Response to review comment #2:

This paper decomposes the global heat transport into basins-scale, large-scale and mesoscale transports. Th main thing is to do so using a spatial filter, rather than the more common time-mean approach. The spatial filter results in different values for eddy heat transports than the time-mean method. However, little discussion is provided o why and how this benefits future work and what the differences really mean. Also, it is stated that this work help parameterization, but it is unclear how exactly. Overall, this paper could potentially be published, but it needs a clearer narrative, fewer figures, and improved use of symbols and names. Also, the importance of the results should be better explained and parts that are doubtful should be removed or require further argumentation.

Thanks very much to the reviewer for their comments. We have described our revisions to the manuscript to address these comments below. Note: The responses to each of the comments are given in red text. Line numbers in the responses refer to the "tracked changes" version of the revised manuscript, so the line numbers are different in the version without tracked changes.

Overall, I would recommend reducing this paper by a couple of figures and with that be more to the point about the most important results here. What is the main story and which figures do you need to prove this?

The key results of the paper, as clarified in the abstract, include:

- The mesoscale-induced divergence of heat from the subtropics towards the equator and the poles, caused by both stationary and transient mesoscale phenomena.
- The differences between spatial and temporal definitions of the eddy flux. The temporal eddy flux does not include stationary mesoscale recirculations, while temporal eddy flux includes long planetary wave contributions but mesoscale flux does not).
- Interannual/decadal variability in the mesoscale contributes to interannual/decadal variability of heat transport primarily in the North Atlantic, tropical Indo-Pacific, and the Southern Ocean (ACC).
- Eddy kinetic energy (EKE) is not a reliable proxy for mesoscale temperature fluxes in energetic high-EKE regions, though it is in some regions of low or moderate EKE.

Figures that do not relate directly to these key points have been removed. This includes Figures 3, 6, 13, and 15 in the previous manuscript. Figure 2 (spectra of meridional velocity at various latitudes) has been moved to Appendix A, in order to allow the manuscript to proceed more quickly to key results while retaining this information to justify the choice of threshold wavelength used. Figure 14 (Indo-Pacific time series) was streamlined to focus only on the situation at 4°N, where the mesoscale heat transport variability is aligned with ENSO. Some figures and their discussion in the text have also been reordered in order to highlight and discuss the key points above.

The concluding discussion in Section 6 has also been revised to clarify the ways in which these key results further the study of mesoscale fluxes and their representation in models:

"By implementing this diagnostic to separate scales in mesoscale-permitting and mesoscale-resolving simulations, it is possible to better understand how stationary and transient mesoscale phenomena induce fluxes across large-scale gradients. Future studies can focus on relating large-scale gradients (in one, two, and three dimensions) to mesoscale diffusivities, and therefore on improving parameterizations of mesoscale impacts from stationary recirculations as well as propagating eddies." (line numbers 563-567)

Rewrite section 2.2 to start with the most common method to define eddies as given in equation 10. From there introduce the spatial filter method. Then explain all the different types of eddies that are used in this paper (including eq 7,8,9) and more importantly, give them clear and distinct names. At this moment mesoscale is used in various different ways and definitions and it becomes very unclear.

Section 2.2 has been reorganized as suggested, starting with the definition of the temporal decomposition, then the discussion of the spatial decomposition that we use. Finally, the decomposition of the mesoscale temperature flux into stationary and time-varying components (eq. 7-9 in the earlier version) is discussed.

I keep finding the MTF confusing as in such contexts M often stand for Meridional and not Mesoscale. Meridional is also part of this study. It distracted me a few times.

Good point: since both meridional and mesoscale are used a lot in the paper it is tempting to use M to abbreviate both of them, but this may be confusing. In the revised version, mesoscale temperature flux is still abbreviated as MTF, but meridional heat transport is abbreviated as "meridional HT".

Specific (but still some major) comments

Eq 5&6: explain the difference.

The equations (now 6 and 7) are complementary to each other; equation (6) is a low-pass filter and equation (7) is a high-pass filter. The sum of the two resulting components is the original unfiltered data. This is now explained in the text:

"The sets of Fourier coefficients V_L and V_M resulting from the two filters sum to the original coefficients V''." (line numbers 159-160).

Figure 4 – the different axis-range for the y-axis is not helpful.

