for instance the reduction in upwelling in this region (Koutavas and Lynch-Stieglitz, 2003) is contradicted by Dubois et al. (2009), who used alkenones (i.e., $U_{37}^{K'}$ ratios) to suggest an upwelling intensification (Zhang et al., 2017). Whilst the $U_{37}^{K'}$ proxy has problems within coastal upwelling sites (Kienast et al., 2012) it does not discount their claim, especially considering that $\delta^{18}\text{O}$ records can themselves be influenced by salinity upon the $\delta^{18}\text{O}_{\text{sw}}$ component (Rincón-Martínez et al., 2011) and the potential influence of $[\text{CO}_3^{2-}]$ upon foraminiferal $\delta^{18}\text{O}_c$ (de Nooijer et al., 2009; Spero et al., 1997; Spero and Lea, 1996). The discrepancies in reconstructed climate between marine cores' is worth noting, as ultimately it is from proxies that inferences are made about past climate (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004). Such inferences have suggest that the past climate of the Pacific region (from the geologically recent too deep time) has been in an: El Niño state (Koutavas et al., 2002; Stott et al., 2002; Koutavas and Lynch-Stieglitz, 2003); permanent El Niño state (Huber and Caballero, 2003) or Super El Niño state (Stott et al., 2002); La Niña state (Andreasen et al., 2001; Beaufort et al., 2001; Martinez et al., 2003); or a different climatic state (Pisias and Mix, 1997; Feldberg and Mix, 2003).

The possibility of a marine sediment archive being able to reconstruct ENSO dynamics comes down to several fundamentals besides whether the signal can or cannot be preserved (*i.e.*, whether the core site has either too low SAR, too high BD or a water depth not conducive to calcite preservation): the time-period captured by the sediment intervals (a combination of SAR and bioturbation); the frequency and intensity of ENSO events; the foraminiferal abundance during ENSO and non-
5 ENSO conditions; as well as what the proxy is recording. There is also the presumption that a particular climate event should be recorded, our Anderson-Darling test for instance highlights that there are locations that cannot discern the difference between El Niño and other climate states whilst for the same time period there are locations where the different climate states can be differentiated. Whilst our analysis is a statistical treatment of the data, each species, and different types of phyto- or zooplankton preserved in ocean sediments, are likely to record the same set of environmental conditions differently
10 (Mix, 2006). This is, in brief, the rationale for the development of FAME, the same climate signal seen through the view of species-specific proxies will give a fractured view constrained by each species particular ecophysiological constraints (Mix, 1987; Roche et al., 2018). A dynamic depth habitat in which the environmental signal becomes a weighted average of the water column can further confound the original signal (Wilke et al., 2006). What can be seen as contradictory reconstructions can therefore be viewed as the prevailing or dominant conditions at a given location at the time when
15 environmental conditions overlap ecological constraints for a given species. Terrestrial records suggest the number of El Niño events per century in the early Holocene (8-6 ka BP) was minimal (Moy et al., 2002), with between 0 and 10 events occurring per century. This dampened ENSO is observed within lake core colour intensity and records driven primarily by precipitation - although like other proxies this can also be interpreted differently, *i.e.* as a large change in the hydrological cycle shifting precipitation away regionally (Trenberth and Otto-Bliesner, 2003). If
20 we assume for now that the number and magnitude of ENSO events was reduced, the relatively low downcore resolution of marine records may not accurately capture the dynamics of such lower amplitude ENSO events using existing methods. The sensitivity and probability of detecting a change in IFA with changes in frequency and amplitude, has been dealt with before (Thirumalai et al., 2013), although without considering bioturbation. The synthesis of pseudo-timeseries to discern the potential distribution for different scenarios, whilst a necessary approximation, is nonetheless one that is free of cause and
25 causality. Modulating a timeseries for events with enhanced or weakened amplitude or fewer or greater number of events assumes in essence that there is limited feedback both regionally (between two sites) and internally within the timeseries (*i.e.*, a process that operates on a higher level). Reconstructions of the past can benefit from inclusion within conceptual frameworks that incorporate both data and modelling studies (e.g., Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). The use of coupled ocean-atmosphere models (e.g., Clement et al., 1999; Zebiak and Cane, 1987); isotope enabled Earth system models (e.g., iCESM; Zhu et al., 2017); or multi-model ensembles with
30 prescribed boundary conditions can be used for the generation of timeseries in which the physics of atmospheric and oceanic circulation are constrained and feedbacks between sites can occur. The perceived failure of several climate models to resolve ENSO adequately, resulting in variable ENSO frequency and amplitude between models, could therefore be used to determine the proxy signal from model derived timeseries at different frequencies and intensities of ENSO. Albeit a

timeseries of variable ENSO that is grounded in ocean-atmosphere coupling. Such analysis could also provide information on a secondary assumption, in which time slices from the same core inherently assume that where a particular oceanographic feature exists now is also where it may have existed before. This gives a somewhat binary view, the feature either occurs or does not occur, and if it occurs then it has either enhanced or weakened. Yet this can (though not always) preclude a scenario

5 in which the feature has shifted. Analysis of the El Niño patterns suggests that there are two types of El Niño that are spatially delineated: the dateline Central Pacific El Niño and the Eastern Pacific El Niño. The expansion, contraction or shift of certain large scale oceanographic features (e.g., Polar Front, Upwelling) during periods of warmer than average (e.g., the last interglacial) or colder than average temperatures (e.g. the LGM) can complicate the comparison of two down core samples, i.e., a static core continuously recording a particular climate event as opposed to a shifting oceanographic regime

10 passing over or beyond a core site (Weyl, 1978). Climate models could therefore also be used to determine applicable core locations for comparison of proxy values with ‘like with like’ oceanographic features (similar to the analysis of Evans et al. (1998) for predicting coral sites), without necessarily the cost of a time-slice project (e.g., CLIMAP, MARGO).

4.2 Limitations of the methods applied and assessment of model uncertainties

For simplicity we have assumed that our model is ‘perfect’, of course that is inaccurate, there are four potential sources of error: the input variables (temperature, salinity and their conversion into $\delta^{18}\text{O}_{\text{sw}}$ and $\delta^{18}\text{O}_{\text{eq}}$); the model’s error with respect to

15 real world values (Roche et al., 2018); the statistical test’s errors (associated Type I – in which attribution of significance is given to an insignificant random event, a false ‘positive’ – and Type II – in which a significant event is attributed to be insignificant, a false ‘negative’ - errors); and reducing the complexities of foraminiferal biology via parameterization. The input variables can have errors associated with both the absolute values of temperature and salinity used here; and the

20 limitation of input values to a single value per month. Whilst it is possible to interpolate to a daily resolution, this is problematic for two reasons: (1) daily temperature records have much more high frequency oscillations than the data here and (2) the lifecycle of a single foraminifera is approximately monthly, therefore by using monthly data it provides an estimate of the average population signal. Conversion of salinity and temperature into $\delta^{18}\text{O}_{\text{sw}}$ and $\delta^{18}\text{O}_{\text{eq}}$ uses a quadratic approximation, one source of error is the unknown influence of carbonate ion concentration on both the Kim and O’Neil

25 (1997) equation and the foraminiferal microenvironment (de Nooijer et al., 2008, 2009; Spero et al., 1997; Spero and DeNiro, 1987; Spero and Lea, 1996) which has further implications due to the upwelling of cool, low pH, waters in the eastern Tropical Pacific (Cole and Tudhope, 2017; Raven et al., 2005). The spatial variability in salinity, particularly within regions underlying the intertropical convergence zone (ITCZ) and the moisture transport from the Caribbean into the eastern Pacific along the topographic low that represents Panama Isthmus, the resultant conversion of salinity to $\delta^{18}\text{O}_{\text{sw}}$ and then

30 $\delta^{18}\text{O}_{\text{eq}}$ may contain further error. If such errors are independent of the absolute value of the variable, i.e. the error on cold temperature is the same and not larger than warm temperatures, then the error terms effectively cancel one another out. A point of note, is that the $\delta^{18}\text{O}$ to $^{\circ}\text{C}$ conversion of Kim and O’Neil (1997) is considered to be marginally larger at the cold end than at the warm end (0.2 ‰ per 1°C to 0.22 ‰ per 1°C) than that originally discerned (O’Neil et al., 1969).

The comparison of the pseudo-Mg/Ca temperature signal produced here (T_c) to a value corresponding to that reconstructed from measurements of Mg/Ca should be done with caution. Computation of pseudo-foraminiferal $\delta^{18}\text{O}$ in FAME is aided by the ability to compute an initial $\delta^{18}\text{O}$ equilibrium value for a given latitude-longitude grid-point and timestep. The weighting of $\delta^{18}\text{O}$ value used in FAME is an approximation of the foraminiferal shell, chambers are generally homogenous in $\delta^{18}\text{O}$ value excluding either terminal features such as crust or gametogenic calcite which can lead to chamber heterogeneities (e.g., Wycech et al., 2018) although this can be approximated with an additional parameter (Roche et al., 2018). The same cannot be said for Mg/Ca, alongside heterogeneities in the shell which may be the result of diurnal processes (terminal features in the computation of $\delta^{18}\text{O}$ are technically simpler to model), there are differences in both sample preparation and measurement techniques. Whilst, the change in Mg/Ca with temperature has been validated (e.g., Elderfield and Ganssen, 2000) the computation of a pseudo-proxy value for and from model parameters remains enigmatic. Construction of a matrix of equilibrium Mg/Ca would ideally be the most logical step in a second generation of the FAME model. Whilst, simply solving the Mg/Ca palaeotemperature equation for an input of T and an output Mg/Ca is a first approximation, as stated previously several other parameters can alter this technique, this includes abiotic effects such as salinity (Allen et al., 2016; Gray et al., 2018; Groeneveld et al., 2008; Kısakürek et al., 2008) or carbonate ion concentration (Allen et al., 2016; Evans et al., 2018; Zeebe and Sanyal, 2002); biotic effects such as diurnal calcification (Eggins et al., 2003; Hori et al., 2018; Sadekov et al., 2008, 2009; Vetter et al., 2013); or additional factors such as sediment (Fallet et al., 2009; Feldmeijer et al., 2013) or specimen (Barker et al., 2003; Greaves et al., 2005) ‘cleaning’ techniques. Given the role of Mg in inhibiting calcium carbonate formation, the manipulation of seawater similar to the modification of the cell’s pH (de Nooijer et al., 2008, 2009) may aid calcification and explain the formation of low-Mg by certain foraminifera (Zeebe and Sanyal, 2002).

Scaling these processes up to a basin-wide model is beyond the remit of this current paper.

Our modelling results also depend upon the species symbiotic nature and potential genotypes. For instance, mixotrophs, those organisms that utilise a mixture of sources for energy and carbon (planktonic foraminifera such as *G. ruber*; and/or *G. sacculifer*) can outcompete heterotrophic (or photoheterotrophic) organisms (planktonic foraminifera such as *Neogloboquadrina pachyderma*; *Neogloboquadrina incompta*) especially in stratified-oligotrophic waters. Whilst FAME uses only the temperature component of FORAMCLIM (Roche et al., 2018) and we have only modelled *N. dutertrei*, it is important to note that there are distinctions between the fundamental niche that FAME computes, *i.e.* the conditions that an organism can survive, and the realised niche, *i.e.* what an organism actually occupies given limiting factors within the environment. Likewise, FAME and FORAMCLIM are based upon the original culture experiments that assumed that both species (*G. bulloides* and *N. dutertrei*) are non-symbiotic or have species associations (see Bird et al., 2018, 2017). A species that hosts symbionts will likely have a restricted temperature that is associated with the temperature tolerance of their symbionts, given that the next generation of a species of planktonic foraminifera must be re-infected with their symbionts. Likewise, cryptic speciation may lead to foraminiferal genotypes exhibiting distinct environmental preferences (Bird et al., 2018, 2017; Darling et al., 2004, 2000, 1999; Huber et al., 1997; Morard et al., 2013; de Vargas et al., 1999, 2002). Incorporation of both a theoretical genotype abundance (Morard et al., 2013) and ecophysiological tolerances of different

genotypes (Bird et al., 2018) within an ecophysiological model could further reduce error within modelling of planktonic foraminiferal habitats, and thus reduce data-model comparison error. For instance, Morard *et al.* (2013) simulated the impact of genotypes upon palaeoceanographic reconstructions (in particular transfer functions) using a theoretical abundance, calculated with a best-fit gaussian response model, depending upon SST later using a similar approach (Morard et al., 2016) to deduce the impact upon $\delta^{18}\text{O}$.

Conclusion



Concentrating on the period spanning the instrumental record by using the FAME module, we forward modelled the species-specific (*i.e.*, *G. ruber*; *G. sacculifer* and *N. dutertrei*) oxygen isotope values ($\delta^{18}\text{O}_c$) and pseudo-Temperature (T_c), computed from ocean reanalysis data. The aim of this study was to determine whether the modelled values from different climate states are statistically different, our results suggest for large expanses of the Tropical Pacific the climate states do have different values. Whilst, the results show that the values between El Nino states and Neutral climate states are statistically different for a large proportion of the Tropical Pacific, the total variance is dominated by the interannual variance for much of the region. Overlaying our computed foraminiferal distributions with the characteristics of the Pacific Ocean we infer that much of the signal recorded in foraminifera corresponds to areas where several processes will alter the preservation of the foraminiferal signal. First, the inferred SAR for much of the region is critically low, and a simulation of bioturbation for different bioturbation depths and SAR, typical for the Pacific indicates that discrete core depths can have a large temporal spread in single foraminifera, possibly precluding the extraction of ENSO-related climate variability. Second, a large proportion of the seafloor lies below the lysocline, the depth at which dissolution of foraminifera becomes apparent. These factors reduce the size of the area available for reconstructions considerably, thus arguably precluding the extraction of a temporally valid palaeoclimate signal using long-standing methods. It is our inference that only at exceptional ocean sediment core sites is it possible to determine the variability in ENSO, which makes it difficult to build a Pacific basin-wide understanding of past ENSO dynamics.

Code and data availability

The ocean reanalysis data used in this paper are available from the Universiteit Hamburg. An open source version of the FAME code is available from Roche et al. (2018). Statistical routines are available as part of the Statistical package of MATLAB R2018a; mapping tools (including the topographic colormap) are part of the Mapping Toolbox. The function to retrieve GEBCO bathymetry (data available at www.gebco.net) from netcdf format, `gebconetcdf(FILE,Wlon,Elon,Slat,Nlat)`, is available from the MATLAB Central File Exchange (<https://mathworks.com/matlabcentral/fileexchange/46669-gebconetcdf-file-wlon-elon-slat-nlat>). The single foraminifera sediment accumulation simulator (SEAMUS) is published in Lougheed (2019), available at <https://doi.org/10.5194/gmd->

[2019-155](https://doi.org/10.5281/zenodo.2554843). A video of the $\delta^{18}\text{O}_{\text{shell}}$ output has been archived online (<https://doi.org/10.5281/zenodo.2554843>, Metcalfe et al., 2019).

Author Contributions

B.M. and D.M.R. designed the study. B.M. analysed the data. B.C.L. processed ocean SAR and depth data, and ran the
5 bioturbation model. B.M. drafted the manuscript with contributions from all authors.

Competing Interests

The authors declare no competing interests.

Acknowledgements

B.M. was supported by a Laboratoire d'excellence (LabEx) of the Institut Pierre-Simon Laplace (Labex L-IPSL), funded by
10 the French Agence Nationale de la Recherche (grant no. ANR-10-LABX-0018). B.M. thanks both LSCE and the VU
University Amsterdam for guest status. D.M.R. is supported by the French agency Centre National de la Recherche
Scientifique (CNRS) and the VU University Amsterdam. This is a contribution to the ACCLIMATE ERC project. The
research leading to these results has received funding to C.W. from the European Research Council under the European
Union's Seventh Framework Programme (FP7/2007-2013 Grant agreement n° 339108). B.C.L. acknowledges Swedish
15 Research Council (Vetenskapsrådet – VR) grant 2018-04992, The Swedish National Infrastructure for Computing (SNIC) at
the Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX) provided computer resources for
running the SEAMUS model. We thank the Universiteit Hamburg for their online access server for ocean reanalysis data.

5 Figure 1. Oceanic Niño Index and the temperature anomaly for a single El Niño event. (Top) Oceanic Niño Index (ONI), sourced
from a 3-month running mean of SST anomalies in ERSST.v5 of the Niño 3.4 region (Huang et al., 2017). Grey vertical bars
represent the periods in which El Niño-like conditions exist using a simple one-month threshold. (Bottom) The sea surface
temperature difference between week beginning 1st December 1997 minus the long-term climatic mean (1971 – 2000) for
December. The 1997 – 1998 El Niño represents an EP-ENSO. The long term monthly climatology, the NOAA optimum
10 interpolation (OI) SST V2, based upon the methodology of Reynolds and Smith (Reynolds and Smith, 1995) using two distinct
climatologies for 1971 - 2000 and 1982 – 2000 (Reynolds et al., 2002). Boxes represent the Niño region: Niño 1 and 2 (0° - -10°S,
90°W - 80°W), Niño 3 (5°N - -5°S, 150°W - 90°W), Niño 3.4 (5°N - -5°S, 170°W - 120°W) and Niño 4 (5°N - -5°S, 160°E - 150°W).

Figure 2. A snapshot of the output of FAME. Each panel represents an individual species (Top Panel *G. sacculifer*; Middle Panel *G. ruber* and Bottom Panel *N. dutertrei*) $\delta^{18}\text{O}$ for a single time step ($t = 696$). The species $\delta^{18}\text{O}$ for each grid-point is based upon the integrating the $\delta^{18}\text{O}_{\text{eq}}$ values using a growth-rate based weighting (FAME; Roche et al., 2018). Values are in per mil (‰ V-PDB).

5

5 Figure 4. Anderson-Darling Results for Input datasets of Temperature and Equilibrium $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{eq}}$). Results of the test in which
input variables underwent the same statistical procedure (see section 2.0) as the modelled data for (A-C) temperature and (D-F)
10 $\delta^{18}\text{O}_{\text{eq}}$ values. Here, model input data was extracted for three fixed depths ([A & D] 5 m; [B & E] 149 m; [C & F] and 235 m)
without any growth weighting applied. Black regions are those grid points in which the null hypothesis (H_0), that the El Niño and
Non- El Niño populations are not statistically different ($\text{FP}_{\text{El Niño}} = \text{FP}_{\text{Non-El Niño}}$), cannot be rejected. Gray regions represent grid
points where the H_1 hypothesis is accepted, therefore the distributions of the foraminiferal population (FP) for El Niño and Non-
El Niño can be said to be unique ($\text{FP}_{\text{El Niño}} \neq \text{FP}_{\text{Non-El Niño}}$). The hatched regions represent areas where the H_1 hypothesis can be
15 accepted, therefore the distributions of the foraminiferal population (FP) for El Niño and Non- El Niño can be said to be unique
($\text{FP}_{\text{El Niño}} \neq \text{FP}_{\text{Non-El Niño}}$), though the difference between the means of tested distribution are less than (A-C) 0.5°C or (D-F) 0.1‰ .
Each panel represents a single depth (5, 149 and 235 m).

15

5 Figure 5. Anderson-Darling Results for modelled FAME- $\delta^{18}\text{O}_{\text{eq}}$: Panels representing locations of where dissimilar and similar
values of FAME modelled species $\delta^{18}\text{O}$ occur between climate states, for (columns) particular species and (rows) particular model
depth cut-off limits. Each panel represents the Anderson-Darling test result, which are plotted with ([A] *Globigerinoides sacculifer*
and [B] *Globigerinoides ruber*) and without ([C] *N. dutertrei*) model derived error. For all panels black areas reflect latitudinal and
10 longitudinal grid points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño
is similar to the Non-El Niño ($\text{FP}_{\text{El Niño}} = \text{FP}_{\text{Non-El Niño}}$). The results in which the H_1 hypothesis is accepted, in which the, therefore
the distributions can be said to be different ($\text{FP}_{\text{El Niño}} \neq \text{FP}_{\text{Non-El Niño}}$), are plotted as either: (A – *G. sacculifer*, B – *G. ruber*) grey and
15 hatched or (C – *N. dutertrei*) solely as white regions. For species with calculatable error, grey regions represent values where the
difference between the two means of the population is greater than species-specific standard deviation of the FAME model and
hatched regions represent those in which the means are less than this standard deviation (Roche et al., 2018). For species without a
calculatable error, the regions are plotted in white. The rows represent the model runs with a depth cut-off limit at: (A-C) 60 m;
(D) 100 m; (E) 200 m; and (F) 400 m.

15 .

5 **Figure 6. Anderson-Darling Results for modelled FAME-T_c:** Panels representing locations of where dissimilar and similar values
of FAME modelled temperature recorded in the calcite shells (T_c) occur between climate states, for (columns) particular species
and (rows) particular model depth cut-off limits. Each panel represents the Anderson-Darling test result, which are plotted with
([A] *Globigerinoides sacculifer* and [B] *Globigerinoides ruber*) and without ([C] *N. dutertrei*) model derived error. For all panels
black areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H₀) and therefore the
foraminiferal population (FP) of the El Niño is similar to the Non-El Niño, and therefore the distribution between the neutral
10 climate and El Nino cannot be said to be different ($FP_{El\ Niño} = FP_{Non-El\ Niño}$). The results in which the H₁ hypothesis is accepted, in
which the distributions can be said to be different ($FP_{El\ Niño} \neq FP_{Non-El\ Niño}$), are plotted as white regions. The rows represent the
model runs with a depth cut-off limit at: (A-C) 60 m; (D) 100 m; (E) 200 m; and (F) 400 m.

5 Figure 7. Combined A-D plots. As figure 5 and figure 6, in that panels represent locations of where dissimilar and similar values
for the two climate states for (a-d) FAME- $\delta^{18}\text{O}_{\text{eq}}$ modelled oxygen isotope values or (e-h) FAME- T_c modelled temperature
recorded in the calcite shells (T_c) occur. Each panel represents the Anderson-Darling test result, the results for *Globigerinoides*
10 *sacculifer*, *Globigerinoides ruber* and *N. dutertrei* are overlaid. For all panels black areas reflect latitudinal and longitudinal grid
points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño is similar to the
Non-El Niño, and therefore the distribution between the neutral climate and El Niño cannot be said to be different ($\text{FP}_{\text{El Niño}} =$
15 $\text{FP}_{\text{Non-El Niño}}$). The results in which the H_1 hypothesis is accepted, in which the distributions can be said to be different ($\text{FP}_{\text{El Niño}} \neq$
 $\text{FP}_{\text{Non-El Niño}}$), are plotted as yellow where the depth is deeper than 3500 m bsl or purple where the depth is shallower than 3500 m
bsl (see Figure 8). Purple locations are where our results suggest that the signal of ENSO has different values and the water depth
allows for preservation – although this purple region will be smaller when inferred SAR is taken into account (see Figure 9). The
rows represent the model runs with a depth cut-off limit at: (A and E) 60 m; (B and F) 100 m; (C and G) 200 m; and (D and H)
where a combination of depths were utilised (Pracht et al., 2019).

15

5 **Figure 8. Bathymetric map of the Tropical Pacific Ocean highlighting the areas above and below the Lysocline and/or Calcite compensation depth (CCD). (A) GEBCO map of height relative to 0 m; (B) same as (A) with location of seamounts plotted (white stars); (C-E) binary colour map of GEBCO data, yellow is values below cut-off depth value ([C] 3500 m below sea-level (bsl); [D] 4000 m bsl; and [E] 4500 m bsl respectively) and purple above the cut-off depth value. The histograms represent the normalised frequency of grid cell height in bins of 250 m wide, yellow is values below cut off value ([C] 3500 m below sea-level (bsl); [D] 4000 m bsl; and [E] 4500 m bsl respectively), purple above cut off value. The grey bins in each histogram are those above 0 m.**

5 Figure 9. Map of the sedimentation rate and oxygen saturation for the Tropical Pacific. (A) Inferred sedimentation rate (Olsen et
2016). White regions represent continental shelf. (B) Oxygen saturation of the bottom grid layer of World Ocean Atlas 2013 (data
from: <https://www.nodc.noaa.gov/cgi-bin/OC5/woa13/woa13oxnu.pl>). (C, E, G) Overlay between water depth and inferred SAR,
Red / Pink: Continental shelf sediments that are (Red) shallower or (Pink) deeper than 3500 mbsl; Gray / White: grid point SAR is
lower than SAR threshold and the seafloor depth is (grey) shallower or (white) deeper than 3500 mbsl; Light Yellow/Gold: Light
yellow represents areas where the SAR is above the threshold but the water depth is deeper than 3500 mbsl in comparison Gold
10 represents areas where the SAR is above the threshold and the water depth is deeper than 3500 mbsl. The ideal locations are
therefore plotted as Gold. Cut-off limits for SAR are (C) ≥ 1 cm kyr⁻¹; (E) ≥ 2 cm kyr⁻¹ and (G) ≥ 5 cm kyr⁻¹, (D, F, H) alongside the
maps the bioturbation simulations for the minimum SAR threshold is plotted (see Figure 10 and Figure 11 for the output of
SEAMUS). Each plot gives the input values of NGRIP (grey) and for each SAR three analysis were performed with different
bioturbation depths (BD) these are (Blue) 5 cm; (Green) 10 cm; and (Orange) 15 cm.

15

5 **Figure 10. Output of the bioturbation model SEAMUS. (A) The unbioturbated input signal, NGRIP (North Greenland Ice Core Project Members, 2004; Rasmussen et al., 2014; Seierstad et al., 2014), used in our simulation of bioturbation for different SAR with SEAMUS (Lougheed, 2019). Sediment mixed layer referred to here as bioturbation depth (BD) is fixed at (B, E, H, K) 5 cm, (C, F, I, L) 10 cm and (D, G, J, M) 15 cm for sedimentation accumulation rates (SAR) of (B-D) 1 cm kyr⁻¹; (E-G) 2 cm kyr⁻¹; (H-J) 5 cm kyr⁻¹ and (K-M) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. In (B-M) grey values represent the unbioturbated input signal, NGRIP. Note, we retain the original units (V-SMOW) of the original timeseries used, no inference between Pacific climate and Greenland is intended by the use of NGRIP (see section 2.7).**

10

- 5 **Figure 11. Histograms of simulated specimen age within the bioturbation depth. The simulated age distribution present within the sediment mixed layer, referred to here as bioturbation depth (BD). BD is fixed at (A, D, G, J) 5 cm, (B, E, H, K) 10 cm and (C, F, I, L) 15 cm for sedimentation accumulation rates (SAR) of (A-C) 1 cm kyr⁻¹; (D-F) 2 cm kyr⁻¹; (G-I) 5 cm kyr⁻¹ and (J-L) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. Note the size of the BD varies, therefore the simulated age distribution comes from a varying ‘core depth’.**

10

FAME genus	FAME species	Model iteration cut-off depth (in m)	Total Standard deviation in ‰			Interannual Standard deviation in ‰			% difference between Total and Interannual			Data reference	Data	Data (age in kyr)	iCESM model Eulerian View (50 m); Standard deviation in ‰	iCESM model Eulerian View (100 m); Standard deviation in ‰	iCESM Lagrangian View; Standard deviation in ‰
			Closest Grid Cell Value	Mean of 3*3 grid	Range of 3*3 grid	Closest Grid Cell Value	Mean of 3*3 grid	Range of 3*3 grid	Closest Grid Cell Value	Mean of 3*3 grid	Range of 3*3 grid						
<i>Globigerinodites</i>	<i>sacculifer</i>	60	0.32	0.33	0.03	0.29	0.29	0.05	0.90	0.89	0.10				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>sacculifer</i>	60	0.38	0.42	0.16	0.31	0.33	0.09	0.82	0.79	0.17				0.53	0.75	0.35
<i>Globigerinodites</i>	<i>ruber</i>	60	0.26	0.27	0.01	0.23	0.23	0.03	0.88	0.88	0.13				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>ruber</i>	60	0.33	0.37	0.13	0.26	0.27	0.08	0.79	0.75	0.20	B	0.51	1.1 or 1.6	0.53	0.75	0.35
<i>Neogloboquadrina</i>	<i>dutertrei</i>	60	0.32	0.32	0.06	0.26	0.26	0.02	0.79	0.82	0.17	A	0.38	1.5	0.40	0.60	0.49
<i>Neogloboquadrina</i>	<i>dutertrei</i>	60	0.41	0.45	0.18	0.33	0.35	0.11	0.81	0.78	0.16	C	0.28	1.6	0.53	0.75	0.35
<i>Globigerinodites</i>	<i>sacculifer</i>	100	0.25	0.26	0.03	0.22	0.22	0.04	0.88	0.87	0.12				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>sacculifer</i>	100	0.33	0.36	0.14	0.28	0.29	0.08	0.84	0.81	0.16				0.53	0.75	0.35
<i>Globigerinodites</i>	<i>ruber</i>	100	0.20	0.21	0.07	0.16	0.16	0.02	0.80	0.80	0.23				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>ruber</i>	100	0.27	0.31	0.11	0.22	0.22	0.08	0.79	0.73	0.23	B	0.51	1.1 or 1.6	0.53	0.75	0.35
<i>Neogloboquadrina</i>	<i>dutertrei</i>	100	0.29	0.29	0.05	0.23	0.23	0.02	0.79	0.82	0.17	A	0.38	1.5	0.40	0.60	0.49
<i>Neogloboquadrina</i>	<i>dutertrei</i>	100	0.40	0.43	0.15	0.33	0.34	0.09	0.83	0.81	0.11	C	0.28	1.6	0.53	0.75	0.35
<i>Globigerinodites</i>	<i>sacculifer</i>	200	0.21	0.22	0.04	0.17	0.18	0.02	0.83	0.81	0.20				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>sacculifer</i>	200	0.28	0.31	0.11	0.23	0.25	0.09	0.83	0.78	0.21				0.53	0.75	0.35
<i>Globigerinodites</i>	<i>ruber</i>	200	0.20	0.20	0.09	0.16	0.16	0.03	0.78	0.79	0.24				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>ruber</i>	200	0.25	0.29	0.10	0.18	0.20	0.07	0.74	0.69	0.27	B	0.51	1.1 or 1.6	0.53	0.75	0.35
<i>Neogloboquadrina</i>	<i>dutertrei</i>	200	0.25	0.25	0.05	0.20	0.20	0.02	0.78	0.81	0.16	A	0.38	1.5	0.40	0.60	0.49
<i>Neogloboquadrina</i>	<i>dutertrei</i>	200	0.35	0.37	0.11	0.30	0.31	0.08	0.85	0.83	0.10	C	0.28	1.6	0.53	0.75	0.35
<i>Globigerinodites</i>	<i>sacculifer</i>	400	0.21	0.22	0.04	0.17	0.18	0.02	0.83	0.81	0.20				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>sacculifer</i>	400	0.28	0.31	0.11	0.23	0.24	0.09	0.83	0.78	0.21				0.53	0.75	0.35
<i>Globigerinodites</i>	<i>ruber</i>	400	0.20	0.20	0.09	0.16	0.16	0.03	0.78	0.79	0.24				0.40	0.60	0.49
<i>Globigerinodites</i>	<i>ruber</i>	400	0.25	0.29	0.10	0.18	0.20	0.07	0.74	0.69	0.27	B	0.51	1.1 or 1.6	0.53	0.75	0.35
<i>Neogloboquadrina</i>	<i>dutertrei</i>	400	0.24	0.23	0.05	0.18	0.19	0.02	0.77	0.80	0.17	A	0.38	1.5	0.40	0.60	0.49
<i>Neogloboquadrina</i>	<i>dutertrei</i>	400	0.34	0.36	0.11	0.29	0.30	0.07	0.85	0.83	0.10	C	0.28	1.6	0.53	0.75	0.35

Data from: A Leduc et al., 2009;

B Koutavas and Joanides, 2012;

C Sadekov et al., 2013

Model (iCESM) values from supplement of Zhu et al., 2017 (converted from variance)

Table 1. Data-model comparison.

References

- Allen, K. A., Hönisch, B., Eggins, S. M., Haynes, L. L., Rosenthal, Y. and Yu, J.: Trace element proxies for surface ocean conditions: A synthesis of culture calibrations with planktic foraminifera, *Geochimica et Cosmochimica Acta*, 193, 197–221, doi:10.1016/j.gca.2016.08.015, 2016.
- 5 An, S.-I. and Bong, H.: Inter-decadal change in El Niño-Southern Oscillation examined with Bjercknes stability index analysis, *Climate Dynamics*, 47(3–4), 967–979, doi:10.1007/s00382-015-2883-8, 2016.
- An, S.-I. and Bong, H.: Feedback process responsible for the suppression of ENSO activity during the mid-Holocene, *Theoretical and Applied Climatology*, 132(3–4), 779–790, doi:10.1007/s00704-017-2117-6, 2018.
- Anderson, L., Abbott, M. B., Finney, B. P. and Edwards, M. E.: Palaeohydrology of the Southwest Yukon Territory, Canada, based on multiproxy analyses of lake sediment cores from a depth transect, *The Holocene*, 15(8), 1172–1183, doi:10.1191/0959683605hl889rp, 2005.
- 10 Anderson, T. W., and Darling, D. A.: A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765-769. <http://dx.doi.org/10.2307/2281537>, 1954
- 15 Andreasen, D.H., Ravelo, A.C., Broccoli, A.J.: Remote forcing at the last glacial maximum in the tropical Pacific Ocean. *Journal of Geophysical Research - Oceans* 106, 879–897, 2001
- Arrhenius, G.: Pelagic Sediments, In: Sears M (Ed) *Oceanography. Publication 67*, American Association Advancement of Science, Washington, D.C., <https://doi.org/10.5962/bhl.title.34806>, 1961
- 20 Asmerom, Y., Polyak, V., Burns, S. and Rasmussen, J.: Solar forcing of Holocene climate: New insights from a speleothem record, southwestern United States, *Geology*, 35(1), 1–4, doi:10.1130/G22865A.1, 2007.
- Balmaseda, M. A., Mogensen, K. and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4, *Quarterly Journal of the Royal Meteorological Society*, 139(674), 1132–1161, 2013.
- 25 Barker, S., Greaves, M. and Elderfield, H.: A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry, *Geochem. Geophys. Geosyst.*, 4(9), 8407, 2003.
- Barron, J. A. and Anderson, L.: Enhanced Late Holocene ENSO/PDO expression along the margins of the eastern North Pacific, *Quaternary International*, 235(1), 3–12, doi:https://doi.org/10.1016/j.quaint.2010.02.026, 2011.
- 30 Barron, J. A., Heusser, L., Herbert, T. and Lyle, M.: High-resolution climatic evolution of coastal northern California during the past 16,000 years, *Paleoceanography*, 18(1), doi:10.1029/2002PA000768, 2003.
- Batiza, R.: Abundances, distribution and sizes of volcanoes in the Pacific Ocean and implications for the origin of non-hotspot volcanoes, *Earth and Planetary Science Letters*, 60, 195-206, 1982
- Bé, A. W. H. and Spero, H. J.: Shell regeneration and biological recovery of planktonic foraminifera after physical injury induced in laboratory culture, *Micropaleontology*, 27(3), 305–316, 1981.
- 35 Bé, A. W. H., Spero, H. J. and Anderson, O. R.: Effects of symbiont elimination and reinfection on the life processes of the planktonic foraminifer *Globigerinoides sacculifer*, *Marine Biology*, 70, 73–86, 1982.

- Beaufort, L., Garidel-Thoron, T., Mix, A.C., Pisias, N.G.: ENSO like forcing on oceanic primary production during the late Pleistocene. *Science* 293, 2440–2444, 2001
- 5 Bemis, B. E., Spero, H. J., Bijma, J. and Lea, D. W.: Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations, *Paleoceanography*, 13(2), 150–160, doi:10.1029/98PA00070, 1998.
- Bennett, J. R., Cumming, B. F., Leavitt, P. R., Chiu, M., Smol, J. P. and Szeicz, J.: Diatom, Pollen, and Chemical Evidence of Postglacial Climatic Change at Big Lake, South-Central British Columbia, Canada, *Quaternary Research*, 55(3), 332–343, doi:10.1006/qres.2001.2227, 2001.
- 10 Benson, L., Kashgarian, M., Rye, R., Lund, S., Paillet, F., Smoot, J., Kester, C., Mensing, S., Meko, D. and Lindström, S.: Holocene multidecadal and multicentennial droughts affecting Northern California and Nevada, *Quaternary Science Reviews*, 21(4), 659–682, doi:https://doi.org/10.1016/S0277-3791(01)00048-8, 2002.
- Berger, W. H.: Foraminiferal Ooze: Solution at Depths, *Science, New Series*, 156(3773), 383–385, 1967.
- 15 Berger W. H.: Planktonic Foraminifera: selective solution and paleoclimatic interpretation. *Deep-Sea Research*, 15, 31-43, 1968
- Berger, W. H.: Planktonic Foraminifera: Differential Production and Expatriation Off Baja California, *Limnology and Oceanography*, 15(2), 183–204, 1970a.
- 20 Berger, W. H.: Planktonic Foraminifera: Selective solution and the lysocline, *Marine Geology*, 8(2), 111–138, doi:10.1016/0025-3227(70)90001-0, 1970b.
- Berger, W. H.: Sedimentation of planktonic foraminifera, *Marine Geology*, 11(5), 325–358, doi:10.1016/0025-3227(71)90035-1, 1971.
- Berger, W. H. and Gardner, J. V.: On the determination of Pleistocene temperatures, *Journal of Foraminiferal Research*, 5(2), 102–113, 1975.
- 25 Berger, W.H., Adelseck, C., and Mayer, L.A.: Distribution of carbonate in surface sediments of the Pacific, *Journal of Geophysical Research*, 81 (15), 2617-2627, doi:10.1029/JC081i015p02617, 1976
- Berger, W. H. and Heath, G. R.: Vertical mixing in pelagic sediments, *Journal of Marine Research*, 26(2), 134–143, 1968.
- 30 Bijma, J., Hönisch, B. and Zeebe, R. E.: Impact of the ocean carbonate chemistry on living foraminiferal shell weight: Comment on “Carbonate ion concentration in glacial-age deep waters of the Caribbean Sea” by W. S. Broecker and E. Clark: COMMENT, *Geochemistry, Geophysics, Geosystems*, 3(11), 1–7, doi:10.1029/2002GC000388, 2002.
- Bird, C., Darling, K. F., Russell, A. D., Davis, C. V., Fehrenbacher, J., Free, A., Wyman, M. and Ngwenya, B. T.: Cyanobacterial endobionts within a major marine planktonic calcifier (*Globigerina bulloides*, Foraminifera) revealed by 16S rRNA metabarcoding, *Biogeosciences*, 14(4), 901–920, doi:10.5194/bg-14-901-2017, 2017.
- 35 Bird, C., Darling, K. F., Russell, A. D., Fehrenbacher, J. S., Davis, C. V., Free, A. and Ngwenya, B. T.: 16S rRNA gene metabarcoding and TEM reveals different ecological strategies within the genus *Neogloboquadrina* (planktonic foraminifer), edited by F. Frontalini, *PLOS ONE*, 13(1), e0191653, doi:10.1371/journal.pone.0191653, 2018.

- Blackman, A. and Somayajulu, B.L.K.: Pacific Pleistocene Cores: Faunal Analyses and Geochronology, *Science*, 154 (3751), 886-889, 1966
- Boltovskoy, D.: The Sedimentary record of Pelagic Biogeography, *Progress in Oceanography*, 34(2-3), 135-160, 1994.
- 5 Boltovskoy, E.: 314. Depth at which foraminifera can survive in sediments, *Contributions from the Cushman Foundation for Foraminiferal Research*, XVII(2), 43-45, 1966.
- Boudreau, B. P. 1998: Mean mixed depth of sediments: The wherefore and the why, *Limnology and Oceanography*, 43 (3), 524-526, 1998.
- 10 Bramlette M.N.: Pelagic sediments. In: Sears M (Ed) *Oceanography*. Publication 67, American Association Advancement of Science, Washington, D.C., <https://doi.org/10.5962/bhl.title.34806>, 1961
- Bramlette, M.N. and Bradley W.H.: Geology and Biology of North Atlantic Deep-Sea Cores between Newfoundland and Ireland: I. Lithology and Geologic Interpretation, U.S. Geological Survey Professional Paper, 196-A, 1-34,
15 <https://pubs.usgs.gov/pp/0196/report.pdf>, 1941
- Brand, W.A., Coplen, T.B., Vogl, J., Rosner, M., and Prohaska, T.: Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure and Applied Chemistry*, 86(3), 425-467, doi:10.1515/pac-2013-1023, 2014
- 20 Branson, O., Redfern, S. A. T., Tyliszczak, T., Sadekov, A., Langer, G., Kimoto, K. and Elderfield, H.: The coordination of Mg in foraminiferal calcite, *Earth and Planetary Science Letters*, 383, 134-141, doi:10.1016/j.epsl.2013.09.037, 2013.
- Branson, O., Bonnin, E. A., Perea, D. E., Spero, H. J., Zhu, Z., Winters, M., Hönisch, B., Russell, A. D., Fehrenbacher, J. S. and Gagnon, A. C.: Nanometer-Scale Chemistry of a Calcite Biomineralization Template: Implications for Skeletal Composition and Nucleation, *Proceedings of the National Academy of Sciences*, 113(46), 12934-12939,
25 doi:10.1073/pnas.1522864113, 2016.
- Caley, T., Roche, D. M., Waelbroeck, C. and Michel, E.: Oxygen stable isotopes during the Last Glacial Maximum climate: perspectives from data-model (iLOVECLIM) comparison, *Climate of the Past*, 10(6), 1939-1955, doi:10.5194/cp-10-1939-2014, 2014.
- 30 Caramanica, A., Quilter, J., Huaman, L., Villanueva, F. and Morales, C. R.: Micro-remains, ENSO, and environmental reconstruction of El Paraíso, Peru, a late preceramic site, *Journal of Archaeological Science: Reports*, 17, 667-677, doi:<https://doi.org/10.1016/j.jasrep.2017.11.026>, 2018.
- Chen, S., Hoffmann, S. S., Lund, D. C., Cobb, K. M., Emile-Geay, J. and Adkins, J. F.: A high-resolution speleothem record of western equatorial Pacific rainfall: Implications for Holocene ENSO evolution, *Earth and Planetary Science Letters*, 442, 61-71, doi:<https://doi.org/10.1016/j.epsl.2016.02.050>, 2016.
- 35 CLIMAP Project Members: The Surface of the Ice-Age Earth, *Science*, 191(4232), 1131-1137, 1976.
- Clement, A.C., Seager, R. and Cane, M.A.: Orbital controls on the El Niño/Southern Oscillation and the tropical climate, *Paleoceanography*, 14 (4), 441-456, 1999
- 40 Clouard, V. and Bonneville, A.: Ages of seamounts, islands, and plateaus on the Pacific plate, *Geological Society of America Special Paper* 388, 71-90, doi: 10.1130/2005.2388(06), 2005

- Cole, J. and Tudhope, A. W.: Reef-Based Reconstructions of Eastern Pacific Climate Variability, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, pp. 535–548, Springer Netherlands, Dordrecht., 2017.
- 5 Conroy, J. L., Overpeck, J. T., Cole, J. E., Shanahan, T. M. and Steinitz-Kannan, M.: Holocene changes in eastern tropical Pacific climate inferred from a Galápagos lake sediment record, *Quaternary Science Reviews*, 27(11), 1166–1180, doi:<https://doi.org/10.1016/j.quascirev.2008.02.015>, 2008.
- Darling, K., Wade, C. M., Steward, I. A., Kroon, D., Dingle, R. and Leigh Brown, A. J.: Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers, *Nature*, 405, 43–47, 2000.
- 10 Darling, K. F., Wade, C. M., Kroon, D., Brown, A. J. L. and Bijma, J.: The Diversity and Distribution of Modern Planktic Foraminiferal Small Subunit Ribosomal RNA Genotypes and their Potential as Tracers of Present and Past Ocean Circulations, *Paleoceanography*, 14(1), 3–12, doi:[10.1029/1998PA900002](https://doi.org/10.1029/1998PA900002), 1999.
- Darling, K. F., Kucera, M., Pudsey, C. J. and Wade, C. M.: Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics, *Proceedings of the National Academy of Sciences*, 101(20), 7657–7662, doi:[10.1073/pnas.0402401101](https://doi.org/10.1073/pnas.0402401101), 2004.
- 15 Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, *Climate of the Past Discussions*, 1–31, doi:[10.5194/cp-2018-13](https://doi.org/10.5194/cp-2018-13), 2018.
- Du, X., Hendy, I. and Schimmelmann, A.: A 9000-year flood history for Southern California: A revised stratigraphy of varved sediments in Santa Barbara Basin, *Marine Geology*, 397, 29–42, doi:<https://doi.org/10.1016/j.margeo.2017.11.014>, 2018.
- 20 Dubois, N., Kienast, M., Normandeau, C. and Herbert, T. D.: Eastern equatorial Pacific cold tongue during the Last Glacial Maximum as seen from alkenone paleothermometry, *Paleoceanography*, 24(4), doi:[10.1029/2009PA001781](https://doi.org/10.1029/2009PA001781), 2009.
- Eggins, S. M., Sadekov, A. and de Deckker, P.: Modulation and daily banding of Mg/Ca in *Orbulina universa* tests by symbiont photosynthesis and respiration: a complication for seawater thermometry, *Earth and Planetary Science Letters*, 225, 411–419, 2003.
- 25 Elderfield, H. and Ganssen, G. M.: Past temperature and $\delta^{18}\text{O}$ of surface ocean waters inferred from foraminiferal Mg/Ca ratios, *Nature*, 405(6785), 442–445, 2000.
- Emiliani, C.: Depth habitats of some species of pelagic foraminifera as indicated by oxygen isotope ratios, *American Journal of Science*, 252(3), 149–158, doi:[10.2475/ajs.252.3.149](https://doi.org/10.2475/ajs.252.3.149), 1954.
- 30 Emiliani, C.: Pleistocene Temperatures, *The Journal of Geology*, 63(6), 538–578, 1955.
- Enzel, Y. and Wells, S. G.: Extracting Holocene paleohydrology and paleoclimatology information from modern extreme flood events: An example from southern California, *Geomorphology*, 19(3), 203–226, doi:[https://doi.org/10.1016/S0169-555X\(97\)00015-9](https://doi.org/10.1016/S0169-555X(97)00015-9), 1997.
- Ericson, D. B. and Wollin, G.: Pleistocene climates and chronology in deep-sea sediments, *Science*, 162, 1227–1234, 1968.

