First, we'd like to thank the reviewer for the time and effort on reviewing our paper. We manage to address all the points from the reviewer and try our best to improve the quality of our paper. Below is our response to the reviewer's comments point by point. We will revise the manuscript accordingly.

(the *original comments* are given by *Italic gray text*, and each follows our response in plain text.)

Response to the major points

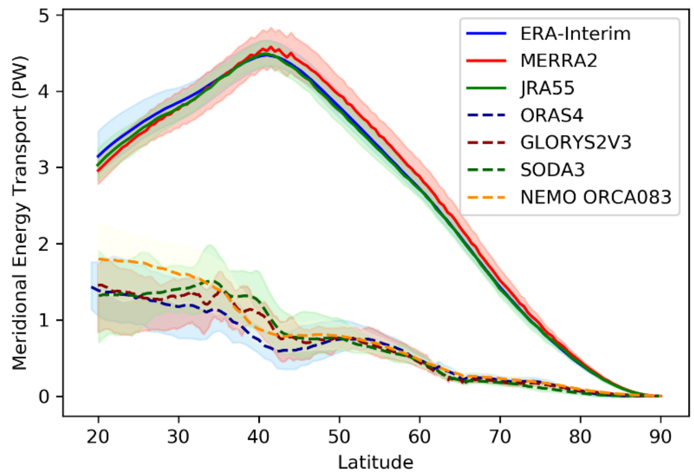
Computation of an energy transport from reanalysis can be difficult due to lack of mass / energy conservation, sampling problems, or numerical schemes different from the original. It seems from figure 2 that MERRA2 and JRA55 suffer from large errors ("noise" on the curves) of the order of O.1-0.2 PW for the long-term mean. Unfortunately, this means these products will be unusable to study the variability of the transport. Indeed, it is evident on Fig. 3 that the interannual variability at 60N is of the same order as this long-term noise.

There are 2 possibilities:

- This problem comes from the reanalysis themselves: the conclusion is then that only ERA-interim is useable. This would be a useful result in itself; but the study of the differences between ERA and the other reanalyses becomes pointless and the paper could instead concentrate more on the impacts on the Arctic.

- This is a problem in the calculations from the reanalysis data, which then has to be solved, and the other results corrected accordingly.

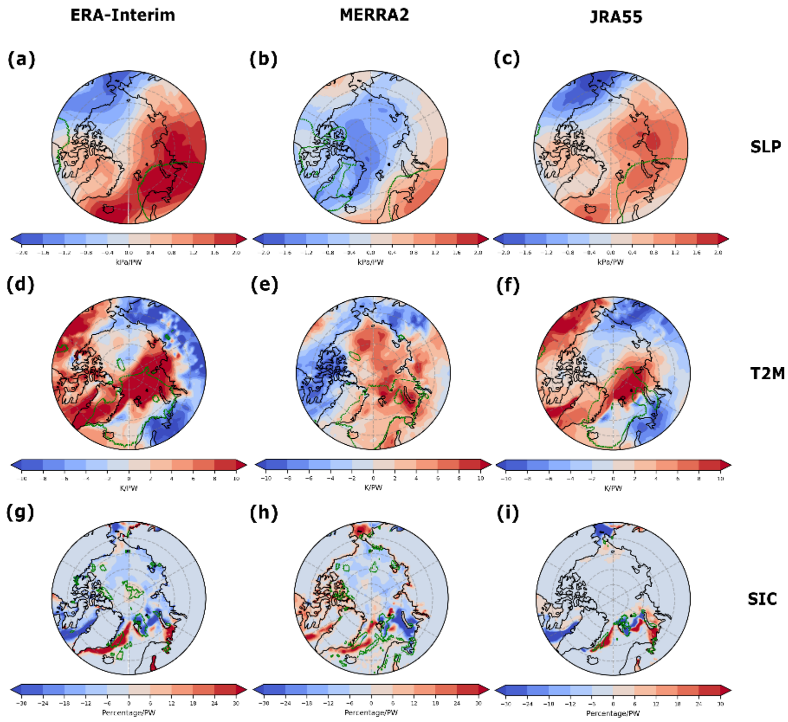
Thanks for the comment. We also have concerns about the noisy annual mean AMET from MERRA2 and JRA55. After thoroughly checking our script and the whole computation procedure, we find that the "noise" in annual mean AMET from MERRA2 and JRA55 in Figure 2 comes from the numerical scheme used in performing the mass correction. The barotropic mass correction method involves calculations of the divergence, the inverse Laplacian and gradients (Trenberth, 1991). We used finite difference (central scheme) method to compute these terms. However, after some tests we noticed that it is more accurate to bring the fields to spectral domain and compute these terms using spherical harmonics. The adjustments to the barotropic wind as results of mass budget correction based on these two numerical methods differ much in the polar regions. Since AMET are very sensitive to the mass budget, the results with these two different numerical schemes can lead to very different results. After recomputing the AMET fluxes from ERA-Interim, MERRA2 and JRA55 with mass correction using spherical harmonics, now the noise in annual mean AMET has been removed. The results are shown in the figure below and will be included in the revised manuscript.



