RESPONSE TO REVIEWER 2

The authors would like to thank Reviewer 2 for the comments and suggestions that have helped to improve the overall quality of the paper. In our answers, we have referred to the line numbers as in the revised manuscript.

General comments:

Much of it is either unsurprising (ORAs are closer to observations than FRMs) or under-explored (what needs to be done in the ORAs to better represent the boundary currents).

We provide a comprehensive analysis showing where the differences between these products come from (i.e. the western boundary currents). The results make sense in relation to the data that have been assimilated. Although the results are perhaps not surprising, they had not be shown with this level of clarity in previous studies. In addition, the consistent impact of assimilation on the interior circulations in multiple products has not been previously shown. The comparison with the FRMs clearly shows that ORAs have additional circulation differences relative to the FRMs in the basin interior.

We agree with the reviewer that there is more work to be done in order to fully understand the reasons for the ORA differences at the western boundary. However, this requires a full study in itself. A different approach making sensitivity experiments in the South Atlantic with a single ORA, changing for example data assimilation configurations and assimilated observations near the western boundary, is needed to make further progress. This is currently being done and will hopefully come out as the future paper.

Just from reading the manuscript, it is not clear how this study is different from the Majumder et al. (2016) paper. This needs to be better highlighted in the introduction section.

Majumder et al. (2016) have used ORAs and observational estimates to study the meridional transports in the Southern Atlantic. However, their study does not go further than showing that the transport magnitudes between ORAs as well as between ORAs and observations show large discrepancies. Motivated by Majudmer et al. (2016), we show here more details of the ORA transport differences and interpret the impact and limitations of the DA schemes in improving the South Atlantic circulation, also through the inclusion of FRMs with distinct spatial resolutions. For example, we clearly demonstrate the dominant role of inter-product spread in velocity, and the limited contribution of inter-model spread in temperature. This is now better highlighted in the text **-** L71 and L83 **-** as below:

L70: “Their result reveals the need for further assessment of the skills and uncertainties of the ORAs in the South Atlantic, such as comparing them with Free-Running Models (FRMs) and evaluating their SAMOC contributions across the eastern, interior, and western boundary regions shown in Fig. 1.”

L82: "Going further than Majumder et al. (2016), we also narrow down these transport differences in an attempt to understand the potential impact (and limitations) of the DA schemes in improving the ORA states in the South Atlantic."

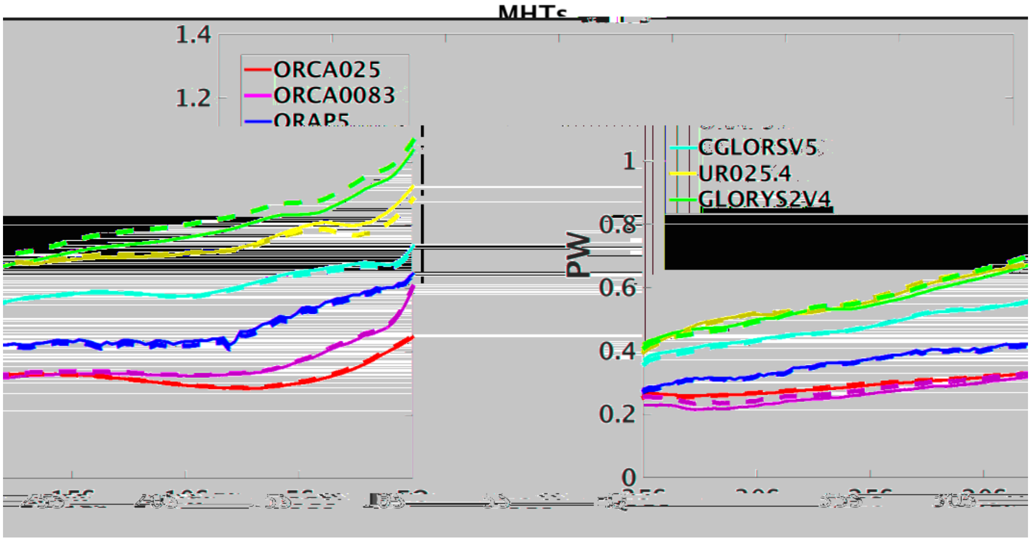
It would be good to discuss the findings, especially that the models so poorly represent the boundary currents, in relation to OSSEs; i.e. which measurements are needed (and where are they needed) to improve the ORAs?