- Ericson, D. B., Ewing, M. and Wollin, G.: The Pleistocene Epoch in Deep-Sea Sediments, *Science, New Series*, 146(3645), 723–732, 1964.
- Evans, D., Müller, W. and Erez, J.: Assessing foraminifera biomineralisation models through trace element data of cultures under variable seawater chemistry, *Geochimica et Cosmochimica Acta*, 236, 198–217, doi:10.1016/j.gca.2018.02.048, 2018.
- 5 Evans, M. N., Kaplan, A. and Cane, M. A.: Optimal sites for coral-based reconstructions of global sea surface temperature, *Paleoceanography*, 13 (5), 502-516, 1998
- Ewing, M., Heezen, B.C., and Ericson, D.: Significance of the Worzel deep-sea ash, *PNAS*, 45, 355, 1959
- 10 Fallet, U., Boer, W., van Assen, C., Greaves, M. and Brummer, G. J. A.: A novel application of wet-oxidation to retrieve carbonates from large, organic-rich samples for application in climate research, *Geochemistry Geophysics Geosystems*, 10(8), 2009.
- Feldberg, M.J., Mix, A.C.: Planktonic foraminifera, sea surface temperatures, and mechanisms of oceanic change in the Peru and south equatorial currents, 0–150 ka BP. *Paleoceanography* 18, 1–16, 2003
- 15 Feldmeijer, W., Metcalfe, B., Scussolini, P. and Arthur, K.: The effect of chemical pretreatment of sediment upon foraminiferal-based proxies, *Geochemistry, Geophysics, Geosystems*, 14(10), 3996–4014, doi:10.1002/ggge.20233, 2013.
- Feldmeijer, W., Metcalfe, B., Brummer, G. J. A. and Ganssen, G. M.: Reconstructing the depth of the permanent thermocline through the morphology and geochemistry of the deep dwelling planktonic foraminifer *Globorotalia truncatulinoides*, *Paleoceanography*, 30(1), 1–22, doi:10.1002/2014PA002687, 2015.
- 20 Filliben, J.J.: The Probability Plot Correlation Coefficient Test for Normality, *Technometrics*, 17 (1), 111-117, 1975
- Ford, H. L., Ravelo, A. C. and Polissar, P. J.: Reduced El Nino-Southern Oscillation during the Last Glacial Maximum, *Science*, 347(6219), 255–258, doi:10.1126/science.1258437, 2015.
- 25 Fraile, I., Schulz, M., Mulitza, S. and Kucera, M.: Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, *Biogeosciences*, 5, 891–911, 2008.
- Fraile, I., Schulz, M., Mulitza, S., Merkel, U., Prange, M. and Paul, A.: Modeling the seasonal distribution of planktonic foraminifera during the Last Glacial Maximum, *Paleoceanography*, 24, doi:10.1029/2008PA001686, 2009.
- Ganssen, G. M., Peeters, F. J. C., Metcalfe, B., Anand, P., Jung, S. J. A., Kroon, D. and Brummer, G.-J. A.: Quantifying sea surface temperature ranges of the Arabian Sea for the past 20 000 years, *Climate of the Past*, 7(4), 1337–1349, doi:10.5194/cp-7-1337-2011, 2011.
- 30 Garidel-Thoron, T. de, Rosenthal, Y., Beaufort, L., Bard, E., Sonzogni, C. and Mix, A. C.: A multiproxy assessment of the western equatorial Pacific hydrography during the last 30 kyr, *Paleoceanography*, 22(3), doi:10.1029/2006PA001269, 2007.
- Gray, W. R., Weldeab, S., Lea, D. W., Rosenthal, Y., Gruber, N., Donner, B. and Fischer, G.: The effects of temperature, salinity, and the carbonate system on Mg/Ca in *Globigerinoides ruber* (white): A global sediment trap calibration, *Earth and Planetary Science Letters*, 482, 607–620, doi:10.1016/j.epsl.2017.11.026, 2018.

- Greaves, M., Barker, S., Daunt, C. and Elderfield, H.: Accuracy, standardization, and interlaboratory calibration standards for foraminiferal Mg/Ca thermometry, *Geochem. Geophys. Geosyst.*, 6(2), Q02D13, 2005.
- 5 Groeneveld, J., Nürnberg, D., Tiedemann, R., Reichert, G.-J., Steph, S., Reuning, L., Crudeli, D. and Mason, P.: Foraminiferal Mg/Ca increase in the Caribbean during the Pliocene: Western Atlantic Warm Pool formation, salinity influence, or diagenetic overprint?, *Geochemistry Geophysics Geosystems*, 9, Q01P23, doi:doi:10.1029/2006GC001564, 2008.
- Hamilton, C. P., Spero, H. J., Bijma, J. and Lea, D. W.: Geochemical investigation of gametogenic calcite addition in the planktonic foraminifera *Orbulina universa*, *Marine Micropaleontology*, 68(3–4), 256–267, 2008.
- 10 Hays, J.D., Saito, T., Opdyke, N., and Burckle, L.H.: Pliocene-Pleistocene Sediments of the Equatorial Pacific: Their Paleomagnetic, Biostratigraphic, and Climatic Record, *Geological Society of America Bulletin*, 80, 1481-1514, 1969
- Hecht, A. D.: A model for determining Pleistocene paleotemperatures from planktonic foraminiferal assemblages, *Micropaleontology*, 19, 68–77, 1973.
- 15 Hendy, I. L., Napier, T. J. and Schimmelmann, A.: From extreme rainfall to drought: 250 years of annually resolved sediment deposition in Santa Barbara Basin, California, *Quaternary International*, 387, 3–12, doi:https://doi.org/10.1016/j.quaint.2015.01.026, 2015.
- Higley, M. C., Conroy, J. L. and Schmitt, S.: Last Millennium Meridional Shifts in Hydroclimate in the Central Tropical Pacific, *Paleoceanography and Paleoclimatology*, 33(4), 354–366, doi:10.1002/2017PA003233, 2018.
- 20 Hillier, J. K.: Pacific seamount volcanism in space and time, *Geophys. J. Int.*, 168, 877-889, doi: 10.1111/j.1365-246X.2006.03250.x, 2007
- Hori, M., Shirai, K., Kimoto, K., Kurasawa, A., Takagi, H., Ishida, A., Takahata, N. and Sano, Y.: Chamber formation and trace element distribution in the calcite walls of laboratory cultured planktonic foraminifera (*Globigerina bulloides* and *Globigerinoides ruber*), *Marine Micropaleontology*, 140, 46–55, doi:10.1016/j.marmicro.2017.12.004, 2018.
- 25 Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S. and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, *Journal of Climate*, 30(20), 8179–8205, doi:10.1175/jcli-d-16-0836.1, 2017.
- Huber, B. T., Bijma, J. and Darling, K.: Cryptic speciation in the living planktonic foraminifer *Globigerinella siphonifera* (d’Orbigny), *Paleobiology*, 23(01), 33–62, doi:10.1017/S0094837300016638, 1997.
- 30 Huber, M. and Caballero, R.: Eocene El Nino: Evidence for Robust Tropical Dynamics in the “Hothouse”, *Science*, 299 (5608), 877-881, 2003
- Hut, G.: Consultants group meeting on stable isotope reference samples for geochemical and hydrological investigations, Int. At. Energy Agency, Vienna., 1987.
- 35 Hutson, W. H.: Transfer functions under no-analog conditions: experiments with Indian Ocean planktonic foraminifera, *Quaternary Research*, 8(3), 355–367, 1977.

- Hutson, W. H.: Applications of transfer functions to Indian Ocean planktonic foraminifera, *Quaternary Research*, 9, 87–112, 1978.
- Hutson, W. H.: Bioturbation of deep-sea sediments: Oxygen isotopes and stratigraphic uncertainty, *Geology*, 8(3), 127–130, doi:10.1130/0091-7613(1980), 1980a.
- 5 Hutson, W. H.: The Agulhas Current during the Late Pleistocene-analysis of modern faunal analogs, *Science*, 207(4426), 64–66, 1980b.
- Imbrie, J. and Kipp, N. G.: A new Micropaleontological method for paleoclimatology: Application to a Late Pleistocene Caribbean core, in *The Late Cenozoic Glacial Ages*, edited by K. K. Turekian, pp. 71–181, Yale University Press, New Haven, Connecticut., 1971.
- 10 Jacob, D. E., Wirth, R., Agbaje, O. B. A., Branson, O. and Eggins, S. M.: Planktic foraminifera form their shells via metastable carbonate phases, *Nature Communications*, 8(1), 1265, doi:10.1038/s41467-017-00955-0, 2017.
- Jones, J. I.: Distribution of living planktonic foraminifera of the West Indies and adjacent waters, PhD Thesis, University Wisconsin-Madison, Madison, Wisconsin USA., 1964.
- 15 Jonkers, L. and Kučera, M.: Quantifying the effect of seasonal and vertical habitat tracking on planktonic foraminifera proxies, *Climate of the Past*, 13(6), 573–586, doi:10.5194/cp-13-573-2017, 2017.
- Kageyama, M., Braconnot, P., Bopp, L., Caubel, A., Foujols, M.-A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., Mariotti, V., Marti, O., Roy, T. and Woillez, M.-N.: Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model—part I: comparing IPSL_CM5A to IPSL_CM4, *Climate Dynamics*, 40(9), 2447–2468, doi:10.1007/s00382-012-1488-8, 2013.
- 20 Keigwin, L. D. and Guilderson, T. P.: Bioturbation artifacts in zero-age sediments, *Paleoceanography*, 24, PA4212, doi: 10.1029/2008PA001727, 2009
- Kienast, M., MacIntyre, G., Dubois, N., Higginson, S., Normandeau, C., Chazen, C. and Herbert, T. D.: Alkenone unsaturation in surface sediments from the eastern equatorial Pacific: Implications for SST reconstructions, *Paleoceanography*, 27(1), doi:10.1029/2011PA002254, 2012.
- 25 Kim, S.-T. and O’Neil, J. R.: Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates, *Geochimica et Cosmochimica Acta*, 61(16), 3461–3475, doi:10.1016/S0016-7037(97)00169-5, 1997.
- Kısakürek, B., Eisenhauer, A., Böhm, F., Garbe-Schönberg, D. and Erez, J.: Controls on shell Mg/Ca and Sr/Ca in cultured planktonic foraminiferan, *Globigerinoides ruber* (white), *Earth and Planetary Science Letters*, 273(3–4), 260–269, doi:10.1016/j.epsl.2008.06.026, 2008.
- 30 Koppers A.A.P. Staudigel H. Pringle M.S. Wijbrans J.R.: Short-lived and discontinuous intraplate volcanism in the South Pacific; hot spots or extensional volcanism, *Geochem. Geophys. Geosys.*, 4 (1089), 2003
- Koutavas, A. and Joanides, S.: El Niño–Southern Oscillation extrema in the Holocene and Last Glacial Maximum, *Paleoceanography*, 27(4), doi:10.1029/2012PA002378, 2012.

- Koutavas, A. and Lynch-Stieglitz, J.: Glacial-interglacial dynamics of the eastern equatorial Pacific cold tongue-Intertropical Convergence Zone system reconstructed from oxygen isotope records: *GLACIAL PACIFIC COLD TONGUE*, *Paleoceanography*, 18(4), n/a-n/a, doi:10.1029/2003PA000894, 2003.
- 5 Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. and Sachs, J. P.: El Niño-Like Pattern in Ice Age Tropical Pacific Sea Surface Temperature, *Science*, 297(5579), 226, doi:10.1126/science.1072376, 2002.
- Koutavas, A., deMenocal, P. B., Olive, G. C. and Lynch-Stieglitz, J.: Mid-Holocene El Niño–Southern Oscillation (ENSO) attenuation revealed by individual foraminifera in eastern tropical Pacific sediments, *Geology*, 34(12), 993, doi:10.1130/G22810A.1, 2006.
- 10 Kretschmer, K., Kucera, M. and Schulz, M.: Modeling the distribution and seasonality of *Neogloboquadrina pachyderma* in the North Atlantic Ocean during Heinrich Stadial 1, *Paleoceanography*, 31(7), 986–1010, doi:10.1002/2015PA002819, 2016.
- Kretschmer, K., Jonkers, L., Kucera, M. and Schulz, M.: Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale, *Biogeosciences Discussions*, 1–37, doi:10.5194/bg-2017-429, 2017.
- 15 Leduc, G., Vidal, L., Cartapanis, O. and Bard, E.: Modes of eastern equatorial Pacific thermocline variability: Implications for ENSO dynamics over the last glacial period: *ENSO DYNAMICS OVER THE LAST 50 KA*, *Paleoceanography*, 24(3), doi:10.1029/2008PA001701, 2009.
- LeGrande, A. N. and Schmidt, G. A.: Global gridded data set of the oxygen isotopic composition in seawater, *Geophysical Research Letters*, 33(12), L12604, 2006.
- Lombard, F., Labeyrie, L., Michel, E., Spero, H. J. and Lea, D. W.: Modelling the temperature dependent growth rates of planktic foraminifera, *Marine Micropalaeontology*, 70, 1–7, 2009.
- 20 Lombard, F., da Rocha, R. E., Bijma, J. and Gattuso, J. P.: Effect of carbonate ion concentration and irradiance on calcification in planktonic foraminifera, *Biogeosciences*, 7(1), 247–255, 2010.
- Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H. and Jorissen, F.: Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, *Biogeosciences*, 8(4), 853–873, doi:10.5194/bg-8-853-2011, 2011.
- 25 Loubere, P., Creamer, W. and Haas, J.: Evolution of the El Niño-Southern Oscillation in the late Holocene and insolation driven change in the tropical annual SST cycle, *Global and Planetary Change*, 100, 129–144, doi:https://doi.org/10.1016/j.gloplacha.2012.10.007, 2013.
- Lougheed, B. C.: SEAMUS (v1.0): a $\Delta 14\text{C}$ -enabled, single-specimen sediment accumulation simulator, *Geoscientific Model Development*, 2019.
- 30 Lougheed, B. C., Metcalfe, B., Ninnemann, U. S. and Wacker, L.: Moving beyond the age depth model paradigm in deep-sea palaeoclimate archives: dual radiocarbon and stable isotope analysis on single foraminifera, *Climate of the Past*, 14(4), 515–526, 2018.
- Löwemark, L.: Importance and Usefulness of Trace fossils and Bioturbation in Paleocyanography, in *Trace Fossils: Concepts, Problems, Prospects*, edited by W. Miller, pp. 413–427, Elsevier, Amsterdam., 2007.

- Löwemark, L. and Grootes, P. M.: Large age differences between planktic foraminifers caused by abundance variations and Zoophycos bioturbation, *Paleoceanography*, 19, PA2001, doi:10.1029/2003PA000949, 2004.
- Löwemark, L., Hong, W.-L., Yui, T.-F. and Hung, G.-W.: A test of different factors influencing the isotopic signal of planktonic foraminifera in surface sediments from the northern South China Sea, *Marine Micropaleontology*, 55(1–2), 49–62, doi:10.1016/j.marmicro.2005.02.004, 2005.
- Löwemark, L., Konstantinou, K. I. and Steinke, S.: Bias in foraminiferal multispecies reconstructions of paleohydrographic conditions caused by foraminiferal abundance variations and bioturbational mixing: A model approach, *Marine Geology*, 256, 101–106, 2008.
- Luz, B.: Stratigraphic and Paleoclimatic Analysis of late Pleistocene Tropical Southeast Pacific Cores, *Quaternary Research*, 3, 56-72, 1973
- Lynts, G. W. and Judd, J. B.: Late Pleistocene paleotemperatures at Tongue of the Ocean, Bahamas, *Science*, 171, 1143–1144, 1971.
- MARGO Project Members*: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, *Nature Geoscience*, 2(2), 127–132, doi:10.1038/ngeo411, 2009.
- McIntyre, A., Kipp, N. G., Bé, A. W. H., Crowley, T., Kellogg, T., Gardner, J. V., Prell, W. L. and Ruddiman, W. F.: Glacial North Atlantic 18,000 years ago: A CLIMAP reconstruction, *Geological Society of America Memoirs*, 145, 43–76, 1976.
- McPhaden, M.J., Zebiak, S.E., and Glantz, M.H.: ENSO as an Integrating Concept in Earth Science, *Science*, 314 (5806), 1740-1745, 2006
- Menard H.W.: *Marine Geology of the Pacific*, McGraw-Hill, New York, 1964
- Metcalf, B., Feldmeijer, W., de Vringer-Picon, M., Brummer, G.-J. A., Peeters, F. J. C. and Ganssen, G. M.: Late Pleistocene glacial–interglacial shell-size–isotope variability in planktonic foraminifera as a function of local hydrography, *Biogeosciences*, 12(15), 4781–4807, doi:10.5194/bg-12-4781-2015, 2015.
- Metcalf, B., Loughheed, B.C., Waelbroeck, C., and Roche D.M.: DATA: Data set, Zenodo, <https://doi.org/10.5281/zenodo.2554843>, 2019
- Metcalf, B., Feldmeijer, W., and Ganssen, G.M.: Oxygen isotope variability of planktonic foraminifera provide clues to past upper ocean seasonal variability, *Paleoceanography and Paleoclimatology*, doi: 10.1029/2018PA003475, 2019
- Mikis, A., Hendry, K. R., Pike, J., Schmidt, D. N., Edgar, K. M., Peck, V., Peeters, F. J. C., Leng, M. J., Meredith, M. P., Todd, C. L., Stammerjohn, S., and Ducklow, H.: Temporal variability in foraminiferal morphology and geochemistry at the West Antarctic Peninsula: a sediment trap study, *Biogeosciences Discuss.*, <https://doi.org/10.5194/bg-2019-19>, in review, 2019.
- Mix, A. C.: The oxygen-isotope record of deglaciation, in: *North America and adjacent oceans during the last deglaciation*, in *The Geology of America*, vol. K-3, edited by W. F. Ruddiman and H. E. J. Wright, pp. 111–135, Geological Society of America, Boulder, Colorado., 1987.
- Mix, A.C.: Running hot and cold in the eastern equatorial Pacific, *Quaternary Science Reviews*, 25, 1147-1149, doi:10.1016/j.quascirev.2006.03.008, 2006

- Morard, R., Quillévéré, F., Escarguel, G., de Garidel-Thoron, T., de Vargas, C. and Kucera, M.: Ecological modeling of the temperature dependence of cryptic species of planktonic Foraminifera in the Southern Hemisphere, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 391, 13–33, doi:10.1016/j.palaeo.2013.05.011, 2013.
- 5 Morard, R., Reinelt, M., Chiessi, C. M., Groeneveld, J. and Kucera, M.: Tracing shifts of oceanic fronts using the cryptic diversity of the planktonic foraminifera *Globorotalia inflata*, *Paleoceanography*, 31(9), 1193–1205, doi:10.1002/2016PA002977, 2016.
- Moy, C. M., Seltzer, G. O., Rodbell, D. T. and Anderson, D. M.: Variability of El Niño/Southern Oscillation activity at millennial timescales during the Holocene epoch, *Nature*, 420, 162, doi:10.1038/nature01194, 2002.
- 10 Mulitza, S., Wolff, T., Pätzold, J., Hale, W. and Wefer, G.: Temperature sensitivity of planktic foraminifera and its influence on the oxygen isotope record, *Marine Micropaleontology*, 33(3–4), 223–240, doi:10.1016/S0377-8398(97)00040-6, 1998.
- Murray, J; Renard, AF: Deep-sea deposits (based on the specimens collected during the voyage of HMS Challenger in the years 1872 to 1876). Report on the scientific results of the voyage of H.M.S. Challenger during the years 1873-76; John Menzies and Co., Edinburgh, United Kingdom, 1891
- 15 Nederbragt, A. J. and Thurow, J. W.: A 6000yr varve record of Holocene climate in Saanich Inlet, British Columbia, from digital sediment colour analysis of ODP Leg 169S cores, *Marine Geology*, 174(1), 95–110, doi:https://doi.org/10.1016/S0025-3227(00)00144-4, 2001.
- Nederbragt, A. J. and Thurow, J. W.: Amplitude of ENSO cycles in the Santa Barbara basin, off California, during the past 15,000 years, *Journal of Quaternary Science*, 20, 447–456, 2006.
- 20 North Greenland Ice Core Project Members: High-resolution record of Northern Hemisphere climate extending into the Last Interglacial period, *Nature*, 431, 147-151, 2004
- de Nooijer, L. J., Toyofuku, T., Oguri, K., Nomaki, H. and Kitazato, H.: Intracellular pH distribution in foraminifera determined by the fluorescent probe HPTS: Intracellular pH of foraminifera, *Limnology and Oceanography: Methods*, 6(11), 25 610–618, doi:10.4319/lom.2008.6.610, 2008.
- de Nooijer, L. J., Toyofuku, T. and Kitazato, H.: Foraminifera promote calcification by elevating their intracellular pH, *Proceedings of the National Academy of Sciences*, 106(36), 15374–15378, doi:10.1073/pnas.0904306106, 2009.
- Nürnberg, D., Bijma, J. and Hemleben, C.: Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures, *Geochimica et Cosmochimica Acta*, 60(5), 803–814, 1996.
- 30 Olson, P., Reynolds, E., Hinnov, L. and Goswami, A.: Variation of ocean sediment thickness with crustal age, *Geochem. Geophys. Geosy.*, 17, 1349–1369, 2016.
- O’Neil, J. R., Clayton, R. N. and Mayeda, T. K.: Oxygen Isotope Fractionation in Divalent Metal Carbonates, *The Journal of Chemical Physics*, 51(12), 5547–5558, 1969.
- Parker, F. L.: Eastern Mediterranean foraminifera, *Rep. Swedish Deep Sea Exped.*, 8(4), 219–283, 1958.
- 35 Parker, F.L. and Berger, W. H.: Faunal and solution patterns of planktonic Foraminifera in surface sediments of the South Pacific, *Deep-Sea Research*, 18, 73-107, 1971

- Patterson, D. T., Prokoph, A., Wright, C., Chang, A. S., Thomson, R. E. and Ware, D. M.: Holocene Solar Variability and Pelagic Fish Productivity in the NE Pacific, *Palaeontologia Electronica*, 7(4), 17, 2004.
- Peeters, F. J. C., Acheson, R., Brummer, G. J. A., de Ruijter, W. P. M., Schneider, R., Ganssen, G. M., Ufkes, E. and D., K.: Vigorous exchange between the Indian and Atlantic oceans at the end of the past five glacial periods, *Nature*, 430, 661–665, 2004.
- Pena, L. D., Cacho, I., Ferretti, P. and Hall, M. A.: El Niño–Southern Oscillation–like variability during glacial terminations and interlatitudinal teleconnections, *Paleoceanography*, 23(3), doi:10.1029/2008PA001620, 2008.
- Peng, T.-H., Broecker, W. S. and Berger, W. H.: Rates of benthic mixing in deep-sea sediment as determined by radioactive tracers, *Quaternary Res.*, 11, 141–149, 1979.
- 10 Phleger, F. B., Parker, F. L. and Pierson, J. F.: North Atlantic Foraminifera, Swedish Deep-Sea Expedition 1947-1948, 7, 3–9, 1953.
- Pisias, N. G. and Mix, A. C.: Aliasing of the geologic record and the search for long-period Milankovitch cycles, *Paleoceanography*, 3(5), 613–619, doi:10.1029/PA003i005p00613, 1988.
- Pisias, N.G., Mix, A.C., 1997. Spatial and temporal oceanographic variability of the eastern equatorial Pacific during the late Pleistocene: evidence from Radiolaria microfossils. *Paleoceanography* 12, 381–393.
- 15 Pracht, H., Metcalfe, B. and Peeters, F. J. C.: Oxygen isotope composition of final chamber of planktic foraminifera provides evidence for vertical migration and depth integrated growth, *Biogeosciences Discussions*, 1–32, doi:10.5194/bg-2018-146, 2018.
- 20 Rasmussen, S.O., M. Bigler, S. Blockley, T. Blunier, B. Buchardt, H. Clausen, I. Cvijanovic, D. Dahl-Jensen, S. Johnsen, H. Fischer, V. Gkinis, M. Guillevic, W. Hoek, J. Lowe, J. Pedro, T. Popp, I. Seierstad, J. Steffensen, A. Svensson, P. Vallelonga, B. Vinther, M. Walker, J.J. Wheatley, M. Winstrup: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy, *Quaternary Science Reviews*, 106, 14-27, 2014
- 25 Raven, J., Caldeira, K., Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C. and Watson, A.: Ocean acidification due to increasing atmospheric carbon dioxide., *The Royal Society*., 2005.
- Reynolds, R. W. and Smith, T. M.: A high-resolution global sea surface temperature climatology, *Journal of Climate*, 8, 1571–1583, 1995.
- Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C. and Wang, W.: An Improved In Situ and Satellite SST Analysis for Climate, *Journal of Climate*, 15(13), 1609–1625, doi:10.1175/1520-0442(2002)015<1609:aiais>2.0.co;2, 2002.
- 30 Rincón-Martínez, D., Steph, S., Lamy, F., Mix, A. and Tiedemann, R.: Tracking the equatorial front in the eastern equatorial Pacific Ocean by the isotopic and faunal composition of planktonic foraminifera, *Marine Micropaleontology*, 79(1), 24–40, doi:10.1016/j.marmicro.2011.01.001, 2011.
- Rink, S., Kühl, M., Bijma, J. and Spero, H. J.: Microsensor studies of photosynthesis and respiration in the symbiotic foraminifer *Orbulina universa*, *Marine Biology*, 131, 583–595, 1998.
- 35

- Roche, D. M., Paillard, D., Caley, T. and Waelbroeck, C.: LGM hosing approach to Heinrich Event 1: results and perspectives from data–model integration using water isotopes, *Quaternary Science Reviews*, 106, 247–261, doi:10.1016/j.quascirev.2014.07.020, 2014.
- 5 Roche, D. M., Waelbroeck, C., Metcalfe, B. and Caley, T.: FAME (v1.0): a simple module to simulate the effect of planktonic foraminifer species-specific habitat on their oxygen isotopic content, *Geoscientific Model Development*, 1–22, doi:10.5194/gmd-2017-251, 2018.
- Rosenthal, Y. and Broccoli, A. J.: In search of Paleo-ENSO, *Science*, 304 (5668), 219-221, 2004
- 10 Roy, T., Lombard, F., Bopp, L. and Gehlen, M.: Projected impacts of climate change and ocean acidification on the global biogeography of planktonic Foraminifera, *Biogeosciences*, 12(10), 2873–2889, doi:10.5194/bg-12-2873-2015, 2015.
- Ruddiman, W. F.: Pleistocene Sedimentation in the Equatorial Atlantic: Stratigraphy and Faunal Paleoclimatology, *Geological Society of America Bulletin*, 82, 283–302, 1971.
- 15 Sadekov, A., Eggins, S. M., De Deckker, P. and Kroon, D.: Uncertainties in seawater thermometry deriving from intratest and intertest Mg/Ca variability in *Globigerinoides ruber*: UNCERTAINTIES Mg/Ca SEAWATER THERMOMETRY, *Paleoceanography*, 23(1), n/a-n/a, doi:10.1029/2007PA001452, 2008.
- Sadekov, A., Eggins, S. M., De Deckker, P., Ninnemann, U., Kuhnt, W. and Bassinot, F.: Surface and subsurface seawater temperature reconstruction using Mg/Ca microanalysis of planktonic foraminifera *Globigerinoides ruber*, *Globigerinoides sacculifer*, and *Pulleniatina obliquiloculata*: SEAWATER TEMPERATURE RECONSTRUCTION, *Paleoceanography*, 24(3), doi:10.1029/2008PA001664, 2009.
- 20 Sadekov, A. Y., Ganeshram, R., Pichevin, L., Berdin, R., McClymont, E., Elderfield, H. and Tudhope, A. W.: Palaeoclimate reconstructions reveal a strong link between El Niño–Southern Oscillation and Tropical Pacific mean state, *Nature Communications*, 4, 2692, doi:10.1038/ncomms3692, 2013.
- 25 Schiebel, R., Barker, Stephen, Lendt, R., Thomas, H. and Bollmann, J.: Planktic foraminiferal dissolution in the twilight zone. *Deep-Sea Research. Part II. Topical studies in oceanography* 54 (5-7), 676-686. 10.1016/j.dsr2.2007.01.009, 2007
- Schott, W.: Zur klimaschichtung der tiefseesedimente im aquatorialen atlantischen ozean, *Geologische Rundschau*, 40(1), 20–31, 1952.
- Schott, W.: Foraminiferenfauun und stratigraphie der tiefsee-sedimente in Nordatlantischen ozean, *Rep. Swedish Deep Sea Exped.*, 7(8), 357–469, 1966.
- 30 van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F.J.C., Biastoch, A., Weijer, W., Turney, C., Paris, C.B., and Zahn R.: Ocean currents generate large footprints in marine palaeoclimate proxies, *Nature Communications*, 6, 6521 doi: 10.1038/ncomms7521, 2015
- 35 Seierstad et al.: Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale $\delta^{18}\text{O}$ gradients with possible Heinrich event imprint, *Quaternary Science Reviews*, 106, 29-46, doi:10.1016/j.quascirev.2014.10.032, 2014
- Skrivanek, A. and Hendy, I. L.: A 500 year climate catch: Pelagic fish scales and paleoproductivity in the Santa Barbara Basin from the Medieval Climate Anomaly to the Little Ice Age (AD 1000–1500), *Quaternary International*, 387, 36–45, doi:https://doi.org/10.1016/j.quaint.2015.07.044, 2015.

- Spero, H. J.: SYMBIOSIS IN THE PLANKTONIC FORAMINIFER, *ORBULINA UNIVERSA* , AND THE ISOLATION OF ITS SYMBIOTIC DINOFLAGELLATE, *GYMNODINIUM BÉII* SP. NOV. ¹, *Journal of Phycology*, 23, 307–317, doi:10.1111/j.1529-8817.1987.tb04139.x, 1987.
- 5 Spero, H. J. and DeNiro, M. J.: The influence of symbiont photosynthesis on the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values of planktonic foraminiferal shell calcite, *Symbiosis*, 4(1–3), 213–228, 1987.
- Spero, H. J. and Lea, D. W.: Experimental determination of stable isotope variability in *Globigerina bulloides*: implications for paleoceanographic reconstructions, *Marine Micropaleontology*, 28(3–4), 231–246, doi:10.1016/0377-8398(96)00003-5, 1996.
- 10 Spero, H. J., Bijma, J., Lea, D. W. and Bemis, B. E.: Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes, *Nature*, 390(6659), 497–500, doi:10.1038/37333, 1997.
- Staines-Urías, F., González-Yajimovich, O. and Beaufort, L.: Reconstruction of past climate variability and ENSO-like fluctuations in the southern Gulf of California (Alfonso Basin) since the last glacial maximum, *Quaternary Research*, 83(3), 488–501, doi:https://doi.org/10.1016/j.yqres.2015.03.007, 2015.
- 15 Steinhardt, J., Cleroux, C, de Nooijer, L.J., Brummer, G.-J.A., Zahn, R., Ganssen, G.M., and Reichert, G.-J.: Reconciling single-chamber Mg / Ca with whole-shell $\delta^{18}\text{O}$ in surface to deep-dwelling planktonic foraminifera from the Mozambique Channel, *Biogeosciences*, 12, 2411-2429, https://doi.org/10.5194/bg-12-2411-2015, 2015
- Stott, L., Poulsen, C., Lund, S. and Thunell, R.: Super ENSO and Global Climate Oscillations at Millennial Time Scales, *Science*, 297(5579), 222, doi:10.1126/science.1071627, 2002.
- 20 Stott, L., Cannariato, K., Thunell, R., Haug, G. H., Koutavas, A. and Lund, S.: Decline of surface temperature and salinity in the western tropical Pacific Ocean in the Holocene epoch, *Nature*, 431(7004), 56–59, doi:10.1038/nature02903, 2004.
- Sverdrup, H.U., Johnson, M.W. and Fleming, R.H.: *The Oceans, Their Physics, Chemistry, and General Biology*. Prentice-Hall, New York, 1942
- 25 Telford, R., Li, C., and Kucera, M.: Mismatch between the depth habitat of planktonic foraminifera and the calibration depth of SST transfer functions may bias reconstructions. *Climate of the Past*, 9, 859-870, https://doi.org/10.5194/cp-9-859-2013, 2013
- 30 Thirumalai, K., Partin, J. W., Jackson, C. S. and Quinn, T. M.: Statistical constraints on El Niño Southern Oscillation reconstructions using individual foraminifera: A sensitivity analysis, *Paleoceanography*, 28(3), 401–412, doi:10.1002/palo.20037, 2013.
- Thunell, R.C., Kier, R.S., and Honjo, S.: Calcite dissolution: an in situ study in the Panama Basin, *Science*, 212 (4495), 659-661, 1981
- 35 Trenberth, K.E. and Otto-Bliesner, B.L.: Toward Integrated Reconstruction of past Climates, *Science*, 300 (5619), 589-591, 2003

- de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W. and Pawlowski, J.: Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces, *Proceedings of the National Academy of Sciences*, 96(6), 2864–2868, doi:10.1073/pnas.96.6.2864, 1999.
- 5 de Vargas, C., Bonzon, M., Rees, N. W., Pawlowski, J. and Zaninetti, L.: A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (dâ€™TMOrbigny), *Marine Micropaleontology*, 45(2), 101–116, 2002.
- Vetter, L., Kozdon, R., Mora, C. I., Eggins, S. M., Valley, J. W., Hönisch, B. and Spero, H. J.: Micron-scale intrashell oxygen isotope variation in cultured planktic foraminifers, *Geochimica et Cosmochimica Acta*, 107, 267–278, doi:10.1016/j.gca.2012.12.046, 2013.
- 10 Wang, J.: The coupling of the Annual Cycle and ENSO over the Tropical Pacific, *Journal of the Atmospheric Sciences*, 51 (8), 1115–1136, [https://doi.org/10.1175/1520-0469\(1994\)051<1115:TCOTAC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<1115:TCOTAC>2.0.CO;2), 1994
- 15 Wang, C., Deser, C., Yu, J.-Y., DiNezio, P. and Clement, A.: El Niño and Southern Oscillation (ENSO): A Review, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, pp. 85–106, Springer Netherlands, Dordrecht., 2017.
- Waterson, A. M., Edgar, K. M., Schmidt, D. N. and Valdes, P. J.: Quantifying the stability of planktic foraminiferal physical niches between the Holocene and Last Glacial Maximum, *Paleoceanography*, doi:10.1002/2016PA002964, 2016.
- 20 Wessel, P.: Sizes and Ages of Seamounts using remote sensing: Implications for Intraplate Volcanism, *Science*, 277, 802–805, 1997
- Wessel P. Lyons S.: Distribution of large Pacific seamounts from Geosat/ERS-1, *Journal of Geophysical Research*, 102, 459–475, 1997
- 25 Weyl, P.K.: Micropaleontology and Ocean Surface Climate, *Science*, 202 (4367), 475–481, 1978
- White, S. M., Ravelo, A. C. and Polissar, P. J.: Dampened El Niño in the Early and Mid-Holocene Due To Insolation-Forced Warming/Deepening of the Thermocline, *Geophysical Research Letters*, 45(1), 316–326, doi:10.1002/2017GL075433, 2018.
- 30 Wilke, I., Peeters, F. J. C., Bickert, T.: The influence of seawater carbonate ion concentration [CO₃²⁻] on the stable isotope composition of planktic foraminifera species *Globorotalia inflata*, *Marine Micropaleontology*, 58(4), 243–258, <https://doi.org/10.1016/j.marmicro.2005.11.005>, 2006
- Williams, D. F.: Late Quarternary fluctuations of the polar front and the subtropical convergence in the southeast Indian Ocean, *Marine Micropaleontology*, 1, 363–375, 1976.
- 35 Williams, D. F. and Johnson, W. C.: Diversity of Recent planktonic foraminifera in the southern Indian Ocean and Late Pleistocene paleotemperatures, *Quarternary Research*, 5, 237–250, 1975.
- Wise S.: Calcite compensation depth. In: *Sedimentology. Encyclopedia of Earth Science*. Springer, Berlin, Heidelberg, doi: 10.1007/3-540-31079-7, 1978

- Wit, J. C., Reichart, G. J. and Ganssen, G. M.: Unmixing of stable isotope signals using single specimen $\delta^{18}\text{O}$ analyses, *Geochemistry, Geophysics, Geosystems*, 14(4), 1312–1320, doi:10.1002/ggge.20101, 2013.
- Wolf-Gladrow, D. A., Bijma, J. and Zeebe, R. E.: Model simulation of the carbonate chemistry in the microenvironment of symbiont bearing foraminifera, *Marine Chemistry*, 64(3), 181–198, doi:10.1016/S0304-4203(98)00074-7, 1999.
- 5 Worzel, J.L.: Extensive deep sea sub-bottom reflections identified as white ash, *PNAS*, 45, 349, 1959
- Wunsch, C.: On sharp spectral lines in the climate record and the millennial peak, *Paleoceanography*, 15(4), 417–424, doi:10.1029/1999PA000468, 2000.
- Wunsch, C. and Gunn, D. E.: A densely sampled core and climate variable aliasing, *Geo-Marine Letters*, 23(1), 64–71, 10 doi:10.1007/s00367-003-0125-2, 2003.
- Wycech, J. B., Kelly, D. C., Kitajima, K., Kozdon, R., Orland, I. J. and Valley, J. W.: Combined effects of gametogenic calcification and dissolution on $\delta^{18}\text{O}$ measurements of the planktic foraminifer *Trilobatus sacculifer*, *Geochem. Geophys. Geosyst.*, doi:10.1029/2018GC007908, 2018.
- Xie, S-P.: On the Genesis of the Equatorial Annual Cycle, *Journal of Climate*, 7, 2008-2013, 1994
- 15 Xie, S-P.: Interaction between the annual and interannual variations in the equatorial Pacific, *Journal of Physical Oceanography*, 25, 1930-1941, 1995
- Žarić, S., Donner, B., Fischer, G., Mulitza, S. and Wefer, G.: Sensitivity of planktic foraminifera to sea surface temperature and export production as derived from sediment trap data, *Marine Micropaleontology*, 55(1–2), 75–105, 20 doi:10.1016/j.marmicro.2005.01.002, 2005.
- Žarić, S., Schulz, M. and Mulitza, S.: Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data, *Biogeosciences*, 3(2), 187–207, doi:10.5194/bg-3-187-2006, 2006.
- Zebiak, S.E. and M.A. Cane: A Model El Nino-Southern Oscillation, *Monthly Weather Review*, **115**, 2262–2278, [https://doi.org/10.1175/15200493\(1987\)115<2262:AMENO>2.0.CO;2](https://doi.org/10.1175/15200493(1987)115<2262:AMENO>2.0.CO;2), 1987
- 25 Zeebe, R. E. and Sanyal, A.: Comparison of two potential strategies of planktonic foraminifera for house building: Mg^{2+} or H^{+} removal?, *Geochimica et Cosmochimica Acta*, 66(7), 1159–1169, doi:10.1016/S0016-7037(01)00852-3, 2002.
- Zeebe, R. E. and Wolf-Gladrow, D. A.: *CO₂ in seawater: Equilibrium, Kinetics, Isotopes*, 1st ed., Elsevier, Amsterdam., 30 2001.
- Zhang, Z., Leduc, G. and Sachs, J. P.: El Niño evolution during the Holocene revealed by a biomarker rain gauge in the Galápagos Islands, *Earth and Planetary Science Letters*, 404, 420–434, doi:https://doi.org/10.1016/j.epsl.2014.07.013, 2014.
- Zhang, S., Li, T., Chang, F., Yu, Z., Xiong, Z., and Wang, H., Correspondence between the ENSO-like state and glacial-interglacial condition during the past 360 kyr, *Chinese Journal of Oceanology and Limnology*, 35 (5), 1018-103, 35 <http://dx.doi.org/10.1007/s00343-017-6082-9>, 2017

Zhu, J., Liu, Z., Brady, E., Otto-Bliesner, B., Zhang, J., Noone, D., Tomas, R., Nusbaumer, J., Wong, T., Jahn, A. and Tabor, C.: Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model, *Geophysical Research Letters*, 44(13), 6984–6992, doi:10.1002/2017GL073406, 2017a.

5 Zhu, Z., Feinberg, J. M., Xie, S., Bourne, M. D., Huang, C., Hu, C. and Cheng, H.: Holocene ENSO-related cyclic storms recorded by magnetic minerals in speleothems of central China, *Proc Natl Acad Sci USA*, 114(5), 852, doi:10.1073/pnas.1610930114, 2017b.

Using a foraminiferal ecology model A proxy modelling approach to test if assess the potential of extracting ENSO signal from tropical Pacific planktonic foraminifera are suitable recorders of ENSO

Brett Metcalfe^{1,2}, Bryan C. Lougheed^{1,3}, Claire Waelbroeck¹ and Didier M. Roche^{1,2}

¹Laboratoire des Sciences du Climat et de l'Environnement, LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

²Earth and Climate Cluster, Department of Earth Sciences, Faculty of Sciences, VU University Amsterdam, de Boelelaan 1085, 1081 HV, Amsterdam, The Netherlands

³Department of Earth Sciences, Uppsala University, Villavägen 16, 75236 Uppsala, Sweden

Correspondence to: Brett Metcalfe (b.metcalfe@vu.nl, brett_metcalfe@outlook.com)

A complete understanding of past El Niño-Southern Oscillation (ENSO) fluctuations is important for the future predictions of regional climate using climate models. One approach to reconstructing past ENSO dynamics uses planktonic foraminifera as recorders of past climate to assess past spatiotemporal changes in upper ocean conditions. In this paper, we utilise a model of planktonic foraminifera populations, *Foraminifera as Modelled Entities* (FAME), to forward model the potential monthly average $\delta^{18}\text{O}_c$ and temperature signal proxy values for *Globigerinoides ruber*, *Globigerinoides sacculifer* and *Neogloboquadrina dutertrei* from input variables covering the period of the instrumental record. We tested whether the modelled foraminifera population $\delta^{18}\text{O}_c$ and T_c associated with El Niño events statistically differ from the values associated with other climate states. For these foraminiferal species provided the assumptions of the model are correct, our results suggest indicate that the values of El Niño events can be differentiated from other climate states using these species. Our model computes the proxy values of foraminifera in the water, suggesting that, in theory, water locations for a large proportionportion of the Tropical Pacific should be suitable for differentiating El Niño events from other climate states. However, in practice it may not be possible to differentiate climate states in the sediment record. Specifically, comparison of our model results with the sedimentological features of the Pacific Ocean shows that a large proportionportion of the hydrographically/ecologically suitable water regions coincide with low sediment accumulation rate at the sea floor and/or regions of sea floor, that lie below the critical threshold water depths for calcite preservation (lysocline and CCD).

