

Authors' Response to Reviews of

Divergent Sensitivities of Apparent Oxygen Utilization to Circulation Changes in the Deep Ocean Across Earth System Models

~~Earth system models overestimate the sensitivity of apparent oxygen utilisation to age change in the deep ocean~~

*Damien Couespel^a, Xabier Davila^a, Nadine Goris^a, Emil Jeansson^a, Siv K. Lauvset^a, Jerry Tjiputra^a

^aNORCE Research AS, Bjerknnes Centre for Climate Research, Bergen, Norway

*Corresponding author: daco@norceresearch.no

Biogeosciences, *EGUsphere* [preprint], <https://doi.org/10.5194/egusphere-2025-2566>

RC: *Reviewers' Comment*, AR: Authors' Response, □ Manuscript Text

We thank the two anonymous referees for their thoughtful comments and suggestions. To revise the manuscript, we conducted additional analyses to assess the uncertainties in the observation-based estimates of the sensitivity of AOU change to age change ($S_{\Delta\text{age}}^{\Delta\text{AOU}}$). The observational dataset analysis now includes a geographical constraint to ensure data points used for computing the trends are within a specified radius. We ran 625 analyses to test the impact of T-S bin size and the radius on the observation-based $S_{\Delta\text{age}}^{\Delta\text{AOU}}$. Additionally, we assessed uncertainty related to data scarcity and the TTD method for estimating age. These new analyses led us to change the title of the manuscript. The revised manuscript emphasizes the methodological differences between the analyses of observational data and ESMs with a new subsection 2.5 and by distinguishing between TTD-mean-age and ideal-age when necessary. The revised results section is divided into three subsections, and the manuscript now includes separate discussion and conclusion sections. Below are our responses to each reviewer's comments, detailing the new analyses and revisions made to the manuscript.

1. Reviewer #1

RC: *Summary: Couespel et al. investigated the relationship between apparent oxygen utilization (AOU) and water mass age using both observational data and five Earth system models (ESMs). They found that ESMs tend to overestimate the sensitivity of deep-ocean AOU to changes in ocean ventilation (water mass age) when compared to observations. This overestimation suggests that the models may be overpredicting the future strengthening of the Biological Carbon Pump (BCP) based on AOU trends. Overall, the manuscript is well-written and presented in a clear, accessible manner. The research question addressed is of significant importance to the scientific community. I appreciate the considerable effort the authors have invested. However, I have major concerns regarding the manuscript in its current form that need to be addressed. I recommend a major revision.*

1.1. Major comments

RC: *I have two major concerns regarding the methodology employed in this study. First, the authors calculate $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ using ideal age and AOU outputs from the models, while for observational data, they use the mean age of IG-TTD and AOU. I assume that the calculations for observations are based on single-tracer*

constrained IG-TTD with assumptions such as $\Delta/\Gamma = 1$ and 100% saturation? Please clarify this in the Methods section. Nevertheless, a recent study from Guo et al. (2025) suggests that the mean age of IG-TTD derived from single-tracer can be biased towards younger ages due to imperfect assumptions about the shape of the transit time distribution and the short atmospheric history of abiotic transient tracers like CFC-12 and SF6. Additionally, the trend in the mean age of these tracers can be influenced by uncertainties in their mixing ratios, potentially producing spurious trends. I recommend calculating $S_{\Delta_{age}}^{\Delta AOU}$ in the models using both AOU and the simulated CFC-12 (or SF6) to ensure methodological consistency. This approach would also allow testing the robustness of the results. To my knowledge, the model you used includes CFC-12 and SF6 simulations, which could be leveraged for this purpose.

AR: As assumed by the reviewer, for the observational data, the mean age from the TTD calculation is estimated using a single tracer, assuming full (100%) saturation when subducted as well as a balance between advection and mixing, $\Delta/\Gamma = 1$. The revised manuscript states:

We use the age product from Jeansson et al. (2021). In this product, measurements of the chlorofluorocarbon CFC-12 from GLODAPv2 (2016) are used with the transit time distribution (TTD) method to compute age, assuming 100 % saturation and a balance between advection and mixing, i.e. $\Delta/\Gamma = 1$.

