Comment on gmd-2021-95
Anonymous Referee #1

Referee comment on "Comparison of ocean heat content estimated using two eddy-resolving hindcast simulations based on OFES1 and OFES2" by Fanglou Liao et al., Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2021-95-RC1, 2021

The manuscript, "Comparison of ocean heat content from two eddy-resolving hindcast simulations with OFES1 and OFES2", by Fanglou Liao, Hua Wang, and Zhiqiang Liu, suggests that OFES 2 does not warm in accordance with upper-ocean observations and the previous reconstruction of OFES 1 from 1990-present. The decomposition of ocean heat content changes into heaving and non-heaving contributions is interesting, but it doesn't conclusively explain what is going on with OFES 2. The manuscript would be improved if it could both document and interpret the dynamical causes of the discrepancy between OFES 1 and 2.

Major issues:

It is interesting that the ocean heat content changes primarily by the change in isopycnal depth. Does the total heat content calculation depend upon the calculation of heat content change by heaving motions? It would be good to provide corroborating evidence regarding the heat content decomposition, such as an independent calculation of total heat content variability.

Abstract: Heat transport is stated to not always be responsible for ocean heat content changes. Doesn't it have to be either the heat transport or the air-sea flux, given thermodynamic energy conservation? Heat content storage will be the residual of these terms. It is puzzling to consider where the thermodynamic energy is transported. The manuscript would benefit with a closed energy budget analysis, which may require the deep ocean and/or the Arctic to accurately assess where the energy goes.

Minor issues:

L29: does this sentence equate objective analysis and ocean reanalysis? They are normally considered to be distinct.
How are ocean heat content changes related to algorithmic changes between OFES 1 and 2?

L65: consider changing "numerical running" to "numerical simulation".

How did the authors justify restricting their analysis to the upper 1400 meters? Their rationale following Emery (2001) and Wunsch (2011) is not compelling. Did these previous works suggest that ocean disequilibrium occurs suddenly at 1400 meters depth? (That would be surprising.) During the time period of interest, i.e., 1950-present, why should the deep ocean be in equilibrium?

For comparison purposes, does the NCEP reanalysis give air-sea conditions every 6 hours, as opposed to 3 hours in OFES 2?

L86: "Validation" is not possible with EN4 as it is also an incomplete and uncertain product.

Figure 1: What happens in the Arctic? What error is incurred by eliminating the Arctic?

Consider changing ZJ/yr into the more relevant quantity of W or watts per square meter of the Earth's surface.

Figure 2 is fascinating, if correct. What is going on with OFES2?

L161: dividing by 56 "years".

What does "SP" stand for?

L242: 10 to the 6th power

Does OFES 2 fit surface data (i.e., SST)?

Table 4 doesn't seem very useful with the inexact metrics for the water-mass source properties.