In this figure (Figure 7 in the revised manuscript) the y-axis range has been standardized for the three panels, ranging from -2 to +2.2 PW. The subsequent Figures 8-10 dealing with meridional heat transport have also had the y-axis ranges standardized.

L160 – This discussion is about the variability, which is as large as the transport itself. Please explain what that implies and what it says about using this filtering method.

This discussion relates to a figure (Figure 8 in the revised manuscript) showing the interannual/decadal variability of the spatial components of meridional HT at various latitudes. It is not clear whether the reviewer's comment refers to the variability of the mesoscale contribution or the total variability. In either case, that the temporal variability is as large as the transport itself does not say much about the filtering method, since the spatial decomposition is applied in the zonal dimension and not the time dimension. The large interannual variability in the Walker circulation (e.g., ENSO) likely accounts for much of the large total (mostly overturning) variability at low latitudes, while the ID variability of the mesoscale contribution are likely to be the result of interannual changes in eddy and TIW activity, as stated in the text:

"The largest peaks in ID mesoscale meridional HT variability occur at 43°-40°S, 3°-4°N, and 32°N, with the latter two peaks explained mostly by contributions from the Indo-Pacific (Figure 8c). These latitudes correspond to the high EKE regions associated with the Agulhas Return Current, Pacific TIWs, and Kuroshio respectively, implying that the very active mesoscale dynamics and interannual variability in these areas result in large MTF variability on ID timescales as well." (line numbers 389-395)

L179 – Why these numbers for s and k0, and what do they mean? Are results sensitive to these choices?

We have provided clarifications in the revised text after eq. (6) and (7):

"The forms of the low-pass (6) and high-pass (7) filters are symmetric and contain a steepness factor s that affects the rate of signal roll-off near the cutoff wavenumber; a higher value approaches a boxcar filter with associated "ringing" effects, while a lower value avoids ringing but at the expense of cutoff precision. In filtering for large-scale and mesoscale v and T, a value of s = 5 was selected as an optimal balance of roll-off between the physical coordinate and spectral wavenumber domains." (line numbers 154-159)

And in the discussion of the smoothing filter: "…for this smoothing the threshold wavenumber k_0 is chosen to be half the threshold wavenumber (twice the threshold wavelength) for the large-scale/mesoscale separation at that latitude. This larger value of the threshold wavelength is used to remove more of the rotational fluxes that occur at the mesoscale. In the smoothing filter the steepness factor s is also set to 2 (rather than 5), because the smoothing effect is more important here than the precision of the wavelength cutoff." (line numbers 264-268)

L182-184 – Are you sure about this statement? If so, what is your evidence for this?

Coarse-resolution models only simulate the large scale explicitly, not the mesoscale. So in order to improve their accuracy, we need to parameterize the effects of the non-resolved dynamics (mesoscale and smaller scales) on the resolved-scale structure.

L188 – why are they opposed?

The figure being discussed (large-scale temperature fluxes) has been removed from the revised manuscript in order to focus on the mesoscale temperature fluxes.

L198 – Why important in eastern boundaries?

Again, the figure being discussed has been removed from the revised manuscript.

L206 – Are you sure? Could it not also be the other way around? How does this work, what is the mechanism?

This discussion was a comparison of maps of the large-scale temperature flux vs. the mesoscale temperature flux. Since the maps of the large-scale temperature flux are no longer in the revised paper, this discussion has been removed from the text.

L220 – Over how long do you average? I guess you come back to that later. Perhaps refer here.

As mentioned towards the end of the revised Section 2.2: "In this analysis the time averages are applied over the full 32 years of model output used." (line number 234)

Fig 8 – Provide Title this is only Indo-Pacific. Perhaps indicate major continents or so in axis. Are all these figures needed to tell you story?

This figure (Figure 5 in the revised manuscript) has been pared down to focus on the cumulative zonally-integrated temperature fluxes that sum up to the heat transport in the basin. The figure caption mentions that this is for the Indo-Pacific basin, and some major features (e.g., currents, marginal seas) have been annotated on the figure panels.

L253 – I find "time-varying" a bad choice, as they are both time varying.

This flux resulting from the temporal decomposition only is now referred to as "all time-varying" to distinguish it from the "high frequency" flux, as well as from the mesoscale time-varying flux MTF_{vary}.

Figure 10 – "mesoscale components" by now I'm so confused as to which components we are dealing with. Temporal, spatial, variability, mean, etc. etc.