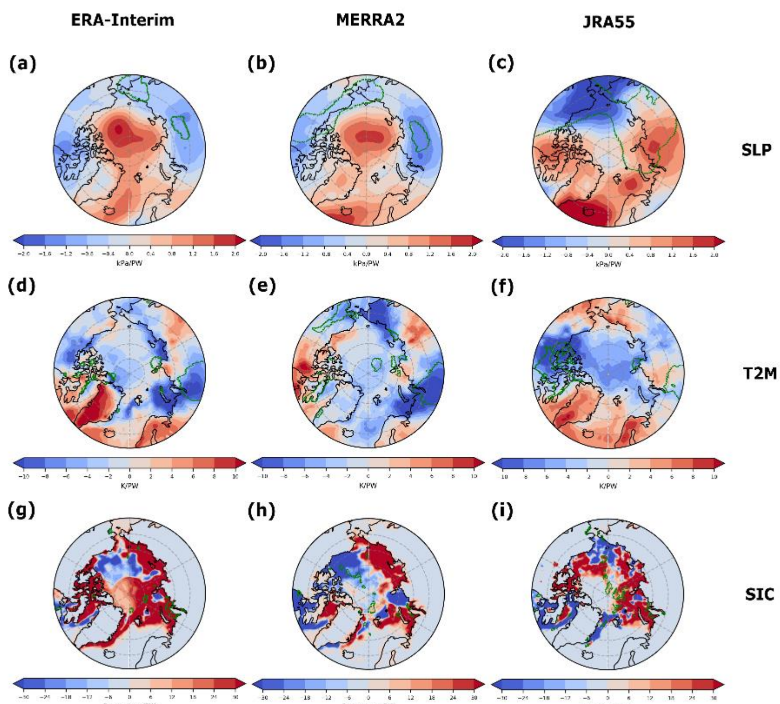
The other results are corrected accordingly. The difference between annual cycles of AMET from ERA-Interim and JRA55 becomes smaller, but the low frequency AMET remains different. Hence the study of the difference between the chosen reanalyses remains informative. As the main point of this paper is to raise a caveat about using low frequency MET from reanalyses, the major results in our paper remains the same. Changes are made to the regressions mainly.

(B) Impacts on the Arctic (section 3.3) The problem of this section is that it gives a few quick examples of regressions of characteristics of the Arctic climate on energy transports, but that they are too short to be really useful. - Show at least the winter and summer seasons, as results could be quite different. - Why show SLP and temperature for AMET, but sea-ice for OMET? Why not show the same variables for both for comparison? (at least sea ice and slp).