In the ESGF database, CFC-12 and SF6 data are only available for IPSL-CM6A-LR and NorESM2-LM. Thus, age trends estimation using the TTD approach cannot be applied to all models. In addition, as noted by the reviewer and demonstrated by multiple studies, the TTD approach introduces uncertainties in water-mass age estimation. Therefore, we consider that it is more reliable to use the ideal-age provided in the models. Nonetheless, we acknowledge the need to quantify the uncertainty related to the TTD approach in the observational estimates of $S_{\Delta_{age}}^{\Delta AOU}$. Following the reviewer's suggestion, we conducted a new analysis (TTD-UNC in the revised manuscript). We (1) computed TTD-mean-age using the TTD method on CFC-12 simulated by NorESM2-LM during the historical period, (2) determined trends in TTD-mean-age, (3) identified a linear relationship between AOU trends and TTD-mean-age trends in Atlantic and Southern dense water masses, and (4) derived $S_{\Delta_{age}}^{\Delta AOU}$ values. $S_{\Delta_{age}}^{\Delta AOU}$ values derived from TTD-mean-age are compared against the reference $S_{\Delta_{age}}^{\Delta AOU}$ values derived from ideal-age to quantify the uncertainty of the TTD method. The new analysis shows that trends in TTD-mean-age are generally stronger than those in ideal-age, resulting in an underestimation of $S_{\Delta_{age}}^{\Delta AOU}$. This, along with uncertainty from data scarcity (addressed in the reply to the second reviewer), compromise the comparability of observation-based $S_{\Delta_{age}}^{\Delta AOU}$ with those derived from ESMs. In the revised manuscript, we have made the following changes:

- Revised the title of the manuscript.
- Included a new subsection (2.5) titled "Evaluation of model-observation comparability", which describes the differences between the methods applied to ESM and the observational dataset, and outlines the new uncertainty analyses.
- Expanded the new subsection 3.3 to describe the results of the observational dataset analysis, accounting for uncertainties.
- Added Table 2 to present the observation-based $S_{\Delta_{age}}^{\Delta AOU}$ and the results of the uncertainty analyses.
- Removed the observation-based $S_{\Delta_{age}}^{\Delta AOU}$ in the revised Fig. 3.
- Expanded the discussion to address the reliability of the observation-based estimate.

- Clarified the distinction between TTD-mean-age derived with the TTD method from observations and the ideal-age tracer computed in the model simulation.

RC: *Second, the authors calculated trends of AOU and age in the TS bins. In this case, the trend of age and AOU contains several pieces of information: (i) the mean state of water parcel ventilation timescales and AOU may differ across various geographical locations within the same temperature-salinity (TS) bin; (ii) the actual temporal change of age and AOU over time. Based on your Figures 2 and S5, where many data points indicate age trends over roughly 10 yr yr⁻¹, I suspect that the first factor—the spatial variability within TS bins—is the primary contributor to the observed age trend. Given this, I wonder if it is appropriate to describe this as a “trend” with units of years per year (yr yr⁻¹). Although it seems it does not affect the meaning of $S_{\Delta\text{age}}^{\Delta\text{AOU}}$, which is somehow like the respiration rate, it would be helpful to clarify this point.*

AR: In the revised manuscript, we included a geographical constraint to compute trends in the observational dataset. This ensures data points used for trends are within a specified radius. The new analysis depends on this radius and the T-S bin size. Following the other reviewer’s suggestion, we ran the analysis 625 times with different random radius and T-S bin size choices to derive a distribution of observation-based $S_{\Delta\text{age}}^{\Delta\text{AOU}}$. The radius was randomly picked 25 times between 500 km and 5000 km. It was not possible to compute enough trends with a radius below 500 km. The observation-based $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ values are now presented as the mean \pm one standard deviation of the $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ distribution. The revised manuscript include the following:

- A description of the geographical constrain:

To ensure that data points are geographically close to each other within each T-S bin, we (i) remove outliers defined as data points geographically further than twice the median distance to the median location and (ii) only keep points that are within a maximum distance from the median location (recomputed without the outliers). [...] Due to the substantial influence of the T-S bin size and the maximum distance, we conduct the analysis of the observational data 625 times with different random choices of these parameters to derive a distribution of the observation-based $S_{\Delta\text{age}}^{\Delta\text{AOU}}$. Temperature/salinity resolutions ranged from 0.026 °C to 0.325 °C and 0.0024 to 0.03, while maximum distance from 500 km to 5000 km.

- A new supplementary Fig. S15 showing the distribution of $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ derived from the observational dataset.

RC: *The title of this study is “Earth system models overestimate the sensitivity of apparent oxygen utilization to age change in the deep ocean,” but I find that there is limited discussion and analysis of this specific point in the main manuscript.*

AR: Both reviewers expressed concerns about potential uncertainties in the observation-based estimation of sensitivity. The new analyses show that comparability between observation and model-based sensitivities is compromised by data scarcity and the use of TTD-mean-age. The title of the revised manuscript is:

Divergent Sensitivities of Apparent Oxygen Utilization to Circulation Changes in the Deep Ocean Across Earth System Models

In the new section 4, we discuss how these divergent sensitivities modulate circulation-driven changes in AOU and the challenges in constraining them, as follows:

Constraining only circulation changes may not be enough to identify the best ESMs at projecting changes in AOU in the interior ocean. The sensitivity of AOU changes to age changes varies substantially between ESMs (Fig. 3), modulating circulation-driven changes in AOU. For instance, the slow down of the Southern and Atlantic MOC in MPI-ESM1.2-LR is stronger compared to MIROC-ES2L (Liu et al., 2023), yet MPI-ESM1.2-LR shows the weakest change in ideal-age-driven AOU trends and MIROC-ES2L the strongest one (Fig. 4). The linear relationship between present and future sensitivity across ESMs is promising, in particular in the Southern dense water-mass which covers the largest portion of the deep ocean. It can, in theory, be used to identify ESMs whose sensitivities are the most consistent with observations in the contemporary period and be used to constrain the sensitivity of the future period. However, at this point in time, we cannot directly constrain the sensitivity following an emergent constraint approach (Bourgeois et al., 2022; Kwiatkowski et al., 2017; Goris et al., 2023) because of the small ESM ensemble available and the uncertainties in the observations based estimates.

The reliability of the $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ estimates from observational data is compromised by the limited number of data points and the usage of the TTD method for estimating age trends. While the scarcity of data and the need to compute trends in T-S space lead to a substantial overestimation when compared to a non-scarce data set, using trends in TTD-mean-age results in an equally strong underestimation when compared to trends computed with ideal-age. While we are confident in our estimate of the first uncertainty, we are more cautious regarding the evaluation of the second uncertainty (TTD method). In the model simulations, in the ocean below 1000 metres, only the Southern Ocean and the North Atlantic are well-ventilated to the degree that the CFC-12 concentrations can provide TTD-mean-age estimates. Thus, we can evaluate trends in TTD-mean-age against trends in ideal-age only in these regions, representing a limited portion of the area covered by the observational dataset. Caution is warranted when using TTD-mean-age to assess changes in ventilation, especially when TTD methods depend on assumptions about the balance between advection and mixing (Δ/Γ). $\Delta/\Gamma = 1$ is a good compromise for the entire ocean but regionally dependent Δ/Γ would lead to more optimal TTD-mean-age when compared to model (He et al., 2018). Approaches employing dual constraint are promising and should be further explored (Guo et al., 2025). While imperfect, the estimates of $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ based on scarce observational data and TTD-mean-age remain the only viable option for comparing $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ derived from ESMs. Given that $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ is to some extent similar to an estimation of the oxygen utilisation rate averaged within the water-masses considered, comparison with prior observational-based OUR estimates is appropriate. In the deep ocean, oxygen utilisation rate estimations typically vary around $0.1 \text{ mmol O}_2 \text{ m}^{-3} \text{ yr}^{-1}$ (Sulpis et al., 2023), substantially higher than the observation-based estimates from our work, yet on the lower end of the $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ range derived from ESMs.

RC: *If I understand correctly, the models show a greater increase in AOU per unit water age increase compared to observations. However, this seems to contrast with findings from previous studies, such as Oschlies et al. (2018), which suggest that models tend to systematically underestimate the observed rates of ocean deoxygenation—mainly driven by increases in AOU—particularly in the deep ocean below 1000 m. It would be helpful to add a paragraph discussing this potential discrepancy in the Discussion section.*

AR: We thank the reviewer for highlighting the relation between our work and deoxygenation studies. The new analyses indicate that comparability of sensitivities is compromised, so we cannot conclude that models overestimate the sensitivity of AOU to circulation changes. However, divergent sensitivities can relate to divergent deoxygenation rates. For a similar circulation slowdown, a model overestimating (underestimating) the sensitivity will have a stronger (weaker) increase in AOU, leading to a stronger (weaker) deoxygenation rate. Constraining these sensitivities would help constrain deoxygenation rates. Given the limited compa-

rability between observations and models, we did not include an extensive discussion on the discrepancy between observed and modelled deoxygenation rates. However, we added the following sentence in the revised conclusion:

Last but not least, constraining simulated changes in oxygen utilisation will be on step toward reconciling simulated and observed rates of current deoxygenation (Ito et al., 2026).

1.2. Minor comments

RC: *Line 56 - 58. The oxygen and age are hardly affected similarly by transport in the real ocean due to the spatial heterogeneity of respiration on isopycnals, as found for idealized isopycnals with prescribed patterns of respiration (Koeve and Kähler, 2016; Guo et al., 2023). “Sufficient oxygen concentration” is not the precondition that AOU and age are linear; instead, the even distribution of respiration rate across spatial is.*

AR: Indeed, thanks for pointing that mistake out. We revised the sentence as follows:

In regions dominated by advection or with an even spatial distribution of the respiration rate, the relation between AOU and age is linear when both are affected similarly by transport (Koeve and Kähler, 2016).