1. Introduction

1.1 El Niño-Southern Oscillation (ENSO)

Predictions of short-term, abrupt changes in regional climate are imperative for improving spatiotemporal precision and accuracy when forecasting future climate. Coupled ocean-atmosphere interactions (wind circulation and sea surface

Style Definition	... [22]
Style Definition	... [21]
Style Definition	... [20]
Style Definition	... [19]
Style Definition	... [18]
Style Definition	... [17]
Style Definition	... [16]
Style Definition	... [15]
Style Definition	... [14]
Style Definition	... [13]
Style Definition	... [12]
Style Definition	... [11]
Style Definition	... [10]
Style Definition	... [9]
Style Definition	... [8]
Style Definition	... [7]
Style Definition	... [6]
Style Definition	... [5]
Style Definition	... [4]
Style Definition	... [3]
Style Definition	... [2]
Formatted	... [23]
Formatted	... [24]
Formatted	... [25]
Formatted	... [26]
Formatted	... [27]
Formatted	... [28]
Formatted	... [29]
Formatted	... [30]
Formatted	... [31]
Formatted	... [32]
Formatted	... [33]
Formatted	... [34]
Formatted	... [35]
Formatted	... [36]
Formatted	... [37]
Formatted	... [38]
Formatted	... [39]
Formatted	... [1]

temperature) in the tropical Pacific, collectively known as the El Niño-Southern Oscillation (ENSO) on interannual timescales and the Pacific Decadal Oscillation on decadal timescales, represent global climate's largest source (Wang et al., 2017) of inter-annual climate variability (Figure 1). Due to ENSO's major socio-economic impacts upon pan-Pacific nations, which, depending on the location, can include flooding, drought and fire risk, it is imperative to have an accurate understanding of both past and future behaviour of ENSO (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). The instrumental record of the past century provides important information (that can be translated into the Southern Oscillation Index; SOI), however, detailed oceanographic observations of the components of ENSO (both the El Niño and Southern Oscillation), such as the Tropical Oceans Global Atmosphere (TOGA; 1985-1994) experiment only provide information from the latter half of the twentieth century (Wang et al., 2017). To acquire longer records, researchers must turn to the geological record using various archives that are available from the (pan-)Pacific region, including: corals (Cole and Tudhope, 2017); foraminifera (Ford et al., 2015; Garidel-Thoron et al., 2007; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Ledue et al., 2009; White et al., 2018); stalagmite (Asmerom et al., 2007; Zhu et al., 2017b); fish detritus (Patterson et al., 2004; Skrivaneck and Hendy, 2015); lake (Anderson et al., 2005; Bennett et al., 2001; Benson et al., 2002; Conroy et al., 2008; Enzel and Wells, 1997; Higley et al., 2018; Loubere et al., 2013; Zhang et al., 2014); terrestrial (Barron et al., 2003; Barron and Anderson, 2011; Caramanica et al., 2018; Hendy et al., 2015; Staines-Urias et al., 2015); and sedimentological parameters (Moy et al., 2002) including varves (Du et al., 2018; Nederbragt and Thurow, 2001, 2006) to reconstruct long-term variations in proxies, linked to climate, that may provide clues to ENSO and its impact upon both regional and global climate. An integrated approach combining palaeoclimate proxies (Ford et al., 2015; Garidel-Thoron et al., 2007; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Ledue et al., 2009; White et al., 2018) and computer models (Zhu et al., 2017a) can help shed light on the triggers of past ENSO events, their magnitude and their spatiotemporal distribution. Yet, the simulation of past ENSO using climate models has been fraught with difficulties due to the associated feedbacks of ENSO upon model boundary conditions (e.g., SST, pCO_2) (Ford et al., 2015). One way to deduce the relative impact and importance of various feedbacks and, in turn, reduce model dependent noise in our predictions, is to compare model output with proxy data (Roche et al., 2017; Zhu et al., 2017a) of inter-annual climate variability (Wang et al., 2017). Due to ENSO's major socio-economic impacts upon pan-Pacific nations, which, depending on the location, can include flooding, drought and fire risk, it is imperative to have an accurate understanding of both past and future behaviour of ENSO (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). The instrumental record of the past century provides important information (that can be translated into the Southern Oscillation Index; SOI), however, detailed oceanographic observations of the components of ENSO (both the El Niño and Southern Oscillation), such as the Tropical Oceans Global Atmosphere (TOGA; 1985-1994) experiment only provide information from the latter half of the twentieth century (Wang et al., 2017). To acquire longer records, researchers must turn to the geological record using various archives that are available from the (pan-)Pacific region. An integrated approach combining palaeoclimate proxies (Ford et al., 2015; Garidel-Thoron et al., 2007; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and

Formatted: Font: Times New Roman

Formatted: Footer1

Lynch-Stieglitz, 2003; Leduc et al., 2009; White et al., 2018) and computer models (Zhu et al., 2017a) can help shed light on the triggers of past ENSO events, their magnitude and their spatiotemporal distribution.

Formatted: Font: Times New Roman

1.2 Foraminiferal Proxies

Formatted: Heading 2

Such an approach, however, requires an abundance of reliable spatiotemporal proxy data from the entire Pacific Ocean. Moreover, such proxy reconstructions are subject to several unknowns, uncertainties and biases. For the specific case of foraminifera populations in the water, it particularly arises from the species-specific ecological niche. The mapping of proxy value to climate value can therefore be skewed, a major factor governing the spatiotemporal distribution of a given planktonic foraminiferal species is the presence of an ideal water temperature. Proxies of past ENSO and Pacific SST (Ford et al., 2015; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; Sadekov et al., 2013; White et al., 2018) are based upon the biomineralisation of the calcite, or a polymorph such as verite (Jacob et al., 2017), shells of foraminifera (Emiliani, 1955; Evans et al., 2018; Zeebe and Wolf-Gladrow, 2001). In general, there are three major types of foraminifera-based palaeoceanographic proxies:

(1) those associated with the faunal composition and their abundance within deep-sea sediments that utilises either a qualitative approach (Phleger et al., 1953; Schott, 1952); a weighted average (Berger and Gardner, 1975; Jones, 1964; Lynts and Judd, 1971); a selected species approach (e.g. coiling direction, or warm-water species presence; Erielson et al., 1964; Erielson and Wollin, 1968; Hutson, 1980b; Parker, 1958; Peeters et al., 2004; Ruddiman, 1971; Schott, 1966); a regression analysis (Hecht, 1973; Imbrie and Kipp, 1971; Williams and Johnson, 1975); or, a transfer function (CLIMAP Project Members, 1976; McIntyre et al., 1976; Williams, 1976; Williams and Johnson, 1975) that compares the down-core records with a dataset of 'modern' values and their associated water-column parameters (Hutson, 1977, 1978);

(2) those associated with the stable oxygen isotope composition of a whole shell analysed either individually (Ganssen et al., 2011; Koutavas et al., 2006; Koutavas and Joanides, 2012; Leduc et al., 2009) or pooled (Garidel-Thoron et al., 2007; Koutavas et al., 2002; Stott et al., 2002, 2004), herein $\delta^{18}\text{O}_c$ (c = calcite), which can be used to reconstruct SST and past oxygen isotope values in seawater $\delta^{18}\text{O}_{sw}$ (sw = seawater) when paired with a proxy that can either reconstruct temperature or salinity;

(3) those associated with trace metal geochemistry (e.g., Ford et al., 2015; Sadekov et al., 2013; Stott et al., 2002, 2004; White et al., 2018), more specifically the natural logarithm of the relative concentration of Mg and Ca ($\ln(\text{Mg}/\text{Ca})$), of the shell, based upon the temperature dependent (Elderfield and Ganssen, 2000; Nürnberg et al., 1996) incorporation and substitution of a Mg cation into the calcite lattice (Branson et al., 2013, 2016).

Formatted: Footer1

The interpretation of these proxies, however, is not straightforward, for example, calibration of foraminiferal assemblage based transfer functions with surface temperatures as opposed to a deeper temperature signal may in fact skew the reconstructed temperature (Telford et al., 2013); $\delta^{18}\text{O}_c$ can be affected by species-specific size effects (Feldmeijer et al., 2015; Metcalfe et al., 2015; Pracht et al., 2018), disequilibria or vital effects, which clouds the accurate reconstruction of past SST and $\delta^{18}\text{O}_{sw}$.

5 There is also no simple bijective function between $\delta^{18}\text{O}_c$ and the oceanic variables $\delta^{18}\text{O}_{sw}$ and temperature used in its calculation, with variability in $\delta^{18}\text{O}_{sw}$ limiting the use of $\delta^{18}\text{O}_c$ as a pure temperature proxy. Likewise, researchers have not been able to discount the impacts of the ambient salinity (Allen et al., 2016; Gray et al., 2018; Groeneveld et al., 2008; Kısakürek et al., 2008) and carbonate ion concentration (Allen et al., 2016; Evans et al., 2018; Zeebe and Sanyal, 2002) on the Mg/Ca content of foraminifera, nor biological effects such as growth banding (Eggins et al., 2003; Hori et al., 2018; Sadokov et al., 2008, 2009; Vetter et al., 2013). Foraminifera are also not passive recorders of environmental conditions such as SST, in that the very ambient environment that researchers wish to reconstruct also modifies the foraminiferal population as well (Mix, 1987; Mulitza et al., 1998). Sensitivity to the variable being reconstructed may increase or decrease the relative contribution of individual events through modulation of the flux to the seafloor, increasing or decreasing the chance of sampling such occurrences (Mix, 1987; Mulitza et al., 1998). Culture experiments have identified temperature (Lombard et al., 2009, 2011), light (Bé et al., 1982; Bé and Spero, 1981; Lombard et al., 2010; Rink et al., 1998; Spero, 1987; Wolf Gladrow et al., 1999), carbonate ion concentration [CO_3^{2-}] (Bijma et al., 2002; Lombard et al., 2010) and ontogenetic changes (Hamilton et al., 2008; Wycech et al., 2018) as variables that drive, alter or induce changes in foraminiferal growth.

Computation of the influence of biological and vital effects upon physiochemical proxies, such as those based on foraminifera should be a fundamental consideration for any accurate data model comparison. Recent attempts at circumnavigating proxy related problems have employed isotope-enabled models (Caley et al., 2014; Roche et al., 2014; Zhu et al., 2017a), proxy models (Dolman and Laepple, 2018; Jonkers and Kučera, 2017; Roche et al., 2018) or uncertainty analysis (Thirumalai et al., 2013; Fraass and Lowery, 2017; Dolman and Laepple, 2018) to predict both the potential $\delta^{18}\text{O}_c$ values in foraminifera and/or the probability of detection of a climatic event. The use of ecophysiological models (Kageyama et al., 2013; Lombard et al., 2009, 2011) can help circumvent some of the problems associated with a purely mathematical approximation (e.g., Caley et al., 2014) of the translation of an ambient signal into a palaeoclimate proxy. They are not limited to foraminifera and can provide an important way to test whether proxies used for palaeoclimate reconstructions are suitable for the given research question. Several studies have investigated the response of planktonic foraminifera from core material or computed pseudo foraminiferal distributions, their proxy values, and the resultant (likely) distribution of these proxy values with respect to ENSO (e.g., Ledue et al., 2009; Thirumalai et al., 2013; Ford et al., 2015; Zhu et al., 2017).

30 4.2 The simulation of past ENSO using climate models has been fraught with difficulties due to ENSO's integration into the climate system and the associated feedbacks of ENSO upon model boundary conditions (e.g., SST, $p\text{CO}_2$) (Ford et al., 2015). One way to deduce the relative impact and importance of various feedbacks and, in turn, reduce model-dependent noise in our predictions, is to compare model output with proxy data such as foraminifera. Such an approach, however, requires an abundance of reliable spatiotemporal proxy data for the entire Pacific Ocean. The reliability of proxy reconstructions are

themselves subject to several unknowns, uncertainties and biases, for instance culture experiments have identified temperature (Lombard et al., 2009, 2011), light (Bé et al., 1982; Bé and Spero, 1981; Lombard et al., 2010; Rink et al., 1998; Spero, 1987; Wolf-Gladrow et al., 1999), carbonate ion concentration ($[CO_3^{2-}]$) (Bijma et al., 2002; Lombard et al., 2010) and ontogenetic changes (Hamilton et al., 2008; Wycech et al., 2018) as variables that drive, alter or induce changes in foraminiferal growth. These variables are important as foraminifera are not passive recorders of environmental conditions such as SST, in that the very ambient environment that researchers wish to reconstruct can modify the foraminiferal population (Mix, 1987; Mulitza et al., 1998). Sensitivity to the variable being reconstructed may increase or decrease the relative contribution of individual ENSO events, due to modulation of the flux to the seafloor, increasing or decreasing the chance of sampling such occurrences, etc. (Mix, 1987; Mulitza et al., 1998). Computation of the influence of biological and vital effects upon physiochemical proxies, such as those based on foraminifera should be a fundamental consideration for any accurate data-model comparison. Recent attempts at circumnavigating proxy related problems have employed isotope-enabled models (Caley et al., 2014; Roche et al., 2014; Zhu et al., 2017a), proxy system models (Evans et al., 2013; Dolman and Laepple, 2018; Jonkers and Kučera, 2017; Roche et al., 2018) or uncertainty analysis (Thirumalai et al., 2013; Fraass and Lowery, 2017; Dolman and Laepple, 2018) to predict both the potential $\delta^{18}O_c$ values in foraminifera and/or the probability of detection of a climatic event. The use of ecophysiological models (Kageyama et al., 2013; Lombard et al., 2009, 2011) can help circumvent some of the problems associated with a purely mathematical approximation (e.g., Caley et al., 2014) of the translation of an ambient signal into a palaeoclimate proxy. They are not limited to foraminifera and can provide an important way to test whether proxies used for palaeoclimate reconstructions are suitable for the given research question. Several studies have investigated the response of planktonic foraminifera from core material or computed pseudo foraminiferal distributions, their proxy values, and the resultant (likely) distribution of these proxy values with respect to ENSO (e.g., Leduc et al., 2009; Thirumalai et al., 2013; Ford et al., 2015; Zhu et al., 2017).

1.3 Aims and Objectives

Here, we investigate whether living planktonic foraminifera can be theoretically used in ENSO reconstructions, differing from previous research (e.g., Thirumalai et al., 2013) by using a foraminiferal growth model, *Foraminifera as modelled entities* (FAME; Roche et al., 2018), to tackle the dynamic seasonal and depth habitat of planktonic foraminifera (Wilke et al., 2006; Steinhardt et al., 2015; Mix, 1987; Mulitza et al., 1998). To be a ~~valid~~^{useful} proxy for the reconstruction of ENSO, the ~~resulting~~ proxy values of populations of planktonic foraminifera associated to different climatic states (i.e., El NiñoNiño, Neutral, La NiñaNiña) should be significantly different from one another. In order to test our research question, 'are the distributions of proxy values associated with El Niño months statistically different from distributions of proxy values associated with neutral or La Niña months?', our methodology follows a forward modelling approach in which the computed values of the temperature recorded by calcite (T_c - a pseudo temperature aimed at mimicking Mg/Ca ~~albeit one uninfluenced by secondary factors~~) and $\delta^{18}O_c$ are assigned to one of these climatological states. This forward modelling approach does not pre-suppose foraminifera can record ENSO variability (~~(i.e., it asks 'Can we detect?')~~ i.e., what), which is done when inverting the core top pooled $\delta^{18}O$

Formatted: Heading 21

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

or individual foraminiferal $\delta^{18}\text{O}$ distributions and using measured values to infer changes in ENSO ('How could we detect?'). Whilst we are principally interested in understanding whether living foraminifera can theoretically reconstruct ENSO (Section 4 and 5), comparison with data requires further analysis. A secondary objective is to compare the output of this approach with secondary factors that further modulate the climatic signal through post-mortem processes. If the foraminifera modulate the original climate signal, then preservation selectively filters which specimens are conserved and bioturbation acts to reorder, thus scrambling the stratigraphic order in which they are recorded by the sediment depth domain, such that the stratigraphic order is no longer directly equivalent to the time domain. Once the sediment is recovered, the researcher acts as a final filter, which is in essence a random picking process (Section 6). We identify regions in the Pacific Ocean where the sedimentation rate (Berger, 1970a, 1971; Boltovskoy, 1994; Loughheed et al., 2018; Olson et al., 2016) may be too low or the water depth (Berger, 1967, 1970b; Boltovskoy, 1966; Loughheed et al., 2018) too deep (causing dissolution of carbonate sediments) thus potentially preventing the capture and preservation of the foraminiferal signal. (Section 7). To aid the reader, only the general methodology is outlined in section 2, with the individual methodologies of each objective (referred to as Experiments 1 to 5) defined in each subsequent section (sections 3 to 7).

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

2. Methods General Methodology

Formatted: Heading 11

2.1 Input variables (Temperature; Salinity and; $\delta^{18}\text{O}_{\text{sw}}$ and $\delta^{18}\text{O}_{\text{eq}}$)

For input variables, temperature and salinity of the ocean reanalysis data product (Universiteit Hamburg, DE) ORA-S4 (Balmaseda et al., 2013) ORAS4 (Balmaseda et al., 2013) were extracted at one-degree resolution for the tropical Pacific (-20°S to 20°N and 120°E to -70°W), with each single grid cell comprised of data for 42 depth intervals (5 – 5300 m water depth) and 696 months (January 1958 – December 2015). For computation of the oxygen isotope of seawater ($\delta^{18}\text{O}_{\text{sw}}$), a global 1-degree grid was generated, and each grid cell was classified as belonging to one of 27 distinct ocean regions, as defined by either societal and scientific agencies, for identifying regional $\delta^{18}\text{O}_{\text{sw}}$ – salinity relationships (LeGrande and Schmidt, 2006). (LeGrande and Schmidt, 2006). Using the $\delta^{18}\text{O}_{\text{sw}}$ database of LeGrande and Schmidt (2006) a regional $\delta^{18}\text{O}_{\text{sw}}$ – salinity relationship was defined, of which the salinity is the salinity measured directly at the isotope sample collection point (included within the database). Two matrices were computed; one giving values of the slope (m) and the other of intercept (c) of the resultant linear regression equations, and these were used as look-up tables to define the monthly $\delta^{18}\text{O}_{\text{sw}}$ from the monthly salinity Ocean reanalysis product ORAS-S4 (Balmaseda et al., 2013), which was used for the calculation of $\delta^{18}\text{O}_{\text{eq}}$, i.e. ORAS4 (Balmaseda et al., 2013), which was used for the calculation of $\delta^{18}\text{O}_{\text{eq}}$, i.e. the expected $\delta^{18}\text{O}$ for foraminiferal calcite formed at a certain temperature (Kim and O'Neil, 1997). (Kim and O'Neil, 1997). The $\delta^{18}\text{O}_{\text{eq}}$ is calculated from a rearranged form of the following temperature equation:

$$T = T_0 - b \cdot (\delta^{18}\text{O}_e - \delta^{18}\text{O}_{\text{sw}}) \cdot (\delta^{18}\text{O}_c - \delta^{18}\text{O}_{\text{sw}}) + a \cdot (\delta^{18}\text{O}_e - \delta^{18}\text{O}_{\text{sw}})^2 \cdot (\delta^{18}\text{O}_c - \delta^{18}\text{O}_{\text{sw}})^2, (1)$$

Specifically, we used the quadratic approximation (Bemis et al., 1998) of Kim and O'Neil (1997), where $T_0 = 16.1$, $a = 0.09$, $b = -4.64$ and converted from V-SMOW to V-PDB using a constant of 0.27 ‰ (Hut, 1987; Roche et al., 2017):

Formatted: Footer1

$\Delta = -b^2$ Specifically, we used the quadratic approximation (Bemis et al., 1998) of Kim and O'Neil (1997), where $T_0 = 16.1$, $a = 0.09$, $b = 4.64$ and converted from V-SMOW to V-PDB using a constant of -0.27 ‰ (Hut, 1987; Roche et al., 2017):

$$\Delta = b^2 - 4a \cdot (T_{\theta} - T_{sw}) \cdot (T_0 - T_{sw}), \quad (2)$$

$$\delta^{18}O_{c,eq} = \frac{-b - \sqrt{\Delta}}{2a} + \delta^{18}O_{sw} - 0.27, \quad (3)$$

5 The difference between the constant of Hut (1987) and the dynamic value (of Brand et al., 2014) is minor, not used.

2.2 Climate classification

Pan-Pacific meteorological agencies differ in their definition of an El Niño (An and Bong, 2016; 2018), with each country's definition reflecting socio-economic factors. Therefore, for simplicity we use the Oceanic Niño Index (ONI), based upon the Niño 3.4 region ($5^{\circ}N$ to $-5^{\circ}S$, $170^{\circ}W$ to $120^{\circ}W$; Supplementary Figure 1) because of the region's importance for interactions between ocean and atmosphere which is a 3-month running mean of SST anomalies in ERSST.v5 (Huang et al., 2017). We utilise a threshold of $\gamma \geq +0.5^{\circ}C$ (where γ is the value of ONI) as a proxy for El Niño, $-0.5^{\circ}C \leq \gamma < +0.5^{\circ}C$ for neutral climate conditions and $\gamma < -0.5^{\circ}C$ for a La Niña in the Oceanic Niño Index. Many meteorological agencies consider that five consecutive months of $\gamma \geq +0.5^{\circ}C$ must occur for the classification of an El Niño event. However, here the only difference is that we consider that any single month falling within our threshold values as representative of El Niño, neutral or La Niña conditions (grey bars in Supplementary Figure 1). This simplification reflects the lifecycle of planktonic foraminifera (~4 weeks) seeing that the population at time step t does not record what happened at $t-1$ or what will happen at $t+1$. As we are producing the mean population growth weighted $\delta^{18}O$ values, the periods when the ONI threshold is exceeded but an El Niño or La Niña event does not occur (i.e., an 'almost' El Niño or 'almost' La Niña) would be indistinguishable from the build-up and subsequent climb-down of actual El Niño and La Niña events when the foraminiferal values are pooled in the sediment. Therefore, these 'almost' El Niño or 'almost' La Niña (months that exceed the threshold) are placed within their respective climatological pools as El Niño or La Niña.

Each time-step for the entirety of the Pacific was classified as one of three climate states (El Niño; Neutral; and La Niña) and the corresponding values at each timestep binned into their respective categories for each grid-point. The binned values are either the input data (Section 3: Experiment 1) or the $\delta^{18}O_c$ and T_c produced by FAME (Section 4: Experiment 2). An Epanechnikov-kernel distribution was first fitted to the binned monthly output of a single climate state (using the fit distribution function `fitdist` of MatLab), the bandwidth varies between grid-points to provide for an optimal kernel distribution (applying the 'default' option of the function in MatLab). The use of a nonparametric representation (i.e., the kernel distribution) to fit the data, as opposed to other types of distribution (e.g., gaussian), represents a trade-off between keeping as many parameters constant; mimicking the underlying dataset for a large number of grid points and avoiding making too many assumptions regarding the structure of the underlying data. The conversion of the data from dataset to distribution may induce some small error by: rounding to whole integers; the use of a $\delta^{18}O_{mid-point}$ which gives an error associated with the bin size (± 0.05 ‰) that

is symmetrical close to the distributions measures of central tendency but asymmetrical at the sides; and finally, the associated rounding error at the bin edges within a histogram ($\pm 0.005\%$). Subsequently, the shape of any two desired distributions can be compared for statistically significant (dis)similarity using an Anderson-Darling test (1954). For each test, comparison is made between all the values of one climatological state and all the values of another climatological state.

5 **3. Experiment 1: Input Parameters**

3.1. Objective

The resultant values produced by FAME are a modulation of the original input climate signal, therefore it is important to determine to what extent our model has altered the signal and if interpretations we garner from FAME depend upon the models growth rates values (Roche et al., 2018). In Experiment 1 we use a basin-wide statistical test to examine whether the temperature or $\delta^{18}\text{O}_{\text{eq}}$ values used as input in FAME for a given El Niño population and a given non-El Niño ('Neutral conditions') population can be expected to be significantly different at any given specific location. Where the two populations exhibit significantly different distributions, ENSO events can potentially be detected by paleoceanographers. However, where the populations do not exhibit significantly different values, then the location represents a poor choice to study ENSO dynamics.

3.2. Methodology (Temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$)

The input datasets of temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$ underwent the following statistical test (Figure 1): for each grid-point and for every timestep, values were extracted from fixed depths of 5, 149 and 235 m (Supplementary Figure 2). These selected values from discrete-depth intervals were placed into their climatological classifications, and the resultant climatic distributions compared with one another using an Anderson-Darling test in order to compare the (dis)similarity of the resultant climatic distributions. Unlike FAME, which integrates over several depth levels using the computed growth rate, the test of the input datasets was with fixed depths without any growth rate weighting, in order to observe the implications of FAME's dynamic depth habitat. The threshold error (*i.e.*, the difference between the means of each distribution) are for temperature $0.5\text{ }^{\circ}\text{C}$ (Figure 1A) and for $\delta^{18}\text{O}_{\text{eq}}$ 0.10% (Figure 1B), these errors should be viewed a guide rather than an implicit rejection of a site.

25 **3.3 Results and Discussion**

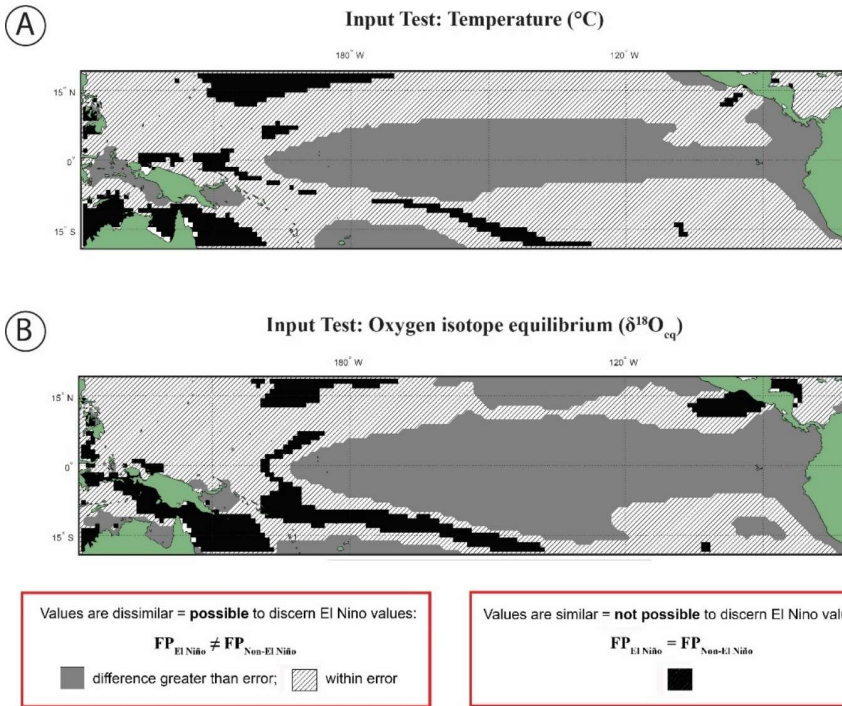


Figure 1 Anderson-Darling Results for Input datasets of Temperature and Equilibrium $\delta^{18}O$ ($\delta^{18}O_{eq}$). Results of the test in which input variables underwent the same statistical procedure (see section 2.0) as the modelled data for (A) temperature and (B) $\delta^{18}O_{eq}$ values. Here, model input data was extracted for a single depth of ~5 m without any growth weighting applied. Black regions are those grid points in which the null hypothesis (H_0), that the El Niño and Non- El Niño (Neutral) foraminifera populations (FP) are not statistically different ($FP_{El Niño} = FP_{Non-El Niño}$), cannot be rejected. Grey regions represent grid points where the H_1 hypothesis is accepted, therefore the distributions of the foraminiferal population for El Niño and Non- El Niño can be said to be unique ($FP_{El Niño} \neq FP_{Non-El Niño}$). The hatched regions represent areas where the H_1 hypothesis can be accepted, therefore the distributions of the foraminiferal population for El Niño and Non- El Niño can be said to be unique ($FP_{El Niño} \neq FP_{Non-El Niño}$), though the difference between the means of tested distribution are less than (A) 0.5°C or (B) 0.1 ‰. For a comparison with three different fixed depths (5; 149; and 235 m) without any growth weighting applied see Supplementary Figure 2.

The results of the Anderson-Darling test performed on the underlying input dataset (temperature and $\delta^{18}O_{eq}$) for each grid point are presented as either black, grey or hashed. Areas where the population distributions of the two climate states are found to be statistically similar have black grid cells. 2.2 Regions in which the difference between the two populations are larger than

Formatted: Font: Times New Roman

Formatted: Footer1

the potential error, are associated with grey, whereas the regions with differences less than the potential error are represented as hashed regions (Figure 1). The results of this fixed depth, non-FAME, test show that the shallowest depths produce populations that are significantly different both in terms of their mean values and their distributions and are thus suitable water locations for recording ENSO dynamics. In the upper panel of Figure 1, the canonical El Niño 3.4 region is clearly visible at 5 m depth. Though there are marked differences and similarities between the Anderson-Darling results for the other depths of the input data (Supplementary Figure 2).

4. Experiment 2: Foraminifera as modelled entities (FAME)

4.1 Objective

In Experiment 2 we run FAME on our two input datasets (temperature and oxygen isotope equilibrium). Data-model comparison studies suffer from an inability to directly compare like with like due to differences in (i) the units used *i.e.*, most proxies reconstructing temperature do not directly give values of temperature in degrees C or K but in their own proxy units (e.g., per mil ‰; mmol/mol; species abundance or ratio) necessitating a conversion; and (ii) scales in the time-depth domain, *i.e.*, models give a wealth of information (multiple depth layers and high resolution time slices). Foraminifera as modelled entities (FAME) was developed as an attempt to reduce the error associated with data-model comparisons by: (i) generating simulated-proxy time-series from a climatic input (a reanalysis dataset or climate model output) that can be compared with age-depth values down core; and (ii) to reduce the model information for a given time-slice into a manageable and relevant value using an integration that would make sense from a biological point of view (Roche et al., 2018), approximating the depth integrated growth of foraminifera (e.g., Pracht et al., 2019; Wilke et al., 2006; Steinhardt et al., 2015). FAME uses the temperature and $\delta^{18}\text{O}_{\text{eq}}$ profiles at each grid cell to compute a time averaged $\delta^{18}\text{O}_{\text{c}}$ and T_{c} for a given species. Using a basin-wide statistical test, we examine whether the $\delta^{18}\text{O}_{\text{c}}$ values of a given El Niño foraminifera population (FP_{EN}) and a given non-El Niño ('Neutral conditions') foraminifera population (FP_{NEU}) can be expected to be significantly different at any given specific location. Where FP_{EN} and FP_{NEU} exhibit significantly different distributions, ENSO events can potentially be detected by paleoceanographers. In cases where FP_{EN} and FP_{NEU} do not exhibit significantly different values, then the chosen species and/or location represent a poor choice to study ENSO dynamics.

4.2 Methodology

4.2.1 FAME Model

The FAME model utilises the temperature-growth rate equations of Lombard *et al.* (2009) to simulate temperature-derived growth rate (Kageyama et al., 2013; Lombard et al., 2009, 2011), this growth rate is then used as a weight to produce a growth rate-weighted proxy value (Roche et al., 2018). The original Lombard et al. (2009, 2011) equations are based upon a synthesis of culture studies, pooled together irrespective of experimental design or rationale, therefore they can be considered to

Formatted: Heading 11

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

conceptually represent the fundamental niche of a given foraminiferal species, *i.e.* the range in environment that the species can survive. The basic structure of FAME is based upon temperature based Michaelis-Menton kinetics to predict growth rate, described in Lombard et al. (2009), without using the parameters (e.g., light, respiration, food) associated with FORAMCLIM (Lombard et al., 2011). The absence of known values or proxy values for the full set of parameters associated with FORAMCLIM has led us to seek a simplified approach in model parameterisation for FAME (Roche et al., 2018). It is important to note that through reducing the complexity of the problem of modelling foraminifera may lead to some deviation between observed and expected values. Our model assumes that temperature provides the dominant signal to the growth of foraminifera and therefore our results should be seen considering this assumption. Other processes may impact species growth such as mixed layer depth and nutrients.

4.2.2 FAME Species selection

Using the MARGO core top $\delta^{18}\text{O}_c$ database (Waelbroeck et al., 2005), Roche *et al.* (2018) validated and computed the optimum depth habitat (the depth habitat that exhibits the strongest correlation when comparing FAME $\delta^{18}\text{O}_c$ and MARGO $\delta^{18}\text{O}_c$) for each species in the MARGO database. Whilst FAME can compute the growth rate of eight foraminiferal species from culture studies (Lombard et al., 2009, 2011; Roche et al., 2018), the limited number of species available for a global core top comparison necessitated a reduction in the number of species modelled (Roche et al., 2018). Here the output of FAME is further restricted to three species that have been the main focus of foraminifera-based studies that have been used to infer ENSO variability, namely the upper ocean dwelling *Globigerinoides sacculifer* and *Globigerinoides ruber*, as well as the thermocline dwelling *Neogloboquadrina dutertrei* (Ford et al., 2015; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Leduc et al., 2009; Sadekov et al., 2013). We use the 1σ values of the observed (MARGO) minus expected (FAME), as computed by Roche et al. (2018) with the MARGO core top $\delta^{18}\text{O}_c$ database, as the potential error associated with the FAME model. The MARGO database does not include *N. Foraminifera* as modelled entities has been developed as a tool for translating, a climatic input (typically a reanalysis dataset or climate model output) into a (simulated) climatic signal, a signal that aims to approximate the depth integrated growth of foraminifera (e.g., *dutertrei* therefore it is not possible to estimate the FAME – MARGO error as can be done with *G. ruber* and *G. sacculifer* (Roche et al., 2018).

4.2.3 FAME Computation

ORAS4 (Praet et al., 2019; Wilke et al., 2006; Steindhardt et al., 2015). Data model comparison studies suffer from an inability to directly compare like with like so that there are differences in (i) the units used *i.e.*, most proxies reconstructing temperature do not give values of temperature in degrees $^{\circ}\text{C}$ or K but in their own proxy units (e.g., per mil ‰; mmol/mol; species abundance or ratio) necessitating a conversion; and (ii) there is a reduction in scales, *i.e.*, models give a wealth of information (multiple depth layers and high resolution time slices) in the time-depth domain. A number of models and modelling studies exist to determine the foraminiferal responses to present (Fraile et al., 2008, 2009; Kageyama et al., 2013; Kretschmer et al., 2017; Lombard et al., 2009, 2011; Roy et al., 2015; Waterson et al., 2016; Žarić et al., 2005, 2006), past (Fraile et al., 2009;

Formatted: Font color: Custom Color(RGB(0,0,10)), Pattern: Clear (White)

Formatted: Footer1

Kretschmer et al., 2016) and future (Roy et al., 2015) climate scenarios, FAME uses the associated temperature and $\delta^{18}\text{O}_{\text{eq}}$ at each grid cell to compute a time averaged $\delta^{18}\text{O}_c$ and T_c for a given species. FAME was produced as an attempt to reduce the error associated with data-model comparisons by (i) generating simulated proxy time series from model runs that can be compared with age-depth values down core; and (ii) to reduce the model information for a given time slice into a manageable and relevant value using an integration that would make sense on a biological point of view (Roche et al., 2018).

The FAME model utilises the temperature-growth rate equations of Lombard et al. (2009) to simulate temperature derived growth rate (Kageyama et al., 2013; Lombard et al., 2009, 2011), this growth rate is then used as a weighing to produce a growth rate weighted proxy value (Roche et al., 2018). ~~The original Lombard et al. (2009, 2011) equations are based upon a synthesis of culture studies, pooled together irrespective of experimental design or rationale, therefore they can be considered to conceptually represent the fundamental niche of a given foraminiferal species, i.e. the range in environment that the species can survive. The basic structure of FAME is based upon temperature based Michaelis-Menton kinetics to predict growth rate, described in Lombard et al. (2009), without using the parameters (e.g., light, respiration, food) associated with FORAMCLIM (Lombard et al., 2011). The absence of known values or proxy values for the full set of parameters associated with FORAMCLIM has led us to an Occam's Razor favoured approach in model parameterisation for FAME (Roche et al., 2018).~~

Although other processes may also impact species such as mixed layer depth and nutrients these variables for now can be set aside, as temperature provides the dominant signal, it is worth noting that in all probability some variance will arise from these processes and deviation between observed and expected values should consider this.

Using the MARGO core top $\delta^{18}\text{O}_c$ database (MARGO Project Members*, 2009), Roche et al. (2018) validated and computed the optimum depth habitat (the depth habitat that exhibits the strongest correlation when comparing FAME $\delta^{18}\text{O}_c$ and MARGO $\delta^{18}\text{O}_c$) for each species in the MARGO database (MARGO Project Members*, 2009). Whilst, both models, FAME and FORAMCLIM, can compute the growth rate of eight foraminiferal species from culture studies (Kageyama et al., 2013; Lombard et al., 2009, 2011; Roche et al., 2017), the limited number of species available for a global core top comparison led to a reduction in the number of species modelled (Roche et al., 2018). Here the output of FAME is further restricted to three species that have been the main focus of foraminifera based studies that have been used to infer ENSO variability, namely the upper ocean dwelling *Globigerinoides saeculifer* and *Globigerinoides ruber*, as well as the thermocline dwelling *Neoglobobulimina dutertrei* (Ford et al., 2015; Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003; Ledue et al., 2009; Sadekov et al., 2013). ~~The MARGO database does not include *N. dutertrei*, meaning that we concentrate our efforts mainly on *G. ruber* and *G. saeculifer*.~~

In this study, ORA-S4 temperature was used as the input variable, (see section 2), with the growth rate computations artificially constrained to arbitrary values of the upper 60; 100 and 200 m to reflect the presence of photosymbiotic algae in the various foraminiferal species and an extreme value of 400 m. The modelled growth rate was used to compute the monthly depth-weighted oxygen isotope distribution for each species, using the aforementioned computed $\delta^{18}\text{O}_{\text{eq}}$ for a given latitudinal and longitudinal grid point (Supplementary Figure 3). No correction for species specific disequilibria, such as vital effect, was applied to the $\delta^{18}\text{O}_{\text{eq}}$ values.

Formatted: Font: Times New Roman

Formatted: Font color: Custom Color(RGB(0,0,10)), Pattern: Clear (White)

Formatted: Footer1

4.2.4 Similar or dissimilar populations

A comparison, for each species, of FAME's predicted growth-weighted $\delta^{18}\text{O}_c$ and T_c distributions associated with each climate event was done using an Anderson-Darling (AD) test. This statistical test can be used to determine whether or not two distributions can be said to come from the same population. The results of this test are presented in the following way: areas where the population distributions of the two climate states are found to be statistically similar have black grid cells in all panels referring to the Anderson-Darling test results (Figure 2; Supplementary Figures 4-6); areas where the populations distributions of two climate states are found to be statistically distinct are shown in white. For plots including the potential error see Supplementary Figures 4 and 5.

10

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

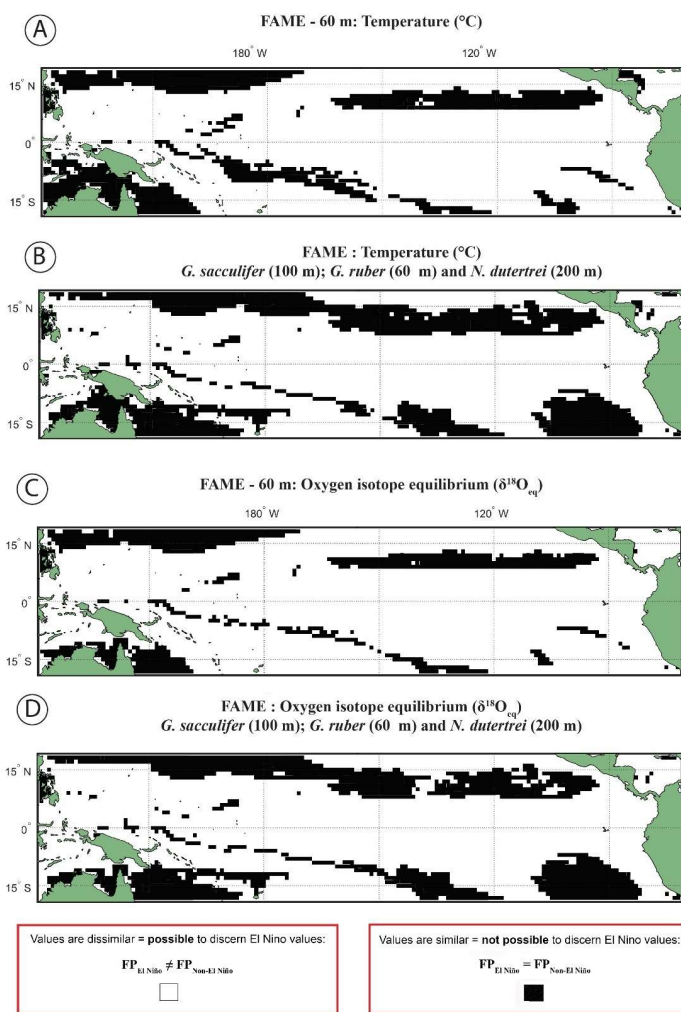


Figure 2 Anderson-Darling results plotted regionally in which species-specific results are overlain. Panels represent water depth locations where dissimilar and similar values for the two climate states for (a-b) FAME- T_c modelled temperature (c-d) FAME- $\delta^{18}O_c$ modelled oxygen isotope values recorded in the calcite shells (T_c) occur. Each panel represents the Anderson-Darling test result, the results for *Globigerinoides sacculifer*, *Globigerinoides ruber* and *N. dutertrei* are overlaid with (A and C) cut-off depth of 60 m and

5

(B and D) species-specific cut-off values. For all panels black areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño is similar to the Non-El Niño, and therefore the distribution between the neutral climate and El Niño cannot be said to be different ($FP_{El\ Niño} = FP_{Non-El\ Niño}$).

5 4.3 Results

Our results show that much of the Pacific Ocean can be considered to have statistically different population between FP_{EN} and FP_{NEU} for both $\delta^{18}O$ and T_c (Figure 2). We consider that the likely cause for such a remarkable result is due to FAME computing a weighted average and, therefore, the lack of a signal found exclusively within the regions demarked in Figure 1 as El Niño regions could represent how the temperature signal is integrated via an extension of the growth rate; growing season and depth habitat of distinct foraminiferal populations. Taking into account the FAME- $\delta^{18}O_c$ error for *G. ruber* and *G. sacculifer*, we have additionally computed regions in which the difference in oxygen isotopes between the two populations is smaller than the aforementioned error (see section 4.2.2) (Hatching in Supplementary Figure 4), i.e. where the mean difference between FP_{EN} and FP_{NEU} is within the error. The hatched regions in Supplementary Figure 4 considerably reduce the areal extent of significant difference between FP_{EN} and FP_{NEU} , with the remaining regions aligning with the El Niño 3.4 region (Supplementary Figure 1). It is important to note that this error relates to the model and in reality, the difference between the climate states could be larger or smaller. No such test was performed on the *N. dutertrei* dataset, because of its absence from the MARGO dataset. To further test the model-driven results and to assess if they are still consistent when the depth limitation is varied, the analysis was rerun with depths of 100, 200 and an extreme value of 400 m (Supplementary Figures 4-6). Whilst it is possible to discern differences between the depths, it is important to note that a large percentage of the tropical Pacific remains accessible to palaeoclimate studies. A shallower depth limitation in the model increases the area for the 'warm' species, suggesting that the influence of a reduced variability in temperature or $\delta^{18}O_{eq}$ with a deeper depth limit causes the differences between FP_{EN} and FP_{NEU} to be reduced. Overlaying the results of the Anderson-Darling test for all three species (Figure 2; Supplementary Figures 4-6) per depth for 60, 100 and 200 m highlights the areas where multi-species comparisons could be made. To account for potential differences in depth habitat we make a combination of shallower depth for *G. ruber* and deeper depths for *G. sacculifer* and *N. dutertrei* (Pracht et al., 2019) in the final panels (-Figure 2B and 2D).

