Ocean Sci. Discuss.,
https://doi.org/10.5194/os-2020-66-AC2, 2020

Ocean Science
Discussions
Open Access

Interactive comment on "Antarctic Bottom Water and North Atlantic Deep Water in CMIP6 models" *by* Céline Heuzé

Céline Heuzé

celine.heuze@gu.se

Received and published: 15 September 2020

The author thanks Reviewer 2 for their scientific comments. The role of the reviewers has been duly acknowledged: L.554-555: "The author thanks the two anonymous reviewers whose comments greatly improved the quality of this manuscript"

The response is organised as follows: first, the comment from the reviewer; then, my answer; finally, when relevant, new or modified text.

The paper presents a comparison of time mean fields between new, CMIP6 models and observations. The focus is the formation and distribution of deep water formed in the North Atlantic and Southern Ocean. The aim (I assume) is to document these biases as a basis for further study. In general, I did not feel that the paper presented much that

Printer-friendly version

Discussion paper

was convincing in terms of new scientific interpretation. [. . .] Nonetheless, the potential lack of new understanding is not necessarily a strong negative as documenting the models can be a useful exercise in and of itself.

Both reviewers raised this important point: the target audience and aim of the paper were not as clearly defined as I thought. This paper is indeed designed as a reference for model users to justify their choice of models for further studies. Although attempts are made at explaining these biases, the emphasis is on quantifying these biases. This sentence was added to the introduction to clarify this objective, lines 51-52: "The primary objective of this paper is to quantify and discuss biases of each model, so that model users can make informed model selections."

In particular many of the reported correlations seem small given that the many of the models are not very independent, which I don't think has been accounted for or even acknowledged.

The lack of independence of the models is acknowledged as early as line 67 (slightly modified from the previous version in response to a comment by reviewer 1): "Furthermore, as some models are not fully independent as they share similar codes (Table 1). . ." In response to the reviewer's comment, this methodological clarification was also added line 68: "To account for this lack of independence, the correlations quoted throughout the text have been verified with different model numbers"

Scientific issues:

L77-78. Why is a different threshold used to the observations? How can you then fairly compare with the observations? Please explain this.

I chose neither threshold. The threshold of the models is the "official" threshold of the CMIP6 procedure; that is, models that wanted to participate in CMIP6 had to use that threshold. Likewise, the threshold used in observations is the one that was chosen by the creators of this observational product. The literature on the impact of one threshold rather than another is plentiful (I even wrote a PhD dissertation chapter on this), and the conclusion is that for detecting spurious modelled deep convection, this difference is not critical. In fact, choosing a larger threshold for the models than for the observations means that we would underestimate deep convection in models. As the objective of this publication is primarily to compare models with each other, the most important is that all models use the same threshold. The following was added to summarise this discussion, lines 80-88: “As is requested for CMIP6, the MLD is then detected as the depth where $\sigma\theta$ differs from that at 10 m depth by more than 0.125 kg m$-$3. [new text associated with the new supplementary figure] Furthermore, a different threshold of 0.03 kgm$-$3 is used in the observational reference (de Boyer Montégut et al., 2004), which could lead to an underestimation of mixed layer depths in the models (as we show in section 3, it does not).”

L224-226. So what can we actually determine or learn from this? Is there a relationship, particularly after controlling for the fact that several of the models are nearly identical (or assimilate observations, which I am surprised is included as it seems fundamentally different to the other models)

Sentence expanded to make the point clearer: “The models that convect the least or not at all tend to be the most accurate. For the CESM2 family, accurate bottom properties and lack of deep convection may both be the result of their overflow parameterisation (Briegleb et al., 2010; Snow et al., 2015). For another model, NorCPM1,the accuracy in all properties may come from its observation assimilation rather than accurate model physics (Counillon et al., 2016).”

L250-251. This link surely only makes sense if the climate sensitivity is driving the DMV. But previously you suggest the logic is the other way around (i.e. larger DMV can sequester more heat and thus reduce climate sensitivity). If the DMV drives the climate sensitivity, why should DMV in the Weddell Sea and SPG themselves be linked? Isn’t it more likely that the DMV in these two regions correlate due to some global model bias?

Printer-friendly version

Discussion paper

Does the DMV impact the climate sensitivity, or does the sensitivity impact the DMV? Less mixing means less heat absorbed by the ocean, so more in the atmosphere and a larger sensitivity. The opposite is true: if the sensitivity is somehow controlled by another "global model bias", a high sensitivity will lead to more ocean surface warming and stratification, and hence less mixing. As is obvious from just these two sentences, what we really have is a feedback loop, and investigating which comes first, or what that other global bias can be, are beyond the scope of this paper. I added a sentence to reflect the point raised by the reviewer lines 277-279: "As already mentioned, no causation can be inferred: deciphering whether global biases in DMV are responsible for the models' sensitivities, or in contrast sensitivities are set by other processes and impact the DMV, is beyond the scope of this analysis"

L364, L367. Are these correlations really robust given the real number of degrees of freedom is likely far fewer than the total number of models

All the correlations have been verified using different model numbers, in particular using only one member per family, and the results remained. This precision has been added to the Methods section, lines 68-69.

Minor issues:

To summarise, all the issues highlighted by the reviewer that impeded the understanding have been corrected. Only the ones for which a response longer than "corrected" was necessary are presented here. I leave it to the copy editor to decide whether grammar rules that we learnt at school can be bent in order to make the manuscript more dynamic and pleasant to read.

