Response to Reviewer 1

Text from reviewers is in black italics with responses in blue.

This paper describes and evaluates the high resolution CCMEP analysis which spans the coastline of Canada and the Arctic. The model and assimilation methodology are introduced in relation to the previous regional system. The performance of the system is assessed by an evaluation of background departure statistics against sea level anomaly and in-situ observations, and compared against the Global Ice Ocean Prediction System.

Firstly let me apologize to the authors for the delay to my review and thank the editor for allowing me sufficient time to complete it. I am glad to be able to have thoroughly read the paper and I feel it very well describes the operational system. Overall I am sure this paper is perfectly suited to GMD and should be published subject to some minor corrections/additions.

I will first give my more general comments, and subsequently give the itemized comments I had whilst reading through the paper.

General comments:
The paper is very well written and was enjoyable to read. For a description of an operational analysis system the authors have the balance just about right in terms of the detailed description within this paper and referencing previous work.

Thank you very much.

In some areas I was not clear whether those references would suffice or if they were differences between the previous system and RIOPSv2. For example the boundary conditions used for the regional system are not well highlighted (I assume they come from GIOPS but this isn't explicitly stated?). Also when the change to model levels was mentioned it highlighted that the vertical domain was not well described and may lead to confusion so a sentence or two would go a long way to help the reader here.

Additional details have been added regarding the open boundary conditions and the statement regarding the model levels was modified to improve clarity.

The section on the online harmonic analysis was far more in depth than the others. I am not familiar with such methods and therefore not sure where the methodology from the literature finishes and the new science (I think related to the sliding window) starts. If this could be highlighted it would be a useful addition.

In response to this comment and a request by Reviewer 2 to include a greater discussion of the performance of the online harmonic analysis scheme several changes have been made. A paragraph has been added to the Introduction to provide better context, including previous studies that assimilate SLA in an ocean model with tidal variations. An additional figure and discussion of the performance of online harmonic analysis filter compared to the well-known T_tide package has been added to Section 3.4.4

More generally what is not discussed is the mismatch between a model that contains tides (needed for the coastal applications) and SLA observations which do not contain tidal signals. However the raw altimetry data would necessarily contain tidal signals, so it would benefit the reader to have a short section on SLA observations and why you aren't using lower level SLA observations.
Assimilating lower level SLA observations that include tides would only be suitable if the intend was to improve the representation of the tides themselves. Since the tidal errors are mainly stationary, this can be treated in advance to separate the tidal and non-tidal signals present in the altimetry data. This also allows use of "standard" altimetry products that include the tidal filtering along with the other processing steps, such as the

dynamic atmospheric correction and long-wave error filtering. A comment to this effect has been added to Section 3.3.3.

I was hoping to see more about sea ice performance but the final paragraph is sufficient to let the reader know why it is not shown. I wonder if this could be expanded upon to hint at how a future system might look to improve sea ice predictions?

As noted by Reviewer 2, sea ice forecasting is outside the main focus of this paper. As such, we don't feel it is appropriate to expand upon this topic in this manuscript. There are other papers that focus directly on sea ice forecasting in RIOPS (e.g. Lemieux et al. 2016; Chikhar et al., 2019).

Finally I would ask that the authors look at the colour scales they are using in their plots. I think a lot of the structures we see are artifacts of the jet colour scheme and the use of perceptually uniform colour scales would remove such artifacts and make the paper more accessible to colour blind readers.

Thank you for the suggestion. In the revised manuscript, Figs. 6-11 have been reproduced using perceptually uniform colour scales. Conclusions and comments in the text regarding these figures remains unchanged.

Detailed comments:
Line 126: I am unfamiliar with these tidal constituents - is this detail necessary and if so should such parameters be described?

Constituents specified are typical and well known within tidal modelling community. However a different number of constituents are used in different contexts. Providing the list of constituents here improves reproducibility of the results.

Line 137: Please comment on the vertical domain - "deep layers (from 500m" suggests these are the deepest layers in your domain, but I would expect the domain to go down to _4000m in the arctic.

Thank you for pointing this out. Indeed, the wording is confusing. The sentence has been modified to clarify that the increased resolution is focused between 250 m and 500 m.

Line 170: This has strong similarities to the ECMWF workflow shown in Browne et al. 2019 Figure 3. ::: Browne, P. A., de Rosnay, P., Zuo, H., Bennett, A., & Dawson, A. (2019). Weakly Coupled Ocean-Atmosphere Data Assimilation in the ECMWF NWP System. Remote Sensing, 11(234), 1–24. https://doi.org/10.3390/rs11030234

We agree with the reviewer that there are strong similarities between the approaches. A comment to this effect has been added to the text.

Lines 209-211: "Another modification required for coupled forecasts was to use 24-h averaged short and long-wave radiation fields to force NEMO-CICE during the analysis cycles such that there is very little diurnal warming present in the ocean analysis". Please can you elaborate on this. Is the system designed to have a strongly damped diurnal cycle in the analysis, or are there reasons for which this was found to be necessary?

