Review of "Influence of initial stratification, wind and sea ice on the modelled oceanic circulation in Nares Strait, northwest Greenland" by L. W. Bergman and C. Heuzé (Manuscript os-2018-122)

This manuscript mainly describes a set of sensitivity experiments with a coupled ocean and sea ice model configuration. As will become clear from my specific comments below, I believe that unfortunately the number of shortcomings is too high for any other outcome than rejection of the submitted manuscript. This is both relevant for the configuration of the experiment, and the interpretation of results. I hope that the list of items has enough details to be helpful to the authors future endeavours in ocean and ice circulation modeling.

1 Major items

1. I find the constructions of atmospheric forcing fields to be very odd. Starting with the "Default" case, this is constructed with a mix of fields from the ERA interim product and winds from a HIRHAM regional product. With such a set of choices, there will be inconsistencies in the air-sea turbulent heat flux (usually parameterized as a function of wind speed and the 2 m air temperature) and moisture (a function of wind speed and humidity). This inconsistency could easily have been dealt with, by applying ERA interim as the only source for atmospheric forcing fields. Next, for the "Wind experiments", the choice of applying winds with a magnitude of up to 30 m/s either locally for the duration of the experiment ("spatially varying" case) or reaching across the entire domain ("temporally varying" case) is not justified: this wind speed is very close to hurricane strength, and completely unrealistic in the present context, even for a sensitivity study. This extreme choice is also in glaring contrast to the modest initial sea ice thicknesses that are specified for the "Sea ice experiments". Furthermore, the manuscript gives the impression that the sensitivity experiments are performed in combination with other atmospheric forcing from ERA interim. Again, this is an unfortunate construction. The authors should have examined how e.g. air temperature typically changes for various wind speed regimes, and applied corresponding spatially uniform forcing fields in the "temporally varying" case. The mixed specification applied by the authors should be avoided in a context of sensitivity experiments.

2. I find little explicit information about the initialization of the experiments, which makes it unnecessary difficult to interpret some of the results. The problem is that with the limited information that is available, it is very difficult to assess how far off the initial state is from a (statistical) equilibrium: I don't find any information

on the initialization of surface elevation or ocean velocities (I'm guessing that they are both set to 0 initially). I also find no information about how swiftly/slowly the model adjusts to the applied forcing at the lateral boundaries and the surface. Some information about the evolution of the (eddy) kinetic energy would have helped. Moreover, I get the impression that the effect of transitioning from an initial state to a state that reflects the applied forcing has completely escaped the authors, as they e.g. many places refer to average results for the entire simulation period (i.e., no consideration of the spin-up; e.g. relevant for Fig. 4, 5). I return to the issue in a couple of the items below.

3. The initialization of the "Stratification experiments" apply specifications for salinity and temperature from the MIMOC climatology. However, three of the five experiments undergo a transition to a less saline regime (Fig. 6). For the "Default" experiment the change is dramatic, with more than a doubling of the freshwater height. However, no analysis of this evolution is presented. Questions that come to mind include (1) to what degree is the evolution due to mis-matching of the initial ocean state and applied forcing (spin-up); (2) what is the source for the additional fresh water that enters the relevant simulations; (3) what is the character of change in stratification when compared to the MIMOC climatology. None of these questions are addressed in the study.