4.4 Discussion

A number of models and modelling studies exist to determine the foraminiferal responses to present (Fraile et al., 2008, 2009; Kageyama et al., 2013; Kretschmer et al., 2017; Lombard et al., 2009, 2011; Roy et al., 2015; Waterson et al., 2016; Žarić et al., 2005, 2006), past (Fraile et al., 2009; Kretschmer et al., 2016) and future (Roy et al., 2015) climate scenarios. Unlike some foraminiferal models, FAME does not include limiting factors such as competition, respiration or predation variables, because no reliable proxy exists for such parameterisation in the geological record, and therefore aspects such as interspecific competition that may limit the niche width of a species are not computed. By identifying the optimum depth habitat, Roche et

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

al. (2018) established the realised niche, *i.e.* the range in environment that the species can be found, for these species for the late Holocene. Unlike some foraminiferal models, FAME does not include limiting factors such as competition, respiration or predation variables as no reliable proxy exists for such parameterisation in the geological record, therefore aspects such as interspecific competition that may limit the niche width of a species are not computed. As these depth constraints (<60 m; <100 m; and <200 m) may induce some variability we opted to include an extreme conservative value of <400 m that grossly exaggerates the potential depth window. It is important to note, however, that as the computation of FAME is based on growth occurring within a temperature window it does not necessarily mean that for a given grid point modelled foraminifera will grow at depths down to 400 m (or whichever cut-off value is used), only that the model in theory can do so (depending if optimal temperature conditions are met to capture the total theoretical niche width.). As the optimised depths computed from the MARGO dataset of Roche *et al.* (2018) are shallower, and upper ocean water is more prone to temperature variability, our approach likely dampens both the modelled $\delta^{18}\text{O}_c$ and T_c . The modelled growth rate was used to compute the monthly depth-weighted oxygen isotope distribution for each species, using the aforementioned computed $\delta^{18}\text{O}_{eq}$ for a given latitudinal and longitudinal grid point (Figures 2 and 3). This was repeated four times, during which the lower depth limit of the growth rate computation was set to 60; 100; 200 and 400 m. No correction for species-specific disequilibria, such as vital effect, was applied to the data. Indeed, the plots testing the input dataset (Section 3; Figure 1) show that our FAME data, in which we allow the possibility for foraminiferal growth down deeper than the depths used in Roche *et al.* (2018), are a conservative estimate.

2.4 Statistical analysis

The tropical Pacific Ocean is divided into four Niño regions based on historical ship tracks, from east to west: Niño 1 and 2 (0° to -10°S, 90°W to 80°W), Niño 3 (5°N to -5°S, 150°W to 90°W), Niño 3.4 (5°N to -5°S, 170°W to 120°W) and Niño 4 (5°N to -5°S, 160°E to 150°W). Pan-Pacific meteorological agencies differ in their definition (An and Bong, 2016, 2018) of an El Niño, with each country's definition reflecting socio-economic factors. Therefore, for simplicity we use the Oceanic Niño Index (ONI), based upon the Niño 3.4 region (because of the region's importance for interactions between ocean and atmosphere) which is a 3-month running mean of SST anomalies in ERSST.v5 (Huang *et al.*, 2017).

Experiment 3: FAME Variance statistics

In Experiment 3 we examine the variance of the $\delta^{18}\text{O}_c$ signal outputted by FAME for *G. sacculifer*. A fundamental problem with proxy records through sampling (Dolman and Laepple, 2018; Piasias and Mix, 1988; Wunsch, 2000; Wunsch and Gunn, 2003) is that they can be confounded by local regional climate, and/or ENSO's teleconnections, that mimic ENSO changes albeit at a different temporal frequency. The results of our Anderson-Darling testing may be unduly influenced by the Pacific decadal variability (PDV), also referred to as the Pacific Decadal Oscillation (PDO) (Pena *et al.*, 2008). In much of the tropical Pacific the ratio of decadal to interannual σSST suggests that they are comparable in magnitude, therefore fluctuations in SST are more obviously apparent outside of the purely canonical regions of ENSO (Wang *et al.*, 2017). It could be that the areas outside of these canonical ENSO regions (Supplementary Figure 1) reflect the PDO (Pena *et al.*, 2008; Wang *et al.*, 2017). The study of ENSO has also focused on whether the variability is entirely in response to ENSO or whether it is dominated by interannual variability (Xie, 1994, 1995; Wang *et al.* 1994, 2010; Thirumalai *et al.*, 2013). Therefore, in order to investigate

Formatted: Font: Times New Roman

Formatted: Footer1

how the signal may respond to a dynamic depth habitat, variance of the climate timeseries at each grid point was computed. As foraminiferal based ENSO studies reliant have used the spread of the individual foraminifera isotope data (either standard deviation $\sigma(\delta^{18}\text{O}_e)$ or its variance) as a measure of the increased variation in SST and, in turn, increased ENSO incidence and/or magnitude (Leduc et al., 2009; Zhu et al., 2017a) this gives us the opportunity to compare our results. For each grid-point both the total variance and the interannual variance ($\sigma^2(\delta^{18}\text{O}_e)$) of the FAME timeseries were computed in order to compare our results with previous studies. For the interannual variance, the computation follows the procedure outlined in Zhu et al. (2017a). We utilise a threshold of $\chi \geq +0.5^\circ\text{C}$ (where χ is the value of ONI) as a proxy for El Niño, $-0.5^\circ\text{C} \leq \chi \leq +0.5^\circ\text{C}$ for neutral climate conditions and $-0.5^\circ\text{C} \leq \chi$ for a La Niña in the Oceanic Niño Index. Many meteorological agencies consider that five consecutive months of $\chi \geq +0.5^\circ\text{C}$ must occur for the classification of an El Niño event. However, here the only difference is that we consider that any single month falling within our threshold values as representative of El Niño, neutral or La Niña conditions (grey bars in Figure 1). This simplification reflects the lifecycle of planktonic foraminifera (~4 weeks) seeing that the population at time step t does not record what happened at $t-1$ or what will happen at $t+1$. The mean monthly climatology is subtracted from the dataset, producing monthly anomalies and a linear trend removed (using the detrend function of MatLab 2019a) – the resultant data was left unfiltered (i.e., Zhu et al., 2017a used a 1-2-1 filter). Comparison between the observed variance of FAME and expected data (Table 1) was done using the nearest grid-cell. However, as foraminifera may drift during their life (van Sebille et al., 2015) a comparison was made with the average variance of a 3 by 3 grid that has the nearest grid-cell to the core location at its centre. A comparison is also made with published iCESM model output for the same core locations (Zhu et al., 2017a).

In a previous study, a Late Holocene sample (~1.5 ka) MD02-2529 (08°12.33'N 84°07.32'W; 1619 m) of *N. dutertrei* individual foraminifera (>250 μm fraction) (Leduc et al., 2009) gave a $\delta^{18}\text{O}$ standard deviation of 0.38 ‰. Here, the full ~60 year time series ($n = 696$) of FAME, gives a standard deviation for all species, of between 0.26 and 0.32 ‰ (<60 m depth); between 0.20 and 0.29 ‰ (<100 m depth); between 0.20 and 0.25 ‰ (<200 m depth); between 0.20 and 0.24 ‰ (<400 m depth) (see Table 1). However, these values can vary if the average of the surrounding grid cells is used (see Table 1). In comparison, the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.4 ‰; Eulerian 100 m: 0.6 ‰; and Lagrangian value of 0.49 ‰. There are three previously analysed samples (Koutavas and Joanides, 2012; Sadekov et al., 2013) located south of core site MD02-2529, these are the Late Holocene (~1.6 ka) samples of V21-30 (01°13'S 89°41'W; 617 m) and (~1.1 ka) V21-29 (01°03'S 89°21'W; 712 m) in which *G. ruber* was measured individually. For these two sites, the measured standard deviation is 0.507 ‰ and 0.510 ‰ for V21-30 and V21-29 respectively (Koutavas and Joanides, 2012). The third core site at a similar location is (~1.6ka) CD38-17P (01°36'04 S 90°25'32W; 2580 m) was not analysed individually, instead replicates of pooled samples of 2 or 3 shells of *N.* As we are producing the mean population growth weighted $\delta^{18}\text{O}$ values, 'almost' El Niño or 'almost' La Niña would be indistinguishable from the build-up and subsequent climb-down of actual El Niño and La Niña events. Therefore, these 'almost' El Niño or 'almost' La Niña are placed within their respective climatological pools as El Niño or La Niña.

Formatted: Font: Times New Roman

Formatted: Footer1

Each time step for the entirety of the Pacific was classified as one of three climate states (El Niño; Neutral; and La Niña), where after the resultant $\delta^{18}\text{O}_e$ and T_e at each timestep produced by FAME for each grid-point were binned into their respective categories. An Epanechnikov kernel distribution was first fitted to the binned monthly output of a single climate state, the bandwidth varies between grid points to provide for an optimal kernel distribution. The use of an Epanechnikov kernel distribution to fit the data, as opposed to other types of distribution, represents a trade-off between keeping as many parameters constant whilst mimicking the underlying dataset for a large number of grid points. The conversion of the data from dataset to distribution may induce some small error induced by: rounding to whole integers; the use of a $\delta^{18}\text{O}_{\text{mid-point}}$ which gives an error associated with the bin size ($\pm 0.05\text{‰}$) that is symmetrical close to the distributions measures of central tendency but asymmetrical at the sides; and finally, the associated rounding error at the bin edges within a histogram ($\pm 0.005\text{‰}$). Subsequently any two desired distributions can be compared for (dis)similarity using an Anderson-Darling test (1954). Here, all values, i.e., the population, associated with a climatological state are compared with the other populations representing the different climatological state, the results plotted here are Neutral climate state vs. El Niño climate state.

2.5 Test of input data (Temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$)

Foraminifera as modelled entities produces a modulated response that seeks to replicate how foraminifera modify the climate signal, several studies have approximated the foraminiferal signal in a different way (e.g., Thirumalai et al., 2013; Zhu et al., 2017a). In order to understand how FAME has altered the signal, and the degree to which the conclusions drawn depend upon the modelled growth rates, the input datasets of the sea water properties (Temperature and calculated $\delta^{18}\text{O}_{\text{eq}}$), underwent a similar statistical test (Figure 4). Unlike FAME, which integrates over several depth levels using the computed growth rate, the test of the input datasets was with fixed depths without any growth rate weighting. These fixed depths are 5, 149 and 235 m, giving a Eulerian view (Zhu et al., 2017a) in which to observe the implications of FAME's dynamic depth habitat. As per the FAME output, each timestep value was placed into its climate state and an Anderson-Darling test performed to compare the (dis)similarity of on the resultant distributions.

2.6 Alternative statistical tests

In order to compare our results with previously published studies using planktonic foraminifera we employed a series of simple statistical tests, mimicking those applied to sediment archives by the palaeoclimate community. A chief parameter that has been employed in previous ENSO proxy work using foraminiferal analysis (more specifically, individual foraminiferal analysis; IFA) is the measure of individual foraminifera downcore standard deviation ($\sigma(\delta^{18}\text{O}_e)$). Increased $\sigma(\delta^{18}\text{O}_e)$ is considered to correlate to increased variation in SST and, in turn, increased ENSO incidence and/or magnitude (Leduc et al., 2009; Zhu et al., 2017a) or increased interannual variance (Thirumalai et al., 2013). The variance ($\sigma^2(\delta^{18}\text{O}_e)$) of the timeseries were computed both as the total variance and as the interannual variance, the latter is computed as outlined in Zhu et al. (2017a). For the interannual variance, the mean monthly climatology is subtracted from the dataset, producing monthly anomalies and a linear trend removed (using the detrend function of MatLab 2019a) — the resultant data was left unfiltered (i.e., Zhu et al., 2017a used a 1-2-1 filter). Four 'picking' experiments were performed, as FAME computes the average value

for a given time step and given the single foraminiferal isotope variance for an equivalent time step (e.g., weeks: Steinhardt et al., 2015) it is more than likely that this computation reduces the real spread in values. Therefore, rather than use the terminology specimen we prefer to use months. Given the complexity in reconstructions of trace metal geochemistry (Elderfield and Ganssen, 2000; Nürnberg et al., 1996): the potential error associated with determining which carbonate phase is first used when foraminifera biomineralise (Jacob et al., 2017); growth band integration; secondary factors (e.g., salinity, carbonate ion) the focus of the picking here has been on the $\delta^{18}\text{O}_e$. Irrespective of which experiment, 60 months were drawn, with replacement, and the number of Monte Carlo iterations is set at 10,000. We assume that the 'picker' is taxonomically well-trained and/or has a procedure in which species can be checked taxonomically post-analysis (e.g. photographing all specimens prior to analysis, Pracht et al. (2019)) and therefore do not include an error that deals with incorrect identification. Although we note that parameterisation of misidentification would be difficult, as it requires understanding of the variability in both standard deviation and absolute values for species co-occurring downcore (Feldmeijer et al., 2015; Metcalfe et al., 2015; 2019). For each run of experiment's (i) to (iii) the drawn months were saved in order to perform (iv):

(i) In Picking *dutertrei* (Sadekov et al., 2013) were made, and these measured values give a standard deviation of 0.28 ‰. The full ~60 year time series (n = 696) of FAME presented here gives a standard deviation for all species, between 0.33 and 0.41 ‰ (<60 m depth); between 0.27 and 0.40 ‰ (<100 m depth); between 0.25 and 0.35 ‰ (<200 m depth); and between 0.25 and 0.34 ‰ (<400 m depth) (see Table 1). Once again, these values can vary if the average of the surrounding grid cells is used (see Table 1). In comparison, the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.53 ‰; Eulerian 100 m: 0.75 ‰; and Lagrangian value of 0.35 ‰.

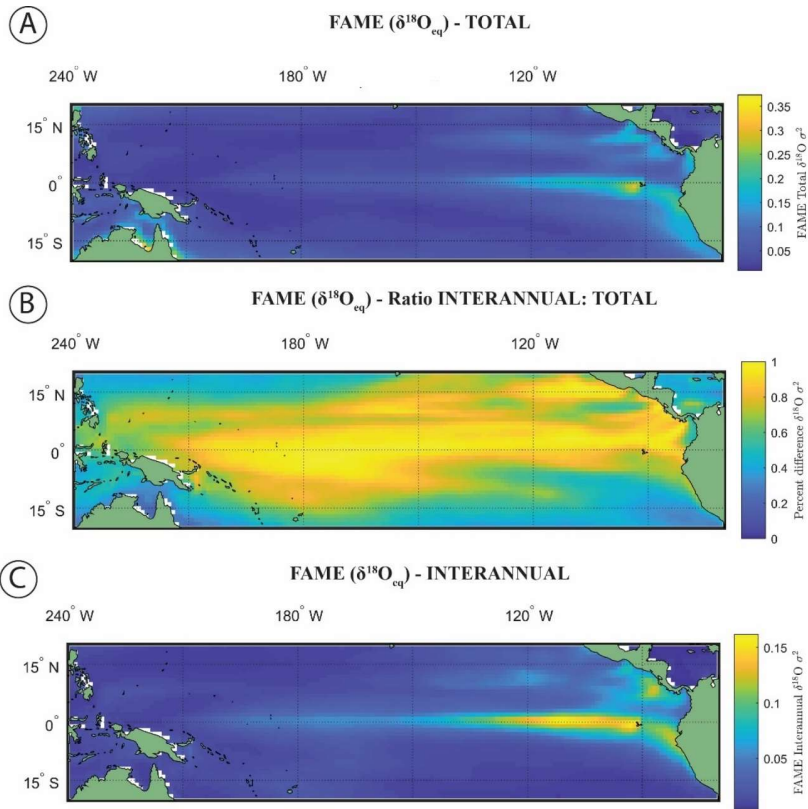


Figure 3 Total variance and Interannual variance. (a) Total variance of *Globigerinoides sacculifer* $\delta^{18}\text{O}_{\text{eq}}$, using FAME- $\delta^{18}\text{O}_{\text{eq}}$ for a cut-off value of 60 m. (b) The ratio of (a) and (c), where (c) is the Interannual variance of the timeseries of (a).

- 5 The use of the variance $\sigma^2(\delta^{18}\text{O}_e)$, or standard deviation $\sigma(\delta^{18}\text{O}_e)$, as an indicator of ENSO is dependent on whether the original climate signal's variance was dominated by interannual variance. Zhu *et al.* (2017) computed the total variance change with and without the annual cycle suggesting that, for some cores the increased assumed ENSO variability at the LGM as deduced by proxy records (Koutavas *et al.*, 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003) may be purely a by-product of the annual cycle or dominated by it. Computing the ratio between the interannual (Figure 3C) and total variance

(Figure 3A) of FAME (Figure 3B; see Table 1) our results have a similarly high ratio of interannual to total variance as iCESM and SODA reanalysis (Carton et al., 2000a, 2000b; Zhu et al., 2017a). Even in regions in the Eastern Equatorial Pacific wherein the ratio reduces, it is still above > 0.5. Although the values of El Niño can be considered significantly different from other climate states (Section 4), our own analysis using the ratio of total to interannual variance also suggests that much of the variance in the simulated foraminiferal signal is dominated by interannual variance. There are differences in the ratio of total to interannual variance between species and in different regions of the tropical Pacific, however, even with a dynamic depth habitat this ratio is still high (Figure 3; Table 1).

6. Experiment 4: FAME Picking Experiment

6.1 Objective

In Experiment 4 we perform a series of picking experiments on our FAME output. One source of potential variation in palaeoceanographic analysis is related to the necessity of picking a finite sample for geochemical analysis. The intention with picking is to produce a robust estimate of the population average without necessarily measuring every individual that constitutes a population, however this can bias the result if either a particular event/seasonal/depth-habitat produces a larger flux of individuals. Several 'picking' experiments were performed to determine the variance between picking iterations.

6.2 Methodology

FAME is not an individual foraminiferal analysis model it instead computes the average value for a given time step, here it is the average of a single month, therefore with respect to terminology what we are in effect picking is individual 'months' rather than individual 'specimens'. Irrespective of which experiment, 60 months were drawn, with replacement, and the number of Monte Carlo iterations is set at 10,000. No attempt to parameterise for misidentification has been done, as although one could assign a random value to a small percentage of the modelled values (conceptually one can argue that misidentification assigns an incorrect value), the assigned value would require knowledge of the values of co-occurring species. Previous work has highlighted the range in and between co-occurring specimens from different species (e.g., Feldmeijer et al., 2015; Metcalfe et al., 2015; 2019). Therefore, the assumption is made that the 'picker' is taxonomically well-trained and/or has a procedure in which species can be checked taxonomically post-analysis, e.g. photographing all specimens prior to analysis (e.g., Pracht et al., 2019). For picking Experiment-1 (Figure 3D), 4A) all grid-points have the same selected months drawn for each iteration of the Monte Carlo were selected and each grid-point was sampled (i.e., there are 10000·60 selected months). This assumes that the picker picks the same months are selected at hypothetical grid point A as they select at grid point B.

Formatted: Font: Times New Roman

Formatted: Footer1

(ii) In Picking Experiment-II (Figure 3E), at each grid-point a Monte Carlo was run (i.e., there are 170·40·10000·60 selected months). This assumes that different months could be selected between hypothetical grid point A and point B.

Formatted: Font: Times New Roman

(iii) In Picking Experiment-III (Figure 3F4C), at each grid-point a Monte Carlo was run using the growth rate weighting for each month (i.e., there are 170·40·10000·60 selected months), this assumes that in periods of higher growth there will be a higher flux of the species and therefore a greater chance of selecting that month. The rationale being that researchers will not only be different months selected pick specimens representing identical time periods between grid point A and point B, but if A and B differ climatologically there may be an over-subscription of ecologically beneficial habitats in one core location compared to the other.

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

(iv) In Picking Experiment-IV (Figures 3G to 3I, Figure 4D and 4E), the second experiment of (ii) was re-run but with the addition of two sources of error: The first error is based upon FAME producing the average value for a given time slice, therefore short-term variability in temperature and/or the spread in the population (i.e., variance in depth of an individual; variance in chamber growth per individual), as evidenced by single foraminiferal analysis of sediment trap samples (e.g., Steinhardt et al., 2015), is potentially lost. Therefore, for each picked month we therefore randomly added between -0.40 and +0.40‰ is added (approximately ±2°C, i.e., for a full range of ~4°C) to the picked month's value (in intervals of 0.02‰, this is approximately ±2°C (i.e., ~4°C)). The second error is the analytical error that an individual measurement will have. Machine measurement error is assumed to lie between -0.12 and 0.12‰ (in intervals of 0.005‰ – the 3rd decimal place is an exaggeration of machine capabilities although it will have repercussions for rounding), the 1σ of within run (as opposed to long-term average) of international stable isotope standards. The intervals of both errors (0.02‰ and 0.005‰) were chosen to give a similar number (n = 41 and 49) of potential randomly selected error for each picked month. Each for this experiment the value assigned to each picked month has their own as a (grid-point specific) randomly selected error value for both of these errors, i.e., each value is the sum of the month picked and their own error. The values for within month variability (Figures 3G, Figure 4D) and machine error (Figure 3H4E) are calculated separately and then combined (Figure 3I4E), as they may have a corresponding or conflicting sign, either 'cancelling' out each other or amplifying the difference.

Formatted: Indent: Left: 0 cm, Right: -0.05 cm

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

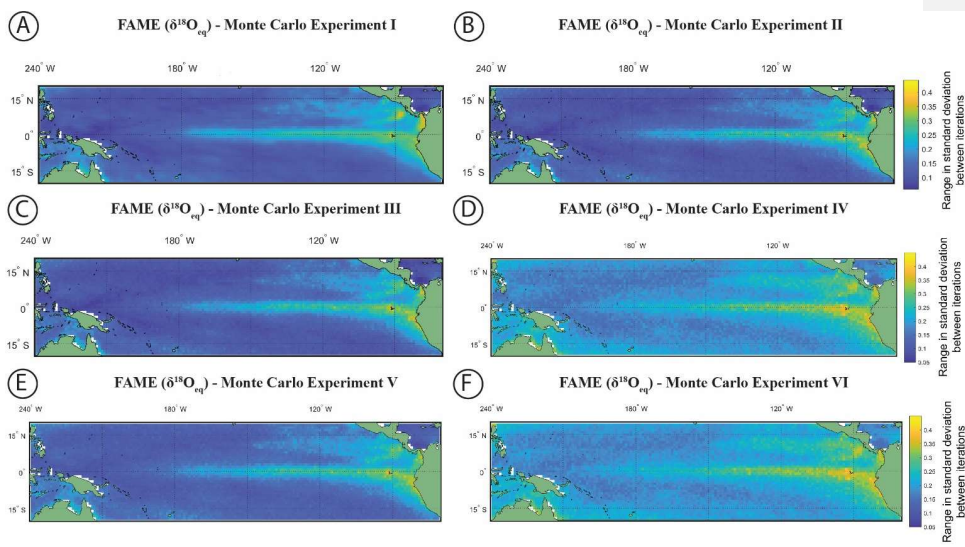


Figure 4 The range in standard deviation of the Monte-Carlo experiments using FAME- $\delta^{18}\text{O}_{\text{eq}}$ *G. sacculifer* with a depth cut-off of 60 m. In (a-f) we plot the range in standard deviation obtained by picking 60 months with replacement with 10,000 iterations, the experiments are as follows: (a) the same months were chosen for all grid-points for each iteration of the Monte-Carlo; (b) each grid-point has its own randomly selected months for each iteration of the Monte-Carlo; (c) the same as (b) but we weight the values by the total amount of growth per month; (d) the months selected for (c) were re-run but a random variability is added to each month (between -0.4 and 0.4 ‰); (e) the months selected for (b) were re-run but a random measurement error is added to each month (between -0.12 and 0.12 ‰); and (f) the months selected for (b) were re-run but the (d) random variability and (e) measurement error were combined.

6.3 Result

The Monte-Carlo experiments (Figure 4A-F) highlight the variation in picking a subset of the months, here 60, from the full timeseries. Given the complexity in reconstructions of trace metal geochemistry (Elderfield and Ganssen, 2000; Nürnberg et al., 1996) the focus of the picking here has been on the $\delta^{18}\text{O}_{\text{eq}}$. The FAME- $\delta^{18}\text{O}_{\text{eq}}$ *G. sacculifer* with a depth cut-off of 60 m is plotted here, the values for each grid point is the range in standard deviation (i.e., the maximum standard deviation minus the minimum standard deviation) between iterations of the Monte-Carlo ($n=10,000$). The range in standard deviations between iterations is plotted instead of the mean of the standard deviations; with increasing n the mean converges toward the sample mean, however as the point of the Monte-Carlo is to generate plausible ‘samples’ it is more important to take into account the range in possible values which would help to establish the potential variability of subsampling. For the most part, regions with high total variance (Figure 4A) also have a larger range in standard deviations between the iterations ‘picked’. It is interesting

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

to note that by changing from the same months picked for each grid-point (Monte-Carlo I: Figure 4A) to varying the months picked between grid-points (Monte-Carlo II: Figure 4B or Monte-Carlo III: Figure 4C) the range goes from 'smooth' to a more noisy dataset. Whilst the values plotted here are not the absolute values (as they are the range in standard deviation for a given grid point for the entire 10,000 iterations), it can be seen that some of the inter-core comparisons could in essence relate to differences in picking, *i.e.* different 'months' picked between grid-points may exacerbate or accentuate differences. Likewise, adding random variability, between -0.4 and 0.4 ‰ (Figure 4D and 4F), may also reduce the differences between areas of high Total variance and low Total variance. Though the values associated with machine error (-0.12 to 0.12 ‰) appear to do little to affect the range (Figure 4E-F). Whilst again the values plotted are not the absolute values, the variability added in an attempt to mimic biological variation of a given time slice increases the range of possible standard deviations in regions with low Total variance (Figure 4D-E). Therefore, understanding the biological variability on shorter timescale (e.g., Steinhardt et al., 2015; Mikis et al., 2019) which, maybe here over exaggerated, may be crucial for understanding discrepancies between cores.

An associated statistical methodology is the graphical summary (as opposed to a numerical summary via a test value) of plotting the quantiles of two probability or the quantiles of sample probability distribution against a theoretical distribution distributions also referred to as a Quantile-Quantile, or Q-Q plot (e.g., Ford et al., 2015; White et al., 2018). A complimentary (*i.e.*, used in association with, not as replacement, Filliben 1975) test metric, the Probability plot correlation coefficient (Filliben, 1975) can be used as a numerical summation of this approach, which bases its rationale on near linearity between the two tested distributions. This graphical technique is not used here for the following reasons, (i) the climatic categories (*i.e.*, El Niño, Neutral, La Niña) imposed upon the data give uneven sized sample distributions requiring an interpolated quantile estimate; and (ii) the large graphical computation required (170-40).

20 2.

7-An approximation of sedimentary archives: Water depth & Sedimentation Rate

Discrete, Experiment 5: Approximation of sediment intervals retrieved from systematically bioturbated deep-sea sediment cores contain foraminifera with ages spanning many centuries (Lougheed et al., 2018; Peng et al., 1979). This is in contrast to other proxies such as corals (Cole and Tudhope, 2017), speleotherms (Chen et al., 2016) and molluscs (Butler et al., 2013; Milano et al., 2017), where distinct time-specific banding is present (true 'time-series' proxies). The ambient signal following translation into a foraminiferal signal within the water is therefore further modulated by several post mortem processes, which include: the latitudinal-longitudinal shift in position of sinking foraminifera—the so-called 'funnel affect' (van Sebille et al., 2015; Deuser et al., 1981); archives dissolution of calcium carbonate either in the water (Schiebel et al., 2007), at the seafloor, or due to pore fluids; and bioturbation. As mentioned, mixing by bioturbation, results in an apparent smoothing of the downcore, discrete-depth multi-specimen signal (Hutson, 1980a; Löwemark, 2007; Löwemark et al., 2005, 2008; Löwemark and Grootes, 2004; Cole and Tudhope, 2017; Mix, 1987), thus leading to the possibility of interpreting single outlying foraminifera values within a specific depth as representing an 'extreme' climate, when they may in fact represent climate from a different time or epoch. This is especially apparent in $\delta^{18}\text{O}_e$ where there is a difference temporally of $\delta^{18}\text{O}_{sw}$ (e.g., the ice volume effect in glacial and interglacial cycles -1.25 ‰) meaning that the same temperature can have radically different $\delta^{18}\text{O}_e$ values, a consequence of this is that a series of high magnitude, but low frequency El Niño events could be disturbed in a discrete-depth record. Therefore, in order to reliably extract short term environmental information from foraminiferal-based proxies, the signal that one is testing or aiming to recover must have a large enough magnitude, be largely unaffected by dissolution (*i.e.*, above the

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Heading 11

Formatted: Footer1

lysocline) so as not to adversely affect the population and the sedimentation rate must be high enough to give sufficient temporal coverage and rule out upwards bioturbation of single foraminifera from significantly different climate periods.

7.1 Objective

- 5 In Experiment 5 we compare our FAME results with bathymetric and sedimentological features of the Tropical Pacific. The preceding analysis has focused upon ~60-year reanalysis data, such a comparable resolution would require a core to have a similar temporal resolution of ~60 years. The hypothetical core should also be above the lysocline to allow for the recovery of a proxy signal equivalent to the original climate signal. At lower sedimentation rates the modification of the original, ambient, climate signal is not limited to just its translation into a foraminiferal proxy signal and the shift in position of sinking
- 10 foraminifera (van Sebille et al., 2015; Deuser et al., 1981) but can also be affected by the dissolution of calcium carbonate either in the water (Schiebel et al., 2007), at the seafloor, or due to pore fluids; and bioturbation. Much of the deep-sea Pacific is both below the lysocline and has a SAR that is very low (e.g., Hays et al., 1969 at $0.96 \pm 0.43 \text{ cm kyr}^{-1}$) although there are regions that satisfy both bathymetry and enhanced sedimentation (e.g., Koutavas and Lynch-Stieglitz, 2003 at $7.20 \pm 2.82 \text{ cm kyr}^{-1}$). In the following section we investigate where in the tropical Pacific it is possible to extract environmental information
- 15 with short frequencies from foraminiferal-based proxies, we consider that a core site must be largely unaffected by dissolution (i.e., above the lysocline) so as not to adversely affect the foraminifer population and the sedimentation rate must be high enough to minimise, as much as possible, the disturbance of the downcore temporal record by bioturbation.

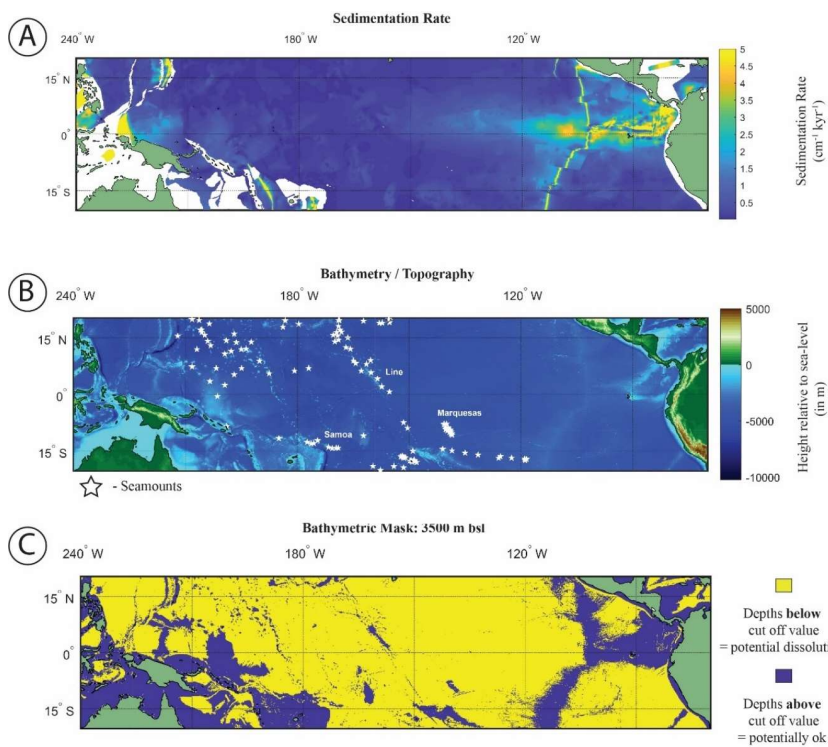


Figure 5 (A) Map of the sedimentation rate and bathymetry of the Tropical Pacific. (A) Inferred sedimentation rate (Olson et al 2016). White regions represent continental shelf. (B) GEBCO map of height relative to 0 m with location of seamounts plotted (white stars). (C) A binary colour map of the GEBCO data, yellow is values below cut-off depth value (3500 m below sea-level (bsl)) and purple above the cut-off depth value. See Supplementary Figure 8 for variation in cut-off values.

7.2 Methodology

7.2.1 Dissolution: Cut-off depth rationale

Whilst the presence of water depths in the ocean lacking calcite-rich sediment was described in the earliest work (e.g., Murray and Renaud, 1891; Sverdrup, 1942), overlaying maps of measured surface sediment carbonate percentage with water depth in the Pacific Ocean led Bramlette (1961) to coin the term ‘compensation depth’ (Wise, 1978). In our first step in consideration of post-mortem signal alteration we focus on dissolution. The lysocline, the depth at which dissolution first becomes apparent (Berger, 1968; 1970), and the Calcite (or Calcium Carbonate) Compensation Depth (This work highlighted the ‘narrow’ depths

of the CCD in the Central Pacific (4-5000 m). Conceptually Berger (1971) placed three levels in the Pacific ocean that were descriptive of the aspects (e.g., chemical, palaeontological and sedimentological) of the calcite budget; the saturation depth, demarking supersaturated from undersaturated; the lysocline, the depth at which dissolution becomes noticeable (Berger 1968, 1971); and compensation depth (Bramlette, 1961), in which supply is compensated through dissolution. Bramlette, 1964 The lysocline and carbonate compensation depth (CCD) vary between the different ocean basins; the modern Atlantic Ocean in which deep water forms has a relatively deep CCD as a by-product of being 'young' well ventilated bottom waters whereas the Pacific Ocean (the final section portion of the global thermohaline circulation conveyor belt) has a shallower CCD.

7.2.2 Dissolution Approximation

Dissolution is approximated by determining if each grid cells depth is above or below the prescribed cut-off value. For much of the equatorial Pacific the lysocline is estimated by a foraminiferal assemblage methodology at ~3800 m (Parker and Berger, 1971), however as the lysocline is where dissolution becomes apparent, ergo it is a sample already visibly degraded, we first set the limit of the water depth mask shallower, at 3500 m bsl. In order to highlight the account for potential variability, two further depths were used as cut-off values: 4000 m bsl and 4500 m bsl these depths represent multiple possible depths under which there is the potential for noticeable dissolution (i.e., lysocline) or complete dissolution (i.e., CCD). The bathymetry of the Pacific was extracted from the General Bathymetric Chart of the Oceans GEBCO 2014 30 arc-second grid (version 20150318, www.gebco.net) between -20°S to 20°N and 120°E to -70°W (Figure 8). Depths of 3500 m below sea-level (bsl), 4000 m bsl and 4500 m bsl are used here as cut-off values, these depths represent multiple possible depths under which there is the potential for noticeable dissolution (i.e., lysocline) or be dissolved www.gebco.net between -20°S to 20°N and 120°E to -70°W (Figure 5B). A compilation of seamounts was also plotted, as these bathymetric features may provide sufficient height to allow preservation of sediment alongside higher sediment accumulation rates (Batiza, 1982; Clouard and Bonneville, 2005; Hillier, 2007; Koppers et al., 2003; Menard, 1964; Wessel and Lyons, 1997). (i.e., CCD). Whilst our intention here is a generalised view to be used as an approximate guide, it is important to note that the Pacific Ocean has the largest proportion globally of >1 km tall seamounts that are smaller than <100 km (Wessel, 1997). Which may have important, relatively shallow water sedimentary sequences, which may also be of sufficient sediment accumulation rate, therefore we supplement the GEBCO bathymetric data with the locations of seamounts. However, whilst there are an estimated 50,000 seamounts in the Pacific that are taller than a km (Menard, 1964; Wessel and Lyons, 1997), only 12,000 have been documented on charts (Batiza, 1982), and approximately 291 have been dated (Koppers et al., 2003; Clouard and Bonneville, 2005; Hillier, 2007). It is these 291, <1% of the estimated seamounts, we have overlain onto the bathymetric data (Figure 8b), although this number is further reduced as we only plot between -20°S and 20°N.

The second step when considering post-mortem signal alteration is the sediment accumulation rate (SAR). We first plot the time-averaged deep-sea SAR (Olson et al., 2016), adapted by Lougheed *et al.* (2018) for the Tropical Pacific (Figure 9). New geochronological tools, such as dual-¹⁴C-^δ¹⁸O measurements on single foraminifera (Lougheed *et al.*, 2018), show that low sedimentation rate cores can have large variances in age between individual foraminifera present within

~~a discrete 1 cm depth interval (Berger and Heath, 1968; Lougheed et al., 2018). In order to model bioturbation, a number of papers~~
7.2.3 Bioturbation

If we factor in the sedimentation rate of the Pacific, which in some regions has been estimated to be lower than 1 cm/ka (Blackman and Somayajulu, 1966; Berger, 1969; Hays et al., 1969; Menard, 1964), then dissolution may become further exacerbated. A secondary factor is bioturbation, systematically bioturbated deep-sea sediment cores can produce discrete sediment intervals with foraminifera that have ages spanning many centuries and/or millennia (Berger and Heath, 1968; Lougheed et al., 2018; Peng et al., 1979). In order to model the effect of bioturbation upon the age distribution of discrete core depths, a number of studies have used a diffusion style approach that reduces the parameters down to sediment mixing intensity accumulation rate (SAR) and sediment mixing depth (herein referred to as bioturbation depth, BD), although this may be an artificial division purely driven by mathematical need rather than biological constraints (Boudreau, 1998). The bioturbation depth BD has been shown to have a global average of 9.8 cm (1 σ : \pm 4.5 cm (1 σ)) that is independent of both water depth and sedimentation rate (Boudreau, 1998), though likely controlled as a result of the energy efficiency of foraging, e.g. deeper burrows may cost more energy to produce than can be offset in extracted food resources, and potential decay in labile food resources with sediment depth. It is not possible to carry out a transient bioturbation model upon the temperature and salinity ocean reanalysis data that we used for FAME, as it only covers half a century of data, whereas thousands of years of input data are required to force a transient bioturbation model.

Following the current available geochronological method (i.e., age-depth method) single specimens that are displaced in depth are assigned the average age of the depth that they were displaced to, which will result in erroneous interpretations of climate variability when analysis such as IFA is applied (Lougheed et al., 2018). To investigate how much temporal signal is integrated into discrete-depth intervals for typical tropical Pacific SAR, we, therefore, utilised (Olson et al., 2016; adapted by Lougheed et al., 2018) the single foraminifera sediment accumulation simulator (SEAMUS, Lougheed, 2019–2020) was utilised to bioturbate, as the input a climate signal (Figure's 9. As it is not possible to carry out a transient bioturbation model with the SAR and BD of the Pacific with only half a century of data (such as the ORAS4 temperature and salinity ocean reanalysis data) a longer highly temporally resolved climate input signal was used, to explore the effect of bioturbation upon a given climate signal. The 0–40,000-year $\delta^{18}O_w$ of NGRIP (North Greenland Ice Core Project Members, 2004; Rasmussen et al., 2014; Seierstad et al., 2014). The ice core time series is an ideal input for a bioturbation simulator, as it represents a highly temporally resolved climate input signal. SEAMUS simulates foraminifera is considered to be a satisfactory replacement signal to simulate a foraminiferal signal in 10-year timesteps. It must be stressed that the use of the NGRIP timeseries here is purely as an input parameter to investigate the effect of bioturbation upon a given an oxygen isotope-based climate signal. It is important to stress that by using NGRIP as an input signal for SEAMUS we are neither implying that tropical Pacific cores should have a signal similar to NGRIP, nor that we are translating the NGRIP signal to the tropical Pacific, or inferring some kind of causal relationship. As we seek to investigate the effect of bioturbation, no attempt has been made to modulate the input signal's absolute values to mimic expected $\delta^{18}O_w$ values and this is why each plot of the synthetic down core time

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

series retains the use of V-SMOW, despite carbonates being required to be V-PDB (Coplen 1995). Keeping all things constant, and varying a

A single parameter was varied whilst all others were kept constant between experiments with SEAMUS, the sediment accumulation rate (SAR) was varied to fixed values of either 1, 2, 5 or 10 cm kyr⁻¹ (that are representative of typical Pacific SAR) and a bioturbation depth (BD). As the oxygen saturation state of the Pacific Ocean bottom waters is above 40 % (Supplementary Figure 9), suggestive that oxygen may not be a limiting factor, values of BD of either 5, 10 or 15 cm were used. These values are based upon the global estimate of BD and its error bounds (Boudreau, 1998). For each experiment, the selected values of SAR and BD were kept constant for the entire SEAMUS model run (i.e., the intensity and magnitude of bioturbation was not varied). In although in reality, SAR and BD may vary temporally depending on local conditions (e.g., food, oxygen). Finally, the FAME results for the three species are overlaid with a water depth mask that highlights whether grid points are above or below 3500 m below sea level (mbsl), to also show seafloor areas under the CCD depth, where carbonate material is not preserved (Berger, 1967, 1970b). A comparison between water depth and time-averaged deep-sea SAR (Olson et al., 2016), adapted by Loughheed et al. (2018) is shown in Figure's 7 and 9. Each experiment was plotted as a histogram of frequency of age of specimen in BD that represent different thicknesses of sediment (5, 10 and 15 cm) and a timeseries using the computed discrete 1 cm depth median age (Figure 9).

7.3 Results

The results of the forward model (Figure 2 and 3) are compared with the input values (Figure 4) discussion

A factor in order to identify regions in which the values are statistically distinct for different climate states (Figure's 5-7). These results are then shown against the water depth (Figure's 7 to 9) and the SAR (Figure's 9-11) for the region. The results utilise Foraminifera as Modelled Entities (FAME; Figures 2, 3, 5, 6, and 7); the original Ocean Reanalysis data with computed $\delta^{18}\text{O}_{\text{eq}}$ (Figure 4); the General Bathymetric Chart of the Oceans (GEBCO; Figures 7 to 9); and the single foraminifera sediment accumulation simulator (SEAMUS; Figures 9 to 11).

3.1 FAME Output: Variance

We compute growth-weighted $\delta^{18}\text{O}_e$ (Figure 5 and 7) and temperature (Figure 6 and 7) distributions for each grid cell in the fifty-eight year simulation using FAME (Roche et al., 2018), constraining the calculation to the Tropical Pacific Ocean (between 20°S to 20°N and 120°E to 70°W). Our model produces 696 individual monthly maps for all three species (Figure 2). While two of the three species (*G. ruber* and *G. sacculifer*) have similar ecologies, they show differences in their resultant $\delta^{18}\text{O}_e$ for the same ocean conditions (Figure 2). A comparison of our computed variance with measured data (Supplementary Table 1) is made, we compare both the value of the nearest grid cell and because of the size of the grid and drift of foraminifera (van Sebille et al., 2015) an average of a 3 by 3 grid in which the nearest grid cell to the core location is in the center. A

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

comparison is made with both the iCESM model output and the core's that match this output (Zhu et al., 2017a). For the Late Holocene sample (-1.5 ka) MD02-2529 (08°12.33'N 84°07.32'W; 1619 m) in which *N. dutertrei* individual foraminifera were analysed from >250 µm (Leduc et al., 2009) giving a calculated standard deviation of measured foraminifera of 0.38 %. Whereas, the full -60 year time series (n = 696) of FAME presented here, gives a standard deviation for all species, at depth cut off 60 m between 0.26 and 0.32 %; at depth cut off 100 m between 0.20 and 0.29 %; at depth cut off 200 m between 0.20 and 0.25 %; and at depth cut off 400 m between 0.20 and 0.24 % (see Table 1). Although these values vary if the average of the surrounding grid cells is used (see Table 1). In comparison the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.4 %; Eulerian 100 m: 0.6 %; and Lagrangian value of 0.49 %. There are three samples (Koutavas and Joanides, 2012; Sadekov et al., 2013) located south of core site MD02-2529, these are the Late Holocene (-1.6 ka) samples of V21-30 (01°13'S 89°41'W; 617 m) and (-1.1 ka) V21-29 (01°03'S 89°21'W; 712 m) in which *G. ruber* was measured individually (Sadekov et al., 2013). For these two sites, the measured standard deviation is 0.507 % and 0.510 % for V21-30 and V21-29 respectively (Koutavas and Joanides, 2012). The third core site at a similar location is (-1.6ka) CD38-17P (01°36'04 S 90°25'32W; 2580 m) was not analysed individually, instead replicates of pooled samples of 2 or 3 shells of *N. dutertrei* (Sadekov et al., 2013) were made these measured values give a standard deviation of 0.28 %. The full -60 year time series (n = 696) of FAME presented here, gives a standard deviation for all species, at depth cut off 60 m between 0.33 and 0.41 %; at depth cut off 100 m between 0.27 and 0.40 %; at depth cut off 200 m between 0.25 and 0.35 %; and at depth cut off 400 m between 0.25 and 0.34 % (see Table 1). Although these values vary if the average of the surrounding grid cells is used (see Table 1). In comparison the iCESM results have the following standard deviation values, for a Eulerian (fixed) depth of 50 m: 0.53 %; Eulerian 100 m: 0.75 %; and Lagrangian value of 0.35 %.

