This study decomposes ocean temperature and salinity changes in a climate change simulation into a component from redistribution of the preindustrial tracer and a component from advection of the perturbed surface boundary condition ("excess" tracer). This partitioning has been done previously for temperature but its application to salinity is novel. Also novel is the method for performing the partitioning which takes advantage of statistical relationships between carbon gradients and those of temperature and salinity. The results are interesting but I have major concerns about the accuracy of the partitioning method, the consistency of the partitioning with the climate change that forces the tracer changes, and representativeness of the results given the large intermodel variation of that climate change forcing.

Major comments in priority order:

- **Accuracy.** The accuracy of the carbon method for distinguishing excess and redistributed tracer is not tested as part of the study. This should be done by comparing the carbon-based estimates with those made with the conventional method of using the perturbation fluxes to force a zero-initialized passive tracer through the experiment. This tracer for temperature is referred to as “passive heat” in Gregory et al (2016). The passive heat experiment produces the “excess” tracer and subtracting this from the active temperature simulation gives the redistributed tracer without the assumptions and approximations that are necessary when using the carbon-based method. The fact that these assumptions and approximations are not really necessary raises an auxiliary question: Why use the carbon-based method at all since it is prone to errors that are not applicable to the zero-initialized passive tracer method?

- **Consistency.** The coupled model flux perturbations that are used here to force the offline ocean model develop interactively with the circulation changes in the coupled climate of HadGEM2-ES. The interaction of ocean circulation and surface heat fluxes has been demonstrated, for example, by Winton et al (2014). A prominent flux/circulation interaction is the downward heat flux perturbation in the subpolar North
Atlantic that forces and responds to the AMOC reduction -- the latter because of reduced heat transport into the region. Therefore it is important, for consistency, that the offline ocean model reproduces the circulation changes in the HadGEM2-ES model that was used to generate the forcing. Otherwise, the flux and circulation perturbations may be inconsistent. This should be assessed for the current study. For example, the overturning stream function and its evolution over time could be compared between HadGEM2-ES and the ocean-only experiments. The temperature and salinity fields and their evolutions should also match HadGEM2-ES.

- **Representativeness.** The results in this paper are presented without accounting for the uncertainty in the excess/redistributed tracer breakdown that stems from uncertainty about climate change induced ocean circulation change. The coupled model that generated the surface flux perturbations (HadGEM2-ES) represents one possible simulation of circulation change - one simulation of AMOC decline, for example. However, models differ significantly in their AMOC decline simulations. Ideally, one would like to make the excess/redistributed tracer calculations for coupled models with a range of AMOC control and decline simulations. If this is not possible, it would be helpful for context to compare the HadGEM2-ES circulation changes that are associated with the perturbation fluxes used here with the circulation changes in other CMIP models.

**Minor comments:**

- **Line 32-33:** Specifically what is meant by "a near linear relationship between the response of the ocean to increasing atmospheric temperatures and CO2 levels". What are the related quantities? Cumulative heat and carbon uptake? Please give a reference for this statement.
- **Line 68-70:** I do not see how the excess/redistributed partition is more developed for carbon than for heat. As far as I know neither is observationally constrained. Please review the observations underpinning this statement.
- **Line 85-87:** The experiment specification is very sketchy. What are the frequencies of heat and freshwater fluxes applied? Is there any surface salinity or temperature restoring? Are any fluxes computed from surface meteorology? Are water fluxes from ice melt included?
- **Line 92 typo:** "...Cnat change Cnat..."
- **Lines 96-154:** There are many untested assumptions and approximations here. A validation strategy is needed (major comment #1).
- **The global excess heat uptake from HadGEM and observations (e.g. Zanna et al 2019) could be plotted on Fig. 1b for validation. How does the historical global excess salinity change compare to observations inferred from historical ice melt?**
- **Lines 200-202:** Is it meaningful to compare global mean excess and redistributed temperature considering that global redistributed temperature is constrained, by energy conservation, to zero?
- **Line 237:** Do circulation metrics also support this statement about settling of the circulation.
- **Line 238-243:** This seems like an argument for proportionality of excess and redistributed heat as both are hypothesized to be proportional to SST change -- redistributed because of the proportionality of SST and AMOC changes.
- **Line 253-254:** The use of the term redistributed here seems inconsistent with the nomenclature of equation 7 which indicates that redistribution requires circulation change.
- **Line 344-348:** How is the model excess salinity consistent with observations of ice
sheet loss? Do the coupled model forcings include ice sheet loss?

References