L1. "Deep water formation is the driver of the global ocean circulation" - on what timescale?

This sentence has already been modified in response to a comment by Reviewer 1.

L5. Large majority - can you be more quantitative?

Sentence modified to ”28 models in the Southern Ocean and all 35 models in the North Atlantic”

L31. Is an accurate representation really needed for climate predictions? Over what time scales are you referring? Of what variable? Predictions in CMIP usually refers to initialised decadal simulations, whereas projections usually refer to century time scale uninitialised simulations.

I am not sure which point the reviewer is trying to make here. I changed ”prediction” to ”projection”, as I meant long term, IPCC-report type results.

Table 1. The horizontal resolution here doesn't take into account any local grid refinement, which could also be noted

They could indeed, but too many cases would need to be considered to fit in the table. Instead, table caption has been modified to ”nominal” grid resolution.

L58. Is this the actual variable name or is it mlotst?

The actual variable name is mlotst. There was a typo.

L58. For the models where you have MLD directly, can you show as a supplementary figure that your method and the online one are equivalent.

This is a very good idea. A new supplementary Figure (A1) has been added to show where and by how much they differ, along with a discussion in the Methods section.

L72. Is the deboyer montegut MLD data just to 2004 or is it updated?

It is updated, as is indicated on the data download website.

L104. “Hardly a third” is colloquial. Please be specific.

The number of models changed since the initial submission. Sentence now reads: ”it is provided by only 18 of the models (from 10 families)”

L173. “(Thin?)” - this made me a bit annoyed. If you want to hypothesize at the reason then please spell this out in a sentence rather than like this

I am confused by the reviewer's reaction as the hypothesis is spelt out in that sentence. I removed the "(thin?)" that brought nothing to the sentence.

L183. Does the pipe physically suck the water in the model physics? Could the description of this parameterization be described more formally?

The following sentence was added lines 196-198: "If the water on the shelf exceeds a critical density, a pipe artificially transports this dense water from the shelves to the deep basin. Without having to cascade, the dense shelf water keeps its properties."

L243. Here (and elsewhere) the referencing is a bit unclear. You've already cited this paper, and here it seems as if you're citing it as a reference showing a link between convection and climate sensitivity here.

I assume the reviewer meant the reference to Zelinka et al. (2020) of (previous version's) line 248. Sentence modified to clarify that I refer to this paper as the source of the sensitivity values I used: "There is a relationship with the climate sensitivities of Zelinka et al. (2020) though"

L257. Please define quantitatively what a "tolerable" bias is?

I do not understand what the reviewer means as the quantitative value is given in the same sentence, two words later.

L289. Is such a small correlation actually significant, especially when accounting for the limited degrees of freedom (i.e. many similar models)

This point has been addressed twice already in this response.

L297. "The question remains". What question? Please specify.

Sentence now reads: "the cause of NorCPM's and other models' bottom density bias in the GIN seas remains unknown"

L321. The AMOC can’t be said to be overestimated given the quoted uncertainty on both the model and observations.

Agreed. Sentence removed.

Interactive comment on Ocean Sci. Discuss., https://doi.org/10.5194/os-2020-66, 2020.

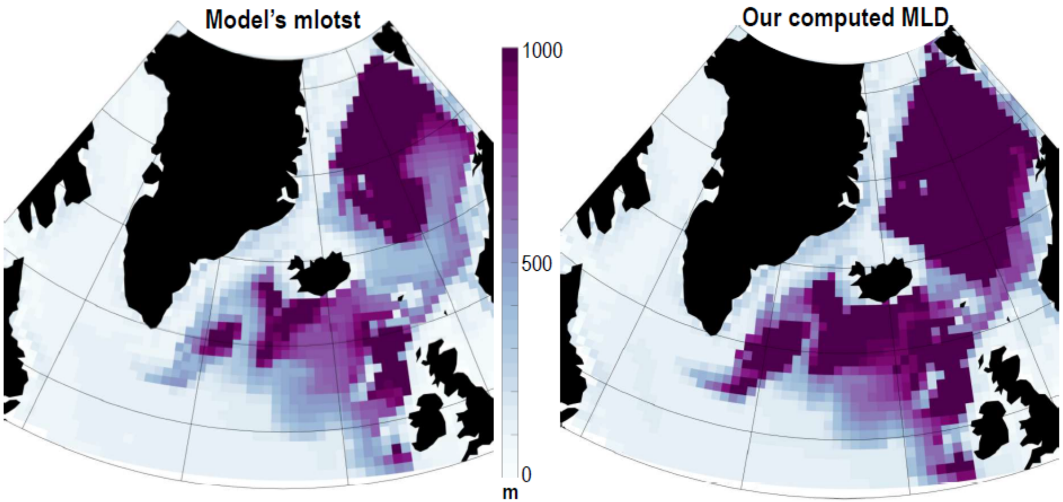

Figure A1. Maximum monthly mixed layer depth in the North Atlantic over 1985-2014 for the model CanESM5: left, using the model output 'mlotst'; right, when computed from the monthly temperature and salinity. Over the entire 30-year period, the root mean square error in the Nordic Seas is 305 m; in the subpolar gyre, 21 m.

Fig. 1. New supplementary figure suggested by the reviewer: comparison of the MLDs

Printer-friendly version

Discussion paper