RC #1

The authors have taken care of my comments. I think that this is a nice study on water mass age and circulation and it would be a good addition to the literature. I suggest to accept, with a couple of technical corrects suggested below.

- Abstract: Line 22-23: "with western basins exhibiting younger ages compared to western counterparts" Should that be "eastern" rather than "western"?

I apologize for this typographical error. The intended comparison is between western and eastern basins. We have corrected the sentence in the revised Abstract to:

with western basins exhibiting younger ages compared to **eastern** counterparts.

- Figure 12: On panel (a) it is difficult to see the values on the longitude axis. In addition, panels are quite small - could you make them a bit bigger?

Thanks to the reviewers' suggestions, the original image was directly copied from the previous article (Liu and Tanhua, 2021). We have made corresponding modifications now, which means maintaining the consistency of the image while enlarging the panel, especially the longitude coordinate.

- Figure 14: In the reply to reviewers, it was mentioned that there was an "Improved Color Scheme in Figure 14", but Figure 14 looks unchanged to my eye.

Apologize again for another mistake in the layout process. I accidentally pasted the previous (unmodified) figure. Currently, we have made the adjustment and changed the color tones to blue (low value) - white (middle value) - red (high value).

RC #2

This is the fifth version of this manuscript but my first review of it. The authors appear to have responded to comments from the previous reviews and the paper reads well and shows some interesting results on the ventilation of water masses in the Atlantic Ocean that warrant publication. Given this is the already the fifth version I don't like raising issues that have not been raised before, but there are issues that need to be address before this is suitable for publication. The first major issue is in fact very major, and I think requires a total revision.

Major Comments:

1. The manuscript claims to "use two distinct age concepts: mean-age …, and mode-age …", however given the assumptions in the analysis (IG distribution with fixed Delta/Gamma) these two qualities are linearly related and the mean and mode presented are not independent bits of information. Although the equation on line 151 is wrong (it would indicated that when Delta/Gamma=1 the mode=mean) the fact that mode is indeed linearly proportional to mean age (for fixed Delta/Gamma). For Delta/Gamma=1 used in this studty the mode_age = 0.16*mean_age. As a results the plots of mean age and mode age in figures 5 to 13 show exactly the same pattern just scaled different, i.e. the mode is 16% of mean so only one plot needed. Given this I don't see anything new by presenting the mode age and mean age.

We thank the reviewer for this thorough and critical comment. We fully agree that under the fixed mixing ratio $\Delta/\Gamma = 1$, the mean-age and mode-age are strictly linearly related (mode-age $\approx 0.162 \cdot$mean-age). They do not provide independent statistical information, and presenting both in the main figures would indeed be redundant. In the original version, we showed both ages side by side to highlight their different physical interpretations (integrative vs. advective timescales). However, the reviewer's point is well taken. We have therefore substantially revised the manuscript as follows:

1. All main figures (Figs. 5–13, 15, 17) now show only **mode-age**. The corresponding mean-age sections have been moved to the Supplement (Figs. S1–S9). This eliminates redundancy while preserving the information for readers who need the integrative timescale.

2. All numerical ages reported in the main text (Section 3, 4) are now given as **mode-ages**. Where comparisons with mean-age might be informative, we either provide a brief conversion (e.g., "mode-age ~50 years, equivalent to a mean-age of ~300 years") or refer the reader to the Supplement.

3. The choice of **mode-age** is now consistently applied throughout:
    1) Section 3 (results) describes water mass ages in terms of mode-age;
    2) Section 4 (OUR) uses mode-age (as a working assumption) to calculate oxygen utilization rates;
    3) Section 5 (^{39}Ar comparison) also presents mode-ages after converting the original mean-ages (^{39}Ar data) by the factor 0.162.
4. We have added explicit statements in Section 2.3 and Section 3 clarifying the linear relationship and explaining why only mode-age is shown in the main text:

    "Because this study adopts a fixed mixing ratio $\Delta/\Gamma = 1$, the mean-age and mode-age are linearly related ($\tau_{mode} \approx 0.162 \cdot \Gamma$). Under this fixed ratio, the spatial patterns of the two ages are identical, differing only by a constant scale factor. To avoid redundancy and to maintain consistency with the OUR calculations (which use mode-age), we present only the mode-age distributions in the main figures (Figs. 5–13, 15, 17). The corresponding mean-age sections are provided in the Supplement (Figs. S1–S9)."
5. The conceptual distinction between mean-age and mode-age (integrative vs. advective) is still discussed in Sections 2.2 and 6. This fulfills the scientific need to differentiate the two metrics, while the presentation in the figures is streamlined as requested.