RC: *Line 78 - 87. I suggest including a table that summarizes key information about the models used in the study.*

AR: Thanks for the suggestion. We added Table 1 describing the ocean, sea-ice and marine biogeochemical components, the ocean grid resolution and the reference papers of the ESMs used in the study.

RC: *Line 88 - 89. Do you mean the model outputs are resampled analogously to the observations?*

AR: No, we meant that we only used data points for which measurements were available for all required variables: temperature, salinity, phosphate, oxygen, as well as estimates of AOU and age. We clarified the text as follows:

We only use data points where measurements of all variables mentioned are available.

RC: *Line 90. It should be GLODAPv2.2023.*

AR: The age product we used in this work (Jeansson et al., 2021), is based on GLODAPv2 (2016) (Key et al., 2015; Olsen et al., 2016; Lauvset et al., 2023). Thus, to maintain consistency, we used the same version of GLODAP even though more recent versions are available. We clarified this point in the revised manuscript as follows:

To have an observational baseline over the recent period, we also conduct an observation-based analysis of the trends in AOU and trends in age using the observational data product GLODAPv2 (2016) (Key et al., 2015; Olsen et al., 2016). [...] In order to be consistent with the age product, we opted for GLODAPv2 (2016), although more recent versions of the observational data product are available (e.g. (Lauvset et al., 2024)).

RC: *Line 129. Could you provide a reason why you exclude the Arctic Ocean?*

AR: We separated the Arctic Ocean from the Atlantic Ocean because the linear relationship between AOU trends and age trends was very different. We clarified this point in the revised manuscript as follows:

For the Southern water-masses, it is necessary to exclude grid-cells located north of 60°N as some Arctic Ocean grid-cells would otherwise be included without being continuously connected to the Southern Ocean. All longitudes are considered for the Southern water-masses. For the Atlantic water-masses, only grid-cells located east of the Drake passage, west of 30°E and south of 80°N are included to exclude grid-cells in the Pacific and Indian Oceans that fulfil the PO*-threshold and to exclude grid-cells in the Arctic Ocean. In the Arctic Ocean, a linear relationship between AOU trends and age trends emerges but with a very different slope than the one found for the Atlantic water-masses: here, AOU trends seem to be much more sensitive to age trends (not shown). Since our focus is on identifying linear relationships representative for most of the deep global ocean, we decided to exclude the Arctic from the analysis.

RC: *Line 131. To clarify, please add a sentence indicating that all longitudes are considered for Southern Ocean water masses.*

AR: Done.

RC: *Line 151. What is the difference between B and ε?*

AR: *B* is the intercept of the linear relationship. It is an output of the linear regression. *B* is the spatially averaged AOU change when there is no change in age. ε is the error of the linear regression in each point. So in each point $B + \varepsilon$ is the change in AOU that is not linearly related to changes in age. We clarified the text as follows:

In this work we intend to express the trends in AOU ($\frac{d\text{AOU}}{dt}$) via trends in age ($\frac{d\text{age}}{dt}$), in each point *X*, as follows:

$$\frac{d\text{AOU}}{dt}(X) = S_{\Delta\text{age}}^{\Delta\text{AOU}} \times \frac{d\text{age}}{dt}(X) + B + \varepsilon(X). \quad (1)$$

We assess the linear relationship between spatial fields of AOU trends and age trends within the previously define water-masses using a linear regression (Virtanen et al., 2020). The slope of the linear regression is the sensitivity of AOU changes to age changes ($S_{\Delta\text{age}}^{\Delta\text{AOU}}$). The intercept of the linear regression, *B*, represents a spatial average of the changes in AOU when there is no change in age. $\varepsilon(X)$ is the error of the linear regression in each point. All together, $B + \varepsilon$ represents the change in AOU that is not linearly related to changes in age, e.g. changes in remineralisation rates.

RC: *Line 161. Jenkins (1982) and Guo et al. (2023) could be included in the citations.*

AR: Indeed. They have been included in the revised manuscript.

RC: *Line 171-180. See my major concern 1.*

AR: See reply to major concern 1.

RC: *Line 195 - 267. I suggest dividing the Results section into several subsections for improved clarity and organization.*

AR: Done. Thanks for the suggestion.

RC: *Line 219 - 221. AOU versus age across water masses can be the reason, but also the change of local biogeochemical processes?*

AR: Indeed, the revised version mention it as follows:

Weaker correlations are potentially related to significant contribution of mixing over advection, spatial variability or local changes of respiration rate, or a partially inappropriate definition of water-masses.

RC: *Figure 2. Are the blue dots representing only values with a significant trend? If so, please clarify this. Otherwise, I am curious why Southern Ocean and Atlantic light/dense water make up around 98% of the deep ocean, yet many points remain grey—are these from the upper 1000 m?*

AR: In Figure 2, the blue shading shows only the values with significant trends belonging to the four water-masses used in the work. The grey background shows the entire ocean, including the non-significant trends and the surface ocean. The revised caption of the figure now reads:

The blue shading shows the number of data point for each bin of ideal-age trends and AOU trends for the Southern and Atlantic light/dense waters, accounting only for grid-points where ideal-age and AOU trends are significant.