The following text has been added to clarify:
"Damping diurnal SST variations in the analysis fields was also found to limit initialization shock in coupled forecasts as the atmospheric analysis was produced using a foundation SST product (Smith et al., 2018)."

Line 215: Smith et al (2015) should be 2016? Or a different paper?

Thanks for pointing out this error. The article was available online in 2015. All references to this article have now been changed to 2016 to avoid confusion.

Line 228: Can you clarify - does this mean you apply the increment from the current cycle to the next cycle (in the first 24 hours)?

Yes, that's correct. As such, the first day of the 7-day IAU run has two increments applied: an increment with a linearly decreasing ramp from the previous cycle, and an increment with a linearly increasing ramp from the current cycle. This helps to make a smooth transition from one cycle to the next.

Figures 3-6: Subplot titles are missing commas, i.e. <SST(r)SST(r+\Delta r)> should be <SST(r),SST(r+\Delta r)>

Perhaps this is a case of different conventions. It is quite common to express an expectation of a product (say A and B) as <AB>. Also, as noted below, it is also quite difficult for us to reproduce these figures as it is computationally time consuming.

Figure 3 - star is not visible, please can you add it on top or produce the image at a higher resolution? In fact maybe just change the text to make it clear the point you are considering is the bottom left corner of the magenta box?

The star has been indicated more clearly.

Figure 3 - this is the first time you mention localisation. Please can you elaborate on why you choose the region that you do? Is the localisation a tensor product of horizontal and vertical localisation functions, or is there no localisation in the vertical? Is it a hard cut off or is it smoothly reduced to zero at the boundary of the "bubble" by a Gaussian or Gaspari-Cohn function? (I suspect you will have a reference for this which would be good, but also nice to know if you have had to make any changes here for RIOPSv2).

There is no localization in the vertical. The horizontal localization is applied using a Gaussian function with the the length scale indicated in the figure. The figure caption has been modified to clarify this point. No changes have been made to the localization.

Figures 3-11,13,14: If possible please replot using a diverging, perceptually uniform colour scale. For instance in Figure 3, <SST(r)T(r+\Delta r)>, the current choice of colourscale appears to show a clear change in correlation around layer 10 whereas I think this is an artifact of the rainbow colour scale. Thyng, Kristen M., et al. "True colors of oceanography: Guidelines for effective and accurate colormap selection." Oceanography 29.3 (2016): 9-13. https://doi.org/10.5670/oceanog.2016.66
Hawkins, Ed. "Scrap rainbow colour scales." Nature 519.7543 (2015): 291-291. https://doi.org/10.1038/519291d

The colour scale has been changed to a perceptually uniform colour scale for Figs. 6-11, as these figures include the main quantitative comparison presented in this paper. As the other figures are used for qualitative comparison we feel the use of the jet colour scale does affect the scientific interpretation of the results. Reproducing Fig. 3 would be quite difficult as the calculations involved are extremely time consuming. The colours are used to provide a qualitative indication of the covariance scales, and as such we feel they are suitable.

Line 278: "model trial field" is not well defined. I suggest replacing with something like "raw model equivalent field"

Text modified as suggested.

Figure 7: Please relabel "CMC" -> "CCMEP" and "Trial" -> "RIOPSv2"

Figure modified as suggested.

Line 299: ", the use of a rotation" -> ", and the use of a rotation"

Modified as suggested (but with the comma removed).

Line 306: Are you using Einstein notation here so there is an implicit summation over the repeated index k? This should be noted. Oh I now see it introduced on line 320 - is it used before or only from line 320?

Text clarified to indicate that Einstein notation used throughout.

Line 311 - the definition of \omega_k is unclear. please look carefully at the wording you use here.

Changed “at” to “is”.

Line 315: Is this a minimisation over only Xˆk ?

Yes. To clarify, we changed the equation as follows:

$$J(X^k) = \frac{1}{2}(A^n - H^n)^* W_{nm}(A^m - H^m)$$

Line 338: Prime symbol is not consistent - MS Word issue?

Thank you for noticing this. Text corrected.

Line 350: "According to the definition of the sliding window weight". Where is this defined? Eqn 3a? Also have you defined explicitly what "its property of normalisation" is?

Paragraph rewritten. Figure added in response to Reviewer 2 to provide more detail regarding the sliding window weight.

Figure 8: Can you clarify, possibly in the caption, that (a) = (b) - (c) - (d)? Again, diverging colour scales would be much better here. In fact the total SSH field (b) is confusing: "Panel (b) shows the instantaneous model SSH field prior to any treatment." but its range is (-1,1). Has this already had the MDT removed? I personally am more familiar with considering the MDT for SLA assimilation when using processed SLA observations that have tidal signals removed, so I would like to see precisely where the MDT is required in your methodology.

The caption has been modified as requested and the figure has been reproduced using diverging colour scales.

The SSH field is the full field, prior to removing the MDT. The MDT is removed as is usually done as part of the observation operator. This is the same for a model without any tides. Perhaps there is some confusion regarding the SLA observations assimilated, as the SLA observations used here are the commonly-used AVISO product that have the tides removed.