We have now included more discussion about model sensitivity studies and possible OSSEs. The text has been changed in L400 as below:

"Observation system simulation experiments (OSSEs) with AMOC trans-basin arrays have shown that the meridional flow strength can be sensitive to the number of hydrographic profiles near the boundaries in both North (e.g. Hirschi et al., 2003; Baehr et al., 2004) and South Atlantic (e.g. Perez et al., 2011). The combined assimilation of open ocean hydrographic observations and the continuous RAPID array western boundary measurements have also been shown to locally improve the AMOC strength at 26.5˚N (Stepanov et al., 2012). This emphasises the role that more systematic observations located at the eastern and western boundaries at several latitudes may play in monitoring the AMOC (Marotzke et al., 1999). In the future, the SAMOC observing system (Ansorge et al., 2014; Hummels et al., 2015), which will provide time series of NBC measurements at the western boundary at 11˚S, could be assimilated into the ORAs to constrain the regions of largest spread in the tropical South Atlantic."

I was surprised to read that the ORAs are much less constrained before the 2000s, yet the authors chose to use the 1997-2010 period for their analysis. Why not start after the Argo era begins then? And why not include later years; 2010 is seven years ago! That would make much more sense to me.

The UR025.4 reanalysis ends in 2010, so we took 2010 as the final year of the chosen period. In addition, the plot below show that meridional transport differences for the periods 1997-2010 (solid lines) and 2000-2010 (dashed lines) are quite small.



Including some years before the Argo period also allows us to show that the ORA interior circulation is better constrained after the Argo begins (see Fig. 13), whereas the overturning spread between them remains steady over time. If we had only used Argo period results we would have had few averaging years. We have modified the paragraph in L153 as below:

"In order to avoid any dynamical spin-up in the early years of the simulation for products starting in the late eighties or early nineties (e.g. UR025.4 and GLORYS2V4), and because UR025.4 ends in 2010, a common time period from 1997 to 2010 is chosen. Despite the fact that subsurface ocean observations are scarcer before the 2000s (i.e. prior to the full deployment of Argo floats), the total meridional transports for the periods 1997-2010 and 2000-2010 do not differ significantly. "

The authors need to state much more clearly which of the observational data they use to assess the skill of the models have gone into the ORAs themselves, i.e. is the XBT-AX18 line, or WOA13, used in the data assimilation process of the ORAs? In other words, are these independent validations?

We agree with the reviewer about this comment and this is now clearly mentioned in a new paragraph starting in L174:

"Of the observational estimates above, the XBT-AX18 line is not independent as it is included in the EN3 and EN4 datasets which are assimilated by the ORAs (see Tab.1). Although WOA13 is not directly assimilated by the ORAs, it uses the same observational information as EN3 and EN4, and so it also cannot be treated as completely independent."

There could be more discussion of the features seen in the Cape Basin. Where do these come from? How are they related to MOC and the Brazil Current? Are they noise or signal?

From checking spatial animations of SLA in the Cape basin as well as maps of Eddy Kinetic Energy for the products, it is clear that there are differences in reproducing the Agulhas rings between the FRMs and ORAs, as well as between the ORAs. This implies that the Cape basin variability impacts the ENS-ORA and ENS-ALL transport spreads (Fig. 9). This may be related to the way SLA is assimilated into the ORAs, see added discussion:

L288: "South of 25˚S the p-OTT contributions to the total MHT are more distributed, with a noticeable contribution from the Agulhas leakage caused by the different intensity and positioning of the Agulhas rings between the products as they travel westward across the Cape basin."

L346: "The different levels of variability in the Agulhas leakage between ORCA025 and ORAs may be attributed to the impacts of the SLA assimilation (Backeberg et al., 2014). However, even between ORAs these Agulhas patterns differ, e.g. the weaker contributions in ORAP5 may be due to smoothing from the super-observation method applied to the altimeter data (Mogensen et al., 2012), as also noted by Masina et al. (2015)."

For a study that focusses so much on boundary currents, it is surprising that there is no mention of how the boundary conditions have been implemented in the models? Are these all the same? Partial slip? What parameters have been used? Is there any relation between the way the boundary conditions are implemented and the skill of the ORAs/FRMs?

We agree with this comment and information about the lateral boundary conditions for all the products are now given in the text:

L121: “Both ORCA0083 and ORCA025 employ a free-slip (no-stress) configuration for the lateral momentum boundary conditions.”