The grey area does not indicate the number of points but rather the range of age and AOU trends in the entire ocean. Typically, trends in the upper 1000 meters are stronger than those in deeper regions, thus covering a larger portion of the figure.

RC: *Figure 4. Please improve the visualization of panel (a), as it currently appears quite cluttered, despite its importance. Are the zonal section panels essential, or could they be simplified or omitted to enhance clarity?*

AR: As the zonal section are extensively discussed in the results section, we choose to keep them. To enhance the clarity, there are two figures in the revised manuscript. The revised Fig. 4 shows the globally integrated trends and the revised Fig. 5 shows the zonally average trends. The revised Fig. 4 has also been simplified to only keep the key informations.

RC: *Line 279 - 290. This section is overall too descriptive. Additionally, as I mentioned in Major concern 1, I am concerned about potential inconsistencies in the methodology when estimating $S_{\Delta\text{age}}^{\Delta\text{AOU}}$ in models versus observations.*

AR: This paragraph has been entirely re-written to account for the results of the uncertainty analyses conducted for the revision.

RC: *Line 299 - 302. Guo et al. (2023) could be cited here. These authors proposed two possible explanations for weak connection between age change and AOU change. First, changes in ocean circulation within a warming climate could alter water mass composition, as different water masses with varying biogeochemical histories are recombined over time—potentially introducing younger yet more oxygen-depleted waters into a given region. Second, local biological activity may influence the AOU signal independently of ventilation. For instance, even if ventilation slightly increases, an increase in local respiration rates could cause AOU to rise, reflecting biological consumption rather than changes in physical mixing.*

AR: Thanks for pointing to this paper. The revised discussion section reads:

While we highlight the importance of change in ideal-age in this water-mass, a substantial portion of the change in AOU is not driven by change in ideal-age in lighter water-masses. Here, changes in export (Henson et al., 2022), spatially variable oxygen utilisation rate (Sulpis et al., 2023) or changes in remineralisation with temperature (Brewer and Peltzer, 2017) can de-correlate changes in AOU from

changes in ideal-age. Furthermore, in a transient climate, the conditions (strong advection over mixing and spatially even respiration) for a linear relationship between AOU trends and age trends can change over time. Changes in circulation can also recombine water masses differently bringing in waters with varying histories; waters can follow different pathways and go through different respiration fields (Guo et al., 2023).

2. Reviewer #2

RC: *This study investigates the relationship between apparent oxygen utilisation (AOU) trends and water mass age trends in the deep ocean using Earth system models and observational data from GLODAPv2. The authors establish linear relationships between these variables to assess the biological carbon pump's sensitivity to circulation changes under contemporary and future climate scenarios. While this work addresses an important aspect of ocean biogeochemistry relevant to climate projections and would be of interest to the journal, several aspects require further clarification or investigation to enhance the paper's overall strength.*

2.1. Major comments

RC: *Water Mass Definition Methodology: The water mass classification using PO* thresholds appears somewhat arbitrary and model-dependent. The authors mention "iterative testing" to refine thresholds but provide insufficient detail about this process. Since most of the results are based on this classification, I believe a more detailed section about the derivation of these thresholds is necessary. I also don't understand why there is a difference between the number of water mass definitions for ESMs and observations (four vs two water masses).*

AR: We revised the PO* thresholds to make them less arbitrary. Following Broecker et al. (1998), the revised thresholds are based on the PO* distribution in the deep subpolar North Atlantic and deep Southern Ocean for each ESM. This approach yields a PO* threshold for each ESM. It does not substantially change the water mass selection but make it more objective and tailored. The revised manuscript now includes:

- The following description of the PO* thresholds derivation :

Broecker et al. (1998) state that the global distribution of PO* has its minimum in the North Atlantic, its maximum in Southern Ocean and that the PO* distribution for deep waters formed in the North Atlantic is very distinct from the distribution for deep waters formed in the Southern Ocean. Based on these statements, we defined the PO*-thresholds for deep ocean water masses individually for each ESM, using PO* averaged on 1982-2013 and applying the following approach: (i) We compute the 95th percentile of the PO* distribution in the deep subpolar North Atlantic, between 40-60°N and 0-70°E. (ii) We compute the 5th percentile of the PO* distribution in the deep Southern Ocean, south of 55°S. (iii) The PO*-threshold is defined as the average between the aforementioned percentiles. We find that Atlantic water-masses have PO* values below the threshold while Southern water-masses have PO* values above the threshold (see supplementary Fig. S1).