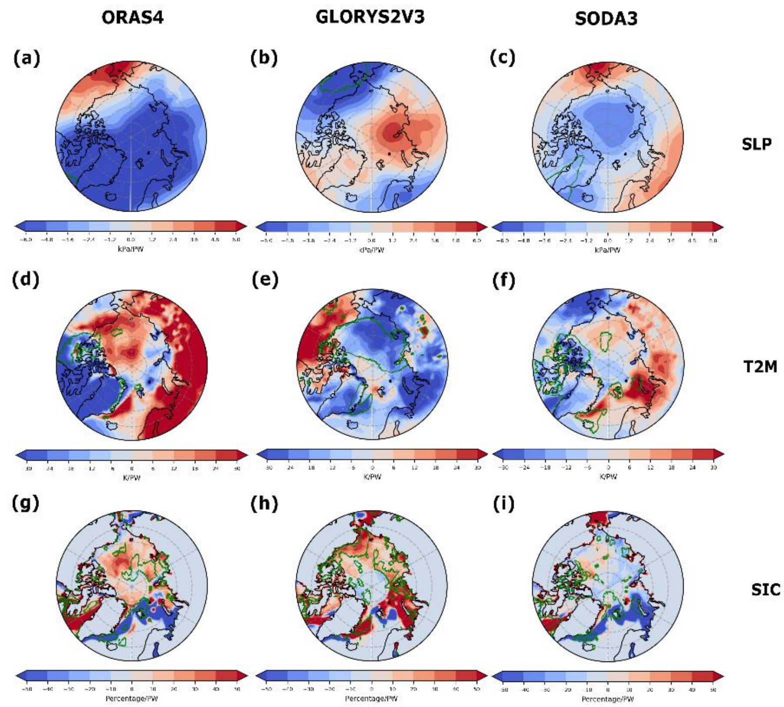
We agree with the reviewer more results are needed to provide more insights about the link between MET and the Arctic and potentially causal relations. Hence, we will include regressions of SLP, SIC for both AMET and OMET in summer and winter and interpret them. In summary they are shown here already and we will further explain these regressions in our revised manuscript:
(the strange-looking colors near the pole for regressions on sea ice in ERA-Interim are due to the quality of sea ice fields in ERA-Interim, see our response to the last small remark.)



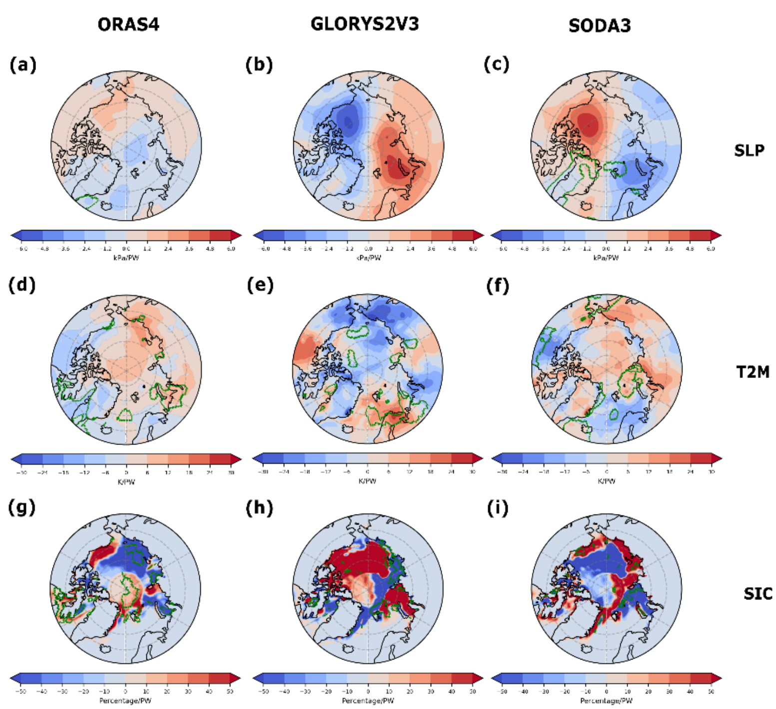
Regression of SLP (a-c), 2 meter temperature (T2M) (d-f) and SIC anomalies on AMET anomalies at 60 N at interannual scale with no time lag in winter (DJF). The monthly mean fields are used here after taking a running mean of 5 years. From left to right, they are the regression of fields on AMET of ERA-Interim, MERRA2 and JRA55. The green contours indicate a significance level of 95%.