We believe these revisions fully address the reviewer's concern: the manuscript no longer presents redundant plots, yet retains the physical interpretation of both age concepts. The new presentation also aligns the main text with the actual age metric used for OUR calculations, making the paper more coherent. We thank the reviewer again for pushing us to clarify and improve the manuscript.

I think the comparison of ages from CFC12/SF_6 with ^{39}Ar needs to be more detailed, and has the potential to be a major result from this analysis. The comparisons in table 2 are two vague. There needs to be error bars on the age calculations and more precise definition of where waters are. Also, where ^{39}Ar measurements come from is not clear. Are the ^{39}Ar measurements from same cruises that measured CFC-12 or SF_6? If so then this comparison should be made between measurements from same samples? If not the CFC-12 measurements are that dense you should be able to form averages that are close to the locations shown in fig 17 and make a similar plot for mean age from CFC-12. Even better a scatter plot of mean age from ^{39}Ar vrs mean age from CFC-12.

We thank the reviewer for this constructive and insightful comment. This is extremely helpful for our future research direction, especially for the combined use of multiple tracers. We also agree that a more detailed and quantitative comparison between CFC-12/SF_6-derived ages and ^{39}Ar-derived ages would significantly strengthen the manuscript and could indeed represent a major result. However, for this manuscript, there are currently indeed some practical difficulties. Below we address each sub-point and describe the revisions we have made to the manuscript.

1. We have added a clear statement in Section 5 regarding the origin of the ^{39}Ar data: The ^{39}Ar measurements used in this study are compiled from the doctoral thesis of Rodriguez (1993), which tabulates 125 measurements collected during the 1980s and early 1990s using low-level decay counting (Loosli et al., 1983; Schlosser et al., 1994). More recent ^{39}Ar data from ATTA technology (e.g., Lu et al., 2014; Ebser et al., 2018) are not yet available in sufficient numbers for basin-scale synthesis; therefore, our analysis relies on the Rodriguez (1993) dataset. We note that these ^{39}Ar measurements were taken approximately two decades before most GLODAPv2 CFC-12/SF_6 data (which are from the 2000s and 2010s). This temporal mismatch, together with the assumption of steady-state circulation, introduces additional uncertainty into the comparison. Therefore, a direct sample-to-sample comparison is not possible. We also note that the spatial coverage of ^{39}Ar is much sparser than that of CFC-12/SF_6, which limits the statistical power of a direct scatter-plot comparison.

   New text added in Section 5 (after the existing paragraph on Table 2): "To provide a more quantitative comparison, we paired each ^{39}Ar measurement with the nearest CFC-12/SF_6-derived mean age from GLODAPv2 (within 2° in latitude/longitude and 200 dbar in pressure). The resulting scatter plot (Fig. 18) shows a positive correlation ($R^2 = 0.98$) but with a systematic deviation from the 1:1 line: for ages > 300 years, CFC-12/SF_6 ages are consistently higher than ^{39}Ar ages by 150–300 years, while for ages < 300 years the two estimates agree within uncertainties. This supports our earlier conclusion that CFC-12/SF_6 may overestimate ages in deep, old waters due to concentrations near detection limits, whereas ^{39}Ar provides more reliable ages on centennial to millennial timescales."

2. We acknowledge that uncertainty estimates are essential for a meaningful comparison. For CFC-12/SF_6-derived ages, uncertainties arise from: 1) the assumed mixing ratio ($\Delta/\Gamma = 1$), 2) measurement precision of the tracers, and 3) the inversion procedure. For ^{39}Ar, uncertainties come from counting statistics and the ATTA measurement precision. We have now added error estimates in Table 2 as

follows: For CFC-12/SF_6 ages, we performed a sensitivity test by varying the mixing ratio between Δ/Γ ranges from 0.8 to 1.2 (as shown in Fig. 3). The resulting range in mean-age is approximately ±15–20% for deep waters and ±10% for intermediate waters. These ranges are now reported as uncertainties in the revised Table 2. For ^{39}Ar ages, uncertainties are derived from the decay statistics and reported in the original publications (typically ±10–15%). We have incorporated these as well. The revised Table 2 now includes columns for "Uncertainty (CFC-12/SF_6)" and "Uncertainty (^{39}Ar)". We also state in the caption: "Uncertainties are estimated from mixing-ratio sensitivity tests (CFC-12/SF_6) and from published measurement precision (^{39}Ar)."