- The supplementary Fig. S1 showing the PO* distributions and PO* threshold for each ESM.

Regarding the number of water masses, we divided the water masses defined by PO* into two density

classes to improve the correlation between trends in AOU and age. Identifying relationships between AOU and age within density classes is common practice (e.g., Sulpis et al. (2023)). Applying the same division to the observational dataset make little because it samples the densest half of the water masses. We compared the density distribution of the data points used here with the density distribution derived from the GLODAPv2.2016b mapped climatologies (Lauvset et al., 2016; Key et al., 2015). We found that 75% and 98% of the data points used here have a density higher than the median density from the mapped climatologies, in the Southern and Atlantic water masses, respectively. The revised manuscript clarify this as follows:

Most of the data points used in this work belong to the densest half of waters originating from the North Atlantic Ocean or the Southern Ocean (98 % and 75 % of the points respectively).

RC: *Linear Relationship Assumption Validity: While the authors acknowledge in the introduction that linear relationships between AOU and age can break down in certain regions and under specific conditions of the BCP. It would be interesting to further discuss under which conditions this linearity assumption fails in a transient climate scenario. For example, it is unclear to me under which conditions this linear relation still applies (regionally or globally) on SSP5-8.5.*

AR: In the revised discussion section, we adress the conditions for the linear relationship as follows:

While we highlight the importance of change in ideal-age in this water-mass, a substantial portion of the change in AOU is not driven by change in ideal-age in lighter water-masses. Here, changes in export (Henson et al., 2022), spatially variable oxygen utilisation rate (Sulpis et al., 2023) or changes in remineralisation with temperature (Brewer and Peltzer, 2017) can de-correlate changes in AOU from changes in ideal-age. Furthermore, in a transient climate, the conditions (strong advection over mixing and spatially even respiration) for a linear relationship between AOU trends and age trends can change over time. Changes in circulation can also recombine water masses differently bringing in waters with varying histories; waters can follow different pathways and go through different respiration fields (Guo et al., 2023). In the lighter water masses and in the Atlantic dense water mass, changes in ideal-age explain a slightly smaller portion of the spatial variability of AOU changes in the future compare to the contemporary period (weaker R^2), suggesting caution when interpreting circulation changes from AOU.

RC: *Drift Correction Methodology: The drift correction procedure using a piControl simulation makes sense, however, it is not specified how the actual years of the piControl are adjusted to fit 1850-2099.*

AR: The piControl years have been chosen based on the information available in the netcdf files indicated which year of the piControl was used to start the historical simulations. The revised manuscript clarify it as follows:

The drifts are estimated for every ocean grid-cell of the ESMs using a linear regression over 250 years of the piControl simulation, starting from the year in the piControl simulation where the historical simulation has been initialised (see supplementary Tab. S1).

RC: *Observational Data Limitations and T-S Binning: The T-S space binning approach for observational data analysis is not well justified. The choice of bin resolution (0.027 and 0.0023) seems arbitrary, and the methodology results in very small sample sizes (only 23 and 118 significant joint trends for Southern and Atlantic water masses). The authors should better emphasise the reason behind using this TS binning.*

AR: We revised the observational data analysis to test sensitivity to the T-S bin size and geographical constraint. The geographical constraint ensures that data points are not too far apart, as suggested by the other reviewer.

We ran the analysis 625 times with random T-S bins and geographical constraints to derive a distribution of $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$. In the revised manuscript, the observation-based $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$ values are presented as the mean \pm one standard deviation from the distribution. The revised manuscript explains this as follows:

Trends in AOU and age are computed within bins in the T-S space (supplementary Fig. S6). To ensure that data points are geographically close to each other within each T-S bin, we (i) remove outliers defined as data points geographically further than twice the median distance to the median location and (ii) only keep points that are within a maximum distance from the median location (recomputed without the outliers). Thus the computation of the trends depends on two choices: the temperature and salinity resolution for the T-S bins and the maximum distance from the median location within each T-S bins. These choices affect: (i) the grouping of comparable measurements into the same T-S bin, regardless of their geographical location, (ii) the number of data points per T-S bin needed to identify significant trends (p-value ≤ 0.05), and (iii) the number of trend estimates (one per T-S bin) required to establish a significant (p-value ≤ 0.05) correlation between AOU trends and age trends. Trends for AOU and age, (dAOU/dt) and (dage/dt) are computed when five or more observations are grouped into a given T-S bin. Due to the substantial influence of the T-S bin size and the maximum distance, we conduct the analysis of the observational data 625 times with different random choices of these parameters to derive a distribution of the observation-based $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$. Temperature/salinity resolutions ranged from 0.026 °C to 0.325 °C and 0.0024 to 0.03, while maximum distance from 500 km to 5000 km.