Regression of SLP (a-c), 2 meter temperature (T2M) (d-f) and SIC anomalies on AMET anomalies at 60 N at interannual scale with no time lag in summer (JJA). The monthly mean fields are used here after taking a running mean of 5 years. From left to right, they are the regression of fields on AMET of ERA-Interim, MERRA2 and JRA55. The green contours indicate a significance level of 95%.



Regression of SLP (a-c), 2 meter temperature (T2M) (d-f) and SIC anomalies on OMET anomalies at 60 N at interannual scale with ocean leading by one month in winter (DJF). The monthly mean fields are used here after taking a running mean of 5 years. From left to right, they are the regression of fields on OMET of ORAS4, GLORYS2V3 and SODA3. The green contours indicate a significance level of 95%.



Regression of SLP (a-c), 2 meter temperature (T2M) (d-f) and SIC anomalies on OMET anomalies at 60 N at interannual scale with ocean leading by one month in summer (JJA). The monthly mean fields are used here after taking a running mean of 5 years. From left to right, they are the regression of fields on OMET of ORAS4, GLORYS2V3 and SODA3. The green contours indicate a significance level of 95%.

Response to the other remarks

p.3 line 24: "it is preferable that they incorporate the latest..."

p.3 line 25: "they better not resemble each other" well, you certainly hope that different reanalyses would be consistent with each other!

Thanks for pointing out this. The sentence is very confusing indeed. Here we want to emphasize we try to include various reanalyses data sets to make the comparison more informative. We have deleted this sentence.

- section 2.3.1 : what is the value of Lv used here ? Lv varies with T in nature, but not necessarily in models... Not sure what's used in reanalyses.

Same as recent studies (e.g. Mayer et. al., 2017, Trenberth & Fasullo, 2018), we use fixed Lv = 2264.67 KJ/Kg. We include this in the method section now.

Mayer, M., Haimberger, L., Edwards, J. M., & Hyder, P. (2017). Toward consistent diagnostics of the coupled atmosphere and ocean energy budgets. Journal of Climate, 30(22), 9225-9246.

Trenberth, K. E., & Fasullo, J. T. (2018). Applications of an updated atmospheric energetics formulation. Journal of Climate, 31(16), 6263-6279.

- p.7 : there may also be issues due to different horizontal advection schemes used in the reanalysis and in the post-treatment.

Thanks for pointing out this. We will include this in the text to be accurate.

- p.8, line 6 : unit should be J/(kgC) or J.C-1.kg-1.

We will correct it.

- p.8 : reference temperature. If the unit of potential temperature Theta is in C, then substracting a reference temp. of 0 C does not accomplish anything. Are you instead converting Theta from K to C to avoid cancellation of large terms problems? This is very unclear.

Reference temperature is used to account for the recirculation and the mass imbalance. Normally people take reference temperature as 0 deg C (e.g. Zheng and Giese, 2009). Since the oceanic reanalyses used here already saved theta in deg C, this actually would not affect the computation. But conceptually a reference temperature should be taken into account.

Zheng, Y. and Giese, B. S.: Ocean heat transport in simple ocean data assimilation: Structure and mechanisms, Journal of Geophysical Research: Oceans, 114, 2009.

- p.9, l25: The differences in resolution are actually small. There must be another explanation to these variations, which are key (main point A)

Now we learn that the noise is due to the choice of finite difference scheme. We will delete this.

