**Climate
of the Past**
Discussions

Interactive comment on "A Late Pleistocene sea level stack" *by* R. M. Spratt and L. E. Lisiecki

Anonymous Referee #2

Received and published: 9 September 2015

I thought this was going to be a great study to consider, but in the end felt disappointed. This study to me seems to be just another example of taking good records that have taken many years to perfect, smear them together in a fairly arbitrary manner, and then running some basic statistics over the top, to try and produce a 'synthesis' with an 'improved signal to noise ratio'. Earlier, work by Kopp et al. 2009 (which by the way was not, at all, an only-coral-based assessment, as suggested in the final paragraph of section 4) did something similar, albeit in a more sophisticated manner and for a shorter time interval, but even that study is blighted by the problem of arbitrary choices of chronological alignment between records, pulling some around to the limits of, or beyond, their stated uncertainties. There have been several other 'syntheses' of late that all use versions of this approach; perhaps it is because of the lure of avoiding the hard graft of working up something original, in favor of writing yet another easy compilation with some statistics to get a potentially well-cited paper.

The original LR04 stack was a game-changer in synthesizing benthic delta-O18 records, but was later found to be blighted by assumptions of synchronicity between records that are made implicitly by use of the Match software (e.g., Skinner and Shackleton 2005 QSR). One of the authors of the current manuscript even had their own paper in 2009 (Lisiecki and Raymo, Paleoceanography 2009) in which this implicitly assumed synchronicity was demonstrated to be flawed. Yet here we see it again, and again without any attempt at proper propagation through all methods and conclusions of the uncertainties and limitations.

The signal matching approach needs to be relegated to history, if it is not backed up by a strong physical rationale, and/or rigorous independent testing, and/or proper uncertainty propagation. Certainly in the way applied in the current study, it is an antiquated approach that is known to be flawed. I suggest that it would be time better spent for the researchers to instead start working on developing independent and testable chronologies for each of the records. In this context, I was surprised that the Red Sea record used is not the most recent version that I have seen (Grant et al. Nature Comms 2014), which has an independent chronological assessment and full probabilistic assessment using the age uncertainties as well as the method uncertainties. That should be used, and then any chronological adjustments needed would need to remain within the limits of that assessment. Similar independent age assessments need to be developed for the other methods; this is where the real challenge lies, and where advancement of sea-level understanding will come from. It is only once that is done, that we come into a position to consider putting the records together (each on their own proper timescales) to evaluate common signals and differences.

Even if we accept the chronological matching as done (though I don't see why we should; see above), then I still remain very worried about the lack of propagation of the legion uncertainties that arise from assumptions and adjustments in the chronologies, through the method and into the final conclusions. I am convinced that the uncertainties around the end product, and any further manipulations based on it, will increase greatly

when this is done. I am particularly worried that the difference between linear and non-linear regressions in section 6 may not be robust when considered relative to fully and properly propagated uncertainties.

A further concern is that the various methods underlying the different sea-level records that are used, are not independent of each other. As such one could wonder if straight PCA is an appropriate analytical tool. After all, we're not just looking at covariances between independent estimates with a common signal, but at covariances between methodologically (partially) related estimates with a common signal. This is not discussed, and there is also no assessment of how the methodological dependence might affect the answers (and their uncertainties). I think this would need some serious thought and discussion too.

Finally, there are some unsupported manipulations, such as the 2 ka lag to the smoothed LR04, as applied in section 6.

I don't think that this study as is does anything to advance understanding and to improve the state of the art. It's the sort of exercise that one might expect from an MSc student, perhaps, but it is not going to help us understand sea-level variability any better than the individual input records. The study would only introduce a false sense of 'understanding' that is flawed because of the (often unspecified) underlying assumptions, uncertainties, and questionable manipulations. It is evident that the real challenge is to get the different records onto their own independent chronologies, and to then compare them statistically (using appropriate statistics and proper uncertainty propagation). That may not be so easy to do, but true understanding doesn't need to come easy. Certainly not when 'easy' is using a known flawed approach. I recommend rejection of the study as is.

———————————————————————