RC: *On the significance of the observed overestimation in ESMs: The authors show clearly that observational values are "on the low side of the ESMs range" which indicates overestimation of the sensitivity of AOU to age change. While they briefly acknowledge limitations in computing estimates for the contemporary period based on observations (lines around 279), I'm not totally convinced that the overestimation in ESMs is not due to having more data. For instance, if the ESMs were sampled at the observational locations and the analysis repeated with the same methodological choices as for observations data (e.g. same water mass definition, same trend calculation approach with the same TS binning), are we going to still going to observe an overestimation ?*

AR: In the revision, we assessed the observational dataset using NorESM2-LM (SAMPLE-UNC analysis in the revised manuscript). Model outputs were sampled according to the spatio-temporal distribution of the observational dataset. As for the observational dataset analysis, we ran 625 analyses with random choices of T-S bins and geographical constraints to derive a $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$ distribution. The mean \pm one standard deviation is compared with the original $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$ (no sampling, trends computed in each grid point) to assess uncertainty from data scarcity and the need for T-S space analysis. SAMPLE-UNC results suggest data scarcity leads to $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$ overestimation. Additionally, as suggested by the other reviewer, we assessed the uncertainty related to the TTD method for estimating age trends (TTD-UNC analysis). Uncertainties in observation-based $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$ compromise comparability with ESM values, preventing conclusions that ESMs overestimate $S_{\Delta_{\text{age}}}^{\Delta_{\text{AOU}}}$. In the revised manuscript, we have made the following changes:

- Revised the title of the manuscript.
- Included a new subsection (2.5) titled "Evaluation of model-observation comparability", which describes the differences between the methods applied to ESM and the observational dataset, and outlines the new uncertainty analyses.
- Expanded the new subsection 3.3 to describe the results of the observational dataset analysis, accounting for uncertainties.

- Added Table 2 to present the observation-based $S_{\Delta_{\text{age}}}^{\Delta\text{AOU}}$ and the results of the uncertainty analyses.
- Removed the observation-based $S_{\Delta_{\text{age}}}^{\Delta\text{AOU}}$ in the revised Fig. 3.
- Expanded the discussion to address the reliability of the observation-based estimate.
- Clarified the distinction between TTD-mean-age derived with the TTD method from observations and the ideal-age tracer computed in the model simulation.

RC: *I hope that the authors will understand my comments in a constructive way, and that I value their work and the time they invested in the preparation of the manuscript. It might be that I have misunderstood something, in this case, if something wasn't clear for me as a reviewer, it is possible that it wouldn't be clear also for the readers.*

AR: We appreciated the constructive comments from the reviewer, that helped the revision.

References

- T. Bourgeois, N. Goris, J. Schwinger, and J. F. Tjiputra. Stratification constrains future heat and carbon uptake in the Southern Ocean between 30°S and 55°S. *Nature Communications*, 13(1):340, Jan. 2022. ISSN 2041-1723. .
- P. G. Brewer and E. T. Peltzer. Depth perception: The need to report ocean biogeochemical rates as functions of temperature, not depth. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2102):20160319, Aug. 2017. .
- W. S. Broecker, S. L. Peacock, S. Walker, R. Weiss, E. Fahrbach, M. Schroeder, U. Mikolajewicz, C. Heinze, R. Key, T.-H. Peng, and S. Rubin. How much deep water is formed in the Southern Ocean? *Journal of Geophysical Research: Oceans*, 103(C8):15833–15843, 1998. ISSN 2156-2202. .
- N. Goris, K. Johannsen, and J. Tjiputra. The emergence of the Gulf Stream and interior western boundary as key regions to constrain the future North Atlantic carbon uptake. *Geoscientific Model Development*, 16(8):2095–2117, Apr. 2023. ISSN 1991-959X. .
- H. Guo, I. Kriest, A. Oschlies, and W. Koeve. Can Oxygen Utilization Rate Be Used to Track the Long-Term Changes of Aerobic Respiration in the Mesopelagic Atlantic Ocean? *Geophysical Research Letters*, 50(13):e2022GL102645, 2023. ISSN 1944-8007. .
- H. Guo, W. Koeve, A. Oschlies, Y.-C. He, T. P. Kemena, L. Gerke, and I. Kriest. Dual-tracer constraints on the inverse Gaussian transit time distribution improve the estimation of water mass ages and their temporal trends in the tropical thermocline. *Ocean Science*, 21(3):1167–1182, July 2025. ISSN 1812-0784. .
- Y.-C. He, J. Tjiputra, H. R. Langehaug, E. Jeansson, Y. Gao, J. Schwinger, and A. Olsen. A Model-Based Evaluation of the Inverse Gaussian Transit-Time Distribution Method for Inferring Anthropogenic Carbon Storage in the Ocean. *Journal of Geophysical Research: Oceans*, 123(3):1777–1800, Mar. 2018. ISSN 2169-9275, 2169-9291. .
- S. A. Henson, C. Laufkötter, S. Leung, S. L. C. Giering, H. I. Palevsky, and E. L. Cavan. Uncertain response of ocean biological carbon export in a changing world. *Nature Geoscience*, 15(4):248–254, Apr. 2022. ISSN 1752-0908. .