- p.10, l10: In ocean models that are not eddy-resolving, there is both an eddy advection (Gent-McWilliams) and a diffusive heat transport, which can be significant compared to the resolved one. How were these incorporated in these analyses? They absolutely need to be taken into account.

Thanks for the comment. In ORAS4, an eddy parameterization scheme from Gent-Mcwilliams (1990) is implemented. The implementation of this eddy parameterization scheme can lead to a big difference in volume transport and heat transport, compared to eddy-permitted models (Stepanov & Haines, 2014). However, in this case the computation of OMET with ORAS4 doesn't include the contribution from eddy-induced velocity as the fields related to the use of eddy advection schemes are not saved by ORAS4. We will include such information in our revised manuscript.

Stepanov, V., & Haines, K. (2014). Mechanisms for AMOC variability simulated by the NEMO model. Ocean Science, 10(4), 645-656.

- section 3.2 and the accompanying figures for the atmosphere is a bit pointless given the low quality except for ERA-interim.

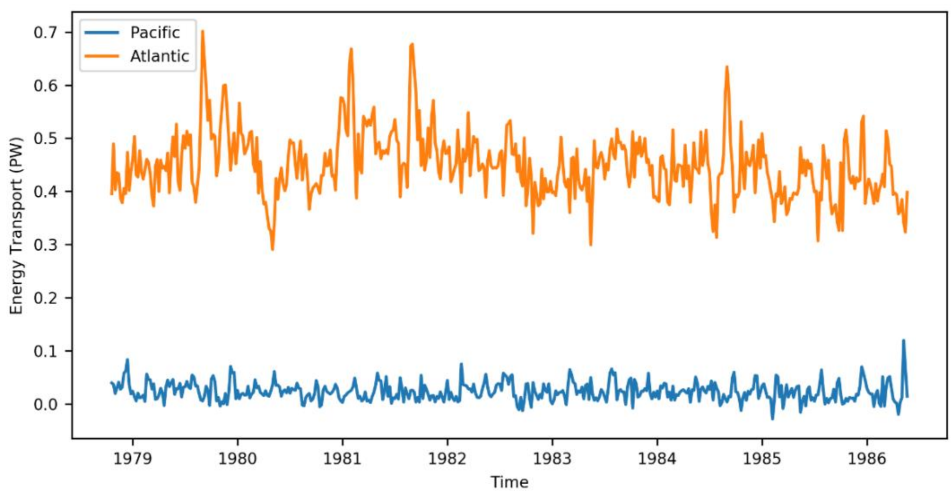
After we updated the results now these comparisons become informative. We will update this part with newly computed fluxes.

- p11, bottom: Are we looking in this section at the total OMET, or only the Atlantic OMET? It would probably be more interesting to look at Atlantic only at 60 N (in line with section 3.3), although knowing the relative roles could be good. Same question for page 13.

Thank you very much for the comment. We are looking at the Atlantic OMET in this part.

Below we show the OMET at 60N in the Pacific and Atlantic from ORAS4. It can be noticed that the OMET in the Atlantic is much larger than that in the Pacific. Note that the Atlantic OMET dominates both the mean value and the variability.

Actually, we also tried the same regressions with total OMET. The results are almost the same.



- p12 : any idea about why ORAS4 seems to work best ?

Given the big differences between chosen oceanic reanalyses as well as many factors that could potentially influence the results, we cannot explain conclusively why ORAS4 seems to work best here.

- p13, l28 : why 5-years timescale? Is it specific to the Arctic? Are the regressions on Fig.13 for 5-yr filtered data, or just year-to-year?

We are focusing on the influence of low frequency signals of OMET on the Arctic. Early studies with climate models also indicate that there are strong correlations between Arctic sea ice and OMET variations at interannual to decadal time scales (Van der Swaluw et al.,2007; Jungclaus and Koenigk, 2010). We take 5-years timescale because reanalysis time scales are relatively short, but we do intend to study variability beyond the annual time scales. Even longer windows for filtering would lead to a too small sample. Figure 13 is the regression with 5 year filtered data. We will add more details to the caption.

