Reviewer 1

This paper presents a comparison of eddy fields in two models: a ROMS configuration and a FESOM configuration. The models seem to agree in general but differ some in the details. The analysis focuses more on eddy statistics and there is little dynamical analysis. The main conclusion is that the two model configurations simulate roughly similar eddy statistics.

I recommend major revision because I think the paper would be much better if it demonstrated a connection between the dynamical analysis (regions of baroclinic or barotropic instability) and the eddy statistics (e.g. generation regions of long-lived eddies). The individual pieces for analyzing this connection mostly exist so this seems reasonable to expect. Otherwise, the paper is well written and easy to understand. Some of the diagnostics used (e.g. T, S anomalies) were not described clearly —these need some major improvement (see below). The figures were nice to look at and illustrate the main points well though some extra panels would be useful (see below). Most of my other comments are minor.

We would like to thank you for the constructive comments, which significantly improved the manuscript. A major comment is about the connection between eddy statistics and dynamics. However, there is no 1:1 correspondence between (a) sites of eddy generation through local instability, (b) the energy conversion as a result of the presence and interaction of an eddy field with the background mean state and (c) the statistics of the diagnosed eddy trajectories (only considering coherent vortices, but not all the filaments and eddy genesis in-between). In the revised version, we discuss the connection between these points (Section 6.1). A deeper analysis, e.g. computing unstable growth rates as done by Isachsen (2015, https://doi.org/10.1002/2014JC010448), would be out of the scope of this paper.

As suggested, we improved the description of the diagnostics. Below are detailed answers to your comments.

1. deformation radius≈4x grid spacing. Is this "eddy resolving" or "eddy permitting"? Some discussion with references to literature would be useful background here.

We added this sentence in the introduction:

“Hallberg (2013) showed that in ocean models, a resolution of two grid points per Rossby radius of deformation can be considered as threshold between "non-eddying" and "eddy-permitting" regimes, and thus higher resolution is needed for a model to be considered as "eddy-resolving".”

2. Line 90: Does Figure 1 show the domain for the ROMS simulation? Is this the region with refined resolution for the FESOM run?

As you also suggested in 19., we added a second panel in Figure 1 showing a map of the Arctic Ocean. The red box indicates the area (76°N-82°N, 10°W-20°E) used for eddy detection. This region comprises the domain of the ROMS simulation (except for very small parts in the north-west and south-west corners, which can be seen in Fig. 2a). It is also the region with refined resolution in FESOM.

3. Line 100: Please list the 4 constraints.

We now listed the 4 constraints in section 2.3.

4. Line 100: Do a and b have units? What is the minimum possible detected eddy size in km? How does this compare to grid spacing?

We added a better description of a and b in section 2.3:
"Parameter a defines over how many grid points the increases in magnitude of v along the EW axes and u along the NS axes are checked, and its unit is grid points. It also defines the minimum size of detectable eddies, which is a-1 grid points. Parameter b defines the size (also in grid points) of the area used to find the local minimum of velocity."

5. Line 167: What was the longest gap you had to interpolate over?

For the WSC time series, the longest gap was 14 days. For the EGC time series, we used the time period Sep 8 2006 - Dec 31 2009 to avoid a long gap in the beginning of 2006. We added this description in the text.

6. Line 162: Did you average the spectra from the 3 moorings?

Yes, we averaged the velocity components u and v over the moorings, and then computed the spectrum from this time series. We added this clarification.

7. Since the eddies in FESOM are weaker (see spectra), should you adjust your criteria to be more appropriate for these weaker eddies? In other words, are you undercounting eddies in FESOM because they are weaker than the thresholds you are using?

We added a new panel in Figure 5 showing a spatial map of eddy kinetic energy (EKE) in both models. As also visible from the seasonal cycle of EKE, FESOM shows a higher

energy level in the WSC than ROMS, and in contrast, the energy level in the EGC is higher in ROMS than in FESOM. This is also reflected in the spectra.
Even if eddies were weaker in FESOM, they are defined by a circular velocity structure which is detected by the algorithm. Thus, the Nencioli method does not depend on the strength of the velocity field, but on the geometry of the flow field (closed contours). Both models have the same resolution (interpolated to the same regular grid), so there is no risk of undercounting.

8. Figure 6: How does this figure compare to a map of EKE averaged over the 3 years? 1. Is the sign or caption wrong in Figure 6c, d? It seems to say there are more anticyclones on the shelves contrary to line 189. 2. Also are the topographic contours the same as in previous figures? They are hard to see and it is hard to identify the1000m isobath referred to in the text.

We added maps of EKE averaged over 2006-2009 to Figure 5. The comparison of the EKE maps and eddy occurrences is discussed now in section 6.1. In fact, Figure 6c,d shows that there are more cyclones on the shelves than anti-cyclones (red color, where bottom topography is shallower than 200 m). We changed the topographic contours and increased the linewidth.

9. Section 4.3: This "eddy intensity" is really a Rossby number (relative vorticity / f). Can you call it that?

Yes, you are right. We take the Rossby number as an index for the eddy intensity. We added the description in section 4.3. This was also done by Kang et al 2014 (https://doi.org/10.1002/jgrc.20318) in their study of Gulf Stream eddies.

10. Section 4.5: These eddy pathways should be interpreted with the help of mean velocity vectors. Jet instabilities generate cyclones on the anticyclonic side of the jet and anticyclones on the cyclonic side of the jet. For the Svalbard Branch for example, one would see anticyclones on the deeper side but cyclones in shallower side which would partially explain the asymmetry in trajectories, i.e. cyclones and anticyclones are being generated in different places. I agree that the long term tendency for anticyclones to cluster in Boreas basin is likely a topographic effect.

