Response to Anonymous Referee #2

*Original reviewer's comments are inserted in black, **A**uthor **R**eplies are added in blue, and **C**hanges made to the **M**anuscript are finally listed in grey, whereby page and line numbers refer to the fully revised version of the manuscript.*

In this study, the authors outline very clearly a compelling Lagrangian analysis of the different sources of water to the NBC, and how their respective property changes are brought about though the South Atlantic. I recommend this study for publication in Ocean Science, though as described below, I think that improvements could be made to the manuscript with the inclusion of more discussion on the big picture implications of their findings.
AR: *Many thanks for your positive feedback. We appreciate your suggestions and believe that including them strengthened the general framing and discussion of our work.*

Main comments:

- Two of the big picture implications that I think would be particularly valuable to discuss are:
  1) the potential implications of the different routes for Stommel's advective salt feedback (or Fov; e.g. Drijfhout et al., 2011, Climate Dynamics). Model's often get this wrong, and the implications for this may be large (e.g. Liu et al., 2017, Science Advances). I wonder, therefore whether, the models could be getting the Fov sign wrong because they are underestimating the fresher DP contribution.
     AR: *We now added one paragraph to the introduction and one paragraph to the conclusions to relate our work to the theories to the salt advection feedback (**please refer to the response to referee#1 for details**). However, we do not think that we can justify a general statement on the relation between the AC/DP partitioning and deficiencies with respect to model's representations of AMOC stability, since (i) most state-of-the art OGCMs have a negative Fov, and also many CMIP5 climate models seem to get the sign right (even though Liu et al., (2017), stated otherwise, cf. Gent (2018)), and (ii) it is currently debated whether Fov is a reliable stability criterion at all (cf. Gent (2018), Cheng (2018)).*
  2) A comparison of the pathways to those produced in more idealized and theoretical studies, such as the recent papers by Spencer Jones and Paola Cessi. This would be useful since those simpler models are the ones we often rely on for clearer diagnoses of the mechanisms at play.
     AR: *We agree that the theoretical considerations by Cessi and Jones (2017) need to be included in the introduction for a thorough and complete review of the existing literature and added respective paragraphs to the manuscript (**please refer to the response to referee#1 for details**).*

- It would also be useful to have more description of the study by Rodrigues et al., (2010), the observations of which are used to validate this work. The authors outline in the introduction that the relative contributions from each source are strongly debated between many studies. Therefore, in order for the reader to accept this study as the most accurate among them, it will require that we agree the comparison to observations is better. However, I only found a three-line description of that observational study (P9L31-25).
  AR: *We understand that the details of the study by Rodrigues et al (2010) may be of interest to the reader, but we are of the opinion that our comparison that focuses on the AC and DP contributions is adequate for the purpose of the manuscript. On the one hand, we give several complementary reasons, why a solution with a non-negligible DP contribution may be the most realistic one (please also see the changes made to the introduction). In fact, to our mind, a good agreement with Rodrigues et al. (2010) alone is not sufficient to accept our study as the most accurate, given the limited spatial and temporal resolution of observational data. On the other hand, a more detailed*

comparison between Rodrigues et al. 2010 and our results in terms of other derived quantities would require complex analysis which are beyond the scope and framing of this study and may be unnecessary, given the detailed model validation performed by us (see method section) and Schwarzkopf et al. (2019).

- Finally, it would help if the authors could provide more discussion of the perceived weaknesses of the experimental setup. While many of the earlier studies did not use high resolution models, this model has quite a short spin-up time and appears to only use interannual forcing fields. E.g. might higher resolution winds allow more water to cross from the AC?
  AR: *It seems our formulations regarding the temporal resolution of the forcing fields has been misleading. The term interannual forcing has been used to contrast the forcing from the employed hindcast spanning the period 1958-2009 to that from a climatological run with no interannual forcing variability. More precisely, the employed atmospheric forcing for the period 1958-2009 from the CORE data set includes 6-hourly atmospheric state variables at 10m height (temperature, humidity and horizontal wind components), daily long and short-wave radiation (prior to 1984 based on a climatological mean annual cycle), and monthly precipitation (prior to 1979 based on a climatological mean annual cycle) as described in the listed reference.*
  Even though the employed spin-up is clearly too short for the deep ocean to reach a stable state, it is rather long for a realistic ocean model configuration at such high resolution. It is fair to assume that at least the upper ocean, which is most relevant for this study, has reached an adequately adjusted state.
  CM: *We reformulated all parts referring to interannually varying atmospheric forcing fields.*
  p.5, ll.24-25: "(…) and subsequently run with forcing from the atmospheric fields of the Coordinated Ocean-Ice Reference Experiments data set version 2 (CORE; Large and Yeager, 2009; Griffies et al., 2009) for the period 1958–2009."
  p.11, ll.29-31: "(…), we used 5-day mean velocity fields of a hindcast experiment, whereas Speich et al. (2001) used monthly means from a climatological experiment. The increase in resolution and allowance for interannual variability most likely lead to (…)"

Other comments:

- P12L21: This is an interesting argument, and the authors have convincingly demonstrated that the two water masses are made more distinct by to their salinity characteristics. However, what I think is probably more important in terms of how they should be labelled is the relative impacts the T and S differences have on density. While the water masses might be more easily delineated by salinity, it does not mean that those salinity differences have as big an impact on density as the temperature differences (e.g. if the salinity range is smaller). Given the nonlinearity of the equation of state, it may not be trivial to fully estimate those impacts, but a rule-of-thumb estimation would still be useful. If it turns out that the temperature differences have a larger impact on density, then the warm- and cold-route terminology would likely remain preferable.
  AR: *We agree that this a very interesting point worth of further dedicated analysis, which are, however, beyond the scope of the current study. Given that the positive temperature anomalies introduced by the inflow of upper waters South of Africa have been suggested to be dampened way faster than the positive salinity anomalies (e.g., Gordon, 1992) we anticipate a larger impact of the salinity difference. The importance of the salinity difference is further supported by the newly included discussion of the salt advection feedback (see comment above). However, overall, the relative importance of the temperature and salinity differences may be dependent on the details of the research question.*
  Gordon, A. L., Weiss, R. F., Smethie, W. M., & Warner, M. J. (1992). Thermocline and intermediate water communication between the south Atlantic and Indian oceans. Journal of Geophysical Research, 97(C5), 7223. https://doi.org/10.1029/92JC00485

CM: *The respective paragraph has been adjusted:*
p.13, ll.19-22: "Hence, we may consider fresh and salty routes as an alternative and more precise terminology, which also accounts for the relative role of the two sources with respect to the salt advection feedback. Yet, dependent on the specific research question, the mean temperature difference between the two may still be of (larger) importance. Therefore, we would recommend referring directly to the geographic origin to avoid ambiguities."

- P13L1-5: I am unclear on these density definitions. Wouldn't these density definitions of surface, central and intermediate waters depend on latitude, and therefore be different for the two sources of water? Some additional description may help.
AR: *The density criteria for separating surface, central, and intermediate waters are – as we do acknowledge in lines 1-6 at page 14 – not uniquely defined. However, the chosen (or very similar) values have been meaningful applied for broader scale analysis in the subtropical and tropical South Atlantic (see references within the manuscript itself). In particular, Antarctic Intermediate Water (AAIW) can be detected in the given density range within the whole South Atlantic basin north of the Subantarctic Front, encompassing the upper limb pathways of the AC as well as DP contribution (cf. section 3.3. of main manuscript and, e.g., Table 2 of Heywood and King, 2002). Everything above the AAIW layer constitutes the upper water layer, which in the subtropical and tropical Atlantic can be further divided in central waters remotely formed by subduction and surface waters under direct influence of local air-sea fluxes. The density level that separates these two layers is indeed latitude dependent, given that the central water range broadens towards the tropics where it includes more and more varieties of subducted surface waters from the subtropics. Hence, we agree that the discussion of surface and central water transformation in the current form may be confusing.*
Heywood, K. J., & King, B. A. (2002). Water masses and baroclinic transports in the South Atlantic and Southern oceans. Journal of Marine Research, 60(5), 639–676. https://doi.org/10.1357/002224002762688687
CM: *We decided to no longer differentiate central and surface waters but instead combine the two into one category termed 'upper waters', which can be clearly separated from intermediate and deep waters by the chosen density criteria within the whole area of interest. The table, as well as Figures 7,8,10 and 11 and corresponding figure captions have been adjusted accordingly and the respective parts in the results section have been re-written.*

- P14L14-17: This is an interesting result and a very nice analysis.
AR: *Thanks*

- P15L32: I don't understand this first sentence.
AR: *We agree that this first sentence may be confusing if one assumes that the terms 'Agulhas leakage' and 'AC contribution to the AMOC' can be used interchangeable. However, in our study, the AC contribution to the AMOC always refers to the contribution of waters with Agulhas origin that make it into the tropics and become part of the NBC. Hence, the question rather is whether we can detect the increase in Agulhas Leakage as an increased contribution of Agulhas waters to the AMOC's upper limb further downstream in the tropics.*
CM: *We now specified that we are referring to the AMOC contribution in the tropics in this particular sentence as well as elsewhere in the manuscript:*
p.16, ll.30-31: "The increase in the AC contribution to the AMOC's upper limb in the tropics is, however, not directly proportional to the increase in Agulhas Leakage, but weaker (1.9 Sv compared to 6.2 Sv, respectively)."

- P16L14: The wording of this sentence could do with some revision.
CM: *We split the sentence into two:*
p.17, ll.11-14: "To do so, we performed Lagrangian analyzes using 5-day mean output from a hindcast experiment (1958–2009) with the high-resolution (1/20°) ocean general circulation model INALT20. We employed the Lagrangian tool ARIANE to calculate O(10^6) advective volume

transport trajectories as well as along-track thermohaline property changes between the two source regions and the North Brazil Current (NBC), which channels the upper limb flow in the tropics."

- P16L27: Here and elsewhere, "evoked" should be something more like 'induced'
  CM: *We exchanged 'evoked' by induced' within the whole manuscript*