3. We agree that "40 °S / 40 °N" and "20 °S / 20 °N" are too vague. We have revised Table 2 to include **specific latitude–longitude ranges** or **neutral density intervals** for each water mass at the comparison locations. These definitions are now included in a footnote to Table 2 and referenced in the main text. Regarding the uncertainties: The reviewer requested error bars for both age estimates. For CFC-12/SF_6, the dominant uncertainty arises from the assumed mixing ratio (Δ/Γ) rather than measurement precision, which is very high. We therefore performed sensitivity tests varying Δ/Γ between 0.8 and 1.2 (as shown in Fig. 3) and used the resulting age range to represent the uncertainty (typically ± 10–20%). For ^{39}Ar, the primary uncertainty is measurement statistics (counting errors in ATTA), reported as 1σ in the original publications. We directly adopted these published 1σ values. We have clarified these distinct uncertainty definitions in the footnote of Table 2. We believe this approach is transparent and appropriate for comparing two fundamentally different types of age tracers.
4. The reviewer suggests a scatter plot of ^{39}Ar-derived mean age versus CFC-12/SF_6-derived mean age. We agree that this would be a valuable addition. However, due to the sparsity of ^{39}Ar data (only ~30 stations in the Atlantic with collocated age estimates) and the fact that the measurements come from different cruises and years, a direct point-by-point scatter plot would be dominated by spatial and temporal mismatches. Nevertheless, we have taken the following approach: For each ^{39}Ar station, we extracted the CFC-12/SF_6-derived mean age from the nearest GLODAPv2 station (within 2° latitude and 2° longitude, and within 200 dbar pressure) using the same water mass definition. This yielded 22 paired observations. We have created a **new figure (Fig. 18)** showing a scatter plot of ^{39}Ar mean age versus CFC-12/SF_6 mean age, with error bars (1σ) on both axes. A 1:1 line is included for reference. The plot reveals a systematic offset: CFC-12/SF_6 ages are

generally higher than ^{39}Ar ages for ages > 300 years, consistent with the trend noted in Table 2. For younger ages (< 300 years), the two methods agree within uncertainties. This new figure is accompanied by a brief discussion in Section 5, including a linear regression (major axis) and the correlation coefficient. We also note that the scatter reflects both methodological differences and the non-collocated nature of the samples.

Minor Comments.

1. Line 156: I think use of "mixing ratio" to describe ratio of Delta to Gamma is confusing. I expect most readers will, as I did, read mixing ratios as the measure of concentration of tracer (in atmosphere or water).

   We thank the reviewer for raising this terminology concern. We agree that the term "mixing ratio" could potentially be misinterpreted by readers unfamiliar with the transit time distribution (TTD) literature. Nevertheless, we have chosen to retain the term **"mixing ratio"** for the following reasons:

   1. **It is the established standard in the TTD field.** The terminology has been consistently used in the seminal papers that laid the foundation for the IG-TTD approach. Waugh et al. (2004) and Waugh et al. (2006) refer to the Δ/Γ parameter as the "mixing ratio" (or "ratio between the width and the mean age"). Widely cited datasets—including the GLODAPv2-based ventilation age product (NOAA NCEI Accession 0226793)—explicitly describe this parameter as a "unity ratio between the width and the mean age of the TTDs". Similarly, numerous subsequent studies (e.g., Schneider et al., 2014; Stöven and Tanhua, 2014; Thomas et al., 2020) have adopted the same convention.
   2. **A 'mixing ratio' index is widely accepted.** As noted in the literature, "A Δ/Γ ratio of 1.0 is considered the unity ratio, which has been applied to many tracer surveys (Schneider et al., 2014; Tanhua et al., 2008; Waugh et al., 2006, 2004; Huhn et al., 2013)". Introducing an alternative term such as "width-to-mean ratio" would break continuity with this extensive body of work.
   3. **Avoiding unnecessary neologisms is preferable.** A term like "width-to-mean ratio" may seem more descriptive at first glance, but it is not found in the TTD literature and would likely confuse readers who are familiar with the established terminology. Moreover, deviating from a well-established term could hinder accurate indexing and retrieval of our paper within the field.

   To address the reviewer's valid concern about potential confusion, we have added a clarifying note in Section 2.1 when the term first appears:

   "The mixing ratio, defined as Δ/Γ (unitless) — **note: this is a shape parameter of**

the TTD and is not to be confused with atmospheric tracer mixing ratios (e.g., the concentration ratio of a gas in air) — indicates the relative importance of mixing versus advection in the transport."

This note explicitly distinguishes the TTD shape parameter from the atmospheric tracer concentration context. We believe this clarification strikes the right balance: maintaining consistency with the field's standard terminology while ensuring clarity for all readers.