The study of ENSO has focused on whether the variability is entirely in response to ENSO or whether it is dominated by interannual variability (Xie, 1994, 1995; Wang et al 1994, 2010), here the interannual (Figure 3C) and total variance (Figure 3A) was computed and a ratio between the two calculated (Figure 3B; see Supplementary Table 1). Like the same analysis of interannual and total variance computed for iCESM and SODA reanalysis (Carton et al., 2000), outlined in Zhu et al. (2017a), there is also high ratio of interannual to total variance in our computed FAME dataset (Figure 3B). Although there are regions in the Eastern Equatorial Pacific wherein this ratio reduces. Despite this reduction, the ratio between total and interannual variance is still above >0.5.

The Monte Carlo experiments (Figure 3D-I) highlight the variation in picking a subset of the months, here 60, from the full timeseries. The FAME $\delta^{18}O_{org}$ *G. sacculifer* with a depth cut off of 60 m is plotted here, the values for each grid point is the range in standard deviation (i.e., the maximum standard deviation minus the minimum standard deviation) between iterations of the Monte Carlo (n = 10,000). The range in standard deviations between iterations is plotted instead of the mean of the standard deviations; with increasing *n* the mean converges toward the sample mean, however as the point of the Monte Carlo is to generate plausible 'samples' it is more important to take into account the range in possible values which would help to establish the potential variability of subsampling. For the most part, regions with high total variance (Figure 3A) also have a larger range in standard deviations between the iterations 'picked'. It is interesting to note that by changing from the same

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

months picked for each grid point (Monte Carlo I: Figure 3D) to varying the months picked between grid points (Monte Carlo II: Figure 3E or Monte Carlo III: 3F) the range goes from 'smooth' to a more noisy dataset. Whilst the values plotted here are not the absolute values (as they are the range in standard deviation for a given grid point for the entire 10,000 iterations), it can be seen that some of the inter-core comparisons could in essence relate to differences in picking, *i.e.* different 'months' picked between grid points may exacerbate or accentuate differences. Likewise, adding random variability, between -0.4 and 0.4 ‰ (Figure 3G and 3I), may also reduce the differences between areas of high Total variance and low Total variance. Though the values associated with machine error (-0.12 to 0.12 ‰) appear to do little to affect the range (Figure 3H and 3J). Whilst again the values plotted are not the absolute values, the variability added in an attempt to mimic biological variation of a given time slice increases the range of possible standard deviations in regions with low Total variance (Figure 3G and 3J). Therefore, understanding the biological variability on shorter timescale (e.g., Steinhardt et al., 2015; Mikis et al., 2019) which, maybe here over-exaggerated, may be crucial for understanding discrepancies between cores.

3.2 FAME Output: Anderson-Darling test

the post-mortem preservation of the oceanographic Using a basin-wide statistical test, we examine whether the $\delta^{18}\text{O}_c$ values of a given El Niño foraminifera population (FP_{EN}) and a given non-El Niño ('Neutral conditions') foraminifera population (FP_{NEU}) can be expected to be significantly different at any given specific location. Where FP_{EN} and FP_{NEU} exhibit significantly different distributions, ENSO events can potentially be detected by paleoceanographers. In cases where FP_{EN} and FP_{NEU} do not exhibit significantly different values, then the chosen species and/or location represent a poor choice to study ENSO dynamics. Each simulation time step was placed into a climate states: identification of timesteps that represent El Niño (EN), Neutral (NEU), and La Niña climate conditions was done using the Oceanic Niño Index (ONI) derivative (Huang et al., 2017) (Figure 1). Comparison, for each species, FAME's predicted growth-weighted $\delta^{18}\text{O}_c$ and T_e distributions associated with each climate event was done using an Anderson-Darling (AD) test. This statistical test can be used to determine whether or not two distributions can be said to come from the same population. The results of this test are presented in the following way, in which there are four criteria: areas where the population distributions of the two climate states are found to be statistically similar have black grid-cells in all panels referring to the Anderson-Darling test results (Figure's 4-7); the results in which the areas where the populations distributions of two climate states are found to be statistically distinct are shown with two distinct colour schemes depending on whether a computable error can be included (Grey and Hashed) or not (White).

For FAME- $\delta^{18}\text{O}_c$ the results where the populations are dissimilar are either plot as grey and hashed for *G. ruber* and *G. sacculifer* or white for *N. dutertrei* (Figure 5). This is because for these two species (*G. ruber* and *G. sacculifer*) we have the possibility to determine how robust these results are. We use the 1σ values of the observed (FAME) minus expected (MARGO), as computed by Roche et al. (2018) with the MARGO core top $\delta^{18}\text{O}_c$ database, as the potential error associated with the FAME model. Regions in which the difference between the two populations are larger than the potential error are associated with grey, whereas those less than the potential error as hashed regions (Figure 5A and B), these errors should be seen as a guide rather than a rejection of a site. Because the MARGO database does not contain *N. dutertrei* we have given the

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

panels concerning this species a separate colour scheme, black represents grid-cells for which the two populations cannot be said to be statistically different, white grid-cells are those in which the two populations can be said to be statistically different (Figure 5C). As we do not have a similar way to calculate the error for T_e , FAME- T_e results are shown (in Figure 6) with this same binary pattern (*i.e.*, white grid-cells are those in which the two populations can be said to be statistically different and black are those in which the two populations can be said to be similar). To reduce the complexity, the overlay of the species Anderson-Darling results (Figure 7) also uses the binary colour scheme (white or black).

Our results show that much of the Pacific Ocean can be considered to have statistically different population between FP_{EN} and FP_{NEU} for both $\delta^{18}O$ (Figure 5) and T_e (Figure 6). We consider that the likely cause for such a remarkable result is due to FAME computing a weighted average and, therefore, the lack of a signal found exclusively within the regions demarked in Figure 1 as El Niño regions could represent how the temperature signal is integrated via an extension of the growth rate; growing season and depth habitat of distinct in foraminiferal populations. Taking into account the FAME- $\delta^{18}O_e$ error for *G. ruber* and *G. sacculifer*, we have computed regions in which the difference in oxygen isotopes between the two populations ($\Delta\delta^{18}O_e$) compared with the AD test is smaller than the aforementioned error (Hatching in Figure 5), *i.e.* where the mean difference between FP_{EN} and FP_{NEU} is within the error. The hatched regions in Figure 5 considerably reduce the areal extent of significant difference between FP_{EN} and FP_{NEU} , with the remaining regions aligning with the El Niño 3.4 region (Figure 1). It is important to note that this error relates to the model and in reality, the difference between the climate states could be larger or smaller. No such test was performed on the *N. dutertrei* dataset, because of its absence from the MARGO dataset. To further test the model-driven results and to assess if they are still consistent when the depth limitation is varied, the analysis was rerun with depths of 100, 200 and an extreme value of 400 m (Figure 5-7). Whilst it is possible to discern differences between the depths, it is important to note that a large percentage of the tropical Pacific remains accessible to palaeoclimate studies. A shallower depth limitation in the model increases the area for the 'warm' species, suggesting that the influence of a reduced variability in temperature or $\delta^{18}O_{eq}$ with a deeper depth limit causes the differences between FP_{EN} and FP_{NEU} to be reduced. Overlaying the results of the Anderson-Darling test for all three species (Figure 7) per depth for 60, 100 and 200 m highlights the areas where multi-species comparisons could be made. To account for potential differences in depth habitat we make a combination of shallower depth for *G. ruber* and deeper depths for *G. sacculifer* and *N. dutertrei* (Pracht et al., 2019) in the final panels (Figures 7D and 7H).

3.3 Test of input parameters (fixed depth: temperature and $\delta^{18}O_{eq}$)

The model-driven results were assessed with the underlying input dataset (temperature and $\delta^{18}O_{eq}$), these underwent the same statistical test (Figure 4), although with fixed depths of 5 m, 149 m and 235 m (see section 2.5). The results for each grid point are presented as either black, grey or hashed. Areas where the population distributions of the two climate states are found to be statistically similar have black grid-cells. Regions in which the difference between the two populations are larger than the potential error are associated with grey, whereas those less than the potential error as hashed regions. The threshold error (*i.e.*, the difference between the means of each distribution) is for temperature (Figure 4A-C) 0.5 °C and for $\delta^{18}O_{eq}$ (Figure 4D-F)

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

0.10 %; these errors should be seen as a guide rather than a rejection of a site. The results of this fixed depth, non-FAME, test show that the shallowest depths produce populations that are significantly different both in terms of their mean values and their distributions. In the upper panel of Figure 4, the shells is whether the shells can be preserved. Irrespective of the bathymetric cut-off value used for the GEBCO bathymetry data it is evident that much of the canonical El Niño 3.4 region is clearly visible at 5 m depth. Whilst differences exist between Anderson-Darling results for the input data (Figure 4) and the FAME $\delta^{18}\text{O}$ (Figure 5) and T_e (Figure 6), for used in oceanography, as well as a large proportion of the Tropical Pacific, is excluded from suitability as a perspective core site (Figure 5B and 5C). Even in regions where bathymetry may be above the cut-off value dissolution may occur. For instance, in regions of high fertility, such as the Eastern Equatorial Pacific, the lysocline was estimated to be present at ~2800 m (Thunell et al., 1981) or ~3000 m (Berger, 1971; Parker and Berger, 1971). In the EEP region the shallower lysocline is accompanied by an equally shallower CCD (located at ~3600 m) for which the high fertility/primary production is considered responsible for its shoaling, lowering the pH through increased CO_2 (Berger et al., 1976). The correspondence between lysocline depth and CCD depth does not hold true for the entirety of the Pacific, plotting a N-S cross-section from 50°N to 50°S Berger (1971) noted that in the Central Equatorial Pacific, the high fertility region generates a larger zone of dissolution resistant facies even with a shoaled lysocline. -close to the Panama isthmus, there are significant similarities between the plots. These plots also show that our FAME data (Figure 5-7), in which we allow foraminiferal growth down deeper than the depths in Roche et al. (2018), are a conservative estimate and thus are on the low-end (Figure 4), to account for potential discrepancies with depth habitats. In the original paper on depth habitats based upon temperatures derived from $\delta^{18}\text{O}_e$, Emiliani (1954) cautioned that the depth habitats obtained would represent a weighted average of the total population, and while foraminiferal depth habitats are likely to vary spatiotemporally, the average depth habitat is skewed toward the dominant signal (Mix, 1987).

3.4 Water depth and SAR

Our analysis uses reanalysis data for the time period 1958-2015, a hypothetical core that had a comparable resolution would essentially be analogous with a sediment core with a rapid sediment accumulation rate (SAR), representing 50 yr cm^{-1} (or 20 cm kyr^{-1}). Based on our analysis, such a hypothetical core could allow for the possible disentanglement of El Niño-related signals from the climatic signal, but only in a best-case scenario involving minimal bioturbation, which is unlikely in the case of oxygenated waters. Extracting the oxygen saturation (SO_2) state, of the Pacific Ocean bottom waters from the Annual Climatology WOA13 give values that are predominantly >40% (Figure 9B). Oxygen saturation is the concentration of Oxygen in a medium against the maximum that can be dissolved in the same medium. Whilst annual variability may exist, it is unlikely that bioturbation would be prevented by low oxygen. Therefore, using a cut-off value that has been a second factor is the sedimentation rate, using a cut off value that has been previously considered sufficiently high enough to outpace bioturbation (e.g., Koutavas and Lynch-Stieglitz, 2003) of 5 cm kyr^{-1} (Figure 9A), it can be demonstrated that much of the Pacific has an inferred lower sedimentation rate (< 5 cm kyr^{-1} ; Figure 9C) than this cut off value. To test the influence of bioturbation, the bioturbation simulator SEAMUS was run using the NGRIP time series. The results of SEAMUS highlight the potential single

Formatted: Font: Times New Roman

Formatted: Footer1

foraminifera depth displacement that low sedimentation rates can result in (Figure 9). Following the current available geochronological method (i.e., age-depth method) such specimens that are displaced in depth are assigned the average age of the depth that they were displaced to, which could result in erroneous interpretations of climate variability when analysis such as IFA is applied (Lougheed et al., 2018). The results of SEAMUS are plotted both as time series of the bioturbated 'NGRIP' signal (Figure 10) and as histograms of the probability of finding a particular pseudo-foraminifera with a given age in the bioturbation depth (Figure 11). As the bioturbation depth varies between 5, 10 and 15 cm for the different simulations of SAR, the histogram in each panel (in Figure 11) represent different thicknesses of sediment, i.e., for Figure 11 panels a, d, g, and j histograms represent data with a BD thickness of 5 cm. Likewise, the timeseries is plotted with the discrete 1 cm depth median age; the median age of the bioturbation depth (Figure 11) is the reason why each timeseries does not 'start' at 0 age (Keigwin and Guilderson, 2009^{5A}).

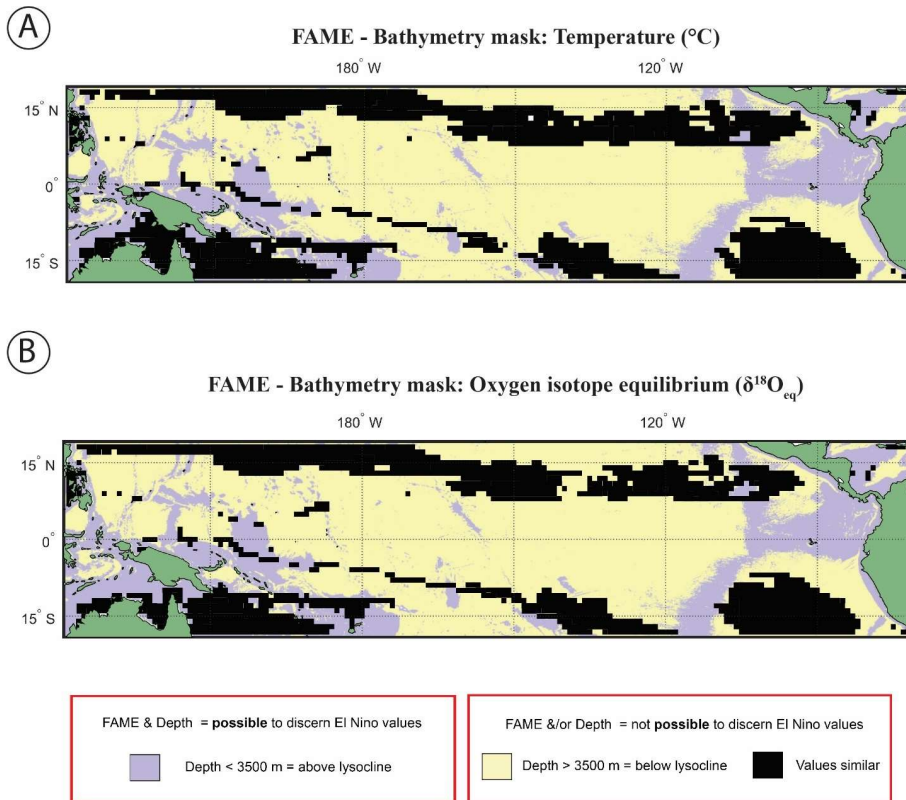


Figure 6 Overlay between bathymetry and FAME results. The variance—results of the FAME Anderson-Darling test for (A) temperature and (B) oxygen isotope values as input. Locations where the H_1 hypothesis can be accepted, i.e. the distributions can be said to be different ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$), are plotted as yellow where the depth is deeper than 3500 m bsl or purple where the depth is shallower than 3500 m bsl (see Figure 2). Purple locations are where our results suggest that the signal of ENSO has different values and the water depth allows for preservation.

Overlaying the water depth and the SAR with the Anderson-Darling results (Figure 6 and Supplementary Figure 7) highlights that of the total area where FP_{EN} is significantly different from FP_{NEU} (i.e. those areas where planktonic foraminiferal flux is suitable for reconstructing past ENSO dynamics), only a small proportion corresponds to areas where the sea floor is both above the CCD (< 3500 mbsl) and SAR is at least 5 cm/ka (Figure 7). However, at certain locations, near islands or seamounts,

Formatted: Font: Times New Roman

Formatted: Footer1

the SAR and water depth may be high enough to allow for a signal to be preserved (Figure 5B) that may not be represented here.

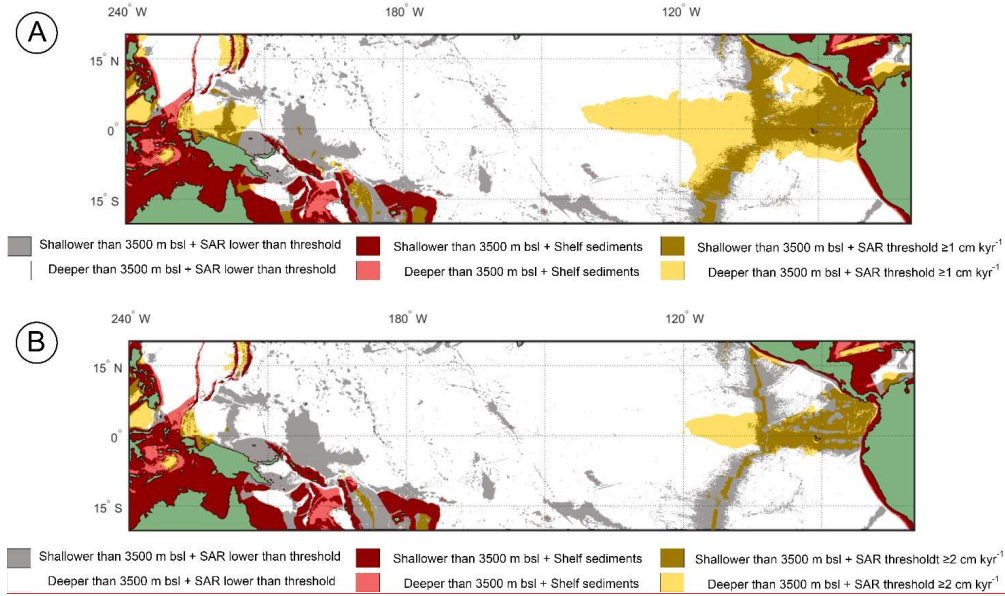


Figure 7 Overlay between water depth and inferred SAR (Olson et al., 2016). Cut-off limits for bathymetry and SAR are 3500 m below sea-level and (A) ≥ 1 cm kyr⁻¹ and (B) ≥ 2 cm kyr⁻¹ respectively. The colours represent the following: Red / Pink: Continental shelf sediments that are (Red) shallower or (Pink) deeper than 3500 mbsl; Grey / White: grid point SAR is lower than SAR threshold and the seafloor depth is (grey) shallower or (white) deeper than 3500 mbsl; Light Yellow/Gold: Light yellow represents areas where the SAR is above the threshold but the water depth is deeper than 3500 mbsl in comparison Gold represents areas where the SAR is above the threshold and the water depth is deeper than 3500 mbsl. The ideal locations are therefore plotted as Gold.

5

10

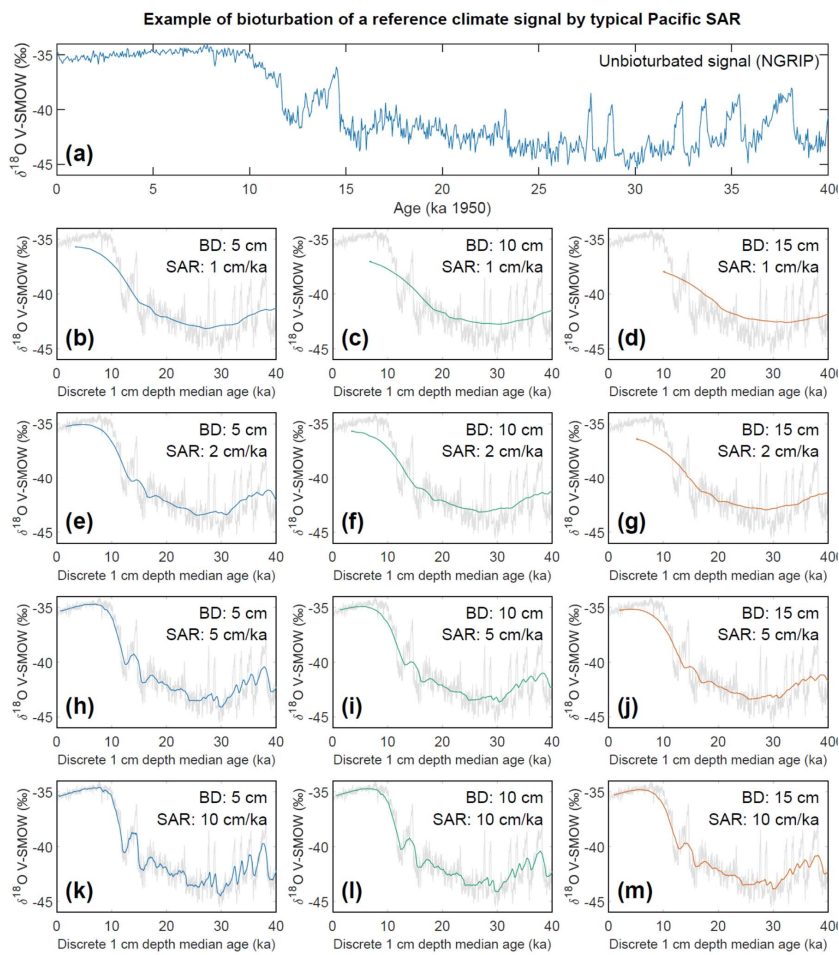


Figure 8 Output of the bioturbation model SEAMUS. (A) The unbioturbated input signal, NGRIP (North Greenland Ice Core Project Members, 2004; Rasmussen et al., 2014; Seierstad et al., 2014), used in our simulation of bioturbation for different SAR with SEAMUS (Loughheed, 2019). Sediment mixed layer referred to here as bioturbation depth (BD) is fixed at (B, E, H, K) 5 cm, (C, F, I, L) 10 cm and (D, G, J, M) 15 cm for sedimentation accumulation rates (SAR) of (B-D) 1 cm kyr⁻¹; (E-G) 2 cm kyr⁻¹; (H-J) 5 cm kyr⁻¹ and (K-M) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. In (B-M) grey values represent the unbioturbated input signal, NGRIP. Note, we retain the original units (V-SMOW) of the original timeseries used, no inference between Pacific climate and Greenland is intended by the use of NGRIP.

The results of the bioturbation simulator SEAMUS, plotted as a time series of the bioturbated 'NGRIP' signal (Figure 8) and as histograms of the probability of finding a particularly pseudo-foraminifera with a given age within the bioturbation depth (Figure 9), highlight the potential single foraminifera depth displacement that occurs with low sedimentation rates (Figure 5). Within a single depth in a core, proxy values largely representsrepresent the integrated time signal for that depth (Figure 11), as opposed to the variance of, the age of specimens within the bioturbation depth may vary from a few to tens of thousands of years (Figure 9). A data-model comparison without sufficient knowledge of bioturbation may equate an integrated proxy signal with a climatic signal for an inferred (or measured) average age for the depths in question. The proxy varianceFor proxies that use an average values (i.e., a pooled foraminiferal signal) or a variance (i.e., individual foraminifera values), the individuals will be based both upon a non-uniform distribution in temporal frequency of specimens, i.e., older specimens are few compared to younger specimens. A large proportion of the specimens in the BD come from years that are 'proximal' (i.e., close to the youngest age) thiswhich may give undue confidence that the probability of picking a specimen from these years is higher, however the long-tail of the distribution means that there is an equally high chance of picking a specimen that has come from several thousand years earlier than the discrete-depth's median age. If we consider for the moment this as picking specimens from a box, there is a high chance of picking from a single box that represents the age you want however there is an equally high chance of picking from numerous boxes with varying age. If the spread in the climatic variable is uniform throughout this time then it can be possible to reproduce a similar signal, although this would not by definition represent the actual spread in the actual climatic variable for a given time, however the spread in the climate variable is unlikely to be constant. With a varying spread in the climatic signal bioturbation can introduce the possibility of spurious interpretations, but it is of course more obvious where the measured distributions over-exaggerate the climate signal (e.g., Whilst the temporal integration involved in bioturbation can be problematic for either age-depth modelling (e.g., Lougheed et al., 2018; Lougheed et al., 2020a) or discrete age measurements (e.g., Lougheed et al., 2020b) it will also integrate the climate signal carried by the individual foraminifera. Wit et al., 2013). Furthermore, if we consider that researchers do not pick as randomly as they profess, there is both a size and preservation bias to specimens selected, and size is not constant down-core (e.g., Metcalfe et al., 2015) we can further introduce bias within the dataset. The SEAMUS output that corresponds with our chosen SAR cut off value of 5 cm kyr⁻¹ (Figure 10 and 11), the lower limit of our mask (Figure 9), is shown in panels in Figures 10H to 10J and Figures 11G to 11I. It is important to note however, that much of the region for which FAME is calculated upon has inferred sedimentation rates lower than this cut off value (Figure 9C to 9H). If for example the spread in a climate variable, such as temperature, is uniform throughout the integrated time (and the abundance at each temperature value is also uniform) then it could be possible to reproduce a similar temperature distribution in bioturbated cores. Although this would not by definition represent the actual spread in the actual climatic variable for a given time. However, the climate signal is unlikely to be constant, integrating a climatic signal bioturbation can therefore introduce artefacts inducing the possibility of spurious interpretations. Of course, identification of spurious datapoints are more obvious where the measured distributions over-exaggerate the climate signal (e.g., Wit et al., 2013). Our simulation of a climate signal reveals (Figure 8) the following: a reduction in signal amplitude with low SAR and/or increasing BD; loss of short

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

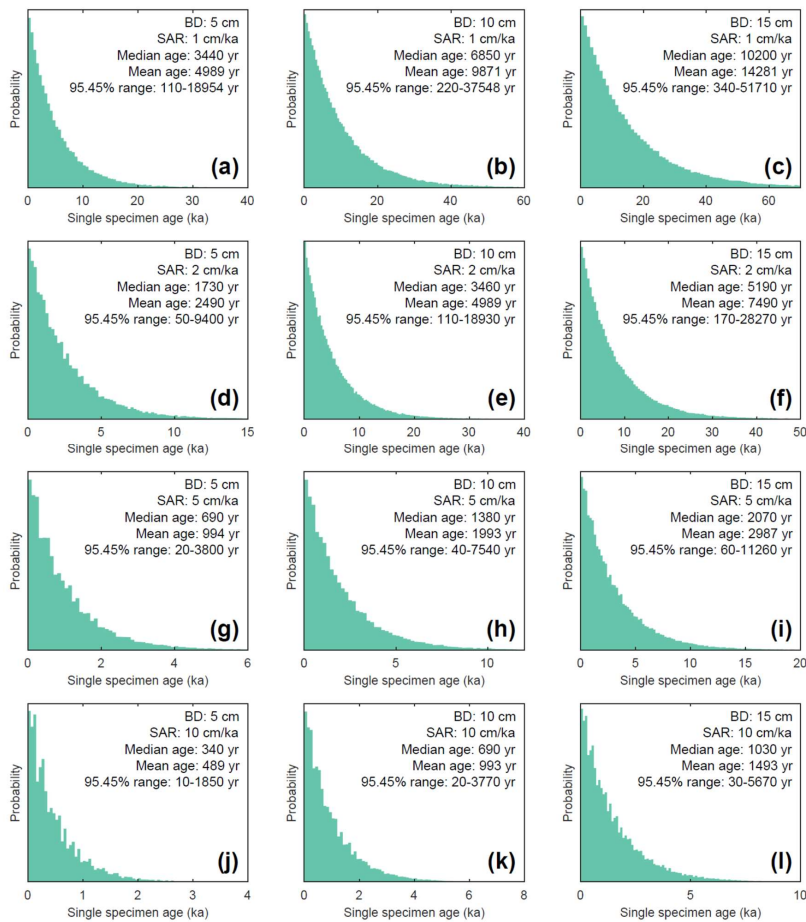
Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

events at low SAR; a shift in the apparent timing of events with increasing BD; and an apparent increasing 'core-top' age with low SAR and increasing BD (Figure 9). The median age of the bioturbation depth (Figure 9) is the reason why each timeseries (Figure 8) does not 'start' at 0 age (Keigwin and Guilderson, 2009).

Simulated single specimen age distribution for typical Pacific SAR



5 **Figure 9** Histograms of simulated specimen age within the bioturbation depth. The simulated age distribution present within the sediment mixed layer, referred to here as bioturbation depth (BD). BD is fixed at (A, D, G, J) 5 cm, (B, E, H, K) 10 cm and (C, F, I, L) 15 cm.

Formatted: Footer1

L) 15 cm for sedimentation accumulation rates (SAR) of (A-C) 1 cm kyr⁻¹; (D-F) 2 cm kyr⁻¹; (G-I) 5 cm kyr⁻¹ and (J-L) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. Note the size of the BD varies, therefore the simulated age distribution comes from a varying 'core depth'.

5

Whilst we are principally interested in understanding whether living foraminifera can theoretically reconstruct ENSO, the results of the sedimentological features, presented here, imply that much of the Pacific Ocean is not suitable for preserving (Figures 5-9) the ENSO signal, despite the possibility of the species of foraminifera in the water having unique values for different climate states (Section 4; Figure 6). In areas where preservation could occur, a hypothetical core could allow for the possible disentanglement of El Niño related signals from the climatic signal, but only in a best-case scenario involving minimal bioturbation, which is unlikely in the case of oxygenated waters. Combined with finite sampling strategies the effects of both dissolution and bioturbation can be further amplified.

10

8. Discussion

8.1 Palaeoceanographic Implications

15

Ecophysiological proxy system models are a mathematical approximation aimed at replicating the proxy signal both as its response to, and modification of, the original target climate signal (e.g., Dees et al., 2015). Linking ecophysiological models to coupled ocean-atmosphere models (e.g., Clement et al., 1999; Zebiak and Cane, 1987); isotope enabled Earth system models (e.g., iCESM; Zhu et al., 2017); or multi-model ensembles with prescribed boundary conditions could be used for the generation of timeseries for testing presumptions in proxy studies. Used a-priori, an explicit forward model can be used to test if it is plausible that the given recording system can record an oceanographic signal to allow robust reconstructions.

20

A critical presumption in proxy studies is embedded in site selection. Sites selected are presumed to be able to (or not) generate a climate signal, the presumptive answer in such studies is either the feature occurs or did not occur, and if it occurs then it has either enhanced or weakened. Such presumption precludes a scenario in which the feature or oceanographic regime has shifted, passing over or beyond a core site (Weyl, 1978), reacting to the expansion, contraction or shift of certain large scale oceanographic features (e.g., Polar Front, Upwelling) during periods of either warmer than average (e.g., the last interglacial) or colder than average temperatures (e.g. glacial maxima). The analysis of recent El Niño patterns suggests that there are two types of spatially delineated El Niño events: the dateline Central Pacific El Niño and the Eastern Pacific El Niño. Here we have highlighted a way of using models to determine the location where the different climate states could be differentiated. More explicit tests using climate models could be used to optimise sampling design, determine applicable core locations for comparison of proxy values with 'like with like' oceanographic features (similar to the analysis of Evans et al. (1998) for predicting coral sites), without necessarily the cost of a time-slice project (e.g., CLIMAP, MARGO).

25

30

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

Another test is whether for the same set of environmental conditions two species can record an identical signal. For species with a dynamic depth habitat in which the environmental signal becomes a weighted average of the water column (e.g., Wilke et al., 2017). An additional factor in the post-mortem preservation of the oceanographic signal in foraminiferal shells is whether the shells can be preserved. The GEBCO bathymetry data is binned into 250 m wide bins, and the data normalised to 1.0. As the data contains both bathymetric and topographic (below and above 0 m), the grey area in each histogram represent > 0 m (Figure 8). Whilst, there are differences depending on the cut-off value (Figures 8C to 8E) much of the canonical El Niño 3.4 region (Wang et al., 2017) used in oceanography (Figure 1) is also excluded from these suitable areas. Overlaying the water depth and the SAR with the Anderson-Darling results (Figure 7) highlights that of the total area where FP_{EN} is significantly different from FP_{NEU} (i.e. those areas where planktonic foraminiferal flux is suitable for reconstructing past ENSO dynamics), only a small proportion corresponds to areas where the sea floor is both above the CCD (< 3500 mbsl) and SAR is at least 5 cm/ka (Figure 9). However, at certain locations, near islands or seamounts, the SAR and water depth may be high enough to allow for a signal to be preserved (Figure 8B) that may not be represented here.

4. Discussion

4.1 From Life to Sedimentary Assemblages

Whilst we are principally interested in understanding whether living foraminifera can theoretically reconstruct ENSO, comparison with data requires an additional analysis. This is because data model comparisons are subjective, nominally supposing that the data is the value to be achieved by the model. However, if the foraminifera modulate the original climate signal, then preservation selectively filters which specimens are conserved whereas bioturbation acts to reorder, transposing the order in which they are recovered from the depth domain. Once the sediment is recovered, the researcher acts as a final filter, which is in essence a random picking. Although technically most researchers will pick whole shells so alongside size selectivity (e.g., Metcalfe et al., 2015) there is also preservation bias associated with picking of foraminifera (e.g. Koutavas and Lynch-Stieglitz, 2003). Whilst the presence of depths in the ocean whereupon calcite is absent from sediments (2006) the likelihood of species recording the same environmental signal becomes less plausible. This is, in brief, the rationale for the development of FAME, the same climate signal seen through the view of species-specific proxies will give a fractured view constrained by each species ecophysiological constraints (Mix, 1987; Roche et al., 2018). FAME is not the first proxy system model, instead it expands upon previous studies that have either approximated a foraminiferal signal either by weighting of ecological (seasonal or depth) preferences or by assuming that foraminifera record a fixed depth in the water column. What can be seen as contradictory proxy reconstructions can therefore be viewed as the prevailing or dominant conditions at a given location at the time when environmental conditions overlap ecological constraints for a given species. Reconstructions of the past climate (LGM-Holocene) of the Pacific have for instance inferred a relatively weaker Walker circulation, a displaced ITCZ and equatorial cooling (Koutavas and Lynch-Stieglitz, 2003); both a reduction (Koutavas and Lynch-Stieglitz, 2003)

and intensification (Dubois et al., 2009) in eastern equatorial Pacific upwelling; and both weakened (Leduc et al., 2009) and strengthened ENSO variability (Koutavas and Joanides, 2012; Sadekov et al., 2013), was described in the earliest work (e.g., Murray and Renaud, 1891; Sverdrup, 1942), overlaying maps of measured surface sediment carbonate percentage with water depth in the Pacific Ocean led Bramlette (1961) to coin the term 'compensation depth' (Wise, 1978). This work highlighted the 'narrow' depths (4-5000 m) in the Central Pacific of the CCD. Conceptually Berger (1971) placed three levels in the Pacific ocean that were descriptive of the aspects (e.g., chemical, palaeontological and sedimentological) of the calcite budget; the saturation depth, demarking supersaturated from undersaturated; the lysocline, the depth at which dissolution becomes noticeable (Berger 1968, 1971); and compensation depth (Bramlette, 1961), in which supply is compensated through dissolution. The aspects of the lysocline was estimated by the faunal assemblages of Parker and Berger (1971, figure's 14 and 15 of that publication), for much of the equatorial Pacific the lysocline is estimated at ~3800 m. As the lysocline is where dissolution becomes apparent, ergo it is a sample already visibly degraded, we therefore set the limit of the water depth mask shallower, at 3500 m bsl. In fact, in regions of high fertility, such as the Eastern Equatorial Pacific, the lysocline was estimated to be present at ~2800 m (Thunell et al., 1981) or ~3000 m (Berger, 1971; Parker and Berger, 1971). For instance, core V21-28 close to the Galapagos Islands (01°05'N, 87°17'W) has a shallower dissolution affect than either of these two values despite being collected from a water depth of 2714 m (Luz, 1973). A comparison between the hydrographic and sedimentary lysocline, using a mooring in the Panama Basin showed that the sedimentary lysocline is a product of where the hydrographic lysocline meets the seafloor (Thunell et al., 1981), therefore, this could lead to dissolution within the water of the settling flux (e.g., Schiebel et al., 2007). In the EEP region the shallower lysocline is accompanied by an equally shallower CCD (located at ~3600 m) for which the highly fertile is considered responsible for its shoaling, lowering the pH through increased CO₂ (Berger et al., 1976). The correspondence between lysocline depth and CCD depth does not hold true for the entirety of the Pacific, plotting a N-S cross section from 50°N to 50°S Berger (1971) noted that in the Central Equatorial Pacific, the high fertility region generates a larger zone of dissolution resistant facies even with a shoaled lysocline. However, a number of the inferences are contentious, for instance the reduction in upwelling in this region (Koutavas and Lynch-Stieglitz, 2003) is contradicted by Dubois et al. (2009), who used alkenones (i.e., U₃₇^{K'} ratios) to suggest an upwelling intensification. Whilst the U₃₇^{K'} proxy has problems within coastal upwelling sites (Kienast et al., 2012) it does not discount their claim, especially considering that δ¹⁸O records can themselves be influenced by salinity upon the δ¹⁸O_{sw} component (Rincón-Martínez et al., 2011) and the potential influence of carbonate ion concentration ([CO₃²⁻]) upon foraminiferal δ¹⁸O_c (de Nooijer et al., 2009; Spero et al., 1997; Spero and Lea, 1996). The discrepancies in reconstructed climate between marine cores are worth noting, as ultimately it is from proxies that inferences are made about past climate (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004). Such inferences have suggested that the past climate of the Pacific region (from the geologically recent too deep time) has been in an: El Niño state (Koutavas et al., 2002; Stott et al., 2002; Koutavas and Lynch-Stieglitz, 2003); permanent El Niño state (Huber and Caballero, 2003); Super El Niño state (Stott et al., 2002); La Niña state (Andreasen et al., 2001; Beaufort et al., 2001; Martínez et al., 2003); or a different climatic state altogether (Pisias and Mix, 1997; Feldberg and Mix, 2003). Ultimately the possibility of a marine sediment archive being able to reconstruct ENSO dynamics comes down

Formatted: Font: Italic

Formatted: Footer1

to several fundamentals besides whether the signal can or cannot be preserved (*i.e.*, whether the core site has either too low SAR, too high BD or a water depth not conducive to calcite preservation): the time-period captured by the sediment intervals (a combination of SAR and bioturbation); the frequency and intensity of ENSO events; the foraminiferal abundance during ENSO and non-ENSO conditions; as well as what the proxy is recording. Reconstructions of the past can benefit from inclusion

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

5 within conceptual frameworks that incorporate both data and modelling studies (e.g., Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). If we factor in the sedimentation rate of the Pacific, which has been estimated to be considerably lower than 1 cm (Blackman and Somayajulu, 1966; Berger, 1969; Menard, 1964), then dissolution may become further exacerbated. The longer a shell remains at the sediment-water interface the greater the prospects for it to be dissolved become, therefore low SAR increases the chance of dissolution (Bramlette, 1961). For instance, in 15 equatorial Pacific cores, below 4000 m, the average SAR was presented (

8 Hays et al., 1969; here calculated) at 0.96 cm kyr^{-1} ($1\sigma \pm 0.43 \text{ cm kyr}^{-1}$). Although there are regions and/or core locations in which the SAR is higher, for instance eight EEP cores shallower than the lysocline depth (Thunell et al., 1981) of ~2800 m were presented by Koutavas and Lynch-Stieglitz (2003) which have an average SAR, calculated, at 7.20 cm kyr^{-1} ($1\sigma \pm 2.82$ 15 cm kyr^{-1}). The average age for these same core's 0 cm core depth is 2184 years ($1\sigma \pm 1521 \text{ yrs}$), whilst it cannot be assumed that there has been no loss during recovery (*i.e.*, core top is not sediment-water interface), a non-zero core top age is expected for both bioturbation (Keigwin and Guilderson, 2009) and dissolution. Alongside, the potential for dissolution there is the also the mixing of ocean sediments by the benthos (Bramlette and Bradley, 1942). For instance, Arrhenius (1961) noted that ash beds present in cores of the EEP (Worzel, 1959; Ewing et al., 1959) had a 2-3 cm layer above and below what should have 20 originally been a sharp boundary in which they estimated that ~50% of the material originated from the other side of the boundary. If one assumes 1 cm kyr^{-1} sedimentation rates, then the range in age of the obviously 6 cm mixed sediments is minimally ~6000 years per cm, comparison with an analogous SEAMUS simulation (bioturbation depth 5 cm; SAR 1cm) highlights the considerable spread in age, placing the 95.45% range between 110 and 18954 years (Figure 11). Much of this temporal variability (either through bioturbation or dissolution) will be hidden, especially when proxy values correspond with 25 the expected values, and more obvious when the values are larger than expected (*e.g.*, Wit et al., 2013). Owing to the lack of absolute variability during the Holocene the apparent confirmation of similarity between proxy values and modern distributions of the '*to be reconstructed*' variable is not a confirmation of proxy reliability. Especially in the tropics wherein seasonal variability is limited. The effects of both bioturbation and dissolution are further amplified when combined with finite sampling strategies. Therefore, the results of the sedimentological features, presented here, imply that much of the Pacific Ocean is not 30 suitable for preserving (Figures 7-9) the ENSO signal, despite the possibility of the species of foraminifera having unique values for different climate states (Figures 4-7). ENSO studies using palaeoceanography have exposed shifts, one can, therefore, question what is being reconstructed in such studies.

Formatted: Footer1

4.2 Palaeoceanographic Implications

4.2.1 Pacific climate reconstructions

One artefact of sampling (Dolman and Laepple, 2018) is the potential occurrence of aliasing (Pisias and Mix, 1988; Wunsch, 2000; Wunsch and Gunn, 2003): a fundamental problem with proxy records is that they can be confounded by local regional climate, and/or ENSO's teleconnections, that mimic ENSO changes albeit at a different temporal frequency. Our own analysis using our FAME $\delta^{18}\text{O}_c$ and T_c output mimics foraminiferal sedimentary archives, pooling several decades worth of data in which the resolution is coarse enough to obscure and prevent individual El Niño events being visible but allowing for some kind of long-term mean state of ENSO activity to be reconstructed (Cole and Tudhope, 2017). The results of our Anderson-Darling test may be unduly influenced by the Pacific decadal variability (PDV), also referred to as the Pacific Decadal Oscillation (PDO) (Pena et al., 2008). In much of the tropical Pacific the ratio of decadal to interannual σSST suggests that they are comparable in magnitude, therefore fluctuations in SST are more obviously apparent outside of the purely canonical regions of ENSO (Wang et al., 2017). It could be that the areas outside of these canonical ENSO regions (Figure 1) reflect the PDO (Pena et al., 2008; Wang et al., 2017). The use of the variance $\sigma^2(\delta^{18}\text{O}_c)$, or standard deviation $\sigma(\delta^{18}\text{O}_c)$, as an indicator of ENSO is dependent on whether the original climate signal's variance was or was not dominated by interannual variance. Zhu et al. (2017) computed the total variance change with and without the annual cycle suggesting that, for some cores the increased assumed ENSO variability at the LGM as deduced by proxy records (Koutavas et al., 2006; Koutavas and Joanides, 2012; Koutavas and Lynch-Stieglitz, 2003) may be purely a by-product of the annual cycle or dominated by it. Although the values of El Niño can be considered significantly different from other climate states, our own analysis using the ratio of total to interannual variance also suggests that much of the variance in the simulated foraminiferal signal is dominated by interannual variance. There are differences in the ratio of total to interannual variance between species and in different regions of the tropical Pacific, however, even with a dynamic depth habitat the ratio is still high (Figure 3; Supplementary Table 1).

