Review of "Comparison of variables from ocean, sea ice, and atmosphere models as forcing data for iceberg drift and deterioration models in the Barents Sea in 2010-2014 and 2020-2021 (Part I)"

General Comments

This paper provides a detailed comparison of variables from ocean, sea ice, and atmosphere models (Topaz4b, Barents-2.5, ERA5, and CARRA) in the Barents Sea for the years 2010–2014 and 2020–2021. It establishes a foundational understanding of how these models differ in their representation of key environmental variables. The analysis is aimed at elucidating the potential implications for iceberg drift and deterioration modeling, which are addressed in the sequel paper.

I really appreciate that the Part I is thorough in its statistical analysis and highlights the interplay between model resolution, data assimilation, and physical parameterizations. but I am not sure if the Part I is appropriate and innovative enough because: (1) the paper's reliance on model intercomparison without significant validation against observational data limits its broader applicability; (2) that compromise the future usage of ocean and ice, as well as the atmosphere datasets forcing in the iceberg model because here I don't see some useful suggestions based on the intercomparison. I understand the observations are quite hard to get, but since the focus period is after 2000, which we had lots of in-situ (buoys and moorings both for ocean and sea ice), weather station, airborne, as well as the satellite observations, which can be used in the observations validation. So, the results don't convince me especially when the author find the big differences between Topaz4b and Barents-2.5, but did not mention which is more accurate in representing Barents Sea region in terms of the current state. This lack of clarity becomes a critical issue when reading Part II, where the authors show that iceberg drift and deterioration are highly sensitive to ocean and sea ice forcing. A practical question then arises: which forcing dataset should be used in iceberg modeling? Part I provides no clear answer, which undermines its practical relevance.

Therefore, I recommend that the paper undergo major revisions, at least bring intensive observation validations, before it can be considered for publication.

Specific Comments:

1. Line 10: "ERA5 show larger wind speeds", should be "ERA5 shows larger wind speeds".
2. Line 17 and Line 374: "Constraint to the physics and observations" → should be "constrained by the physics and observations"
3. Section 3.1.1: The projection of low-resolution datasets (e.g., 12.3 km in Topaz) onto finer grids using nearest-neighbor interpolation may introduce artifacts, particularly in regions with complex bathymetry. Wouldn't a more refined method, such as inverse distance weighting or bilinear interpolation, yield better results? Additionally, the temporal upsampling of Topaz from daily to 2-hourly intervals might affect the accuracy of comparisons and subsequent model outputs. The authors should critically discuss these potential issues.
4. Temporal resampling: I am uncertain whether the adopted 2-hourly temporal resolution adequately captures short-term dynamics, especially those relevant for iceberg drift. Rapid environmental changes, particularly in regions influenced by tides or strong winds, may not be well-represented at this frequency.
5. Figure 5: The caption for Figure 5 lacks a reference to the color legend, which makes the figure less intuitive for readers.

6. ERA5 and CARRA comparison: The comparison between ERA5 and CARRA seems unnecessary, as CARRA is partly forced by ERA5 surface fields. The dependence between the two datasets inherently limits the differences, as also demonstrated in Part II, where the authors show that atmospheric forcing has a relatively small impact on iceberg drift and deterioration. The practical utility of this comparison is unclear.

In general, while the authors have done an excellent job of identifying systematic differences between the datasets, the paper's reliance on intercomparison without validation is a significant limitation. This issue is compounded by the lack of clear, practical recommendations for selecting forcing datasets for iceberg modeling. Given the substantial sensitivity of iceberg behavior to ocean and sea ice conditions demonstrated in Part II, Part I should ideally offer stronger guidance on which model or dataset is more suitable for specific applications. Including observational validation and addressing potential artifacts from resampling methods would significantly improve the paper's credibility and practical relevance.