Line 477/478. Regarding the Gulf stream errors, would you expect the SEEK filter methodology to be able to effectively constrain a region which such variability given the error covariances are coming solely from a climatology and therefore may be far too smooth there? It is nice to see the errors reduced in RIOPSv2.

This is an important point. The covariances do not come from a climatology, but rather are constructed from sub-monthly anomalies from a 10-year model run. As such, each anomaly (or error mode) is at roughly the same resolution as the model (apart from SST for which there is additional filtering as described in the text. The approach was developed with the intent to constrain the mesoscale variability (See Lellouche et al., 2013 and references therein). That being said, GIOPS is only of eddy-permitting resolution and would resolve less well the structures in the Gulf Stream region, whereas RIOPS resolves these features better. The text has been modified to clarify this point.

Line 485: The difference in errors in the Laptev sea is intriging. Could you produce a comparison plot of the two different MDTs used?

Both RIOPS and GIOPS use the same MDT field. The confusion is probably due to the word "field" being erroneously used in the plural. This has been corrected in the revised manuscript.

Line 488: We should see errors here in salinity too if that is the case? [Refer to my later comment on Figures 11-13]

Unfortunately, there aren`t any in situ observations in the areas for which we see the largest biases (Hudson's Bay, northern Laptev Sea, mouth of Lena River). When activating the bias correction scheme we did note that the mean SLA innovations increased suggesting that water mass errors do contribute to this signal.

Line 489-491: Surely this is entirely due to ice cover - polar orbiting satellites should have much richer coverage in the arctic due to their much more regular return period (if you exclude the polar data gap).

While the reduced number of observations for Sentinel3, Altika and Cryosat2 are due to sea ice cover, observations from Jason3 are also included, which do not cover the Arctic Ocean. As such we feel the statement in the manuscript is correct: "…due to satellite orbits and ice coverage, many fewer observations are present over the Arctic Ocean…"

Section 4.2.2/Figure 10. Given the broadly similar structures and different scales between Gulf stream errors and other regions it would be nice to see a difference plot and/or a normalised difference plot of 10(a) and (b). This might be too difficult to produce in a reasonable amount of time/effort, but may prove enlightening.

Thank you for the suggestion, but we find using a normalized difference plot over-emphasizes small differences, whereas we prefer to focus on the main patterns and areas of error.

Figures 11-13: They appear to be missing the polar sector from 45E to 180E. Is this a plotting error? Are there no in-situ obs in the Hudson bay? I was hoping to see salinity departures which would corroborate interpretation of errors around the mouth of the Lena river that you associated to the fresh water fluxes.

We fully agree with the reviewer that in situ observations in Hudson Bay, at the mouth of the Lena River and more generally throughout the Arctic would be of significant value, in particular for helping to corrected diagnose and improve SLA errors. Unfortunately, there are no observations available in these specific regions in the dataset we use. The availability of in situ data in the Arctic is a significant issue (for a review see Smith et al., 2019).

figure 11 - Salinity worse in the Baltic sea - are there problems in brackish waters?

The problem appears quite localized and different than what is seen, for example, in the Gulf of St. Lawrence. As such, we don't feel it's a generalized problem but rather likely associated with errors in freshwater runoff applied in the Baltic Sea.

Line 522: "upper 50 m of the water column" did you mean 500m as in the plot?

Thank you for pointing this out. Figure 12 is incorrectly referenced in this context. The error is indeed localized in the upper 50 m as indicated in the text. The text has been modified to indicate that this is "not shown" (rather than referencing Fig. 12).

Line 540 - should be Figure 14?

Yes. Thank you for catching this oversight.

figure 15(b) - are units km?

Yes. Units for other subplots have also been added.

Line 545: please refer to table 3 for the RMS numbers too.

A reference to Table 3 was added after the comment regarding the RMS SLA innovations.

Table 3: Is this R or Rˆ2 you are listing?

As indicated in the Table caption, the correlation values (r) are provided, not the proportion of explained variance (r^2).

Line 546/546: "likely representing the delay in the analysis system in adjusting to SLA errors". Could it not be due to the non-stationary location of the front/eddies? And looking only at observations with a 10 day return period you are not capturing the evolution of the model between overpasses?

We agree with the reviewer that there may be others causes of the differences noted in the figure. In addition to the theory proposed by the reviewer, they may also be associated with anomalies in the observations. We have removed this speculation from the manuscript.

Line 553 - Figure 15 -> Figure 14. Also subsequent references to Fig 16 -> Fig 15.
Figure 15: Caption is for Figure 14
Figure 16 not present but referenced.

Our sincerest apologies for the oversight related to the missing caption and incorrect figure references and any confusion it may have caused in reviewing our manuscript. In response to comments by Reviewer 2, additional figures have been removed and all figures numbers and references have been updated.

Line 622 "OPP" -> "YOPP"?

No. OPP is correct. The acronym has been expanded in the revised manuscript as Ocean Protection Plan.