This manuscript uses two models of different resolutions to assess the relative contributions of the sedimentation pump and subduction pump in the North Atlantic. However, model validation is inadequate, and discrepancies between the present estimates and those from previous studies are not fully addressed. The main findings are therefore not sufficiently compelling, and it is difficult to determine whether differences from earlier work reflect new finding or stem from model deficiencies and methodological artifacts. As a result, it is unclear about what new insights offered by this study relative to previous investigations.

Model validation:

1. Validating the model solely against surface observations is inadequate. Although the authors only present carbon export flux across the variable mixed layer depth, the subsurface and deep ocean should also be validated for at least two reasons. First, the calculation of mixed-layer pump and subduction pump uses model outputs both at and below the surface, as the mixed-layer pump incorporates entrainment/detrainment and the subduction pump is the net flux of subduction and obduction. Second, a lack of careful validation for subsurface and deep-ocean substaintially increases overfitting risks. The model may reproduce realistic POC concentrations by incorrect underlying mechanisms, such as biased sinking velocity, that is a key parameter governing sedimentation flux.
2. The authors performed validation only at the bioregional scale. Since the mixed layer pump and eddy subduction pump are localized and exhibit strong spatial heterogeneity, spatial patterns of mixed layer depth, eddy intensity, POC, and DOC should also be validated.
3. The observed DOC concentrations are up to 10 times higher than the modelled values. The authors attribute these large discrepancies to the model's omission of labile and refractory DOC, mismatches in bioregion definitions between the model and observations, and differences in temporal coverage. However, the authors have not evaluated whether these factors can adequately explain the DOC model–observation misfit. For example, the authors need to report the typical fractional contribution of semi-labile DOC to quantify how much of the bias stems from excluding labile and refractory DOC. They should also discuss how this omission affects subduction flux estimates. In addition, the authors should compare model outputs against observations using consistent bioregion boundaries to quantify the impact of different bioregion definitions. Regarding temporal coverage, given that the coarse-resolution model runs until 2022, the DOC comparison should be restricted to the overlapping time period available from this coarse-resolution model.
4. There is no description of the observational products used for model validation. Are these climatological or monthly datasets? What methods were applied to generate these gridded products — interpolation or machine learning? What measurement techniques were used to obtain these observations? What particle size range is included in the POC measurements? Such information is essential to understand the model–observation misfit.

5. Which modelled pools are compared against the observed POC? In some studies, observed POC is compared with the combined pool of simulated phytoplankton, zooplankton, POC, and GOC. The authors should clarify this point. Beyond POC concentration, they should also validate the modelled POC fluxes.

Comparison with previous estimates:

The authors attribute discrepancies between their estimates and previous studies to differences in the definition of export horizons (mixed layer depth versus fixed depth horizons) and in the definition of export flux (e.g., the subduction flux in Resplandy et al. did not include horizontal advection). Given the limited model validation, the possibility that flux discrepancies arise from model deficiencies cannot be ruled out. Furthermore, the authors have not attempted to quantify the impacts of these proposed explanations. Since the required variables are available within their model, they should recalculate carbon export fluxes using the fixed depth horizons and carbon-export definitions adopted in earlier studies to test their hypotheis.

Collectively, due to inadequate validation and a lack of careful attribution of discrepancies between this study and previous studies, it is difficult to determine whether these differences reflect new findings, or stem from model biases and methodological artifacts. It therefore remains unclear what new insights this study provides.

Other comments:

The authors compare two models with different resolutions and conclude that model resolution is important. However, these two models adopt different diffusivity and viscosity parameterizations. Because these parameterizations are critical for representing mesoscale and submesoscale features that is important for subduction pump, it is difficult to determine whether the differences between the two models stem from resolution or from these parameterizations.