Section 2.2 has been reorganized as suggested by the reviewer, with clearer terminology introduced in that section. Hopefully this resolves the confusion.

L261 – Very loose statement about Kelvin and Rossby waves. Explain the physics and provide references.

Text has been added explaining that long Kelvin and Rossby waves "are time-varying with frequencies ranging from intraseasonal to interannual, but with spatial wavelengths spanning an entire ocean basin (e.g., Boulanger and Menkes 1999; McPhaden and Yu 1999)". (line numbers 412-414)

L262 – This is also a loose stamen. It may be clear from the results, but it is not explained why, and why one is preferred over the other, and for what situation such preference might be true or false. In other words, it is not yet shows that spatial filter is better than temporal, nor explained exactly way the reason is behind al the differences. Only shown it is different.

Is the reviewer referring to the statement that was in lines 261-262 itself, or asking a more general question about why the spatial decomposition would be considered preferable to the temporal decomposition? The statement that was in lines 261-262 has been refined slightly to specify that long planetary waves (especially in the tropics) have very different characteristics from mesoscale eddies.

As for the latter point, the motivation for a spatial decomposition is stated in Section 1:

"…these temporal decomposition methods may conflate the contributions of large-scale and mesoscale circulations, as gyres and planetary waves have temporal covariances between v and T. Moreover, the effects of stationary mesoscale features (e.g., recirculation gyres) are not included in the temporal eddy meridional HT contributions. Since spatial resolution prevents climate models from explicitly simulating mesoscale ocean dynamics, accurate representations of the ocean depend on parameterizations (e.g., Gent and McWilliams, 1990; Eden and Greatbatch, 2008; Marshall et al., 2012) that must take into account stationary as well as transient mesoscale fluxes." (line numbers ?)

Figure 12 – Which mesoscale flow is shown and how much it is smoothed?

The zonal smoothing filter is the same one used in earlier maps of the time-mean mesoscale temperature flux; this has been clarified in the figure caption (Figure 11 in the revised manuscript): "zonally smoothed using the same smoothing filter as in Figures 2-4".

L280 - I'm not convinced about the need of the paragraph starting at L280 and the associated Figure 13. Isn't the variation only 10-20% of the total at the peaks only. Is this important? Perhaps it is, but it is not clear here.

This discussion has been removed from the manuscript along with the figure that it is associated with.

Figure 14 – Clarify in title that OT is removed in lower panels.

This has been clarified in the figure caption (Figure 13 in the revised manuscript): "(b) Same as (a), but with the overturning component of meridional HT variability removed".

L317 – How substantial is the variability? To what are we comparing this and how much is that percentagewise or what is the correlation factor?

The R_{MTF} (as mapped in Figure 12 in the revised manuscript) reflects both the amplitude of the variability and its correlation with the total heat transport at that latitude, as indicated in eq. (11). In addition to the discussion of RMTF, the text also mentions two regions where the correlation itself is particularly high: "…there are two regions where the local MTF and basin-integrated total meridional HT are also highly positively correlated with C(loc,tot) > 0.5; the Pacific TIW region north of the equator, and the lee of the Kerguelen Plateau. The implication is that mesoscale processes in these two regions have a direct impact on the meridional HT at their respective latitudes, without much compensation from the overturning or large-scale contributions." (line numbers 504-507)

Section 5.2

I'm not convinced that this section is correct. It is based on a very big IF in L340. Check papers on changing Rossby radius of deformation and temperature on short and large scales. They both vary strongly with y. I'm not convinced this part is very meaningful, without more evidence that this could work.

The reviewer is correct about the "very big IF", as mixing length and meridional temperature gradient may indeed vary greatly. Yet Figure 14 in the revised manuscript shows that, in spite of these caveats, EKE explains 50% or more of the interannual/decadal heat transport variance in some moderate eddying regions (subtropical eddy bands, southern excursions of the ACC). Hence the results in Section 5.2 indicate where the assumption that mixing length and temperature gradients are fairly consistent in time is valid, and where the assumption is not valid (energetic western boundary currents).

A sentence has been added to the text to emphasize this: "Figure 14 does however provide an indication of areas of the ocean where surface EKE might be a decent proxy for the MTF, in contrast with the more energetic regions where mixing length and gradient variability can interfere with the EKE-MTF relationship." (line numbers 540-542)