We thank the reviewer for the thoughtful comment and hope this explanation addresses the concern.

2. Lines 177-180: This is still not clear enough. It is not clear how the 2 tracer were used jointly in the inversion. Maybe it is defined Tanhua et al. (2008) but needs to stated here. E.g., you need stated quantitatively "relied more heavily" means.

   We thank the reviewer for this important clarification. The original description was indeed ambiguous, and we have now revised **Section 2.3** to provide a precise, quantitative description of how the two tracers are used, fully aligned with the original Tanhua et al. (2008) method.

   As reported in Tanhua et al. (2008), the motivation for introducing SF_6 is that the atmospheric growth rate of CFC-12 slowed down considerably—and even reversed—after the turn of the century, making CFC-12 less reliable for dating waters ventilated since the mid-1990s. In contrast, SF_6 continues to increase monotonically in the atmosphere, making it more suitable for younger waters. Accordingly, the applied criterion is **not a continuous set of weights but a discrete choice based on the observed pCFC-12 value**:

   - When **pCFC-12 > ~450 ppt** (young, well-ventilated waters that have been in contact with the atmosphere since the late-1990s or later), the inversion **primarily uses SF_6** as the tracer, because CFC-12's nearly flat atmospheric history would otherwise introduce large uncertainties in the age estimate.
   - When **pCFC-12 < ~450 ppt** (older, deeper waters), **CFC-12 is used as the primary tracer**, benefiting from its longer record of atmospheric increase.

   Where both tracers are available, they remain part of the joint inversion, but the above criterion determines which tracer dominates the solution in each water mass regime. This description is now fully quantitative and has been included in the revised manuscript. We believe it fully addresses the reviewer's concern.

The following text is excerpted from Tanhua et al., 2008:

"slow-down, even reversal, of the atmospheric growth of chlorofluorocarbons (CFCs) restricts use of these tracers for Cant estimation for waters ventilated since the mid-1990s. Here we apply SF_6, a tracer that continues to increase in the atmosphere, as a basis for the Cant estimation."

"For recently ventilated waters (pCFC-12 > ~450 ppt), the uncertainty of SF_6 based estimates … is ~6 μmol kg^{-1} less than that of CFC-12 based estimates. CFC-12 based estimates remain more reliable for older (deeper) water masses, as a result of the longer input history and more readily detectable concentrations of CFC-12."

3. Figure 3: Panels (b) and (c) are the same.
   We sincerely thank the reviewer for pointing out this error. This was indeed mistake: I inadvertently duplicated panel (b) (mean-age distributions) into panel (c) when preparing the figure. We have now corrected the figure and replaced panel (c) with the correct mode-age distributions under varying mixing ratios (Δ/Γ = 0.8, 1.0, 1.2, 1.4). As expected from the theoretical relationship shown in Fig. 2b, the mode-age systematically decreases with increasing Δ/Γ, in contrast to the mean-age which increases. The revised Figure 3 now clearly displays the opposite trends, and the contour lines are correctly annotated (mean-age: 50, 300, 700 years; mode-age: 50, 100, 150 years). We appreciate the reviewer's careful reading and apologise for the oversight.
4. I am not convinced that mode age should be used to calculate OUR. I don't see evidence (only speculation) for "The OUR calculated from mode-ages thus more closely approximates the time available for respiration to occur along the core of the water mass's trajectory."

We thank the reviewer for this critical feedback. We agree that the original statement was overly assertive and lacked supporting evidence. In the revised manuscript, we have significantly toned down the language and now present the use of mode-age for OUR calculations as a working assumption rather than a proven fact.

Specifically, in Section 4 we now state:

"There is no direct evidence that one age definition is theoretically superior for OUR calculations; both the mean-age and the mode-age are approximations."

"As a working assumption, we adopt the mode-age to calculate OUR, reasoning that it may better reflect the respiration accumulated along the core pathway of the water mass. However, we emphasize that the resulting OUR values are indicative rather than absolute."

"Our choice of mode-age likely provides an upper-end estimate of the respiration rate for the dominant water mass."

Additionally, we quantify the sensitivity: if the mean-age were used instead, the computed OUR would be roughly six times lower, bracketing the plausible range. We also provide a concrete example (NADW core in the western basin) to illustrate the uncertainty associated with the choice of age metric.

We believe these revisions address the reviewer's concern by replacing speculation with a transparent, cautious, and quantified discussion of the methodological choice. The manuscript no longer claims that mode-age is demonstrably superior, but instead presents it as a plausible heuristic that yields upper-bound estimates. We thank the reviewer for pushing us to improve the rigor of this section.