4. The discussion of results for sea ice thickness is suffering from omissions as well as a possibly erroneous deduction. The deduction I have in mind is that when the authors identify the "2-layer" experiment as having a larger heat content and thinner sea ice than the two "Tilt" simulations, they attribute the difference in sea ice to advection. My problem with this is (1) their justification: they point to Fig. 4 and 5, but neither is strictly relevant (Fig. 4 shows volume transport with no information of the vertical distribution of momentum and thus not the surface velocity; Fig 5 shows the vertical distribution, but only for one selected cross-section, and while differences are evident below 100 m depth, the near-surface results appear to be similar); (2) the vertically integrated heat content, which is largest in the "2-layer" experiment, may not be relevant to the ice budget since with the initial heat anomaly in the salty lower layer this heat may be trapped in the lower layer and thus not available for surface heat fluxes and freezing/melting of ice (no depiction is provided for the vertical distribution of temperature; while very speculative, I note that there is a "fresh anomaly" layer above 100 m in Fig. 10f, and this can conceivably be a fingerprint of water masses advected from the open boundary which might insulate warm waters below from interacting with sea ice). Regarding omissions, I find that the attribution of differences in sea ice thickness to two processes is too narrow. I would suggest considering another three items in this list: stratification (from what I just explained), surface boundary conditions (the mis-match issue in my initial item above), and open boundary conditions (sponge layers may lead to accumulation near and/or inside the sponge, I believe).

5. I read this manuscript mainly as a description of a set of sensitivity experiment which has a number of simplifications, short-comings and imposed perturbations that are possibly, if not likely, non-representative for the actual state in the Nares Strait. While the sparse observations are of interest to potential readers, the issues I just mentioned points to low relevance of direct comparisons between model results and observations. So I think that the paper would have been better, and more to the point, if the observed state was briefly mentioned in the Introduction or the Conclusion section, and not have a full (sub-)section (3.5) devoted to a model-data comparison. I also note that the guidelines for the extent of Ocean Science Technical Notes is "a few pages only", so the submitted manuscript is a bit long.

2 Selected minor items

1. The authors mention a subsurface jet with a width of 5-10 km (p.2, l.4) while the model resolution across the strait is 2.5 km. Are the model results likely to be impacted by limitations due to resolution?

2. The authors refer to one of their configurations as "homogeneous 2-layer ocean" (p.4, l.21), but both from the following text and from Fig. 3 it is clear that this configuration is inhomogeneous, but (apparently) with no initial pressure gradients.

3. An experiment is included with an ocean-only configuration, without the possibility for ice to freeze at the surface (p.5, l.5). I would suggest to remove this case from the study, since the forced absence of ice might lead to extreme cooling of the surface and excessive, un-physical overturning (possibly by means of towering vertical mixing in the model's numerical formulation).

4. In Eq. (1) and (2), h_1 and h_2 are not defined.

5. On p.7, l.6-8 the authors note that the un-tilted experiments failed to create a realistic current structure. They are also referred to as "too salty" (p.9, l.11). Yet, these are the experiments which retain a freshwater height closed to their initialization by MIMOC climatology. So is the MIMOC climatology too salty? Regarding the tilted experiments, which are not explicitly referred to as unrealistic by the authors but are departing from the initialization from MIMOC climatology, do they have a realistic water mass distribution and structure?

6. On p.8, l.6-8the authors state that with their imposed spatially varying winds, the ice drifts faster southwards in the region with strongest winds. However, my interpretation of the wind forcing in this experiment is that the stress is directed from east towards west. In that case, the Ekman transport is northwards, so why does the ice drift southwards? (I note that Fig. 5c has strong currents near the surface directed southwards, so I'm probably mis-interpreting the direction of the wind stress.)

7. P.8, l.23; “forcing files” should probably be corrected to “forcing fields” (the arrangement of forcing into various files are not relevant to anyone but the authors).

8. On. p.9, l.12-14 the authors attribute missing heat to near-surface conditions that are too cold by about one degree. How deep does the one degree anomaly need to extend in order to account for all of the missing heat?

9. P.9, l.26; I believe “at one point” should be rewritten as “at a few locations”.

10. The y axis values in Fig. 7 are positive, yet, the water masses are initialized with negative and 0 degree temperatures (p.4, l.24-27). But then, by the definition in Eq. (2) the heat content becomes negative.

11. The y axis label is “Average summer heat content”; is “summer” correct here?

12. Table 2: I would have liked a line with the results from the MIMOC climatology.