Further comments intended to clarify our work towards all reviewers and the Topical Editor

Having received reviews from all three anonymous referees we would like to make some general comments before posting a line by line response to referees #2 and #3.

We note that all three referees have submitted very similar reviews. We invested quite some effort in responding to the points raised by the first referee as quickly as possible, so that referees #2 and #3 would be able to consult our response before submitting their own personal reviews. It seems, however, that these two referees submitted very similar reviews as referee #1, which would suggest that referees #2 and #3 have not taken our response to referee #1 into account. It is unfortunate that the full potential of the Climate of the Past Discussions forum format was not exploited. It would indeed have been much more profitable to discuss the comments of referee #2 & #3 if they had been written in the light of our initial response. We also note that even though we strongly disagree in our initial response with the first reviewer comment, at time of writing, there has been no further comments from the reviewers on the discussion page on this first response although the discussion period is not ended. In the interest of stimulating such a discussion, we will briefly synthesise and address the three main criticisms raised by all three reviewers here and our arguments against them. There are several other, minor criticisms/corrections raised by the referees which we find very valid, for which we thank the reviewers and which we will gladly address. The main criticisms raised by the three reviews can be distilled into the following four points:

(1) The study is a forward model and, therefore, does not inverse model ENSO recorded by sediments.

This is correct. Our study uses 60 years of observed monthly climate data to forward model planktonic foraminifera populations accounting for habitat water depths and growth season for the entire Pacific Ocean, allowing us to investigate whether or not ENSO dynamics are recorded by foraminifera populations *in the water*. We think that this investigation is in itself a fundamentally interesting subject for research/publication, since it is the first step/prerequisite to investigate whether it is possible to distinguish ENSO from non-ENSO conditions from the planktonic foraminifer fossil record. This question has not been investigated before and is, therefore, new and has the potential for interesting conclusions.

Our model approach is indeed not an inverse model. It is also not a sediment model. The manuscript that we have submitted is a forward model of planktonic foraminifera populations in the water. Clearly, the reading of the referees' comments shows that they would have preferred if we had submitted a completely different manuscript; as they have reviewed our manuscript as if it were yet another study on the inverse problem of detecting a change in ENSO from a change in the distribution of sediment archive planktonic foraminifer $\delta^{18}O$. We are disappointed and strongly disagree with being reviewed in the view of a different manuscript that could have been written, and not evaluated based on the actual manuscript submitted and the science contained therein. We feel this is against the basic principles of a standard review process and strongly object to this treatment of our work.

(2) The FAME output in this study has not been tested against core tops / sediment records, etc.

This is correct. In the present manuscript, we have indeed used a 60 years monthly record of climate input to run our FAME foraminifera population model. Core tops and sediment records in the Pacific integrate foraminifera populations from centuries/millennia of time, and are therefore totally unsuitable to compare to our model output in the context of evaluating the impact of ENSO. The best thing that we could think of to compare our FAME output to, in the interest of this particular study, would be data from monthly plankton tows spanning multiple decades, but we couldn't find any. There is thus simply no data we know of that can be used meaningfully in the context of our current work on ENSO. We note, however, that the FAME model itself has been validated against

core tops in Roche et al. (2018), including core tops from the Pacific Ocean. The claim that FAME has not been validated in the Pacific Ocean is therefore unfounded.

(3) The statistics are not applicable to palaeo records

This is correct. We did indeed not select a statistical test on the basis of it being applicable to palaeo records. We chose the statistical test (Anderson Darling test) that was most suitable for our purpose and hypothesis: to examine the (dis)similarity of different high-resolution probability distributions produced by our FAME output driven by 60 years of monthly observational data. This approach allows us to use the full potential of our model output to directly test our hypothesis. The standard deviation (as suggested by all three reviewers), is unsuitable for comparing the dissimilarity of distributions, as it can return spurious results: for example, it is possible that two distributions of differing shape could have the same standard deviation, whereas an A-D test would detect a difference. The research question and its associated hypotheses are the central tenant upon which we based our experimental design. We note that we are not developing a statistical toolbox for palaeorecords and readers should of course always assess, on a case by case basis, statistics that are appropriate for their own studies.

(4) The authors have not run a bioturbation model.

This is correct. Since we are not modelling sediments in this manuscript, we did not carry out a transient bioturbation model. It would not be suitable to carry out a transient bioturbation model upon 60 years' worth of input data, seeing as bioturbation models require, due to the interplay between sediment accumulation rate (SAR) and bioturbation depth, many millennia of data/spin up time in order to produce valid output.

We were, however, curious as to whether the parts of the Pacific Ocean where FAME predicts that foraminifera in the water can record ENSO dynamics coincide with high SAR sediments. Therefore, as an extra, we included a rough map of regions of the Pacific Ocean where SAR < 5 cm/ka. The reviewers asked why 5 cm/ka was used as a cut-off point. We intentionally used a very generous cut-off point of 5 cm/ka and assumed that the reader would understand that 5 cm/ka is a very slow SAR that would severely hamper the detection of average ENSO dynamics for specific time periods (i.e. less than one millennia). It is not, however, necessary to carry out a transient bioturbation model to understand the time periods integrated by a 5 cm/ka SAR. We can simply carry out a calculation following established understanding of the influence of bioturbation upon the age dispersal of single foraminifera (Berger and Heath, 1968; Berger and Johnson, 1978; Berger and Killingley, 1982), the same understanding that is included in transient bioturbation models themselves (e.g. Trauth, 2013; Dolman and Laepple, 2018; Lougheed et al., 2018). In such a case, assuming the common bioturbation depth of 10 cm (Peng et al., 1979; Trauth et al., 1997; Boudreau, 1998), we can calculate that the 1 sigma age value of foraminifera contained within a single cm of a 5 cm/ka core is:

10 [cm] / (5 [cm/ka] /1000) = 2000 years (from which follows that 2 sigma = 4000 years).

Our map may of course miss limited parts of the Pacific Ocean with SAR higher than 5 cm/ka. We will mention this and/or highlight such places in the final version of the manuscript.

Kind Regards,

Brett Metcalfe

Bryan Lougheed

Claire Waelbroeck

Didier Roche