4.2.2 The use of models in reconstructions

Reconstructions of the past climate of the Pacific have inferred a relatively weaker Walker circulation, a displaced ITCZ and equatorial cooling (Koutavas and Lynch-Stieglitz, 2003); both a reduction (Koutavas and Lynch-Stieglitz, 2003) and intensification (Dubois et al., 2009) in eastern equatorial Pacific upwelling; and both weakened (Ledue et al., 2009) and strengthened ENSO variability (Koutavas and Joanides, 2012; Sadekov et al., 2013). However, a number of the inferences are contentious, for instance the reduction in upwelling in this region (Koutavas and Lynch-Stieglitz, 2003) is contradicted by Dubois et al. (2009), who used alkenones (*i.e.*, U_{37}^{Kl} ratios) to suggest an upwelling intensification (Zhang et al., 2017). Whilst the U_{37}^{Kl} proxy has problems within coastal upwelling sites (Kienast et al., 2012) it does not discount their claim, especially considering that $\delta^{18}\text{O}$ records can themselves be influenced by salinity upon the $\delta^{18}\text{O}_{sw}$ component (Rincón-Martínez et al., 2011) and the potential influence of $[\text{CO}_2]$ upon foraminiferal $\delta^{18}\text{O}_c$ (de Nooijer et al., 2009; Spero et al., 1997; Spero and Lea, 1996). The discrepancies in reconstructed climate between marine cores is worth noting, as ultimately it is from proxies that inferences are made about past climate (Trenberth and Otto-Bliesner, 2003; Rosenthal and Broecker, 2004). Such

Formatted: Font: Italic

Formatted: Footer1

inferences have suggest that the past climate of the Pacific region (from the geologically recent too deep time) has been in an El Niño state (Koutavas et al., 2002; Stott et al., 2002; Koutavas and Lynch-Stieglitz, 2003); permanent El Niño state (Huber and Caballero, 2003) or Super El Niño state (Stott et al., 2002); La Niña state (Andreasen et al., 2001; Beaufort et al., 2001; Martinez et al., 2003); or a different climatic state (Pisias and Mix, 1997; Feldberg and Mix, 2003).

5 The possibility of a marine sediment archive being able to reconstruct ENSO dynamics comes down to several fundamentals besides whether the signal can or cannot be preserved (*i.e.*, whether the core site has either too low SAR, too high BD or a water depth not conducive to calcite preservation); the time period captured by the sediment intervals (a combination of SAR and bioturbation); the frequency and intensity of ENSO events; the foraminiferal abundance during ENSO and non-ENSO conditions; as well as what the proxy is recording. There is also the presumption that a particular climate event should be
10 recorded, our Anderson-Darling test for instance highlights that there are locations that cannot discern the difference between El Niño and other climate states whilst for the same time period there are locations where the different climate states can be differentiated. Whilst our analysis is a statistical treatment of the data, each species, and different types of phyto or zooplankton preserved in ocean sediments, are likely to record the same set of environmental conditions differently (Mix, 2006). This is, in brief, the rationale for the development of FAME, the same climate signal seen through the view of species-specific proxies will give a fractured view constrained by each species particular ecophysiological constraints (Mix, 1987; Roche et al., 2018). A dynamic depth habitat in which the environmental signal becomes a weighted average of the water column can further confound the original signal (Wilke et al., 2006). What can be seen as contradictory reconstructions can therefore be viewed as the prevailing or dominant conditions at a given location at the time when environmental conditions overlap ecological constraints for a given species.

20 Terrestrial records suggest the number of El Niño events per century in the early Holocene (8-6 ka BP) was minimal (Moy et al., 2002), with between 0 and 10 events occurring per century. This dampened ENSO is observed within lake core colour intensity and records driven primarily by precipitation—although like other proxies this can also be interpreted differently, *i.e.* as a large change in the hydrological cycle shifting precipitation away regionally (Trenberth and Otto-Bliesner, 2003). If we assume for now that the number and magnitude of ENSO events was reduced, the relatively low downcore resolution of marine records may not accurately capture the dynamics of such lower amplitude ENSO events using existing methods. The sensitivity and probability of detecting a change in IFA with changes in frequency and amplitude, has been dealt with before (Thirumalai et al., 2013), although without considering bioturbation. The synthesis of pseudo-timeseries to discern the potential distribution for different scenarios, whilst a necessary approximation, is nonetheless one that is free of cause and causality. Modulating a timeseries for events with enhanced or weakened amplitude or fewer or greater number of events assumes in essence that there
25 is limited feedback both regionally (between two sites) and internally within the timeseries (*i.e.*, a process that operates on a higher level). Reconstructions of the past can benefit from inclusion within conceptual frameworks that incorporate both data and modelling studies (e.g., Trenberth and Otto-Bliesner, 2003; Rosenthal and Broccoli, 2004; McPhaden et al., 2006). The use of coupled ocean-atmosphere models (e.g., Clement et al., 1999; Zebiak and Cane, 1997); isotope-enabled Earth system models (e.g., iCESM; Zhu et al., 2017); or multi-model ensembles with prescribed boundary conditions can be used for the

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

generation of timeseries in which the physics of atmospheric and oceanic circulation are constrained and feedbacks between sites can occur. The perceived failure of several climate models to resolve ENSO adequately, resulting in variable ENSO frequency and amplitude between models, could therefore be used to determine the proxy signal from model derived timeseries at different frequencies and intensities of ENSO. Albeit a timeseries of variable ENSO that is grounded in ocean atmosphere coupling. Such analysis could also provide information on a secondary assumption, in which time slices from the same core inherently assume that where a particular oceanographic feature exists now is also where it may have existed before. This gives a somewhat binary view, the feature either occurs or does not occur, and if it occurs then it has either enhanced or weakened. Yet this can (though not always) preclude a scenario in which the feature has shifted. Analysis of the El Niño patterns suggests that there are two types of El Niño that are spatially delineated: the dateline Central Pacific El Niño and the Eastern Pacific El Niño. The expansion, contraction or shift of certain large scale oceanographic features (e.g., Polar Front, Upwelling) during periods of warmer than average (e.g., the last interglacial) or colder than average temperatures (e.g. the LGM) can complicate the comparison of two down core samples, i.e., a static core continuously recording a particular climate event as opposed to a shifting oceanographic regime passing over or beyond a core site (Weyl, 1978). Climate models could therefore also be used to determine applicable core locations for comparison of proxy values with 'like with like' oceanographic features (similar to the analysis of Evans et al. (1998) for predicting coral sites), without necessarily the cost of a time slice project (e.g., CLIMAP, MARGO).

4.2 Limitations of the methods applied and assessment of model uncertainties

For simplicity we have assumed that our model is 'perfect', of course that is inaccurate, there are four potential sources of error: the input variables (temperature, salinity and their conversion into $\delta^{18}\text{O}_{\text{sw}}$ and $\delta^{18}\text{O}_{\text{eq}}$); the model's error with respect to real world values (Roche et al., 2018); (Roche et al., 2018); the statistical test's errors (associated Type I – in which attribution of significance is given to an insignificant random event, a false 'positive' – and Type II – in which a significant event is attributed to be insignificant, a false 'negative' - errors); and reducing the complexities of foraminiferal biology via parameterization. The input variables can have errors associated with both the absolute values of temperature and salinity used here, and the limitation of input values to a single value per month. Whilst it is possible to interpolate to a daily resolution, this is problematic for two reasons: (1) daily temperature records have much more high frequency oscillations than the data here and (2) the lifecycle of a single foraminifera is approximately monthly, therefore by using monthly data it provides an estimate of the average population signal. Conversion of salinity and temperature into $\delta^{18}\text{O}_{\text{sw}}$ and $\delta^{18}\text{O}_{\text{eq}}$ uses a quadratic approximation, one source of error is the unknown influence of carbonate ion concentration on both the Kim and O'Neil (1997) equation and the foraminiferal microenvironment (de Nooijer et al., 2008, 2009; Spero et al., 1997; Spero and DeNiro, 1987; Spero and Lea, 1996) which has further implications due to the upwelling of cool, low pH, waters in the eastern Tropical Pacific (Cole and Tudhope, 2017; Raven et al., 2005); (de Nooijer et al., 2008, 2009; Spero et al., 1997; Spero and DeNiro, 1987; Spero and Lea, 1996) which has implications due to the upwelling of cool, low pH, waters in the eastern Tropical Pacific (Cole and Tudhope, 2017; Raven et al., 2005). The spatial variability in salinity, particularly within regions underlying the

Formatted: Font: Times New Roman

Formatted: Heading 21

Formatted: Footer1

intertropical convergence zone (ITCZ) and the moisture transport from the Caribbean into the eastern Pacific along the topographic low that represents Panama Isthmus, the resultant conversion of salinity to $\delta^{18}\text{O}_{\text{sw}}$ and then $\delta^{18}\text{O}_{\text{eq}}$ may contain further error. If such errors are independent of the absolute value of the variable, *i.e.* the error on cold temperature is the same and not larger than warm temperatures, then the error terms effectively cancel one another out. A point of note, is that the $\delta^{18}\text{O}$ to $^{\circ}\text{C}$ conversion of Kim and O'Neil (1997) is considered to be marginally larger at the cold end than at the warm end (0.2 ‰ per 1°C to 0.22 ‰ per 1°C) than that originally discerned (O'Neil et al., 1969). (O'Neil et al., 1969).

The comparison of the pseudo-Mg/Ca temperature signal produced here (T_c) to a value corresponding to that reconstructed from measurements of Mg/Ca should be done with caution. Computation of pseudo-foraminiferal $\delta^{18}\text{O}$ in FAME is aided by the ability to compute an initial $\delta^{18}\text{O}$ equilibrium value for a given latitude-longitude grid-point and timestep. The weighting of $\delta^{18}\text{O}$ value used in FAME is an approximation of the foraminiferal shell, chambers are generally homogenous in $\delta^{18}\text{O}$ value excluding either terminal features such as crust or gametogenic calcite which can lead to chamber heterogeneities (e.g., Wycech et al., 2018) although this can be approximated with an additional parameter (Roche et al., 2018). The same cannot be said for Mg/Ca, alongside heterogeneities in the shell which may be the result of diurnal processes (terminal features in the computation of $\delta^{18}\text{O}$ are technically simpler to model), there are differences in both sample preparation and measurement techniques. Whilst, the change in Mg/Ca with temperature has been validated (e.g., Elderfield and Ganssen, 2000) the computation of a pseudo-proxy value for and from model parameters remains enigmatic. Construction of a matrix of equilibrium Mg/Ca would ideally be the most logical step in a second generation of the FAME model. Whilst, simply solving the Mg/Ca palaeotemperature equation for an input of T and an output Mg/Ca is a first approximation, as stated previously several other parameters can alter this technique, this includes abiotic effects such as salinity (Allen et al., 2016; Gray et al., 2018; Groeneveld et al., 2008; Kısakürek et al., 2008) or carbonate ion concentration (Allen et al., 2016; Evans et al., 2018; Zeebe and Sanyal, 2002); biotic effects such as diurnal calcification (Eggins et al., 2003; Hori et al., 2018; Sadekov et al., 2008, 2009; Vetter et al., 2013); or additional factors such as sediment (Fallet et al., 2009; Feldmeijer et al., 2013) or specimen (Barker et al., 2003; Greaves et al., 2005) 'cleaning' techniques. Given the role of Mg in inhibiting calcium carbonate formation, the manipulation of seawater similar to the modification of the cell's pH (de Nooijer et al., 2008, 2009) may aid calcification and explain the formation of low Mg by certain foraminifera (Zeebe and Sanyal, 2002). The weighting of $\delta^{18}\text{O}$ value used in FAME is an approximation of the foraminiferal shell, chambers being generally homogenous in $\delta^{18}\text{O}$ value, excluding either terminal features such as crust or gametogenic calcite which can lead to chamber heterogeneities (e.g., Wycech et al., 2018), although the latter can be approximated with an additional parameter (Roche et al., 2018). The same cannot be said for Mg/Ca, alongside heterogeneities in the shell which may be the result of diurnal processes, there are differences in both sample preparation and measurement techniques. Whilst the change in Mg/Ca with temperature has been validated (e.g., Elderfield and Ganssen, 2000) the computation of a pseudo-proxy value for and from model parameters remains enigmatic. Construction of a matrix of equilibrium Mg/Ca would ideally be the most logical step in a second generation of the FAME model. Whilst simply solving the Mg/Ca palaeotemperature equation for an input of T and an output Mg/Ca is a first approximation, as stated previously

several other parameters can alter this technique, this includes abiotic effects such as salinity (Allen et al., 2016; Gray et al., 2018; Groeneveld et al., 2008; Kısakürek et al., 2008) or carbonate ion concentration (Allen et al., 2016; Evans et al., 2018; Zeebe and Sanyal, 2002); biotic effects such as diurnal calcification (Eggins et al., 2003; Hori et al., 2018; Sadekov et al., 2008, 2009; Vetter et al., 2013); or additional factors such as sediment (Fallet et al., 2009; Feldmeijer et al., 2013) or specimen (Barker et al., 2003; Greaves et al., 2005) 'cleaning' techniques. Given the role of Mg in inhibiting calcium carbonate formation, the manipulation of seawater similar to the modification of the cell's pH (de Nooijer et al., 2008, 2009) may aid calcification and explain the formation of low-Mg by certain foraminifera (Zeebe and Sanyal, 2002). Scaling these processes up to a basin-wide model is beyond the remit of this current paper. ▲

Our modelling results also depend upon the species symbiotic nature and potential genotypes. For instance, mixotrophs, those organisms that utilise a mixture of sources for energy and carbon (planktonic foraminifera such as *G. ruber*; and/or *G. sacculifer*) can outcompete heterotrophic (or photoheterotrophic) organisms (planktonic foraminifera such as *Neogloboquadrina pachyderma*; *Neogloboquadrina incompta*) especially in stratified-oligotrophic waters. Whilst FAME uses only the temperature component of FORAMCLIM (Roche et al., 2018) and we have only modelled *N. dutertrei*, it is important to note that there are distinctions between the fundamental niche that FAME computes, *i.e.* the conditions that an organism can survive, and the realised niche, *i.e.* what an organism actually occupies given limiting factors within the environment. Likewise, FAME and FORAMCLIM are based upon the original culture experiments that assumed that both species (*G. bulloides* and *N. dutertrei*) are non-symbiotic or have species associations (see Bird et al., 2018, 2017). A species that hosts symbionts will likely have a restricted temperature that is associated with the temperature tolerance of their symbionts, given that the next generation of a species of planktonic foraminifera must be re-infected with their symbionts. Likewise, cryptic speciation may lead to foraminiferal genotypes exhibiting distinct environmental preferences (Bird et al., 2018, 2017; Darling et al., 2004, 2000, 1999; Huber et al., 1997; Morard et al., 2013; de Vargas et al., 1999, 2002). Incorporation of both a theoretical genotype abundance (Morard et al., 2013) and ecophysiological tolerances of different genotypes (Bird et al., 2018) within an ecophysiological model could further reduce error within modelling of planktonic foraminiferal habitats, and thus reduce data-model comparison error. For instance, Morard *et al.* (2013) simulated the impact of genotypes upon palaeoceanographic reconstructions (in particular transfer functions) using a theoretical abundance, calculated with a best-fit gaussian response model, depending upon SST later using a similar approach (Morard et al., 2016) to deduce the impact upon $\delta^{18}\text{O}$.

Our modelling results also depend upon the species symbiotic nature and potential genotypes. For instance, mixotrophs, those organisms that utilise a mixture of sources for energy and carbon (planktonic foraminifera such as *G. ruber*; and/or *G. sacculifer*) can outcompete heterotrophic (or photoheterotrophic) organisms (planktonic foraminifera such as *Neogloboquadrina pachyderma*; *Neogloboquadrina incompta*) especially in stratified-oligotrophic waters. Whilst FAME uses only the temperature component of FORAMCLIM (Roche et al., 2018) it is important to note that there are distinctions between the fundamental niche that FAME computes, *i.e.* the conditions that an organism can survive, and the realised niche, *i.e.* what

Formatted: Font: Times New Roman

Formatted: Footer1

an organism actually occupies given limiting factors within the environment. As FORAMCLIM and therefore FAME are based upon culture experiments, new observations highlight symbiotic or species associations (see Bird et al., 2018, 2017). A species that hosts symbionts will likely have a restricted temperature that is associated with the temperature tolerance of their symbionts. Likewise, cryptic speciation may lead to foraminiferal genotypes exhibiting distinct environmental preferences (Bird et al., 2018, 2017; Darling et al., 2004, 2000, 1999; Huber et al., 1997; Morard et al., 2013; de Vargas et al., 1999, 2002). Incorporation of both a theoretical genotype abundance (Morard et al., 2013) and ecophysiological tolerances of different genotypes (Bird et al., 2018) within an ecophysiological model could further reduce error within modelling of planktonic foraminiferal habitats, and thus reduce data-model comparison error. For instance, Morard et al. (2013) simulated the impact of genotypes upon palaeoceanographic reconstructions (in particular transfer functions) using a theoretical abundance, calculated with a best-fit gaussian response model, depending upon SST later using a similar approach (Morard et al., 2016) to deduce the impact upon $\delta^{18}\text{O}$.

Conclusion

Concentrating on the period spanning the instrumental record by using the FAME module, we forward modelled the species-specific (i.e., *G. ruber*; *G. sacculifer* and *N. dutertrei*) oxygen isotope values ($\delta^{18}\text{O}_c$) and pseudo-Temperature (T_c), computed from ocean reanalysis data using the temperature driven FAME module. The aim of this study was to determine whether the modelled values from different climate states are statistically different. If our assumptions are correct, including the reduction in Foraminiferal complexity and the choice of generic distribution (i.e., kernel) to the fit the data prior to performing an Anderson-Darling test, our results suggest for large expanses of the Tropical Pacific the climate states do have different values. Whilst, the results show that the values between El Niño states and Neutral climate states are statistically different for a large proportion of the Tropical Pacific, the total variance is dominated by the interannual variance for much of the region. Overlaying our computed foraminiferal distributions with the characteristics of the Pacific Ocean we infer that much of the signal recorded in foraminifer region available for reconstructions corresponds to areas where several processes will alter the preservation of the foraminiferal signal. First, the inferred SAR for much of the region is critically low, and a simulation of bioturbation for different bioturbation depths and SAR, typical for the Pacific indicates that discrete core depths can have a large temporal spread in single foraminifera, possibly precluding the extraction of ENSO-related climate variability. Second, a large proportion of the seafloor lies below the lysocline, the depth at which dissolution of foraminifera becomes apparent. These factors reduce the size of the area available for reconstructions considerably, thus arguably precluding the extraction of a temporally valid palaeoclimate signal using long-standing methods. It is our inference that only at exceptional ocean sediment core sites is it possible to determine the variability in ENSO based on planktonic foraminifer measurements, which makes it difficult to build a Pacific basin-wide understanding of past ENSO dynamics.

Formatted: Heading 11

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Font: Times New Roman

Formatted: Footer1

Code and data availability

The ocean reanalysis data used in this paper are available from the Universiteit Hamburg. An open source version of the FAME code is available from Roche et al. (2018). Statistical routines are available as part of the Statistical package of MATLAB R2018a; mapping tools (including the topographic colormap) are part of the Mapping Toolbox. The function to retrieve GEBCO bathymetry (data available at www.gebco.net/www.gebco.net) from netcdf format, `gebconetcd(FILE,Wlon,Elon,Slat,Nlat)`, is available from the MATLAB Central File Exchange (<https://mathworks.com/matlabcentral/fileexchange/46669-gebconetcd-file-wlon-elon-slat-nlat><https://mathworks.com/matlabcentral/fileexchange/46669-gebconetcd-file-wlon-elon-slat-nlat>). The single foraminifera sediment accumulation simulator (SEAMUS) is published in Lougheed (2019,2020), available at <https://doi.org/10.5194/gmd-2019-155><https://doi.org/10.5194/gmd-2019-155>. A video of the $\delta^{18}\text{O}_{\text{shell}}$ output has been archived online (<https://doi.org/10.5281/zenodo.2554843><https://doi.org/10.5281/zenodo.2554843>, Metcalfe et al., 2019).

Author Contributions

B.M. and D.M.R. designed the study. B.M. analysed the data. B.C.L. processed ocean SAR-and depth data, and ran the bioturbation model. B.M. drafted the manuscript with contributions from all authors.

Competing Interests

The authors declare no competing interests.

Acknowledgements

B.M. was supported by a Laboratoire d'excellence (LabEx) of the Institut Pierre-Simon Laplace (Labex L-IPSL), funded by the French Agence Nationale de la Recherche (grant no. ANR-10-LABX-0018). B.M. thanks both LSCE and the VU University Amsterdam for guest status. D.M.R. is supported by the French agency Centre National de la Recherche Scientifique (CNRS) and the VU University Amsterdam. This is a contribution to the ACCLIMATE ERC project. The research leading to these results has received funding to C.W. from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013 Grant agreement n° 339108). B.C.L. acknowledges Swedish Research Council (Vetenskapsrådet – VR) grant 2018-04992, The Swedish National Infrastructure for Computing (SNIC) at the Uppsala Multidisciplinary Centre for Advanced Computational Science (UPPMAX) provided computer resources for running the SEAMUS model. We thank the Universiteit Hamburg for their online access server for ocean reanalysis data.

Formatted: Heading 11

Formatted: Font: Times New Roman

Formatted: Heading 11

Formatted: Heading 11

Formatted: Heading 11

Formatted: Footer1

5 Figure 1. Oceanic Niño Index and the temperature anomaly for a single El Niño event. (Top) Oceanic Niño Index (ONI), sourced
from a 3-month running mean of SST anomalies in ERSST.v5 of the Niño 3.4 region (Huang et al., 2017). Grey vertical bars represent
the periods in which El Niño-like conditions exist using a simple one-month threshold. (Bottom) The sea surface temperature
10 difference between week beginning 1st December 1997 minus the long-term climatic mean (1971–2000) for December. The 1997–
1998 El Niño represents an EP-ENSO. The long-term monthly climatology, the NOAA optimum interpolation (OI) SST V2, based
upon the methodology of Reynolds and Smith (Reynolds and Smith, 1995) using two distinct climatologies for 1971–2000 and 1982
–2000 (Reynolds et al., 2002). Boxes represent the Niño region: Niño 1 and 2 (0° – 10° S, 90° W– 80° W), Niño 3 (5° N– 5° S, 150° W–
 90° W), Niño 3.4 (5° N– 5° S, 170° W– 120° W) and Niño 4 (5° N– 5° S, 160° E– 150° W).

Figure 2. A snapshot of the output of FAME. Each panel represents an individual species (Top Panel *G. sacculifer*; Middle Panel *G. ruber* and Bottom Panel *N. dutertrei*) $\delta^{18}\text{O}$ for a single time step ($t = 696$). The species $\delta^{18}\text{O}$ for each grid-point is based upon the integrating the $\delta^{18}\text{O}_{\text{eq}}$ values using a growth-rate based weighting (FAME; Roche et al., 2018). Values are in per mil (‰ V-PDB).

5

Figure 3-Figure Captions

Figure 10 Anderson-Darling Results for Input datasets of Temperature and Equilibrium $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{eq}}$). Results of the test in which input variables underwent the same statistical procedure (see section 2.0) as the modelled data for (A) temperature and (B) $\delta^{18}\text{O}_{\text{eq}}$ values. Here, model input data was extracted for a single depth of ~5 m without any growth weighting applied. Black regions are those grid points in which the null hypothesis (H_0), that the El Niño and Non- El Niño (Neutral) foraminifera populations (FP) are not statistically different ($FP_{\text{El Niño}} = FP_{\text{Non-El Niño}}$), cannot be rejected. Grey regions represent grid points where the H_1 hypothesis is accepted, therefore the distributions of the foraminiferal population for El Niño and Non- El Niño can be said to be unique ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$). The hatched regions represent areas where the H_1 hypothesis can be accepted, therefore the distributions of the foraminiferal population for El Niño and Non- El Niño can be said to be unique ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$), though the difference between the means of tested distribution are less than (A) 0.5°C or (B) 0.1 ‰. For a comparison with three different fixed depths (5; 149; and 235 m) without any growth weighting applied see Supplementary Figure 2.

Figure 11 Anderson-Darling results plotted regionally in which species-specific results are overlain. Panels represent water depth locations where dissimilar and similar values for the two climate states for (a-b) FAME-T, modelled temperature (c-d) FAME- $\delta^{18}\text{O}_{\text{c}}$ modelled oxygen isotope values recorded in the calcite shells (Tc) occur. Each panel represents the Anderson-Darling test result, the results for *Globigerinoides sacculifer*, *Globigerinoides ruber* and *N. dutertrei* are overlaid with (A and C) cut-off depth of 60 m and (B and D) species-specific cut-off values. For all panels black areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño is similar to the Non-El Niño, and therefore the distribution between the neutral climate and El Niño cannot be said to be different ($FP_{\text{El Niño}} = FP_{\text{Non-El Niño}}$).

Figure 12 Total variance and Interannual variance and the range in standard deviation of the Monte-Carlo experiments. (a) Total variance of *Globigerinoides sacculifer* $\delta^{18}\text{O}_{\text{c}}$, using FAME- $\delta^{18}\text{O}_{\text{eq}}$ for a cut-off value of 60 m. (b) The ratio of (a) and (c), where (c) is the Interannual variance of the timeseries of (a).

Figure 13 The range in standard deviation of the Monte-Carlo experiments using FAME- $\delta^{18}\text{O}_{\text{c}}$ *G. sacculifer* with a depth cut-off of 60 m. In (d-i) we plot the range in standard deviation obtained by picking 60 months with replacement with 10,000 iterations, the experiments are as follows: (da) the same months were chosen for all grid-points for each iteration of the Monte-Carlo; (eb) each grid-point has its own randomly selected months for each iteration of the Monte-Carlo; (fc) the same as (eb) but we weight the values by the total amount of growth per month; (gd) the months selected for (ec) were re-run but a random variability is added to each month (between -0.4 and 0.4 ‰); (he) the months selected for (eb) were re-run but a random measurement error is added to each month (between -0.12 and 0.12 ‰); and (if) the months selected for (eb) were re-run but the (gd) random variability and (he) measurement error were combined. Note the scale change between (d-f) and (g-i).

5 **Figure 4. Anderson-Darling Results for Input datasets of Temperature and Equilibrium $\delta^{18}\text{O}$ ($\delta^{18}\text{O}_{\text{eq}}$).** Results of the test in which
input variables underwent the same statistical procedure (see section 2.0) as the modelled data for (A-C) temperature and (D-F)
 $\delta^{18}\text{O}_{\text{eq}}$ values. Here, model input data was extracted for three fixed depths ([A & D] 5 m; [B & E] 149 m; [C & F] and 235 m) without
10 any growth weighting applied. Black regions are those grid points in which the null hypothesis (H_0), that the El Niño and Non-El
Niño populations are not statistically different ($FP_{\text{El Niño}} = FP_{\text{Non-El Niño}}$), cannot be rejected. Gray regions represent grid points where
the H_1 hypothesis is accepted, therefore the distributions of the foraminiferal population (FP) for El Niño and Non-El Niño can be
said to be unique ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$). The hatched regions represent areas where the H_1 hypothesis can be accepted, therefore
15 the distributions of the foraminiferal population (FP) for El Niño and Non-El Niño can be said to be unique ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$),
though the difference between the means of tested distribution are less than (A-C) 0.5°C or (D-F) 0.1‰. Each panel represents a
single depth (5, 149 and 235 m).

5
10
15
Figure 5.—Anderson-Darling Results for modelled FAME- $\delta^{18}\text{O}_{\text{eq}}$. Panels representing locations of where dissimilar and similar values of FAME-modelled species $\delta^{18}\text{O}$ occur between climate states, for (columns) particular species and (rows) particular model depth cut-off limits. Each panel represents the Anderson-Darling test result, which are plotted with (A) *Globigerinoides sacculifer* and (B) *Globigerinoides ruber* and without (C) *N. dutertrei* model derived error. For all panels black areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño is similar to the Non-El Niño ($\text{FP}_{\text{El Niño}} = \text{FP}_{\text{Non-El Niño}}$). The results in which the H_1 hypothesis is accepted, in which the, therefore the distributions can be said to be different ($\text{FP}_{\text{El Niño}} \neq \text{FP}_{\text{Non-El Niño}}$), are plotted as either: (A—*G. sacculifer*, B—*G. ruber*) grey and hatched or (C—*N. dutertrei*) solely as white regions. For species with calculatable error, grey regions represent values where the difference between the two means of the population is greater than species-specific standard deviation of the FAME model and hatched regions represent those in which the means are less than this standard deviation (Roche et al., 2018). For species without a calculatable error, the regions are plotted in white. The rows represent the model runs with a depth cut-off limit at: (A-C) 60 m; (D) 100 m; (E) 200 m; and (F) 400 m.

5 Figure 6. Anderson-Darling Results for modelled FAME T_c. Panels representing locations of where dissimilar and similar values of
FAME modelled temperature recorded in the calcite shells (T_c) occur between climate states, for (columns) particular species and
(rows) particular model depth cut-off limits. Each panel represents the Anderson-Darling test result, which are plotted with ([A]
10 *Globigerinoides sacculifer* and [B] *Globigerinoides ruber*) and without ([C] *N. dutertrei*) model derived error. For all panels black
areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H₀) and therefore the foraminiferal
population (FP) of the El Niño is similar to the Non-El Niño, and therefore the distribution between the neutral climate and El Niño
cannot be said to be different ($FP_{El Niño} = FP_{Non-El Niño}$). The results in which the H₁ hypothesis is accepted, in which the distributions
can be said to be different ($FP_{El Niño} \neq FP_{Non-El Niño}$), are plotted as white regions. The rows represent the model runs with a depth cut-
off limit at: (A-C) 60 m; (D) 100 m; (E) 200 m; and (F) 400 m.

Figure 7. Combined A-D plots. As figure 5 and figure 6, in that panels represent locations of where dissimilar and similar values for the two climate states for (a-d) FAME- $\delta^{18}\text{O}_{\text{eq}}$ modelled oxygen isotope values or (

5 **Figure 14 (A) Map of the sedimentation rate and bathymetry of the Tropical Pacific. (A) Inferred sedimentation rate (Olson et 2016). White regions represent continental shelf. (B) GEBCO map of height relative to 0 m with location of seamounts plotted (white stars). (C) A binary colour map of the GEBCO data, yellow is values below cut-off depth value (3500 m below sea-level (bsl)) and purple above the cut-off depth value. See Supplementary Figure 8 for variation in cut-off values.**

10 **Figure 15 Overlay between bathymetry and FAME results. The results of the FAME Anderson-Darling test for (A) temperature and (B) oxygen isotope values as input. Locations where the H_1 hypothesis can be accepted, i.e. h) FAME-T, modelled temperature recorded in the calcite shells (Tc) occur. Each panel represents the Anderson-Darling test result, the results for *Globigerinoides saeculifer*, *Globigerinoides ruber* and *N. dutertrei* are overlaid. For all panels black areas reflect latitudinal and longitudinal grid points that failed to reject the null hypothesis (H_0) and therefore the foraminiferal population (FP) of the El Niño is similar to the Non-El Niño, and therefore the distribution between the neutral climate and El Niño cannot be said to be different ($FP_{\text{El Niño}} = FP_{\text{Non-El Niño}}$). The results in which the H_1 hypothesis is accepted, in which, the distributions can be said to be different ($FP_{\text{El Niño}} \neq FP_{\text{Non-El Niño}}$), are plotted as yellow where the depth is deeper than 3500 m bsl or purple where the depth is shallower than 3500 m bsl (see Figure 82). Purple locations are where our results suggest that the signal of ENSO has different values and the water depth allows for preservation—although this purple region will be smaller when inferred SAR is taken into account (see Figure 9). The rows represent the model runs with a depth cut-off limit at: (A and E) 60 m; (B and F) 100 m; (C and G) 200 m; and (D and H) where a combination of depths were utilised (Pracht et al., 2019).**

15

20

Formatted: Font: 10 pt, Not Bold

Formatted: Justified, Line spacing: 1.5 lines

Formatted: Footer1

5

Figure 8. Bathymetric map of the Tropical Pacific Ocean highlighting the areas above and below the Lysocline and/or Calcite compensation depth (CCD). (A) GEBCO map of height relative to 0 m; (B) same as (A) with location of seamounts plotted (white stars); (C-E) binary colour map of GEBCO data, yellow is values below cut-off depth value ([C] 3500 m below sea level (bsl); [D] 4000 m bsl; and [E] 4500 m bsl respectively) and purple above the cut-off depth value. The histograms represent the normalised frequency of grid-cell height in bins of 250 m wide, yellow is values below cut-off value ([C] 3500 m below sea level (bsl); [D] 4000 m bsl; and [E] 4500 m bsl respectively), purple above cut-off value. The grey bins in each histogram are those above 0 m.

- 5 **Figure 9.** Map of the sedimentation rate and oxygen saturation for the Tropical Pacific. (A) Inferred sedimentation rate (Olsen et al. 2016). **White regions represent continental shelf.** (B) Oxygen saturation of the bottom grid layer of World Ocean Atlas 2013 (data from: <https://www.nodc.noaa.gov/cgi-bin/OC5/woa13/woa13oxnu.pl>). (C, E, G) Overlay between water depth and inferred SAR, (Olsen et al., 2016). Cut-off limits for bathymetry and SAR are 3500 m below sea-level and (A) $\geq 1 \text{ cm kyr}^{-1}$ and (B) $\geq 2 \text{ cm kyr}^{-1}$ respectively. The colours represent the following: Red / Pink: Continental shelf sediments that are (Red) shallower or (Pink) deeper than 3500 mbsl; Gray/Grey / White: grid point SAR is lower than SAR threshold and the seafloor depth is (grey) shallower or (white) deeper than 3500 mbsl; Light Yellow/Gold: Light yellow represents areas where the SAR is above the threshold but the water depth is deeper than 3500 mbsl in comparison Gold represents areas where the SAR is above the threshold and the water depth is deeper than 3500 mbsl. The ideal locations are therefore plotted as Gold. Cut-off limits for SAR are (C) $\geq 1 \text{ cm kyr}^{-1}$; (E) $\geq 2 \text{ cm kyr}^{-1}$ and (G) $\geq 5 \text{ cm kyr}^{-1}$. (D, F, H) alongside the maps the bioturbation simulations for the minimum SAR threshold is plotted (see
- 10
- 15 **Figure 10 and Figure 11** for the output of SEAMUS). Each plot gives the input values of NGRIP (grey) and for each SAR three analysis were performed with different bioturbation depths (BD) these are (Blue) 5 cm; (Green) 10 cm; and (Orange) 15 cm.

5 **Figure 10.17** Output of the bioturbation model SEAMUS. (A) The unbioturbated input signal, NGRIP (North Greenland Ice Core
Project Members, 2004; Rasmussen et al., 2014; Seierstad et al., 2014), used in our simulation of bioturbation for different SAR with
SEAMUS (Lougheed, 2019). Sediment mixed layer referred to here as bioturbation depth (BD) is fixed at (B, E, H, K) 5 cm, (C, F,
10 I, L) 10 cm and (D, G, J, M) 15 cm for sedimentation accumulation rates (SAR) of (B-D) 1 cm kyr⁻¹; (E-G) 2 cm kyr⁻¹; (H-J) 5 cm
kyr⁻¹ and (K-M) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. In (B-M) grey values represent the
unbioturbated input signal, NGRIP. Note, we retain the original units (V-SMOW) of the original timeseries used, no inference
between Pacific climate and Greenland is intended by the use of NGRIP [\(see section 2.7\)](#).

5 **Figure 14.18** Histograms of simulated specimen age within the bioturbation depth. The simulated age distribution present within the sediment mixed layer, referred to here as bioturbation depth (BD). BD is fixed at (A, D, G, J) 5 cm, (B, E, H, K) 10 cm and (C, F, I, L) 15 cm for sedimentation accumulation rates (SAR) of (A-C) 1 cm kyr⁻¹; (D-F) 2 cm kyr⁻¹; (G-I) 5 cm kyr⁻¹ and (J-L) 10 cm kyr⁻¹. The output is plotted as the discrete 1 cm depth median age. Note the size of the BD varies, therefore the simulated age distribution comes from a varying 'core depth'.

10

Formatted: Normal

Formatted: Footer1

Table Caption

Table 1. Data-model comparison.

5

References

Allen, K. A., Hönisch, B., Eggins, S. M., Haynes, L. L., Rosenthal, V. and Yu, J.: Trace element proxies for surface ocean conditions: A synthesis of culture calibrations with planktic foraminifera, *Geochimica et Cosmochimica Acta*, 193, 197–221, doi:10.1016/j.gca.2016.08.015, 2016.

- 5 An, S.-I. and Bong, H.: Inter-decadal change in El Niño–Southern Oscillation examined with Bjerknes stability index analysis, *Climate Dynamics*, 47(3–4), 967–979, doi:10.1007/s00382-015-2883-8, 2016.

An, S.-I. and Bong, H.: Feedback process responsible for the suppression of ENSO activity during the mid-Holocene, *Theoretical and Applied Climatology*, 132(3–4), 779–790, doi:10.1007/s00704-017-2117-6, 2018.

- 10 Anderson, L., Abbott, M. B., Finney, B. P. and Edwards, M. E.: Palaeohydrology of the Southwest Yukon Territory, Canada, based on multiproxy analyses of lake sediment cores from a depth transect, *The Holocene*, 15(8), 1172–1183, doi:10.1191/0959683605hl889rp, 2005.

Anderson, T. W., and Darling, D. A.: A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765–769. <http://dx.doi.org/10.2307/2281537>, 1954

- 15 Andreasen, D.H., Ravelo, A.C., Broccoli, A.J.: Remote forcing at the last glacial maximum in the tropical Pacific Ocean. *Journal of Geophysical Research—Oceans* 106, 879–897, 2001

Arrhenius, G.: Pelagic Sediments, In: Sears M (Ed) *Oceanography*. Publication 67, American Association Advancement of Science, Washington, D.C., <https://doi.org/10.5962/bhl.title.34806>, 1961

- 20 Asmerom, Y., Polyak, V., Burns, S. and Rasmussen, J.: Solar forcing of Holocene climate: New insights from a speleothem record, southwestern United States, *Geology*, 35(1), 1–4, doi:10.1130/G22865A.1, 2007.

Balmaseda, M. A., Mogensen, K. and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4, *Quarterly Journal of the Royal Meteorological Society*, 139(674), 1132–1161, 2013.

- 25 Barker, S., Greaves, M. and Elderfield, H.: A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry, *Geochem. Geophys. Geosyst.*, 4(9), 8407, 2003.

Barron, J. A. and Anderson, L.: Enhanced Late Holocene ENSO/PDO expression along the margins of the eastern North Pacific, *Quaternary International*, 235(1), 3–12, doi:<https://doi.org/10.1016/j.quaint.2010.02.026>, 2011.

- 30 Barron, J. A., Heusser, L., Herbert, T. and Lyle, M.: High-resolution climatic evolution of coastal northern California during the past 16,000 years, *Paleoceanography*, 18(1), doi:10.1029/2002PA000768, 2003.

Batiza, R.: Abundances, distribution and sizes of volcanoes in the Pacific Ocean and implications for the origin of non-hotspot volcanoes, *Earth and Planetary Science Letters*, 60, 195–206, 1982

- 35 Bé, A. W. H. and Spero, H. J.: Shell regeneration and biological recovery of planktonic foraminifera after physical injury induced in laboratory culture, *Micropaleontology*, 27(3), 305–316, 1981.

Bé, A. W. H., Spero, H. J. and Anderson, O. R.: Effects of symbiont elimination and reinfection on the life processes of the planktonic foraminifer *Globigerinoides sacculifer*, *Marine Biology*, 70, 73–86, 1982.

Formatted: Heading 11

Formatted: Footer1

- Beaufort, L., Garidel-Thoron, T., Mix, A.C., Pisias, N.G.: ENSO-like forcing on oceanic primary production during the late Pleistocene. *Science*, 293, 2440–2444, 2001.
- 5 Bemis, B. E., Spero, H. J., Bijma, J. and Lea, D. W.: Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. *Paleoceanography*, 13(2), 150–160, doi:10.1029/98PA00070, 1998.
- Bennett, J. R., Cumming, B. F., Leavitt, P. R., Chiu, M., Smol, J. P. and Szeicz, J.: Diatom, Pollen, and Chemical Evidence of Postglacial Climatic Change at Big Lake, South-Central British Columbia, Canada. *Quaternary Research*, 55(3), 332–343, doi:10.1006/qres.2001.2227, 2001.
- 10 Benson, L., Kashgarian, M., Rye, R., Lund, S., Paillet, F., Smoot, J., Kester, C., Mensing, S., Meko, D. and Lindström, S.: Holocene multidecadal and multicentennial droughts affecting Northern California and Nevada. *Quaternary Science Reviews*, 21(4), 659–682, doi:https://doi.org/10.1016/S0277-3791(01)00048-8, 2002.
- Berger, W. H.: Foraminiferal Ooze: Solution at Depths. *Science, New Series*, 156(3773), 383–385, 1967.
- 15 Berger, W. H.: Planktonic Foraminifera: selective solution and paleoclimatic interpretation. *Deep-Sea Research*, 15, 31–43, 1968.
- Berger, W. H.: Planktonic Foraminifera: Differential Production and Expatriation Off Baja California. *Limnology and Oceanography*, 15(2), 183–204, 1970a.
- 20 Berger, W. H.: Planktonic Foraminifera: Selective solution and the lysocline. *Marine Geology*, 8(2), 111–138, doi:10.1016/0025-3227(70)90001-0, 1970b.
- Berger, W. H.: Sedimentation of planktonic foraminifera. *Marine Geology*, 11(5), 325–358, doi:10.1016/0025-3227(71)90035-1, 1971.
- Berger, W. H. and Gardner, J. V.: On the determination of Pleistocene temperatures. *Journal of Foraminiferal Research*, 5(2), 102–113, 1975.
- 25 Berger, W. H., Adelseck, C., and Mayer, L.A.: Distribution of carbonate in surface sediments of the Pacific. *Journal of Geophysical Research*, 81 (15), 2617–2627, doi:10.1029/JC081i015p02617, 1976.
- Berger, W. H. and Heath, G. R.: Vertical mixing in pelagic sediments. *Journal of Marine Research*, 26(2), 134–143, 1968.
- 30 Bijma, J., Hönisch, B. and Zeebe, R. E.: Impact of the ocean carbonate chemistry on living foraminiferal shell weight: Comment on “Carbonate ion concentration in glacial-age deep waters of the Caribbean Sea” by W. S. Broecker and E. Clark: COMMENT. *Geochemistry, Geophysics, Geosystems*, 3(11), 1–7, doi:10.1029/2002GC000388, 2002.
- Bird, C., Darling, K. F., Russell, A. D., Davis, C. V., Fehrenbacher, J., Free, A., Wyman, M. and Ngwenya, B. T.: Cyanobacterial endobionts within a major marine planktonic calcifier (*Globigerina bulloides*, Foraminifera) revealed by 16S rRNA metabarcoding. *Biogeosciences*, 14(4), 901–920, doi:10.5194/bg-14-901-2017, 2017.
- 35 Bird, C., Darling, K. F., Russell, A. D., Fehrenbacher, J. S., Davis, C. V., Free, A. and Ngwenya, B. T.: 16S rRNA gene metabarcoding and TEM reveals different ecological strategies within the genus *Neogloboquadrina* (planktonic foraminifer), edited by F. Frontalini. *PLOS ONE*, 13(1), e0191653, doi:10.1371/journal.pone.0191653, 2018.