Thank you for this explanation. We included an extended version of explanation that resorts to the fact that there is a front between different water masses: “Along the Svalbard coast, the Svalbard Coastal Current transports cold and fresh waters northward, close to the salty and warm AW which is carried northward by the WSC a little offshore. The meandering between the two water masses, light water on the eastern side and

denser water on the western side (roughly indicated by the 200 m isobath) leads to the generation of cyclones on the eastern side and anti-cyclones on the western (offshore) side."
The situation is comparable to the Gulf Stream, where dense (cold) water is located on the northern side of the main path, and light (warm) water is located south of the main path. This leads to the generation of cyclones on the southern side, and anticyclones on the northern side of the main path. We added this description in section 4.5.

11. Line 255: How do you define the vertical extent?

For every eddy center location (of the eddy detected in 100 m depth), we compute relative vorticity, temperature and salinity in the whole water column at that location, i.e. in every model layer. We added this description in the text: "...by calculating relative vorticity/f at the location of the eddy centres..."

12. Line 255: I also didn't understand how you define T, S anomalies. Are these T, S at all detected eddy center locations minus a monthly mean at the eddy center? Is this a climatological monthly mean or a monthly mean for a particular year the eddy was tracked? This definition should be clearly written.

Right, we compute T and S at the detected eddy center locations, and then subtract the monthly mean value at this location of the particular month. Description added in section 4.6.

13. Figure 8,9: Why is there a large difference in number of cyclone tracks between ROMS and FESOM in "FS Central South"

Indeed, more cyclones are generated in FESOM in "FS Central South" than in ROMS (269 and 107 generated cyclones in FESOM and ROMS, respectively, during the years 2006-2009). Not only cyclones, but also anticyclones are more abundant in FESOM than in ROMS in this region. This is consistent with the fact that baroclinic energy conversion in this region is stronger in FESOM (Fig. 12).

14. Lines 285-290: There are other regions with large conversions (e.g. just north and south of 77N) that are stable in Figure 12b.

- Would this disappear if you calculated the barotropic PV gradient properly accounting for u also?

We computed the barotropic PV gradient by also taking into account u. In particular, we defined segments of length 0.25° along the 1000 m isobath and approximated the along-stream velocity as the velocity tangential to the segments. The across-stream derivative was computed in the normal direction of the segments. However, the difference to the term d^2v/dx^2 is only minor, and does not reveal visible difference.

Nonetheless, there are small regions where the barotropic PV gradient changes sign near 77°N, which are maybe hard to see in Figure 12b. We modified the colormap slightly, and hope that the figure is more clear now.

· Did you use depth-averaged and monthly mean v? or monthly mean v at 100m?

We used depth-averaged monthly mean v velocity. We added this description in section 5.1.

· Why did you only use FESOM for Figure 12b?

Since the energy transfer terms in Figure 11a,b are very similar, we do not expect significant differences. For simplicity and also practically to reduce the workload of coauthors on the ROMS side, we decided to show it only for FESOM.

15. "EPE" — this should really be APE (available potential energy) since eddies are deriving their KE from the APE of the /mean/ state (Vallis 2006 textbook).

Corrected.

16. Line 315: This is a nice demonstration that the flow is baroclincally unstable but is it unstable in the regions necessary to explain the eddy tracks in Figures 8,9? Some connection between the instability analysis and the eddy tracks is needed.

Eddy tracks are determined by both eddy generation and eddy movement. The latter as depicted in Figures 8,9 makes it hard to link the tracks to eddy generation associated with instability. Instead, we can compare the number of detected eddies (Figure 6a,b) with the maps of EKE (Figure 5) and the energy conversion rates (Figure 12). We added a subsection (Section 6.1) where we discuss this connection:

"Although baroclinic instability is the main driver of mesoscale eddy variability, the connection between eddy occurrences and EKE as well as the APE to EKE conversion rate (w'b') is very non-local. For one thing, the eddies form as the result of nonlinear evolution of baroclinic instability waves and jet meanders. For another, mean circulation transports and modifies all eddy-like features, moving them away from the sites where their 'seeds' originally appeared. As a result, the observed pattern of eddy occurrences (Figure 6a,b) differ from the EKE and w'b' distributions (Figures 5 and 12)...."

17. Lines 345, 346: I'm not sure there was any discussion of "EKE→MKE conversion associated with steepening isopycnals". Please describe this in more detail. The "energy backscatter" parameterizations being currently developed would be useful literature to reference (e.g.Jansen et al 2015)

We reformulated the last paragraph in the discussion, and hope to explain it better now. Thanks for pointing us to the study by Jansen et al. We also cite the papers by Juricke et al, which describe the implementation of the method by Jansen et al. into FESOM and its application to realistic test cases.

18. Discuss sensitivity of eddy maps to choice of 100m depth

We added a justification of the choice of the depth of 100 m:
"We decided to choose the depth of 100 m because both main water masses of the Fram Strait, AW and PW, are present at this depth (e.g. Wekerle et al., 2017, their Figure 9). In addition, the maximum relative vorticity is also located close to this depth for different regions of the Fram Strait (Figure 10)."

19. Figure 1: Showing a second panel with a larger region would be nice for readers not familiar with this place.

Thanks for this suggestion. We added a second panel in Figure 1 showing the Arctic Ocean and our study region.

20. Figure 4: It is more conventional to show logarithm axis scales instead of showing log10(quantity) on a linear scale. Why did you choose to not do that? Please mark1/30 days since that's the frequency you chose to separate "mean" and "eddies"

As you suggested, we changed Figure 4 and now show logarithm axis scales. We also marked the frequency at 1/30 days.