L148: “On lateral boundaries, UR025.4 and ORAP5 adopt a free-slip configuration whereas CGLORSV5 and GLORYS2V4 employ a partial-slip condition. In the latter, the constant of proportionality (α) between the tangential stress and the tangential velocity is defined as 0.5 for both products.”

Some further discussion has been added in L394 with respect to the lateral boundary conditions:

“It is noteworthy that the lateral boundary conditions in the ORAs and FRMs vary between free-slip (α=0) and partial-slip (α=0.5). However, there is no clear correspondence between the choice of lateral boundary conditions and the strength of the western boundary transports, with free-slip products (e.g. UR025.4) having similar transports to partial-slip products (e.g. GLORYS2V4).”

Minor issues:

- line 14: Start the abstract with a sentence about why the South Atlantic is relevant? The introduction section starts with a few good paragraphs, and these might be summarised at the beginning of the abstract?

The abstract now starts with the following sentence:

“The Meridional Heat Transport (MHT) of the South Atlantic plays a key role in the global heat budget: it is the only equatorward basin-scale ocean heat transport and it sets the northward direction of the global cross-equatorial transport.”

We also thought we could better improve the abstract as a whole, making it easier to understand for a broader number of readers. So you will find a few changes and text rearrangements there.

- line 49: Add ‘possibly’ before ‘leading’ ?

Done.

- line 65: Explicitly state that most ORAs lack dynamical consistency?

We have removed the word "dynamically" from the ORA definition in L64. We think this is enough to meet the reviewer's requirement.

- line 113: I thought most NEMO models were z*?

For all the NEMO-based products used here, the code versions with fixed z-levels and partial cell topography are used. There are additional challenges in implementing DA (especially altimetry) with z*.

- line 119: What do 'W', 'SAT' and 'SAH' stand for?

These acronyms are defined in L106 and L107:

"The ERA-Interim reanalysis provides Winds (W) at 10 m, Surface Air Temperature (SAT) and Surface Air Humidity (SAH) at 2 m, daily Radiative Fluxes (RF) and Precipitation (P) fields…"

- line 263: Why is this interpretation made? What is the motivation for this?

There is a clear line of spread in the p-$\bar{v}T$ right against the western boundary (Fig. 7f). This might indicate that local response to small temperature changes on the western boundary slope may largely determine the meridional transport variability in ocean models through changes in the western boundary current velocities (p-$v\bar{T}$; Fig. 7d), as already stated by Bingham and Hughes (2009) for the North Atlantic (as discussed in the paper in L390).

- Figure 2: How is 'spread' defined here? Highest-lowest? Standard deviation?

The spread is defined as the standard deviation between the products. To make this point clear, we have added this information to the caption of Fig. 2:

"(a) The AMOC strength ψ_{max} (Sv) averaged over 1997-2010 as a function of latitude, and (b) its spread (Sv) defined as the standard deviation of the ENS-ALL and ENS-ORA."

- Figure 3: Why does the size of the horizontal lines on the whiskers vary? What does this mean?

The horizontal lines on the whiskers do not mean anything relevant. However, the size of these horizontal lines in Fig. 3c were adjusted to be the same size.

- Figure 5: How is the standard error defined here?

It is defined as the standard deviation divided by the square root of the length of the monthly time series. The caption of Fig. 5 is now changed to include the following information:

“The black bars represent the standard errors where the size of the sample is defined as the length of the monthly time series.”

- Figure 8: What is the relevance at this line of 0.7 correlation? Is this significance level? How calculated?

We have removed the 0.7 line from the figure to avoid any kind of misunderstanding. In Fig. 8, we compute Pearson correlations calculated with 95% significance level on a monthly time scale. This information is now added to the caption of Fig. 8:

“The monthly Pearson correlation between the SAMOC strength and the MHT as a function of latitude for 1997-2010, calculated with significance level of 95%. “

- Figure 11: What do the dashed circles represent above the top two bar charts?

It was an attempt to say that in the tropical South Atlantic the northward upper western boundary has part of its transports to compensate for the southward deep western boundary flow (solid circles), and other part to compensate for the southward upper interior flow (dashed circles). However, we have removed the dashed circles in order to avoid any kind of misunderstanding.

- line 29: The word ‘right’ is confusing here, as it might also be interpreted as the right (as opposed to left) side of the plots?

We have removed the entire sentence of L26 since we have changed the abstract as a whole.

- line 81: ‘Focusing on’ and line 412: ‘near the boundaries’

Done.

- line 267: Is this correlation R or R^2?

R.