- Blackman, A. and Somayajulu, B.L.K.: Pacific Pleistocene Cores: Faunal Analyses and Geochronology, *Science*, 154 (3751), 886-889, 1966
- Boltovskoy, D.: The Sedimentary record of Pelagic Biogeography, *Progress in Oceanography*, 34(2-3), 135-160, 1994.
- 5 Boltovskoy, E.: 314. Depth at which foraminifera can survive in sediments, *Contributions from the Cushman Foundation for Foraminiferal Research*, XVII(2), 43-45, 1966.
- Boudreau, B. P.: 1998: Mean mixed depth of sediments: The wherefore and the why, *Limnology and Oceanography*, 43 (3), 524-526, 1998.
- 10 Bramlette M.N.: Pelagic sediments. In: Sears M (Ed) *Oceanography. Publication 67, American Association Advancement of Science*, Washington, D.C., <https://doi.org/10.5962/bhl.title.34806>, 1961
- Bramlette, M.N. and Bradley W.H.: Geology and Biology of North Atlantic Deep-Sea Cores between Newfoundland and Ireland: I. Lithology and Geologic Interpretation, U.S. Geological Survey Professional Paper, 196 A, 1-34, <https://pubs.usgs.gov/pp/0196/report.pdf>, 1941
- 15 Brand, W.A., Coplen, T.B., Vogl, J., Rosner, M., and Prohaska, T.: Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure and Applied Chemistry*, 86(3), 425-467, doi:10.1515/pac-2013-1023, 2014
- 20 Branson, O., Redfern, S. A. T., Tyliszczak, T., Sadekov, A., Langer, G., Kimoto, K. and Elderfield, H.: The coordination of Mg in foraminiferal calcite, *Earth and Planetary Science Letters*, 383, 134-141, doi:10.1016/j.epsl.2013.09.037, 2013.
- Branson, O., Bonnin, E. A., Perea, D. E., Spero, H. J., Zhu, Z., Winters, M., Hönisch, B., Russell, A. D., Fehrenbacher, J. S. and Gagnon, A. C.: Nanometer-Scale Chemistry of a Calcite Biomineralization Template: Implications for Skeletal Composition and Nucleation, *Proceedings of the National Academy of Sciences*, 113(46), 12934-12939, doi:10.1073/pnas.1522864113, 2016.
- 25 Caley, T., Roche, D. M., Waelbroeck, C. and Michel, E.: Oxygen stable isotopes during the Last Glacial Maximum climate: perspectives from data-model (iLOVECLIM) comparison, *Climate of the Past*, 10(6), 1939-1955, doi:10.5194/cp-10-1939-2014, 2014.
- Caramanica, A., Quilter, J., Huaman, L., Villanueva, F. and Morales, C. R.: Micro-remains, ENSO, and environmental reconstruction of El Paraiso, Peru, a late preceramic site, *Journal of Archaeological Science: Reports*, 17, 667-677, doi:<https://doi.org/10.1016/j.jasrep.2017.11.026>, 2018.
- 30 Chen, S., Hoffmann, S. S., Lund, D. C., Cobb, K. M., Emile-Geay, J. and Adkins, J. F.: A high-resolution speleothem record of western equatorial Pacific rainfall: Implications for Holocene ENSO evolution, *Earth and Planetary Science Letters*, 442, 61-71, doi:<https://doi.org/10.1016/j.epsl.2016.02.050>, 2016.
- CLIMAP Project Members: The Surface of the Ice-Age Earth, *Science*, 191(4232), 1131-1137, 1976.
- 35 Clement, A.C., Seager, R. and Cane, M.A.: Orbital controls on the El Niño/Southern Oscillation and the tropical climate, *Paleoceanography*, 14 (4), 441-456, 1999
- 40 Clouard, V. and Bonneville, A.: Ages of seamounts, islands, and plateaus on the Pacific plate, *Geological Society of America Special Paper* 388, 71-90, doi: 10.1130/2005.2388(06), 2005

- Cole, J. and Tudhope, A. W.: Reef Based Reconstructions of Eastern Pacific Climate Variability, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, pp. 535–548, Springer Netherlands, Dordrecht, 2017.
- 5 Conroy, J. L., Overpeck, J. T., Cole, J. E., Shanahan, T. M. and Steinitz-Kannan, M.: Holocene changes in eastern tropical Pacific climate inferred from a Galápagos lake sediment record, *Quaternary Science Reviews*, 27(11), 1166–1180, doi:<https://doi.org/10.1016/j.quascirev.2008.02.015>, 2008.
- Darling, K., Wade, C. M., Steward, I. A., Kroon, D., Dingle, R. and Leigh Brown, A. J.: Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers, *Nature*, 405, 43–47, 2000.
- 10 Darling, K. F., Wade, C. M., Kroon, D., Brown, A. J. L. and Bijma, J.: The Diversity and Distribution of Modern Planktic Foraminiferal Small Subunit Ribosomal RNA Genotypes and their Potential as Tracers of Present and Past Ocean Circulations, *Paleoceanography*, 14(1), 3–12, doi:[10.1029/1998PA900002](https://doi.org/10.1029/1998PA900002), 1999.
- Darling, K. F., Kucera, M., Pudsey, C. J. and Wade, C. M.: Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics, *Proceedings of the National Academy of Sciences*, 101(20), 7657–7662, doi:[10.1073/pnas.0402401101](https://doi.org/10.1073/pnas.0402401101), 2004.
- 15 Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, *Climate of the Past Discussions*, 1–31, doi:[10.5194/ep-2018-13](https://doi.org/10.5194/ep-2018-13), 2018.
- Du, X., Hendy, I. and Schimmelmann, A.: A 9000-year flood history for Southern California: A revised stratigraphy of varved sediments in Santa Barbara Basin, *Marine Geology*, 397, 29–42, doi:<https://doi.org/10.1016/j.margeo.2017.11.014>, 2018.
- 20 Dubois, N., Kienast, M., Normandeau, C. and Herbert, T. D.: Eastern equatorial Pacific cold tongue during the Last Glacial Maximum as seen from alkenone paleothermometry, *Paleoceanography*, 24(4), doi:[10.1029/2009PA001781](https://doi.org/10.1029/2009PA001781), 2009.
- Eggins, S. M., Sadokov, A. and de Deckker, P.: Modulation and daily banding of Mg/Ca in *Orbulina universa* tests by symbiont photosynthesis and respiration: a complication for seawater thermometry, *Earth and Planetary Science Letters*, 225, 411–419, 2003.
- 25 Elderfield, H. and Ganssen, G. M.: Past temperature and $\delta^{18}\text{O}$ of surface ocean waters inferred from foraminiferal Mg/Ca ratios, *Nature*, 405(6785), 442–445, 2000.
- Emiliani, C.: Depth habitats of some species of pelagic foraminifera as indicated by oxygen isotope ratios, *American Journal of Science*, 252(3), 149–158, doi:[10.2475/ajs.252.3.149](https://doi.org/10.2475/ajs.252.3.149), 1954.
- Emiliani, C.: Pleistocene Temperatures, *The Journal of Geology*, 63(6), 538–578, 1955.
- 30 Enzel, Y. and Wells, S. G.: Extracting Holocene paleohydrology and paleoclimatology information from modern extreme flood events: An example from southern California, *Geomorphology*, 19(3), 203–226, doi:[https://doi.org/10.1016/S0169-555X\(97\)00015-9](https://doi.org/10.1016/S0169-555X(97)00015-9), 1997.
- Eriksen, D. B. and Wollin, G.: Pleistocene climates and chronology in deep-sea sediments, *Science*, 162, 1227–1234, 1968.
- Eriksen, D. B., Ewing, M. and Wollin, G.: The Pleistocene Epoch in Deep-Sea Sediments, *Science, New Series*, 146(3645), 723–732, 1964.

- Evans, D., Müller, W. and Erez, J.: Assessing foraminifera biomineralisation models through trace element data of cultures under variable seawater chemistry, *Geochimica et Cosmochimica Acta*, 236, 198–217, doi:10.1016/j.gca.2018.02.048, 2018.
- Evans, M. N., Kaplan, A. and Cane, M. A.: Optimal sites for coral-based reconstructions of global sea surface temperature, *Paleoceanography*, 13 (5), 502–516, 1998.
- 5 Ewing, M., Heezen, B.C., and Eriksen, D.: Significance of the Worzel deep-sea ash, *PNAS*, 45, 355, 1959
- Fallet, U., Boer, W., van Assen, C., Greaves, M. and Brummer, G. J. A.: A novel application of wet-oxidation to retrieve carbonates from large, organic-rich samples for application in climate research, *Geochemistry-Geophysics-Geosystems*, 10(8), 2009.
- 10 Feldberg, M.J., Mix, A.C.: Planktonic foraminifera, sea surface temperatures, and mechanisms of oceanic change in the Peru and south equatorial currents, 0–150 ka BP. *Paleoceanography* 18, 1–16, 2003
- Feldmeijer, W., Metcalfe, B., Scussolini, P. and Arthur, K.: The effect of chemical pretreatment of sediment upon foraminiferal-based proxies, *Geochemistry, Geophysics, Geosystems*, 14(10), 3996–4014, doi:10.1002/ggge.20233, 2013.
- 15 Feldmeijer, W., Metcalfe, B., Brummer, G. J. A. and Ganssen, G. M.: Reconstructing the depth of the permanent thermocline through the morphology and geochemistry of the deep-dwelling planktonic foraminifer *Globorotalia truncatulinoides*, *Paleoceanography*, 30(1), 1–22, doi:10.1002/2014PA002687, 2015.
- Filliben, J.J.: The Probability-Plot Correlation Coefficient Test for Normality, *Technometrics*, 17 (1), 111–117, 1975
- 20 Ford, H. L., Ravelo, A. C. and Polissar, P. J.: Reduced El Niño-Southern Oscillation during the Last Glacial Maximum, *Science*, 347(6219), 255–258, doi:10.1126/science.1258437, 2015.
- Fraile, I., Schulz, M., Mulitza, S. and Kucera, M.: Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, *Biogeosciences*, 5, 891–911, 2008.
- 25 Fraile, I., Schulz, M., Mulitza, S., Merkel, U., Prange, M. and Paul, A.: Modeling the seasonal distribution of planktonic foraminifera during the Last Glacial Maximum, *Paleoceanography*, 24, doi:10.1029/2008PA001686, 2009.
- Ganssen, G. M., Peeters, F. J. C., Metcalfe, B., Anand, P., Jung, S. J. A., Kroon, D. and Brummer, G. J. A.: Quantifying sea surface temperature ranges of the Arabian Sea for the past 20 000 years, *Climate of the Past*, 7(4), 1337–1349, doi:10.5194/cp-7-1337-2011, 2011.
- 30 Garidel-Thoron, T. de, Rosenthal, Y., Beaufort, L., Bard, E., Sonzogni, C. and Mix, A. C.: A multiproxy assessment of the western equatorial Pacific hydrography during the last 30 kyr, *Paleoceanography*, 22(3), doi:10.1029/2006PA001269, 2007.
- Gray, W. R., Weldeab, S., Lea, D. W., Rosenthal, Y., Gruber, N., Donner, B. and Fischer, G.: The effects of temperature, salinity, and the carbonate system on Mg/Ca in *Globigerinoides ruber* (white): A global sediment trap calibration, *Earth and Planetary Science Letters*, 482, 607–620, doi:10.1016/j.epsl.2017.11.026, 2018.
- 35 Greaves, M., Barker, S., Daunt, C. and Elderfield, H.: Accuracy, standardization, and interlaboratory calibration standards for foraminiferal Mg/Ca thermometry, *Geochem.-Geophys.-Geosyst.*, 6(2), Q02D13, 2005.

- Groeneveld, J., Nürnberg, D., Tiedemann, R., Reichert, G. J., Steph, S., Reuning, L., Crudeli, D. and Mason, P.: Foraminiferal Mg/Ca increase in the Caribbean during the Pliocene: Western Atlantic Warm Pool formation, salinity influence, or diagenetic overprint?, *Geochemistry Geophysics Geosystems*, 9, Q01P23, doi:doi:10.1029/2006GC001564, 2008.
- 5 Hamilton, C. P., Spero, H. J., Bijma, J. and Lea, D. W.: Geochemical investigation of gametogenic calcite addition in the planktonic foraminifera *Orbulina universa*, *Marine Micropaleontology*, 68(3–4), 256–267, 2008.
- Hays, J.D., Saito, T., Opdyke, N., and Burekle, L.H.: Pliocene-Pleistocene Sediments of the Equatorial Pacific: Their Paleomagnetic, Biostratigraphic, and Climatic Record, *Geological Society of America Bulletin*, 80, 1481–1514, 1969
- 10 Hecht, A. D.: A model for determining Pleistocene paleotemperatures from planktonic foraminiferal assemblages, *Micropaleontology*, 19, 68–77, 1973.
- Hendy, I. L., Napiet, T. J. and Schimmelmann, A.: From extreme rainfall to drought: 250 years of annually resolved sediment deposition in Santa Barbara Basin, California, *Quaternary International*, 387, 3–12, doi:https://doi.org/10.1016/j.quaint.2015.01.026, 2015.
- 15 Higley, M. C., Conroy, J. L. and Schmitt, S.: Last Millennium Meridional Shifts in Hydroclimate in the Central Tropical Pacific, *Paleoceanography and Paleoclimatology*, 33(4), 354–366, doi:10.1002/2017PA003233, 2018.
- Hillier, J. K.: Pacific seamount volcanism in space and time, *Geophys. J. Int.*, 168, 877–889, doi: 10.1111/j.1365-246X.2006.03250.x, 2007
- 20 Hori, M., Shirai, K., Kimoto, K., Kurasawa, A., Takagi, H., Ishida, A., Takahata, N. and Sano, Y.: Chamber formation and trace element distribution in the calcite walls of laboratory cultured planktonic foraminifera (*Globigerina bulloides* and *Globigerinoides ruber*), *Marine Micropaleontology*, 140, 46–55, doi:10.1016/j.marmicro.2017.12.004, 2018.
- Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S. and Zhang, H. M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Interecomparisons, *Journal of Climate*, 30(20), 8179–8205, doi:10.1175/jcli-d-16-0836.1, 2017.
- 25 Huber, B. T., Bijma, J. and Darling, K.: Cryptic speciation in the living planktonic foraminifer *Globigerinella siphonifera* (d'Orbigny), *Paleobiology*, 23(01), 33–62, doi:10.1017/S0094837300016638, 1997.
- Huber, M. and Caballero, R.: Eocene El Niño: Evidence for Robust Tropical Dynamics in the “Hothouse”, *Science*, 299 (5608), 877–881, 2003
- 30 Hut, G.: Consultants group meeting on stable isotope reference samples for geochemical and hydrological investigations, Int. At. Energy Agency, Vienna., 1987.
- Hutson, W. H.: Transfer functions under no-analog conditions: experiments with Indian Ocean planktonic foraminifera, *Quaternary Research*, 8(3), 355–367, 1977.
- 35 Hutson, W. H.: Applications of transfer functions to Indian Ocean planktonic foraminifera, *Quaternary Research*, 9, 87–112, 1978.
- Hutson, W. H.: Bioturbation of deep-sea sediments: Oxygen isotopes and stratigraphic uncertainty, *Geology*, 8(3), 127–130, doi:10.1130/0091-7613(1980), 1980a.

- Hutson, W. H.: The Agulhas Current during the Late Pleistocene—analysis of modern faunal analogs, *Science*, 207(4426), 64–66, 1980b.
- Imbrie, J. and Kipp, N. G.: A new Micropaleontological method for paleoclimatology: Application to a Late Pleistocene Caribbean core, in *The Late Cenozoic Glacial Ages*, edited by K. K. Turekian, pp. 71–181, Yale University Press, New Haven, Connecticut., 1971.
- Jacob, D. E., Wirth, R., Agbaje, O. B. A., Branson, O. and Eggins, S. M.: Planktic foraminifera form their shells via metastable carbonate phases, *Nature Communications*, 8(1), 1265, doi:10.1038/s41467-017-00955-0, 2017.
- Jones, J. I.: Distribution of living planktonic foraminifera of the West Indies and adjacent waters, PhD Thesis, University Wisconsin-Madison, Madison, Wisconsin USA., 1964.
- Jonkers, L. and Kučera, M.: Quantifying the effect of seasonal and vertical habitat tracking on planktonic foraminifera proxies, *Climate of the Past*, 13(6), 573–586, doi:10.5194/ep-13-573-2017, 2017.
- Kageyama, M., Braconnot, P., Bopp, L., Caubel, A., Foujols, M. A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., Mariotti, V., Marti, O., Roy, T. and Woillez, M. N.: Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model—part I: comparing IPSL_CM5A to IPSL_CM4, *Climate Dynamics*, 40(9), 2447–2468, doi:10.1007/s00382-012-1488-8, 2013.
- Keigwin, L. D. and Guilderson, T. P.: Bioturbation artifacts in zero-age sediments, *Paleoceanography*, 24, PA4212, doi:10.1029/2008PA001727, 2009.
- Kienast, M., MacIntyre, G., Dubois, N., Higginson, S., Normandeau, C., Chazen, C. and Herbert, T. D.: Alkenone unsaturation in surface sediments from the eastern equatorial Pacific: Implications for SST reconstructions, *Paleoceanography*, 27(1), doi:10.1029/2011PA002254, 2012.
- Kim, S.-T. and O’Neil, J. R.: Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates, *Geochimica et Cosmochimica Acta*, 61(16), 3461–3475, doi:10.1016/S0016-7037(97)00169-5, 1997.
- Kisakürek, B., Eisenhauer, A., Böhm, F., Garbe-Schönberg, D. and Erez, J.: Controls on shell Mg/Ca and Sr/Ca in cultured planktonic foraminiferan, *Globigerinoides ruber* (white), *Earth and Planetary Science Letters*, 273(3–4), 260–269, doi:10.1016/j.epsl.2008.06.026, 2008.
- Koppers A.A.P., Staudigel H., Pringle M.S., Wijbrans J.R.: Short lived and discontinuous intraplate volcanism in the South Pacific; hot spots or extensional volcanism, *Geochem. Geophys. Geosys.*, 4 (1089), 2003.
- Koutavas, A. and Joannides, S.: El Niño–Southern Oscillation extrema in the Holocene and Last Glacial Maximum, *Paleoceanography*, 27(4), doi:10.1029/2012PA002378, 2012.
- Koutavas, A. and Lynch-Stieglitz, J.: Glacial-interglacial dynamics of the eastern equatorial Pacific cold tongue–Intertropical Convergence Zone system reconstructed from oxygen isotope records: GLACIAL–PACIFIC–COLD–TONGUE, *Paleoceanography*, 18(4), n/a–n/a, doi:10.1029/2003PA000894, 2003.
- Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. and Sachs, J. P.: El Niño-Like Pattern in Ice-Age Tropical Pacific Sea Surface Temperature, *Science*, 297(5579), 226, doi:10.1126/science.1072376, 2002.

- Koutavas, A., deMenocal, P. B., Olive, G. C. and Lynch-Stieglitz, J.: Mid-Holocene El Niño–Southern Oscillation (ENSO) attenuation revealed by individual foraminifera in eastern tropical Pacific sediments, *Geology*, 34(12), 993, doi:10.1130/G22810A.1, 2006.
- 5 Kretschmer, K., Kucera, M. and Schulz, M.: Modeling the distribution and seasonality of *Neogloboquadrina pachyderma* in the North Atlantic Ocean during Heinrich Stadial 1, *Paleoceanography*, 31(7), 986–1010, doi:10.1002/2015PA002819, 2016.
- Kretschmer, K., Jonkers, L., Kucera, M. and Schulz, M.: Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale, *Biogeosciences Discussions*, 1–37, doi:10.5194/bg-2017-429, 2017.
- Leduc, G., Vidal, L., Cartapanis, O. and Bard, E.: Modes of eastern equatorial Pacific thermocline variability: Implications for ENSO dynamics over the last glacial period: ENSO DYNAMICS OVER THE LAST 50 KA, *Paleoceanography*, 24(3), doi:10.1029/2008PA001701, 2009.
- 10 LeGrande, A. N. and Schmidt, G. A.: Global gridded data set of the oxygen isotopic composition in seawater, *Geophysical Research Letters*, 33(12), L12604, 2006.
- Lombard, F., Labeyrie, L., Michel, E., Spero, H. J. and Lea, D. W.: Modelling the temperature dependent growth rates of planktic foraminifera, *Marine Micropalaeontology*, 70, 1–7, 2009.
- 15 Lombard, F., da Rocha, R. E., Bijma, J. and Gattuso, J. P.: Effect of carbonate ion concentration and irradiance on calcification in planktonic foraminifera, *Biogeosciences*, 7(1), 247–255, 2010.
- Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailliau, S., Howa, H. and Jorissen, F.: Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, *Biogeosciences*, 8(4), 853–873, doi:10.5194/bg-8-853-2011, 2011.
- 20 Loubere, P., Creamer, W. and Haas, J.: Evolution of the El Niño–Southern Oscillation in the late Holocene and insolation driven change in the tropical annual SST cycle, *Global and Planetary Change*, 100, 129–144, doi:https://doi.org/10.1016/j.gloplacha.2012.10.007, 2013.
- Lougheed, B. C.: SEAMUS (v1.0): a $\Delta 14C$ -enabled, single-specimen sediment accumulation simulator, *Geoscientific Model Development*, 2019.
- Lougheed, B. C., Metcalfe, B., Ninnemann, U. S. and Wacker, L.: Moving beyond the age depth model paradigm in deep-sea palaeoclimate archives: dual radiocarbon and stable isotope analysis on single foraminifera, *Climate of the Past*, 14(4), 515–526, 2018.
- 25 Löwemark, L.: Importance and Usefulness of Trace fossils and Bioturbation in Paleoceanography, in *Trace Fossils: Concepts, Problems, Prospects*, edited by W. Miller, pp. 413–427, Elsevier, Amsterdam., 2007.
- Löwemark, L. and Grootes, P. M.: Large age differences between planktic foraminifera caused by abundance variations and Zoophycos bioturbation, *Paleoceanography*, 19, PA2001, doi:10.1029/2003PA000949, 2004.
- 30 Löwemark, L., Hong, W.-L., Yui, T.-F. and Hung, G.-W.: A test of different factors influencing the isotopic signal of planktonic foraminifera in surface sediments from the northern South China Sea, *Marine Micropaleontology*, 55(1–2), 49–62, doi:10.1016/j.marmicro.2005.02.004, 2005.

Löwemark, L., Konstantinou, K. I. and Steinke, S.: Bias in foraminiferal multispecies reconstructions of paleohydrographic conditions caused by foraminiferal abundance variations and bioturbational mixing: A model approach, *Marine Geology*, 256, 101–106, 2008.

5 Luz, B.: Stratigraphic and Paleoclimatic Analysis of late Pleistocene Tropical Southeast Pacific Cores, *Quaternary Research*, 3, 56–72, 1973

Lynts, G. W. and Judd, J. B.: Late Pleistocene paleotemperatures at Tongue of the Ocean, Bahamas, *Science*, 171, 1143–1144, 1971.

10 MARGO Project Members*: Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum, *Nature Geoscience*, 2(2), 127–132, doi:10.1038/ngeo411, 2009.

McIntyre, A., Kipp, N. G., Bé, A. W. H., Crowley, T., Kellogg, T., Gardner, J. V., Prell, W. L. and Ruddiman, W. F.: Glacial North Atlantic 18,000 years ago: A CLIMAP reconstruction, *Geological Society of America Memoirs*, 145, 43–76, 1976.

McPhaden, M. J., Zebiak, S. E., and Glantz, M. H.: ENSO as an Integrating Concept in Earth Science, *Science*, 314 (5806), 1740–1745, 2006

15 Menard H. W.: *Marine Geology of the Pacific*, McGraw-Hill, New York, 1964

Metcalfe, B., Feldmeijer, W., de Vringer Picon, M., Brummer, G. J. A., Peeters, F. J. C. and Ganssen, G. M.: Late Pleistocene glacial-interglacial shell size isotope variability in planktonic foraminifera as a function of local hydrography, *Biogeosciences*, 12(15), 4781–4807, doi:10.5194/bg-12-4781-2015, 2015.

20 Metcalfe, B., Lougheed, B. C., Waelbroeck, C., and Roche D. M.: DATA: Data set, Zenodo, <https://doi.org/10.5281/zenodo.2554843>, 2019

Metcalfe, B., Feldmeijer, W., and Ganssen, G. M.: Oxygen isotope variability of planktonic foraminifera provide clues to past upper ocean seasonal variability, *Paleoceanography and Paleoclimatology*, doi: 10.1029/2018PA003475, 2019

25 Mikis, A., Hendry, K. R., Pike, J., Schmidt, D. N., Edgar, K. M., Peek, V., Peeters, F. J. C., Leng, M. J., Meredith, M. P., Todd, C. L., Stammerjohn, S., and Ducklow, H.: Temporal variability in foraminiferal morphology and geochemistry at the West Antarctic Peninsula: a sediment trap study, *Biogeosciences Discuss.*, <https://doi.org/10.5194/bg-2019-19>, in review, 2019.

30 Mix, A. C.: The oxygen isotope record of deglaciation, in: *North America and adjacent oceans during the last deglaciation*, in *The Geology of America*, vol. K-3, edited by W. F. Ruddiman and H. E. J. Wright, pp. 111–135, Geological Society of America, Boulder, Colorado., 1987.

35 Mix, A. C.: Running hot and cold in the eastern equatorial Pacific, *Quaternary Science Reviews*, 25, 1147–1149, doi:10.1016/j.quascirev.2006.03.008, 2006

Morard, R., Quillévéré, F., Escarguel, G., de Garidel-Thoron, T., de Vargas, C. and Kucera, M.: Ecological modeling of the temperature dependence of cryptic species of planktonic Foraminifera in the Southern Hemisphere, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 391, 13–33, doi:10.1016/j.palaeo.2013.05.011, 2013.

40 Morard, R., Reinelt, M., Chiessi, C. M., Groeneveld, J. and Kucera, M.: Tracing shifts of oceanic fronts using the cryptic diversity of the planktonic foraminifera *Globorotalia inflata*, *Paleoceanography*, 31(9), 1193–1205, doi:10.1002/2016PA002977, 2016.

- Moy, C. M., Seltzer, G. O., Rodbell, D. T. and Anderson, D. M.: Variability of El Niño/Southern Oscillation activity at millennial timescales during the Holocene epoch, *Nature*, 420, 162, doi:10.1038/nature01194, 2002.
- Mulitza, S., Wolff, T., Pätzold, J., Hale, W. and Wefer, G.: Temperature sensitivity of planktic foraminifera and its influence on the oxygen isotope record, *Marine Micropaleontology*, 33(3–4), 223–240, doi:10.1016/S0377-8398(97)00040-6, 1998.
- 5 Murray, J.; Renard, A.F.: Deep-sea deposits (based on the specimens collected during the voyage of HMS Challenger in the years 1872 to 1876). Report on the scientific results of the voyage of H.M.S. Challenger during the years 1873–76; John Menzies and Co., Edinburgh, United Kingdom, 1891
- Nederbragt, A. J. and Thurow, J. W.: A 6000yr varve record of Holocene climate in Saanich Inlet, British Columbia, from digital sediment colour analysis of ODP Leg 169S cores, *Marine Geology*, 174(1), 95–110, doi:https://doi.org/10.1016/S0025-3227(00)00144-4, 2001.
- 10 Nederbragt, A. J. and Thurow, J. W.: Amplitude of ENSO cycles in the Santa Barbara basin, off California, during the past 15,000 years, *Journal of Quaternary Science*, 20, 447–456, 2006.
- North Greenland Ice Core Project Members: High-resolution record of Northern Hemisphere climate extending into the Last Interglacial period, *Nature*, 431, 147–151, 2004
- 15 de Nooijer, L. J., Toyofuku, T., Oguri, K., Nomaki, H. and Kitazato, H.: Intracellular pH distribution in foraminifera determined by the fluorescent probe HPTS: Intracellular pH of foraminifera, *Limnology and Oceanography: Methods*, 6(11), 610–618, doi:10.4319/lom.2008.6.610, 2008.
- 20 de Nooijer, L. J., Toyofuku, T. and Kitazato, H.: Foraminifera promote calcification by elevating their intracellular pH, *Proceedings of the National Academy of Sciences*, 106(36), 15374–15378, doi:10.1073/pnas.0904306106, 2009.
- Nürnberg, D., Bijma, J. and Hemleben, C.: Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures, *Geochimica et Cosmochimica Acta*, 60(5), 803–814, 1996.
- 25 Olson, P., Reynolds, E., Hinnov, L. and Goswami, A.: Variation of ocean sediment thickness with crustal age, *Geochem. Geophys. Geosy.*, 17, 1349–1369, 2016.
- O'Neil, J. R., Clayton, R. N. and Mayeda, T. K.: Oxygen Isotope Fractionation in Divalent Metal Carbonates, *The Journal of Chemical Physics*, 51(12), 5547–5558, 1969.
- Parker, F. L.: Eastern Mediterranean foraminifera, *Rep. Swedish Deep-Sea Exped.*, 8(4), 219–283, 1958.
- Parker, F.L. and Berger, W. H.: Faunal and solution patterns of planktonic Foraminifera in surface sediments of the South Pacific, *Deep-Sea Research*, 18, 73–107, 1971
- 30 Patterson, D. T., Prokoph, A., Wright, C., Chang, A. S., Thomson, R. E. and Ware, D. M.: Holocene Solar Variability and Pelagic Fish Productivity in the NE Pacific, *Palaeontologia Electronica*, 7(4), 17, 2004.
- Peeters, F. J. C., Acheson, R., Brummer, G. J. A., de Ruijter, W. P. M., Schneider, R., Ganssen, G. M., Ufkes, E. and D., K.: Vigorous exchange between the Indian and Atlantic oceans at the end of the past five glacial periods, *Nature*, 430, 661–665, 2004.
- 35

- Pena, L. D., Cacho, I., Ferretti, P. and Hall, M. A.: El Niño–Southern Oscillation–like variability during glacial terminations and interlatitudinal teleconnections, *Paleoceanography*, 23(3), doi:10.1029/2008PA001620, 2008.
- Peng, T. H., Broecker, W. S. and Berger, W. H.: Rates of benthic mixing in deep-sea sediment as determined by radioactive tracers, *Quaternary Res.*, 11, 141–149, 1979.
- 5 Phleger, F. B., Parker, F. L. and Pierson, J. F.: North Atlantic Foraminifera, Swedish Deep-Sea Expedition 1947–1948, 7, 3–9, 1953.
- Pisias, N. G. and Mix, A. C.: Aliasing of the geologic record and the search for long-period Milankovitch cycles, *Paleoceanography*, 3(5), 613–619, doi:10.1029/PA0034005p00613, 1988.
- Pisias, N.G., Mix, A.C., 1997. Spatial and temporal oceanographic variability of the eastern equatorial Pacific during the late
10 Pleistocene: evidence from Radiolaria microfossils. *Paleoceanography* 12, 381–393.
- Praecht, H., Metcalfe, B. and Peeters, F. J. C.: Oxygen isotope composition of final chamber of planktic foraminifera provides evidence for vertical migration and depth integrated growth, *Biogeosciences Discussions*, 1–32, doi:10.5194/bg-2018-146, 2018.
- 15 Rasmussen, S.O., M. Bigler, S. Blockley, T. Blunier, B. Buchardt, H. Clausen, I. Cvijanovic, D. Dahl-Jensen, S. Johnsen, H. Fischer, V. Gkinis, M. Guillevic, W. Hoek, J. Lowe, J. Pedro, T. Popp, I. Seierstad, J. Steffensen, A. Svensson, P. Vallelonga, B. Vinther, M. Walker, J.J. Wheatley, M. Winstrup: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy, *Quaternary Science Reviews*, 106, 14–27, 2014
- 20 Raven, J., Caldeira, K., Elderfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C. and Watson, A.: Ocean acidification due to increasing atmospheric carbon dioxide., *The Royal Society*,, 2005.
- Reynolds, R. W. and Smith, T. M.: A high-resolution global sea surface temperature climatology, *Journal of Climate*, 8, 1571–1583, 1995.
- Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C. and Wang, W.: An Improved In Situ and Satellite SST Analysis for Climate, *Journal of Climate*, 15(13), 1609–1625, doi:10.1175/1520-0442(2002)015<1609:aaisas>2.0.co;2, 2002.
- Rincón-Martínez, D., Steph, S., Lamy, F., Mix, A. and Tiedemann, R.: Tracking the equatorial front in the eastern equatorial Pacific Ocean by the isotopic and faunal composition of planktonic foraminifera, *Marine Micropaleontology*, 79(1), 24–40, doi:10.1016/j.marmicro.2011.01.001, 2011.
- 30 Rink, S., Kühl, M., Bijma, J. and Spero, H. J.: Microsensor studies of photosynthesis and respiration in the symbiotic foraminifer *Orbulina universa*, *Marine Biology*, 131, 583–595, 1998.
- Roehe, D. M., Paillard, D., Caley, T. and Waelbroeck, C.: LGM hosing approach to Heinrich Event 1: results and perspectives from data-model integration using water isotopes, *Quaternary Science Reviews*, 106, 247–261, doi:10.1016/j.quascirev.2014.07.020, 2014.
- 35 Roehe, D. M., Waelbroeck, C., Metcalfe, B. and Caley, T.: FAME (v1.0): a simple module to simulate the effect of planktonic foraminifer species-specific habitat on their oxygen isotopic content, *Geoscientific Model Development*, 1–22, doi:10.5194/gmd-2017-251, 2018.
- Rosenthal, Y. and Broccoli, A. J.: In search of Paleo-ENSO, *Science*, 304 (5668), 219–221, 2004

Roy, T., Lombard, F., Bopp, L. and Gehlen, M.: Projected impacts of climate change and ocean acidification on the global biogeography of planktonic Foraminifera, *Biogeosciences*, 12(10), 2873–2889, doi:10.5194/bg-12-2873-2015, 2015.

5 Ruddiman, W. F.: Pleistocene Sedimentation in the Equatorial Atlantic: Stratigraphy and Faunal Paleoclimatology, *Geological Society of America Bulletin*, 82, 283–302, 1971.

Sadekov, A., Eggins, S. M., De Deckker, P. and Kroon, D.: Uncertainties in seawater thermometry deriving from intratest and intertest Mg/Ca variability in *Globigerinoides ruber*: UNCERTAINTIES Mg/Ca SEAWATER THERMOMETRY, *Paleoceanography*, 23(1), n/a–n/a, doi:10.1029/2007PA001452, 2008.

10 Sadekov, A., Eggins, S. M., De Deckker, P., Ninnemann, U., Kuhnt, W. and Bassinot, F.: Surface and subsurface seawater temperature reconstruction using Mg/Ca microanalysis of planktonic foraminifera *Globigerinoides ruber*, *Globigerinoides saeculifer*, and *Pulleniatina obliquiloculata*: SEAWATER TEMPERATURE RECONSTRUCTION, *Paleoceanography*, 24(3), doi:10.1029/2008PA001664, 2009.

15 Sadekov, A. Y., Ganeshram, R., Pichevin, L., Berdin, R., McClymont, E., Elderfield, H. and Tudhope, A. W.: Palaeoclimate reconstructions reveal a strong link between El Niño–Southern Oscillation and Tropical Pacific mean state, *Nature Communications*, 4, 2692, doi:10.1038/ncomms3692, 2013.

Schiebel, R., Barker, Stephen, Lentz, R., Thomas, H. and Bollmann, J.: Planktic foraminiferal dissolution in the twilight zone: Deep-Sea Research. Part II. Topical studies in oceanography 54 (5–7), 676–686. 10.1016/j.dsr2.2007.01.009, 2007

20 Schott, W.: Zur klimaschichtung der tiefseesedimente im aquatorialen atlantischen ozean, *Geologische Rundschau*, 40(1), 20–31, 1952.

Schott, W.: Foraminiferenfauna und stratigraphie der tiefsee-sedimente in Nordatlantischen ozean, *Rep. Swedish Deep Sea Exped.*, 7(8), 357–469, 1966.

25 van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F.J.C., Biastoch, A., Weijer, W., Turney, C., Paris, C.B., and Zahn R.: Ocean currents generate large footprints in marine palaeoclimate proxies, *Nature Communications*, 6, 6521 doi: 10.1038/ncomms7521, 2015

30 Seierstad et al.: Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale $\delta^{18}\text{O}$ gradients with possible Heinrich event imprint, *Quaternary Science Reviews*, 106, 29–46, doi:10.1016/j.quascirev.2014.10.032, 2014

Skrivanek, A. and Hendy, I. L.: A 500 year climate catch: Pelagic fish scales and paleoproductivity in the Santa Barbara Basin from the Medieval Climate Anomaly to the Little Ice Age (AD 1000–1500), *Quaternary International*, 387, 36–45, doi:https://doi.org/10.1016/j.quaint.2015.07.044, 2015.

35 Spero, H. J.: SYMBIOSIS IN THE PLANKTONIC FORAMINIFER, *ORBULINA UNIVERSA*, AND THE ISOLATION OF ITS SYMBIOTIC DINOFLAGELLATE, *GYMNODINIUM BÉH* SP. NOV. [†], *Journal of Phyeology*, 23, 307–317, doi:10.1111/j.1529-8817.1987.tb04139.x, 1987.

Spero, H. J. and DeNiro, M. J.: The influence of symbiont photosynthesis on the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values of planktonic foraminiferal shell calcite, *Symbiosis*, 4(1–3), 213–228, 1987.

- Spero, H. J. and Lea, D. W.: Experimental determination of stable isotope variability in *Globigerina bulloides*: implications for paleoceanographic reconstructions, *Marine Micropaleontology*, 28(3–4), 231–246, doi:10.1016/0377-8398(96)00003-5, 1996.
- 5 Spero, H. J., Bijma, J., Lea, D. W. and Bemis, B. E.: Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes, *Nature*, 390(6659), 497–500, doi:10.1038/37333, 1997.
- Staines-Urías, F., González-Yajimovich, O. and Beaufort, L.: Reconstruction of past climate variability and ENSO-like fluctuations in the southern Gulf of California (Alfonso Basin) since the last glacial maximum, *Quaternary Research*, 83(3), 488–501, doi:https://doi.org/10.1016/j.yqres.2015.03.007, 2015.
- Steinhardt, J., Cleroux, C., de Nooijer, L.J., Brummer, G. J.A., Zahn, R., Ganssen, G.M., and Reichert, G. J.: Reconciling 10 single-chamber Mg/Ca with whole-shell $\delta^{18}\text{O}$ in surface to deep-dwelling planktonic foraminifera from the Mozambique Channel, *Biogeosciences*, 12, 2411–2429, https://doi.org/10.5194/bg-12-2411-2015, 2015
- Stott, L., Poulsen, C., Lund, S. and Thunell, R.: Super ENSO and Global Climate Oscillations at Millennial Time Scales, *Science*, 297(5579), 222, doi:10.1126/science.1071627, 2002.
- 15 Stott, L., Cannariato, K., Thunell, R., Haug, G. H., Koutavas, A. and Lund, S.: Decline of surface temperature and salinity in the western tropical Pacific Ocean in the Holocene epoch, *Nature*, 431(7004), 56–59, doi:10.1038/nature02903, 2004.
- Sverdrup, H.U., Johnson, M.W. and Fleming, R.H.: *The Oceans, Their Physics, Chemistry, and General Biology*. Prentice-Hall, New York, 1942
- 20 Telford, R., Li, C., and Kucera, M.: Mismatch between the depth habitat of planktonic foraminifera and the calibration depth of SST transfer functions may bias reconstructions. *Climate of the Past*, 9, 859–870, https://doi.org/10.5194/cp-9-859-2013, 2013
- Thirumalai, K., Partin, J. W., Jackson, C. S. and Quinn, T. M.: Statistical constraints on El Niño Southern Oscillation reconstructions using individual foraminifera: A sensitivity analysis, *Paleoceanography*, 28(3), 401–412, doi:10.1002/palo.20037, 2013.
- Thunell, R.C., Kier, R.S., and Honjo, S.: Calcite dissolution: an in situ study in the Panama Basin, *Science*, 212 (4495), 659–661, 1981
- 30 Trenberth, K.E. and Otto-Bliesner, B.L.: Toward Integrated Reconstruction of past Climates, *Science*, 300 (5619), 589–591, 2003
- de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W. and Pawlowski, J.: Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces, *Proceedings of the National Academy of Sciences*, 96(6), 2864–2868, doi:10.1073/pnas.96.6.2864, 1999.
- 35 de Vargas, C., Bonzon, M., Rees, N. W., Pawlowski, J. and Zaninetti, L.: A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (dâ€™Orbigny), *Marine Micropaleontology*, 45(2), 101–116, 2002.

- Vetter, L., Kozdon, R., Mora, C. I., Eggins, S. M., Valley, J. W., Hönisch, B. and Spero, H. J.: Micron-scale intrashell oxygen isotope variation in cultured planktic foraminifers, *Geochimica et Cosmochimica Acta*, 107, 267–278, doi:10.1016/j.gca.2012.12.046, 2013.
- 5 Wang, J.: The coupling of the Annual Cycle and ENSO over the Tropical Pacific, *Journal of the Atmospheric Sciences*, 51(8), 1115–1136, [https://doi.org/10.1175/1520-0469\(1994\)051<1115:TCOTAC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1994)051<1115:TCOTAC>2.0.CO;2), 1994
- Wang, C., Deser, C., Yu, J. Y., DiNezio, P. and Clement, A.: El Niño and Southern Oscillation (ENSO): A Review, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, pp. 85–106, Springer Netherlands, Dordrecht., 2017.
- 10 Waterson, A. M., Edgar, K. M., Schmidt, D. N. and Valdes, P. J.: Quantifying the stability of planktic foraminiferal physical niches between the Holocene and Last Glacial Maximum, *Paleoceanography*, doi:10.1002/2016PA002964, 2016.
- Wessel, P.: Sizes and Ages of Seamounts using remote sensing: Implications for Intraplate Volcanism, *Science*, 277, 802–805, 1997
- 15 Wessel, P., Lyons, S.: Distribution of large Pacific seamounts from Geosat/ERS-1, *Journal of Geophysical Research*, 102, 459–475, 1997
- Weyl, P.K.: Micropaleontology and Ocean Surface Climate, *Science*, 202 (4367), 475–481, 1978
- 20 White, S. M., Ravelo, A. C. and Polissar, P. J.: Dampened El Niño in the Early and Mid-Holocene Due To Insolation-Forced Warming/Deepening of the Thermocline, *Geophysical Research Letters*, 45(1), 316–326, doi:10.1002/2017GL075433, 2018.
- Wilke, I., Peeters, F. J. C., Bickert, T.: The influence of seawater carbonate ion concentration $[CO_3^{2-}]$ on the stable isotope composition of planktic foraminifera species *Globorotalia inflata*, *Marine Micropaleontology*, 58(4), 243–258, <https://doi.org/10.1016/j.marmicro.2005.11.005>, 2006
- 25 Williams, D. F.: Late Quarternary fluctuations of the polar front and the subtropical convergence in the southeast Indian Ocean, *Marine Micropaleontology*, 1, 363–375, 1976.
- Williams, D. F. and Johnson, W. C.: Diversity of Recent planktonic foraminifera in the southern Indian Ocean and Late Pleistocene paleotemperatures, *Quarternary Research*, 5, 237–250, 1975.
- 30 Wise, S.: Calcite compensation depth. In: *Sedimentology. Encyclopedia of Earth Science*. Springer, Berlin, Heidelberg, doi: 10.1007/3-540-31079-7, 1978
- Wit, J. C., Reichert, G. J. and Ganssen, G. M.: Unmixing of stable isotope signals using single specimen $\delta^{18}O$ analyses, *Geochemistry, Geophysics, Geosystems*, 14(4), 1312–1320, doi:10.1002/ggge.20101, 2013.
- 35 Wolf-Gladrow, D. A., Bijma, J. and Zeebe, R. E.: Model simulation of the carbonate chemistry in the microenvironment of symbiont-bearing foraminifera, *Marine Chemistry*, 64(3), 181–198, doi:10.1016/S0304-4203(98)00074-7, 1999.
- Worzel, J.L.: Extensive deep-sea sub-bottom reflections identified as white ash, *PNAS*, 45, 349, 1959

- Wunsch, C.: On sharp spectral lines in the climate record and the millennial peak, *Paleoceanography*, 15(4), 417–424, doi:10.1029/1999PA000468, 2000.
- Wunsch, C. and Gunn, D. E.: A densely sampled core and climate variable aliasing, *Geo-Marine Letters*, 23(1), 64–71, doi:10.1007/s00367-003-0125-2, 2003.
- 5 Wyceeh, J. B., Kelly, D. C., Kitajima, K., Kozdon, R., Orland, I. J. and Valley, J. W.: Combined effects of gametogenic calcification and dissolution on $\delta^{18}\text{O}$ measurements of the planktic foraminifer *Trilobatus sacculifer*, *Geochem. Geophys. Geosyst.*, doi:10.1029/2018GC007908, 2018.
- Xie, S. P.: On the Genesis of the Equatorial Annual Cycle, *Journal of Climate*, 7, 2008–2013, 1994
- 10 Xie, S. P.: Interaction between the annual and interannual variations in the equatorial Pacific, *Journal of Physical Oceanography*, 25, 1930–1941, 1995
- Žarić, S., Donner, B., Fischer, G., Mulitza, S. and Wefer, G.: Sensitivity of planktic foraminifera to sea surface temperature and export production as derived from sediment trap data, *Marine Micropaleontology*, 55(1–2), 75–105, doi:10.1016/j.marmicro.2005.01.002, 2005.
- 15 Žarić, S., Schulz, M. and Mulitza, S.: Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data, *Biogeosciences*, 3(2), 187–207, doi:10.5194/bg-3-187-2006, 2006.
- Zebiak, S. E. and M. A. Cane: A Model El Niño Southern Oscillation, *Monthly Weather Review*, 115, 2262–2278, [https://doi.org/10.1175/15200493\(1987\)115<2262:AMENO>2.0.CO;2](https://doi.org/10.1175/15200493(1987)115<2262:AMENO>2.0.CO;2), 1987
- 20 Zeebe, R. E. and Sanyal, A.: Comparison of two potential strategies of planktonic foraminifera for house building: Mg^{2+} or H^{+} removal?, *Geochimica et Cosmochimica Acta*, 66(7), 1159–1169, doi:10.1016/S0016-7037(01)00852-3, 2002.
- Zeebe, R. E. and Wolf-Gladrow, D. A.: *CO₂ in seawater: Equilibrium, Kinetics, Isotopes*, 1st ed., Elsevier, Amsterdam, 2001.
- 25 Zhang, Z., Leduc, G. and Sachs, J. P.: El Niño evolution during the Holocene revealed by a biomarker rain gauge in the Galápagos Islands, *Earth and Planetary Science Letters*, 404, 420–434, doi:https://doi.org/10.1016/j.epsl.2014.07.013, 2014.
- Zhang, S., Li, T., Chang, F., Yu, Z., Xiong, Z., and Wang, H.: Correspondence between the ENSO-like state and glacial-interglacial condition during the past 360 kyr, *Chinese Journal of Oceanology and Limnology*, 35 (5), 1018–103, <http://dx.doi.org/10.1007/s00343-017-6082-9>, 2017
- 30 Zhu, J., Liu, Z., Brady, E., Otto-Bliesner, B., Zhang, J., Noone, D., Tomas, R., Nusbaumer, J., Wong, T., Jahn, A. and Tabor, C.: Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model, *Geophysical Research Letters*, 44(13), 6984–6992, doi:10.1002/2017GL073406, 2017a.
- 35 Zhu, Z., Feinberg, J. M., Xie, S., Bourne, M. D., Huang, C., Hu, C. and Cheng, H.: Holocene ENSO-related cyclic storms recorded by magnetic minerals in speleothems of central China, *Proc Natl Acad Sci USA*, 114(5), 852, doi:10.1073/pnas.1610930114, 2017b. Allen, K. A., Hönisch, B., Eggins, S. M., Haynes, L. L., Rosenthal, Y. and Yu, J.: Trace element proxies for surface ocean conditions: A synthesis of culture calibrations with planktic foraminifera, *Geochimica et Cosmochimica Acta*, 193, 197–221, doi:10.1016/j.gca.2016.08.015, 2016.