Van der Swaluw, E., Drijfhout, S., and Hazeleger, W.: Bjerknes compensation at high northern latitudes: The ocean forcing the atmosphere, Journal of climate, 20, 6023–6032, 2007.

Jungclaus, J. H. and Koenigk, T.: Low-frequency variability of the arctic climate: the role of oceanic and atmospheric heat transport variations, Climate dynamics, 34, 265–279, 2010.

- Figure 1 :this is hardly commented in the text: seasonal cycle, low contribution of LH (especially to the seasonal cycle)... By the way, I would replace this figure by either the components of the time-mean transport as a function of latitude, or by the mean seasonal cycle at 60 N of the different components. (With figures for the interannual variance if needed)

Thank you for the comment. We agree that the figure is not quantitatively clear for the seasonal cycles. We will replace it with mean AMET at 60N as a function of month.

- Figure 2: The dashed-line is built using annual means?

Yes, the dashed-lines are also annual means. Here we use solid lines and dashed lines to show AMET and OMET, respectively.

- Figure 3: panel b) needs confidence intervals, based on interannual variance. For panel a), do the standard deviations include the seasonal cycle or are they for interannual anomalies?

We will include confidence intervals for figure 3b.

In figure 3a the standard deviations include the seasonal cycle. (The standard deviations for interannual anomalies are included in figure 3b)

- Figure 4: not sure what is the point of these figures, apart from showing that the high-res analyses have eddies? It would be easier to copare maps integrated to the same low resolution, or time-means, and also to see the GM components.

As these oceanic reanalyses use different curvilinear grid, to integrate them to the same low resolution can introduce errors. We agree that the figures here are a bit pointless. We will delete these figures and only explain these in the text.

- Figure 6 could easily be replaced by the figures given in the text. Note that latent heat transport may not contribute much to the differences because it’s low to begin with, but also because it’s concentrated in the near-surface layers, so does not suffer too much from slight mass flux imbalances.

Thanks for your comment. We will replace these figures by texts.

- Figure 11, caption: I guess it is "interannual" time-scale? (i.e. year-to-year variability)

Here we use annual scale to refer to the 1-year filter we applied to the data. By saying interannual scale, we mean a 5-year filter is applied. Sorry for the confusion.

- Figures 11-13: consistent time-scales and variables, with different seasons would be nicer. Also, the green shading masks the color underneath, making it hard to read. Same-color shading maybe?

We will include consistent time scales for the regressions of AMET and OMET on SLP and SIC in summer and winter. We will change the green shading to contour lines to make it easy to read.

- Figure 13b: there are strange-looking colors (opposite the rest of the Arctic...) near the pole in all 3 plots (abrupts changes of sign...) Is this an artefact?

No, this is not an artefact. There are some quality-issues with sea ice in ERA-Interim, which can be inferred from an evaluation of reanalyses data sets concerning near-surface fields (Lindsay et. al., 2014), as ERA-Interim air-sea flux fields account for local sea ice concentration (IFS DOCUMENTATION – Cy31r1, Chapters 3 and 7). Such strange colors near the pole do not exist in figures with sea ice fields from MERRA2 and JRA55 (see response to major point 2). That's why a 'disc' appears close to the pole in some plots. We will add a note to this issue in the text.

Lindsay, R., Wensnahan, M., Schweiger, A., & Zhang, J. (2014). Evaluation of seven different atmospheric reanalysis products in the Arctic. Journal of Climate, 27(7), 2588-2606.

Again, thank you so much for the comments. We are sincerely gratitude to every point the reviewer made to improve this paper. Thanks a lot!

With best regards

Yang Liu, Ben Moat, Jisk Attema and Wilco Hazeleger