## Reviewer 2

The present paper reports results from a sensitivity study on the ability of two different numerical circulation models (ROMS and FESOM) in reproducing comparable eddy dynamic properties in terms of generation sites and propagation pathway in Fram Strait, a region characterized by a complex bathymetric configuration. Main difference between these models lies on their numerics and formulations, which includes numerical grid discretization, horizontal and vertical mesh resolution, parameterizations, and coverage (i.e. global vs. regional models).

The study is generally well written, and the analysis performed are appropriate. The figures are neat and clean. The discussion of the results is superficial. The authors should inspect more thoroughly the results presented. Nevertheless, the study stands as a good contribution, as it adds knowledge to the region. However, unfortunately it doesn't add new knowledge to the subject field whether in ocean modelling or in geophysical fluid dynamics. If it does, then perhaps it has not been clearly presented.

Find below few/minor points which the authors should address before its acceptance for publication.

We would like to thank you for your helpful comments. As you suggested, we improved the Discussion section. In particular, we added subsections discussing the connection between eddy occurrences and EKE (Section 6.1), the differences between models and observations (Section 6.2), and implications for contributing to future model development (Section 6.3). Below, we addressed your comments point by point.

The reference for ROMS in the introduction section may not be the most appropriate here!

In the introduction section, we added references to the papers by Shchepetkin and McWilliams (2005) and Budgell (2005).

Line 105 the sentence seems incomplete. In part it reads as "we also the Okubo-Weiss". Do you mean "we also used the Okubo-Weiss"?

Corrected.

Are there any criteria on the choice of the number of days taken as threshold to discard eddies with lifetime lesser than 3 days? Please note that the caption in Figure 10 and in other parts of the manuscript indicate a lifetime of 30 days. Is there any inconsistency??

We added a justification of the choice of the 3 days in section 2.3:
"We decided to use a threshold of 3 days mainly because the temporal resolution of the model output data is daily, and the eddy should form a track. This also helps to make sure that the eddies detected are real and not an over-detection due to uncertainties in the detection method. Eddies with a lifetime of at least three days are also required when computing the translation velocity needed to compute the eddy nonlinearity parameter, for which centred differences are used."
We analyzed all detected eddies (with lifetime>2 days) and discussed them in section 4.1, 4.2, 4.3 and 4.4. However, in sections 4.5 and 4.6 we decided to focus on long-lived eddies only. Analysis of pathways only makes sense for long-lived eddies. Otherwise the displacements are too short to be informative. In both sections, we made it clear that we only analyse "eddies with lifetime of more than 30 days".

Paragraph 130. What do you mean by eddy detected by experts?? Please make it clear.

Chaigneau et al. (2008) compared two methods of eddy detection, the 'winding-angle' method and Okubo-Weiss method. To validate their results, they asked oceanographic experts to detect eddies from sea level anomaly maps. Compared to the eddies detected by experts, the Okubo-Weiss method over-detected eddies. To avoid confusion, we just removed this sentence in the new version.

Did you use any filtering on the field of Okubo-Weiss parameter?

No, we do not use any filtering to compute the Okubo-Weiss parameter.

Perhaps the differences between the model maps shown in Figures 2 and 3 should be quantified to highlight geographical/spatial sites where the models converge and where they diverge (e.g: r=ROMS – FESOM for the parameters presented in Fig2 and Fig3).

Since the thermo-haline properties are not the main focus of this paper, we decided not to show the difference between ROMS and FESOM in T and S. Instead, we added contours of the 1°C and 2°C isotherms and the 34 and 35 isohalines, so that simulated T and S in both models can be more easily compared. Note that Figure 2 shows snapshots in time, so it would not make sense to plot the difference.

Under the Model assessment section, paragraph 155, the authors indicate the model simulates similar spatial distribution of the water masses. However, no T/S diagrams have been shown. Perhaps replacing the term "water-masses" with "thermo-haline properties" would be more appropriate. The same is true wherever this term appears in this manuscript.

As suggested, we added T/S diagrams for ROMS and FESOM in Figure 3. The differences in the simulated thermohaline properties are better illustrated with it. Model differences are described now in more detail in section 3. We also exchanged the term “water-masses” with the term “thermo-haline properties” where it is appropriate.

Lines 170 – 173, is the difference only related to the effects of tides simulated in ROMS? Could not be also somehow related to differences in the surface forcing fields between the models?

This is correct. The surface atmospheric forcing as well plays a role. We added this in Section 3.

Lines 180 – 184: Are the eddy statistics computed in a Lagrangian or Eulerian frame of reference? Please clarify this important aspect.

We added this clarification in Section 4.2:
“Eddy properties such as their radius are determined at the locations where they are detected. In this sense, the eddy statistics are computed in a Lagrangian framework.”

An important eddy property which could be included in this study is the eddy non-linearity parameter. This parameter would give a good insight on the eddy’s coherence and ability to trap and transport material along their pathway of propagation.

Thanks for this suggestion. We added a new section (4.7) to analyse the eddy nonlinearity. Chelton et al. (2011) describe three different parameters to study eddy non-linearity, the advective nonlinearity parameter, the quasi-geostrophic nonlinearity parameter and the upper-layer thickness nonlinearity parameter. Here we decided to focus on the advective nonlinearity parameter, defined as the ratio of maximum rotational speed and translation speed.

Lines 248 – 249. Are the tides the only difference between the models? What about the vertical discretization of the water column??

We changed the sentence to:
”One of the many differences between the models...”

There are much more differences, which are discussed in Section 7. In this paragraph, we just speculate that tides are the main reason for the difference in eddy occurrences on the Yermak plateau in the models.

In section vertical extension and hydrographic properties: Are the values of vorticity in the eddy centers, a single point values or are averaged values within the eddies? Please make this aspect clear.

The relative vorticity in the water column is computed in the eddy center, and thus it is a single point value for every depth layer. We added in section 4.6: "...by calculating relative vorticity/f at the location of the eddy centres..."

Line 279: is the superscript number after (Figure 11a, b)1 meant to be there?

The superscript refers to a footnote.