- An, S.-I. and Bong, H.: Inter-decadal change in El Niño-Southern Oscillation examined with Bjerknes stability index analysis. *Climate Dynamics*, 47(3–4), 967–979, doi:10.1007/s00382-015-2883-8, 2016.
- An, S.-I. and Bong, H.: Feedback process responsible for the suppression of ENSO activity during the mid-Holocene. *Theoretical and Applied Climatology*, 132(3–4), 779–790, doi:10.1007/s00704-017-2117-6, 2018.
- 5 Anderson, T. W., and Darling, D. A.: A test of goodness of fit. *Journal of the American Statistical Association*, 49, 765-769, <http://dx.doi.org/10.2307/2281537>, 1954
- Andreasen, D.H., Ravelo, A.C., Broccoli, A.J.: Remote forcing at the last glacial maximum in the tropical Pacific Ocean. *Journal of Geophysical Research - Oceans* 106, 879–897, 2001
- 10 Balmaseda, M. A., Mogensen, K. and Weaver, A. T.: Evaluation of the ECMWF ocean reanalysis system ORAS4. *Quarterly Journal of the Royal Meteorological Society*, 139(674), 1132–1161, 2013.
- Barker, S., Greaves, M. and Elderfield, H.: A study of cleaning procedures used for foraminiferal Mg/Ca paleothermometry. *Geochem. Geophys. Geosyst.*, 4(9), 8407, 2003.
- 15 Batiza, R.: Abundances, distribution and sizes of volcanoes in the Pacific Ocean and implications for the origin of non-hotspot volcanoes. *Earth and Planetary Science Letters*, 60, 195-206, 1982
- Bé, A. W. H. and Spero, H. J.: Shell regeneration and biological recovery of planktonic foraminifera after physical injury induced in laboratory culture. *Micropaleontology*, 27(3), 305–316, 1981.
- 20 Bé, A. W. H., Spero, H. J. and Anderson, O. R.: Effects of symbiont elimination and reinfection on the life processes of the planktonic foraminifer *Globigerinoides sacculifer*. *Marine Biology*, 70, 73–86, 1982.
- Beaufort, L., Garidel-Thoron, T., Mix, A.C., Pisias, N.G.: ENSO like forcing on oceanic primary production during the late Pleistocene. *Science* 293, 2440–2444, 2001
- 25 Bemis, B. E., Spero, H. J., Bijma, J. and Lea, D. W.: Reevaluation of the oxygen isotopic composition of planktonic foraminifera: Experimental results and revised paleotemperature equations. *Paleoceanography*, 13(2), 150–160, doi:10.1029/98PA00070, 1998.
- Berger W. H.: Planktonic Foraminifera: selective solution and paleoclimatic interpretation. *Deep-Sea Research*, 15, 31–43, 1968
- 30 Berger, W. H.: Sedimentation of planktonic foraminifera. *Marine Geology*, 11(5), 325–358, doi:10.1016/0025-3227(71)90035-1, 1971.
- Berger, W.H., Adelseck, C., and Mayer, L.A.: Distribution of carbonate in surface sediments of the Pacific. *Journal of Geophysical Research*, 81 (15), 2617-2627, doi:10.1029/JC081i015p02617, 1976
- 35 Berger, W. H. and Heath, G. R.: Vertical mixing in pelagic sediments. *Journal of Marine Research*, 26(2), 134–143, 1968.
- Bijma, J., Hönisch, B. and Zeebe, R. E.: Impact of the ocean carbonate chemistry on living foraminiferal shell weight: Comment on “Carbonate ion concentration in glacial-age deep waters of the Caribbean Sea” by W. S. Broecker and E. Clark: COMMENT, *Geochemistry, Geophysics, Geosystems*, 3(11), 1–7, doi:10.1029/2002GC000388, 2002.

- Bird, C., Darling, K. F., Russell, A. D., Davis, C. V., Fehrenbacher, J., Free, A., Wyman, M. and Ngwenya, B. T.: Cyanobacterial endobionts within a major marine planktonic calcifier (*Globigerina bulloides*, Foraminifera) revealed by 16S rRNA metabarcoding, *Biogeosciences*, 14(4), 901–920, doi:10.5194/bg-14-901-2017, 2017.
- 5 Bird, C., Darling, K. F., Russell, A. D., Fehrenbacher, J. S., Davis, C. V., Free, A. and Ngwenya, B. T.: 16S rRNA gene metabarcoding and TEM reveals different ecological strategies within the genus *Neogloboquadrina* (planktonic foraminifer), edited by F. Frontalini, *PLOS ONE*, 13(1), e0191653, doi:10.1371/journal.pone.0191653, 2018.
- Blackman, A. and Somayajulu, B.L.K.: Pacific Pleistocene Cores: Faunal Analyses and Geochronology, *Science*, 154 (3751), 886-889, 1966
- 10 Boudreau, B. P. 1998: Mean mixed depth of sediments: The wherefore and the why, *Limnology and Oceanography*, 43 (3), 524-526, 1998.
- Bramlette M.N.: Pelagic sediments. In: Sears M (Ed) *Oceanography*, Publication 67, American Association Advancement of Science, Washington, D.C., <https://doi.org/10.5962/bhl.title.34806>, 1961
- 15 Brand, W.A., Coplen, T.B., Vogl, J., Rosner, M., and Prohaska, T.: Assessment of international reference materials for isotope-ratio analysis (IUPAC Technical Report). *Pure and Applied Chemistry*, 86(3), 425-467, doi:10.1515/pac-2013-1023, 2014
- 20 Caley, T., Roche, D. M., Waelbroeck, C. and Michel, E.: Oxygen stable isotopes during the Last Glacial Maximum climate: perspectives from data–model (iLOVECLIM) comparison, *Climate of the Past*, 10(6), 1939–1955, doi:10.5194/cp-10-1939-2014, 2014.
- Carton, J.A., Chepurin, G.A., Cao, X. and Giese, B.S.: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950-1995, Part 1: methodology, *J. Phys. Oceanogr.*, 30, 294-309, 2000a
- 25 Carton, J.A., Chepurin, G.A., and Cao, X.: A Simple Ocean Data Assimilation analysis of the global upper ocean 1950-1995 Part 2: results, *J. Phys. Oceanogr.*, 30, 311-326, 2000b
- Clement, A.C., Seager, R. and Cane, M.A.: Orbital controls on the El Niño /Southern Oscillation and the tropical climate, *Paleoceanography*, 14 (4), 441-456, 1999
- 30 Clouard, V. and Bonneville, A.: Ages of seamounts, islands, and plateaus on the Pacific plate, *Geological Society of America Special Paper* 388, 71-90, doi: 10.1130/2005.2388(06), 2005
- Coplen, T.B.: Reporting of stable carbon, hydrogen, and oxygen isotopic abundances, Reference and intercomparison materials for stable isotopes of light elements, IAEA, Vienna, pp 159
- 35 Darling, K., Wade, C. M., Steward, I. A., Kroon, D., Dingle, R. and Leigh Brown, A. J.: Molecular evidence for genetic mixing of Arctic and Antarctic subpolar populations of planktonic foraminifers, *Nature*, 405, 43–47, 2000.
- Darling, K. F., Wade, C. M., Kroon, D., Brown, A. J. L. and Bijma, J.: The Diversity and Distribution of Modern Planktic Foraminiferal Small Subunit Ribosomal RNA Genotypes and their Potential as Tracers of Present and Past Ocean Circulations, *Paleoceanography*, 14(1), 3–12, doi:10.1029/1998PA900002, 1999.
- 40 Darling, K. F., Kucera, M., Pudsey, C. J. and Wade, C. M.: Molecular evidence links cryptic diversification in polar planktonic protists to Quaternary climate dynamics, *Proceedings of the National Academy of Sciences*, 101(20), 7657–7662, doi:10.1073/pnas.0402401101, 2004.

- Deuser, W.G., Ross, E.H., and Anderson, R.F.: Seasonality in the supply of sediment to the deep Sargasso Sea and implications for the rapid transfer of matter to the deep ocean, *Deep Sea Research Part A: Oceanographic Research Papers*, 28 (5), 495-505, doi: 10.1016/0198-0149(81)90140-0, 1981
- 5 Dolman, A. M. and Laepple, T.: Sedproxy: a forward model for sediment archived climate proxies, *Climate of the Past Discussions*, 1–31, doi:10.5194/cp-2018-13, 2018.
- Dubois, N., Kienast, M., Normandeau, C. and Herbert, T. D.: Eastern equatorial Pacific cold tongue during the Last Glacial Maximum as seen from alkenone paleothermometry, *Paleoceanography*, 24(4), doi:10.1029/2009PA001781, 2009.
- 10 Eggins, S. M., Sadekov, A. and de Deckker, P.: Modulation and daily banding of Mg/Ca in *Orbulina universa* tests by symbiont photosynthesis and respiration: a complication for seawater thermometry, *Earth and Planetary Science Letters*, 225, 411–419, 2003.
- Elderfield, H. and Ganssen, G. M.: Past temperature and $\delta^{18}\text{O}$ of surface ocean waters inferred from foraminiferal Mg/Ca ratios, *Nature*, 405(6785), 442–445, 2000.
- 15 Evans, M. N., Kaplan, A. and Cane, M. A.: Optimal sites for coral-based reconstructions of global sea surface temperature, *Paleoceanography*, 13 (5), 502-516, 1998
- Evans, M. N., Tolwinski-Ward, S. E., Thompson D. M., and Anchukaitis K. J.: Applications of proxy system modeling in high resolution paleoclimatology, *Quaternary Science Reviews*, doi:10.1016/j.quascirev.2013.05.024, 2013
- 20 Fallet, U., Boer, W., van Assen, C., Greaves, M. and Brummer, G. J. A.: A novel application of wet-oxidation to retrieve carbonates from large, organic-rich samples for application in climate research, *Geochemistry Geophysics Geosystems*, 10(8), 2009.
- Feldberg, M.J., Mix, A.C.: Planktonic foraminifera, sea surface temperatures, and mechanisms of oceanic change in the Peru and south equatorial currents, 0–150 ka BP. *Paleoceanography* 18, 1–16, 2003
- 25 Feldmeijer, W., Metcalfe, B., Scussolini, P. and Arthur, K.: The effect of chemical pretreatment of sediment upon foraminiferal-based proxies, *Geochemistry, Geophysics, Geosystems*, 14(10), 3996–4014, doi:10.1002/ggge.20233, 2013.
- Feldmeijer, W., Metcalfe, B., Brummer, G. J. A. and Ganssen, G. M.: Reconstructing the depth of the permanent thermocline through the morphology and geochemistry of the deep dwelling planktonic foraminifer *Globorotalia truncatulinoides*, *Paleoceanography*, 30(1), 1–22, doi:10.1002/2014PA002687, 2015.
- 30 Ford, H. L., Ravelo, A. C. and Polissar, P. J.: Reduced El Niño-Southern Oscillation during the Last Glacial Maximum, *Science*, 347(6219), 255–258, doi:10.1126/science.1258437, 2015.
- Fraass, A. J., and Lowery, C. M.: Defining uncertainty and error in planktic foraminiferal oxygen isotope measurements, *Paleoceanography*, 32, 104– 122, doi:10.1002/2016PA003035, 2017
- 35 Fraile, I., Schulz, M., Mulitza, S. and Kucera, M.: Predicting the global distribution of planktonic foraminifera using a dynamic ecosystem model, *Biogeosciences*, 5, 891–911, 2008.
- Fraile, I., Schulz, M., Mulitza, S., Merkel, U., Prange, M. and Paul, A.: Modeling the seasonal distribution of planktonic foraminifera during the Last Glacial Maximum, *Paleoceanography*, 24, doi:10.1029/2008PA001686, 2009.

- Garidel-Thoron, T. de, Rosenthal, Y., Beaufort, L., Bard, E., Sonzogni, C. and Mix, A. C.: A multiproxy assessment of the western equatorial Pacific hydrography during the last 30 kyr, *Paleoceanography*, 22(3), doi:10.1029/2006PA001269, 2007.
- Gray, W. R., Weldeab, S., Lea, D. W., Rosenthal, Y., Gruber, N., Donner, B. and Fischer, G.: The effects of temperature, salinity, and the carbonate system on Mg/Ca in *Globigerinoides ruber* (white): A global sediment trap calibration, *Earth and Planetary Science Letters*, 482, 607–620, doi:10.1016/j.epsl.2017.11.026, 2018.
- Greaves, M., Barker, S., Daunt, C. and Elderfield, H.: Accuracy, standardization, and interlaboratory calibration standards for foraminiferal Mg/Ca thermometry, *Geochem. Geophys. Geosyst.*, 6(2), Q02D13, 2005.
- Groeneveld, J., Nürnberg, D., Tiedemann, R., Reichert, G.-J., Steph, S., Reuning, L., Crudeli, D. and Mason, P.: Foraminiferal Mg/Ca increase in the Caribbean during the Pliocene: Western Atlantic Warm Pool formation, salinity influence, or diagenetic overprint?, *Geochemistry Geophysics Geosystems*, 9, Q01P23, doi:doi:10.1029/2006GC001564, 2008.
- Hamilton, C. P., Spero, H. J., Bijma, J. and Lea, D. W.: Geochemical investigation of gametogenic calcite addition in the planktonic foraminifera *Orbulina universa*, *Marine Micropaleontology*, 68(3–4), 256–267, 2008.
- Hays, J.D., Saito, T., Opdyke, N., and Burckle, L.H.: Pliocene-Pleistocene Sediments of the Equatorial Pacific: Their Paleomagnetic, Biostratigraphic, and Climatic Record, *Geological Society of America Bulletin*, 80, 1481-1514, 1969.
- Hillier, J. K.: Pacific seamount volcanism in space and time, *Geophys. J. Int.*, 168, 877-889, doi: 10.1111/j.1365-246X.2006.03250.x, 2007.
- Hori, M., Shirai, K., Kimoto, K., Kurasawa, A., Takagi, H., Ishida, A., Takahata, N. and Sano, Y.: Chamber formation and trace element distribution in the calcite walls of laboratory cultured planktonic foraminifera (*Globigerina bulloides* and *Globigerinoides ruber*), *Marine Micropaleontology*, 140, 46–55, doi:10.1016/j.marmicro.2017.12.004, 2018.
- Huang, B., Thorne, P. W., Banzon, V. F., Boyer, T., Chepurin, G., Lawrimore, J. H., Menne, M. J., Smith, T. M., Vose, R. S. and Zhang, H.-M.: Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons, *Journal of Climate*, 30(20), 8179–8205, doi:10.1175/jcli-d-16-0836.1, 2017.
- Huber, B. T., Bijma, J. and Darling, K.: Cryptic speciation in the living planktonic foraminifer *Globigerinella siphonifera* (d'Orbigny), *Paleobiology*, 23(01), 33–62, doi:10.1017/S0094837300016638, 1997.
- Huber, M. and Caballero, R.: Eocene El Niño: Evidence for Robust Tropical Dynamics in the “Hothouse”, *Science*, 299 (5608), 877-881, 2003.
- Hut, G.: Consultants group meeting on stable isotope reference samples for geochemical and hydrological investigations, *Int. At. Energy Agency, Vienna*, 1987.
- Jacob, D. E., Wirth, R., Agbaje, O. B. A., Branson, O. and Eggins, S. M.: Planktic foraminifera form their shells via metastable carbonate phases, *Nature Communications*, 8(1), 1265, doi:10.1038/s41467-017-00955-0, 2017.
- Kageyama, M., Braconnot, P., Bopp, L., Caubel, A., Foujols, M.-A., Guilyardi, E., Khodri, M., Lloyd, J., Lombard, F., Mariotti, V., Marti, O., Roy, T. and Woillez, M.-N.: Mid-Holocene and Last Glacial Maximum climate simulations with the IPSL model—part I: comparing IPSL_CM5A to IPSL_CM4, *Climate Dynamics*, 40(9), 2447–2468, doi:10.1007/s00382-012-1488-8, 2013.
- Keigwin, L. D. and Guilderson, T. P.: Bioturbation artifacts in zero-age sediments, *Paleoceanography*, 24, PA4212, doi: 10.1029/2008PA001727, 2009.

- Kienast, M., MacIntyre, G., Dubois, N., Higginson, S., Normandeau, C., Chazen, C. and Herbert, T. D.: Alkenone unsaturation in surface sediments from the eastern equatorial Pacific: Implications for SST reconstructions, *Paleoceanography*, 27(1), doi:10.1029/2011PA002254, 2012.
- 5 [Kim, S.-T. and O'Neil, J. R.: Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates, *Geochimica et Cosmochimica Acta*, 61\(16\), 3461–3475, doi:10.1016/S0016-7037\(97\)00169-5, 1997.](#)
- [Kisakürek, B., Eisenhauer, A., Böhm, F., Garbe-Schönberg, D. and Erez, J.: Controls on shell Mg/Ca and Sr/Ca in cultured planktonic foraminiferan, *Globigerinoides ruber* \(white\), *Earth and Planetary Science Letters*, 273\(3–4\), 260–269, doi:10.1016/j.epsl.2008.06.026, 2008.](#)
- 10 [Koppers A.A.P., Staudigel H., Pringle M.S., Wijbrans J.R.: Short-lived and discontinuous intraplate volcanism in the South Pacific: hot spots or extensional volcanism, *Geochem. Geophys.*, 4 \(1089\), 2003](#)
- [Koutavas, A. and Joanides, S.: El Niño–Southern Oscillation extrema in the Holocene and Last Glacial Maximum, *Paleoceanography*, 27\(4\), doi:10.1029/2012PA002378, 2012.](#)
- 15 [Koutavas, A. and Lynch-Stieglitz, J.: Glacial-interglacial dynamics of the eastern equatorial Pacific cold tongue-Intertropical Convergence Zone system reconstructed from oxygen isotope records: GLACIAL PACIFIC COLD TONGUE, *Paleoceanography*, 18\(4\), n/a-n/a, doi:10.1029/2003PA000894, 2003.](#)
- [Koutavas, A., Lynch-Stieglitz, J., Marchitto, T. M. and Sachs, J. P.: El Niño-Like Pattern in Ice Age Tropical Pacific Sea Surface Temperature, *Science*, 297\(5579\), 226, doi:10.1126/science.1072376, 2002.](#)
- 20 [Koutavas, A., deMenocal, P. B., Olive, G. C. and Lynch-Stieglitz, J.: Mid-Holocene El Niño–Southern Oscillation \(ENSO\) attenuation revealed by individual foraminifera in eastern tropical Pacific sediments, *Geology*, 34\(12\), 993, doi:10.1130/G22810A.1, 2006.](#)
- [Kretschmer, K., Kucera, M. and Schulz, M.: Modeling the distribution and seasonality of *Neogloboquadrina pachyderma* in the North Atlantic Ocean during Heinrich Stadial 1, *Paleoceanography*, 31\(7\), 986–1010, doi:10.1002/2015PA002819, 2016.](#)
- 25 [Kretschmer, K., Jonkers, L., Kucera, M. and Schulz, M.: Modeling seasonal and vertical habitats of planktonic foraminifera on a global scale, *Biogeosciences Discussions*, 1–37, doi:10.5194/bg-2017-429, 2017.](#)
- [Leduc, G., Vidal, L., Cartapanis, O. and Bard, E.: Modes of eastern equatorial Pacific thermocline variability: Implications for ENSO dynamics over the last glacial period: ENSO DYNAMICS OVER THE LAST 50 KA, *Paleoceanography*, 24\(3\), doi:10.1029/2008PA001701, 2009.](#)
- 30 [LeGrande, A. N. and Schmidt, G. A.: Global gridded data set of the oxygen isotopic composition in seawater, *Geophysical Research Letters*, 33\(12\), L12604, 2006.](#)
- [Lombard, F., Labeyrie, L., Michel, E., Spero, H. J. and Lea, D. W.: Modelling the temperature dependent growth rates of planktic foraminifera, *Marine Micropalaeontology*, 70, 1–7, 2009.](#)
- 35 [Lombard, F., da Rocha, R. E., Bijma, J. and Gattuso, J. P.: Effect of carbonate ion concentration and irradiance on calcification in planktonic foraminifera, *Biogeosciences*, 7\(1\), 247–255, 2010.](#)

Lombard, F., Labeyrie, L., Michel, E., Bopp, L., Cortijo, E., Retailleau, S., Howa, H. and Jorissen, F.: Modelling planktic foraminifer growth and distribution using an ecophysiological multi-species approach, *Biogeosciences*, 8(4), 853–873, doi:10.5194/bg-8-853-2011, 2011.

Lougheed, B. C.: SEAMUS (v1.0): a $\Delta^{14}\text{C}$ -enabled, single-specimen sediment accumulation simulator, *Geoscientific Model Development*, 2019.

5 Lougheed, B. C., Metcalfe, B., Ninnemann, U. S. and Wacker, L.: Moving beyond the age depth model paradigm in deep-sea palaeoclimate archives: dual radiocarbon and stable isotope analysis on single foraminifera, *Climate of the Past*, 14(4), 515–526, 2018.

Martínez, I., Keigwin, L., Barrows, T. T., Yokoyama, Y., and Southon, J.: La Niña-like conditions in the eastern equatorial Pacific and a stronger Choco jet in the northern Andes during the last glaciation, *Paleoceanography*, 18, 1033, doi:10.1029/2002PA000877, 2, doi: 10.1029/2002PA000877, 2003

10 McPhaden, M.J., Zebiak, S.E., and Glantz, M.H.: ENSO as an Integrating Concept in Earth Science, *Science*, 314 (5806), 1740-1745, 2006

Menard H.W.: *Marine Geology of the Pacific*, McGraw-Hill, New York, 1964

15 Metcalfe, B., Feldmeijer, W., de Vringer-Picon, M., Brummer, G.-J. A., Peeters, F. J. C. and Ganssen, G. M.: Late Pleistocene glacial–interglacial shell-size–isotope variability in planktonic foraminifera as a function of local hydrography, *Biogeosciences*, 12(15), 4781–4807, doi:10.5194/bg-12-4781-2015, 2015.

20 Metcalfe, B., Lougheed, B.C., Waelbroeck, C., and Roche D.M.: DATA: Data set, Zenodo, <https://doi.org/10.5281/zenodo.2554843>, 2019

Metcalfe, B., Feldmeijer, W., and Ganssen, G.M.: Oxygen isotope variability of planktonic foraminifera provide clues to past upper ocean seasonal variability, *Paleoceanography and Paleoclimatology*, doi: 10.1029/2018PA003475, 2019

25 Mix, A. C.: The oxygen-isotope record of deglaciation, in: *North America and adjacent oceans during the last deglaciation*, in *The Geology of America*, vol. K-3, edited by W. F. Ruddiman and H. E. J. Wright, pp. 111–135, Geological Society of America, Boulder, Colorado., 1987.

30 Morard, R., Quillévéré, F., Escarguel, G., de Garidel-Thoron, T., de Vargas, C. and Kucera, M.: Ecological modeling of the temperature dependence of cryptic species of planktonic Foraminifera in the Southern Hemisphere, *Palaeogeography, Palaeoclimatology, Palaeoecology*, 391, 13–33, doi:10.1016/j.palaeo.2013.05.011, 2013.

Morard, R., Reinelt, M., Chiessi, C. M., Groeneveld, J. and Kucera, M.: Tracing shifts of oceanic fronts using the cryptic diversity of the planktonic foraminifera *Globorotalia inflata*, *Paleoceanography*, 31(9), 1193–1205, doi:10.1002/2016PA002977, 2016.

35 Mulitza, S., Wolff, T., Pätzold, J., Hale, W. and Wefer, G.: Temperature sensitivity of planktic foraminifera and its influence on the oxygen isotope record, *Marine Micropaleontology*, 33(3–4), 223–240, doi:10.1016/S0377-8398(97)00040-6, 1998.

Murray, J; Renard, AF: Deep-sea deposits (based on the specimens collected during the voyage of HMS Challenger in the years 1872 to 1876). Report on the scientific results of the voyage of H.M.S. Challenger during the years 1873-76; John Menzies and Co., Edinburgh, United Kingdom, 1891

North Greenland Ice Core Project Members: High-resolution record of Northern Hemisphere climate extending into the Last Interglacial period, *Nature*, 431, 147-151, 2004

5 de Nooijer, L. J., Toyofuku, T., Oguri, K., Nomaki, H. and Kitazato, H.: Intracellular pH distribution in foraminifera determined by the fluorescent probe HPTS: Intracellular pH of foraminifera, *Limnology and Oceanography: Methods*, 6(11), 610–618, doi:10.4319/lom.2008.6.610, 2008.

de Nooijer, L. J., Toyofuku, T. and Kitazato, H.: Foraminifera promote calcification by elevating their intracellular pH, *Proceedings of the National Academy of Sciences*, 106(36), 15374–15378, doi:10.1073/pnas.0904306106, 2009.

10 Nürnberg, D., Bijma, J. and Hemleben, C.: Assessing the reliability of magnesium in foraminiferal calcite as a proxy for water mass temperatures, *Geochimica et Cosmochimica Acta*, 60(5), 803–814, 1996.

Olson, P., Reynolds, E., Hinnov, L. and Goswami, A.: Variation of ocean sediment thickness with crustal age, *Geochem. Geophys. Geosy.*, 17, 1349–1369, 2016.

O'Neil, J. R., Clayton, R. N. and Maveda, T. K.: Oxygen Isotope Fractionation in Divalent Metal Carbonates, *The Journal of Chemical Physics*, 51(12), 5547–5558, 1969.

15 Parker, F.L. and Berger, W. H.: Faunal and solution patterns of planktonic Foraminifera in surface sediments of the South Pacific, *Deep-Sea Research*, 18, 73-107, 1971

Pena, L. D., Cacho, I., Ferretti, P. and Hall, M. A.: El Niño–Southern Oscillation–like variability during glacial terminations and interlatitudinal teleconnections, *Paleoceanography*, 23(3), doi:10.1029/2008PA001620, 2008.

20 Peng, T.-H., Broecker, W. S. and Berger, W. H.: Rates of benthic mixing in deep-sea sediment as determined by radioactive tracers, *Quaternary Res.*, 11, 141–149, 1979.

Pisias, N. G. and Mix, A. C.: Aliasing of the geologic record and the search for long-period Milankovitch cycles, *Paleoceanography*, 3(5), 613–619, doi:10.1029/PA003i005p00613, 1988.

25 Pisias, N.G., Mix, A.C., 1997. Spatial and temporal oceanographic variability of the eastern equatorial Pacific during the late Pleistocene: evidence from Radiolaria microfossils. *Paleoceanography* 12, 381–393.

Pracht, H., Metcalfe, B. and Peeters, F. J. C.: Oxygen isotope composition of final chamber of planktic foraminifera provides evidence for vertical migration and depth integrated growth, *Biogeosciences*, 1–32, doi:10.5194/bg-2018-146, 2019.

30 Rasmussen, S.O., M. Bigler, S. Blockley, T. Blunier, B. Buchardt, H. Clausen, I. Cvijanovic, D. Dahl-Jensen, S. Johnsen, H. Fischer, V. Gkinis, M. Guillevic, W. Hoek, J. Lowe, J. Pedro, T. Popp, I. Seierstad, J. Steffensen, A. Svensson, P. Vallelonga, B. Vinther, M. Walker, J.J. Wheatley, M. Winstrup: A stratigraphic framework for abrupt climatic changes during the Last Glacial period based on three synchronized Greenland ice-core records: refining and extending the INTIMATE event stratigraphy, *Quaternary Science Reviews*, 106, 14-27, 2014

35 Rincón-Martínez, D., Steph, S., Lamy, F., Mix, A. and Tiedemann, R.: Tracking the equatorial front in the eastern equatorial Pacific Ocean by the isotopic and faunal composition of planktonic foraminifera, *Marine Micropaleontology*, 79(1), 24–40, doi:10.1016/j.marmicro.2011.01.001, 2011.

Rink, S., Kühl, M., Bijma, J. and Spero, H. J.: Microsensor studies of photosynthesis and respiration in the symbiotic foraminifer *Orbulina universa*, *Marine Biology*, 131, 583–595, 1998.

- Roche, D. M., Paillard, D., Caley, T. and Waelbroeck, C.: LGM hosing approach to Heinrich Event 1: results and perspectives from data-model integration using water isotopes, *Quaternary Science Reviews*, 106, 247–261, doi:10.1016/j.quascirev.2014.07.020, 2014.
- 5 Roche, D. M., Waelbroeck, C., Metcalfe, B. and Caley, T.: FAME (v1.0): a simple module to simulate the effect of planktonic foraminifer species-specific habitat on their oxygen isotopic content, *Geoscientific Model Development*, 1–22, doi:10.5194/gmd-2017-251, 2018.
- Rosenthal, Y. and Broccoli, A. J.: In search of Paleo-ENSO, *Science*, 304 (5668), 219–221, 2004
- 10 Roy, T., Lombard, F., Bopp, L. and Gehlen, M.: Projected impacts of climate change and ocean acidification on the global biogeography of planktonic Foraminifera, *Biogeosciences*, 12(10), 2873–2889, doi:10.5194/bg-12-2873-2015, 2015.
- Sadekov, A., Eggins, S. M., De Deckker, P. and Kroon, D.: Uncertainties in seawater thermometry deriving from intratest and intertest Mg/Ca variability in *Globigerinoides ruber*: Uncertainties Mg/Ca seawater thermometry *Paleoceanography*, 23(1), n/a-n/a, doi:10.1029/2007PA001452, 2008.
- 15 Sadekov, A., Eggins, S. M., De Deckker, P., Ninnemann, U., Kuhnt, W. and Bassinot, F.: Surface and subsurface seawater temperature reconstruction using Mg/Ca microanalysis of planktonic foraminifera *Globigerinoides ruber*, *Globigerinoides sacculifer*, and *Pulleniatina obliquiloculata*: Seawater temperature reconstruction, *Paleoceanography*, 24(3), doi:10.1029/2008PA001664, 2009.
- 20 Sadekov, A. Y., Ganeshram, R., Pichevin, L., Berdin, R., McClymont, E., Elderfield, H. and Tudhope, A. W.: Palaeoclimate reconstructions reveal a strong link between El Niño-Southern Oscillation and Tropical Pacific mean state, *Nature Communications*, 4, 2692, doi:10.1038/ncomms3692, 2013.
- Schiebel, R., Barker, Stephen, Lendt, R., Thomas, H. and Bollmann, J.: Planktic foraminiferal dissolution in the twilight zone. *Deep-Sea Research. Part II. Topical studies in oceanography* 54 (5-7), 676–686. 10.1016/j.dsr2.2007.01.009, 2007
- 25 van Sebille, E., Scussolini, P., Durgadoo, J. V., Peeters, F.J.C., Biastoch, A., Weijer, W., Turney, C., Paris, C.B., and Zahn R.: Ocean currents generate large footprints in marine palaeoclimate proxies, *Nature Communications*, 6, 6521 doi:10.1038/ncomms7521, 2015
- 30 Seierstad et al.: Consistently dated records from the Greenland GRIP, GISP2 and NGRIP ice cores for the past 104 ka reveal regional millennial-scale $\delta^{18}\text{O}$ gradients with possible Heinrich event imprint, *Quaternary Science Reviews*, 106, 29–46, doi:10.1016/j.quascirev.2014.10.032, 2014
- Spero, H. J. and DeNiro, M. J.: The influence of symbiont photosynthesis on the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ values of planktonic foraminiferal shell calcite, *Symbiosis*, 4(1–3), 213–228, 1987.
- 35 Spero, H. J. and Lea, D. W.: Experimental determination of stable isotope variability in *Globigerina bulloides*: implications for paleoceanographic reconstructions, *Marine Micropaleontology*, 28(3–4), 231–246, doi:10.1016/0377-8398(96)00003-5, 1996.
- Spero, H. J., Bijma, J., Lea, D. W. and Bemis, B. E.: Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes, *Nature*, 390(6659), 497–500, doi:10.1038/37333, 1997.
- 40 Steinhardt, J., Cleroux, C. de Nooijer, L.J., Brummer, G.-J.A., Zahn, R., Ganssen, G.M., and Reichert, G.-J.: Reconciling single-chamber Mg / Ca with whole-shell $\delta^{18}\text{O}$ in surface to deep-dwelling planktonic foraminifera from the Mozambique Channel, *Biogeosciences*, 12, 2411–2429, https://doi.org/10.5194/bg-12-2411-2015, 2015

- Stott, L., Poulsen, C., Lund, S. and Thunell, R.: Super ENSO and Global Climate Oscillations at Millennial Time Scales, *Science*, 297(5579), 222, doi:10.1126/science.1071627, 2002.
- 5 Sverdrup, H.U., Johnson, M.W. and Fleming, R.H.: *The Oceans, Their Physics, Chemistry, and General Biology*. Prentice-Hall, New York, 1942
- Thirumalai, K., Partin, J. W., Jackson, C. S. and Quinn, T. M.: Statistical constraints on El Niño Southern Oscillation reconstructions using individual foraminifera: A sensitivity analysis, *Paleoceanography*, 28(3), 401–412, doi:10.1002/palo.20037, 2013.
- 10 Thunell, R.C., Kier, R.S., and Honjo, S.: Calcite dissolution: an in-situ study in the Panama Basin, *Science*, 212 (4495), 659–661, 1981
- Trenberth, K.E. and Otto-Bliessner, B.L.: Toward Integrated Reconstruction of past Climates, *Science*, 300 (5619), 589–591, 2003
- 15 de Vargas, C., Norris, R., Zaninetti, L., Gibb, S. W. and Pawlowski, J.: Molecular evidence of cryptic speciation in planktonic foraminifers and their relation to oceanic provinces, *Proceedings of the National Academy of Sciences*, 96(6), 2864–2868, doi:10.1073/pnas.96.6.2864, 1999.
- de Vargas, C., Bonzon, M., Rees, N. W., Pawlowski, J. and Zaninetti, L.: A molecular approach to biodiversity and biogeography in the planktonic foraminifer *Globigerinella siphonifera* (dâ€™Orbigny), *Marine Micropaleontology*, 45(2), 101–116, 2002.
- Vetter, L., Kozdon, R., Mora, C. I., Eggins, S. M., Valley, J. W., Hönisch, B. and Spero, H. J.: Micron-scale intrashell oxygen isotope variation in cultured planktic foraminifers, *Geochimica et Cosmochimica Acta*, 107, 267–278, doi:10.1016/j.gca.2012.12.046, 2013.
- 25 Waelbroeck, C., et al.: A global compilation of late Holocene planktonic foraminiferal $\delta^{18}\text{O}$: relationship between surface water temperature and $\delta^{18}\text{O}$, *Quaternary Science Reviews*, 24: 853–868, 2005
- Wang, C., Deser, C., Yu, J.-Y., DiNezio, P. and Clement, A.: El Niño and Southern Oscillation (ENSO): A Review, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, pp. 85–106, Springer Netherlands, Dordrecht., 2017.
- 30 Waterson, A. M., Edgar, K. M., Schmidt, D. N. and Valdes, P. J.: Quantifying the stability of planktic foraminiferal physical niches between the Holocene and Last Glacial Maximum, *Paleoceanography*, doi:10.1002/2016PA002964, 2016.
- Wessel P. and Lyons S.: Distribution of large Pacific seamounts from Geosat/ERS-1, *Journal of Geophysical Research*, 102, 459–475, 1997
- 35 Weyl, P.K.: Micropaleontology and Ocean Surface Climate, *Science*, 202 (4367), 475–481, 1978
- White, S. M., Ravelo, A. C. and Polissar, P. J.: Dampened El Niño in the Early and Mid-Holocene Due to Insolation-Forced Warming/Deepening of the Thermocline, *Geophysical Research Letters*, 45(1), 316–326, doi:10.1002/2017GL075433, 2018.
- 40 Wilke, I., Peeters, F. J. C., Bickert, T.: The influence of seawater carbonate ion concentration $[\text{CO}_3^{2-}]$ on the stable isotope composition of planktic foraminifera species *Globorotalia inflata*, *Marine Micropaleontology*, 58(4), 243–258, https://doi.org/10.1016/j.marmicro.2005.11.005, 2006

- Wit, J. C., Reichart, G. J. and Ganssen, G. M.: Unmixing of stable isotope signals using single specimen $\delta^{18}\text{O}$ analyses, *Geochemistry, Geophysics, Geosystems*, 14(4), 1312–1320, doi:10.1002/ggge.20101, 2013.
- 5 Wolf-Gladrow, D. A., Bijma, J. and Zeebe, R. E.: Model simulation of the carbonate chemistry in the microenvironment of symbiont bearing foraminifera, *Marine Chemistry*, 64(3), 181–198, doi:10.1016/S0304-4203(98)00074-7, 1999.
- Wunsch, C.: On sharp spectral lines in the climate record and the millennial peak, *Paleoceanography*, 15(4), 417–424, doi:10.1029/1999PA000468, 2000.
- Wunsch, C. and Gunn, D. E.: A densely sampled core and climate variable aliasing, *Geo-Marine Letters*, 23(1), 64–71, doi:10.1007/s00367-003-0125-2, 2003.
- 10 Wycech, J. B., Kelly, D. C., Kitajima, K., Kozdon, R., Orland, I. J. and Valley, J. W.: Combined effects of gametogenic calcification and dissolution on $\delta^{18}\text{O}$ measurements of the planktic foraminifer *Trilobatus sacculifer*, *Geochem. Geophys. Geosyst.*, doi:10.1029/2018GC007908, 2018.
- Xie, S-P.: On the Genesis of the Equatorial Annual Cycle, *Journal of Climate*, 7, 2008-2013, 1994
- 15 Xie, S-P.: Interaction between the annual and interannual variations in the equatorial Pacific, *Journal of Physical Oceanography*, 25, 1930-1941, 1995
- Žarić, S., Donner, B., Fischer, G., Mulitza, S. and Wefer, G.: Sensitivity of planktic foraminifera to sea surface temperature and export production as derived from sediment trap data, *Marine Micropaleontology*, 55(1–2), 75–105, doi:10.1016/j.marmicro.2005.01.002, 2005.
- 20 Žarić, S., Schulz, M. and Mulitza, S.: Global prediction of planktic foraminiferal fluxes from hydrographic and productivity data, *Biogeosciences*, 3(2), 187–207, doi:10.5194/bg-3-187-2006, 2006.
- Zebiak, S.E. and M.A. Cane: A Model El Niño -Southern Oscillation, *Monthly Weather Review*, 115, 2262–2278, https://doi.org/10.1175/15200493(1987)115<2262:AMENO>2.0.CO;2, 1987
- 25 Zeebe, R. E. and Sanyal, A.: Comparison of two potential strategies of planktonic foraminifera for house building: Mg^{2+} or H^+ removal? *Geochimica et Cosmochimica Acta*, 66(7), 1159–1169, doi:10.1016/S0016-7037(01)00852-3, 2002.
- Zhu, J., Liu, Z., Brady, E., Otto-Bliesner, B., Zhang, J., Noone, D., Tomas, R., Nusbaumer, J., Wong, T., Jahn, A. and Tabor, C.: Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model, *Geophysical Research Letters*, 44(13), 6984–6992, doi:10.1002/2017GL073406, 2017.
- 30

Formatted: Font: 10 pt

Formatted: Footer1

Page 1: [1] Formatted	Brett Metcalfe	21/02/2020 18:26:00
Footer1		
Page 1: [2] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Header: Line spacing: single, Tab stops: 7.96 cm, Centered + 15.92 cm, Right + Not at 8 cm + 16 cm		
Page 1: [3] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Footnote Text: Font: (Default) Times New Roman, Justified		
Page 1: [4] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Default: Adjust space between Latin and Asian text, Adjust space between Asian text and numbers		
Page 1: [5] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Bibliography		
Page 1: [6] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Revision		
Page 1: [7] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
EndNote Bibliography: Check spelling and grammar		
Page 1: [8] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
EndNote Bibliography Title: Check spelling and grammar		
Page 1: [9] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Comment Subject		
Page 1: [10] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Comment Text		
Page 1: [11] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Equation		
Page 1: [12] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Balloon Text		
Page 1: [13] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
List Paragraph		
Page 1: [14] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Copernicus_Word_template		
Page 1: [15] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Name		
Page 1: [16] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Kontakt		
Page 1: [17] Style Definition	Brett Metcalfe	21/02/2020 18:26:00
Bullets: No bullets or numbering		
Page 1: [18] Style Definition	Brett Metcalfe	21/02/2020 18:26:00

Betreff

Page 1: [19] Style Definition **Brett Metcalfe** **21/02/2020 18:26:00**

Heading 4: Font: (Default) +Headings (Times New Roman), Not Bold, Italic, Font color: Accent 1, Space Before: 2 pt, Keep lines together

Page 1: [20] Style Definition **Brett Metcalfe** **21/02/2020 18:26:00**

Heading 3: Font: (Default) +Headings (Times New Roman), 12 pt, Not Bold, Font color: Accent 1, Space Before: 2 pt, After: 0 pt, Line spacing: 1.5 lines, Keep lines together

Page 1: [21] Style Definition **Brett Metcalfe** **21/02/2020 18:26:00**

Heading 2: Font: (Default) +Headings (Times New Roman), 13 pt, Not Bold, Font color: Accent 1, Space Before: 2 pt, After: 0 pt, Line spacing: 1.5 lines, Keep lines together

Page 1: [22] Style Definition **Brett Metcalfe** **21/02/2020 18:26:00**

Heading 1: Font: (Default) +Headings (Times New Roman), 16 pt, Not Bold, Font color: Accent 1, Space Before: 12 pt, After: 0 pt, Line spacing: 1.5 lines, Keep lines together

Page 1: [23] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Width: 21 cm

Page 1: [24] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [25] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [26] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [27] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [28] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [29] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [30] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [31] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [32] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [33] Formatted **Brett Metcalfe** **21/02/2020 18:25:00**

Font: Times New Roman

Page 1: [34] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Font: Times New Roman

Page 1: [35] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Font: Times New Roman

Page 1: [36] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Font: Times New Roman

Page 1: [37] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Font: Times New Roman, Not Bold

Page 1: [38] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Heading 11

Page 1: [39] Formatted	Brett Metcalfe	21/02/2020 18:25:00
-------------------------------	-----------------------	----------------------------

Font: Times New Roman