- T. Ito, Y. Takano, Y. A. Eddebbbar, J. F. Tiputra, Z. Wang, S. Minobe, L. Cheng, J. Du, and Y. Abe. Are Simulated Ocean Deoxygenation Rates Consistent with the Observational Reconstructions? *Annual Review Earth and Planetary Sciences*, Jan. 2026. .
- E. Jeansson, R. Steinfeldt, and T. Toste. Water mass ages based on GLODAPv2 data product (NCEI Accession 0226793). NOAA National Centers for Environmental Information., 2021.
- W. J. Jenkins. Oxygen utilization rates in North Atlantic subtropical gyre and primary production in oligotrophic systems. *Nature*, 300(5889):246–248, Nov. 1982. ISSN 1476-4687. .
- R. M. Key, A. Olsen, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Perez, and T. Suzuki. Global Ocean Data Analysis Project, Version 2 (GLODAPv2), ORNL/CDIAC-162, NDP-093. Technical report, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee., 2015.
- W. Koeve and P. Kähler. Oxygen utilization rate (OUR) underestimates ocean respiration: A model study. *Global Biogeochemical Cycles*, 30(8):1166–1182, 2016. ISSN 1944-9224. .
- L. Kwiatkowski, L. Bopp, O. Aumont, P. Ciais, P. M. Cox, C. Laufkötter, Y. Li, and R. Sférian. Emergent constraints on projections of declining primary production in the tropical oceans. *Nature Climate Change*, 7(5):355–358, May 2017. ISSN 17586798. .
- S. K. Lauvset, R. M. Key, A. Olsen, S. van Heuven, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Perez, T. Suzuki, and S. Watelet. A new global interior ocean mapped climatology: The $1^\circ \times 1^\circ$ GLODAP version 2. *Earth System Science Data*, 8(2):325–340, Aug. 2016. ISSN 1866-3508. .
- S. K. Lauvset, R. M. Key, A. Olsen, S. M. A. C. van Heuven, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Pérez, T. Suzuki, and S. Watelet. A new global interior ocean mapped climatology: The $1^\circ \times 1^\circ$ GLODAP version 2 from 1972-01-01 to 2013-12-31 (NCEI Accession 0286118). NOAA National Centers for Environmental Information. Dataset., 2023.
- S. K. Lauvset, N. Lange, T. Tanhua, H. C. Bittig, A. Olsen, A. Kozyr, M. Álvarez, K. Azetsu-Scott, P. J. Brown, B. R. Carter, L. Cotrim da Cunha, M. Hoppema, M. P. Humphreys, M. Ishii, E. Jeansson, A. Murata, J. D. Müller, F. F. Pérez, C. Schirnick, R. Steinfeldt, T. Suzuki, A. Ulfsbo, A. Velo, R. J. Woosley, and R. M. Key. The annual update GLODAPv2.2023: The global interior ocean biogeochemical data product. *Earth System Science Data*, 16(4):2047–2072, Apr. 2024. ISSN 1866-3508. .
- Y. Liu, J. K. Moore, F. Primeau, and W. L. Wang. Reduced CO₂ uptake and growing nutrient sequestration from slowing overturning circulation. *Nature Climate Change*, 13(1):83–90, Jan. 2023. ISSN 1758-6798. .
- A. Olsen, R. M. Key, S. van Heuven, S. K. Lauvset, A. Velo, X. Lin, C. Schirnick, A. Kozyr, T. Tanhua, M. Hoppema, S. Jutterström, R. Steinfeldt, E. Jeansson, M. Ishii, F. F. Pérez, and T. Suzuki. The Global Ocean Data Analysis Project version 2 (GLODAPv2) – an internally consistent data product for the world ocean. *Earth System Science Data*, 8(2):297–323, Aug. 2016. ISSN 1866-3508. .
- A. Oschlies, P. Brandt, L. Stramma, and S. Schmidtko. Drivers and mechanisms of ocean deoxygenation. *Nature Geoscience*, 11(7):467–473, 2018. .

- O. Sulpis, D. S. Trossman, M. Holzer, E. Jeansson, S. K. Lauvset, and J. J. Middelburg. Respiration Patterns in the Dark Ocean. *Global Biogeochemical Cycles*, 37(8):e2023GB007747, 2023. ISSN 1944-9224. .
- P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, and P. van Mulbregt. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17(3):261–272, Mar. 2020. ISSN 1548-7105. .