We would like to thank **Reviewer # 2** for their thorough review of our manuscript and thoughtful suggestions. Below, we respond to the more significant suggestions, in the order they appear. The reviewer's comment is in italics, while our **Response** follows.

This study proposes a new data set of steric sea level and ocean heat content (OHC) based on Argo data for the 0-2000m depth range and 2003-2023 time span. From my understanding, the two main differences with other existing Argo-based steric data sets are: (1) the use of a 'geodetic' steric climatology as reference-or first guess- (defined as the difference between the altimetry-based global mean sea level/GMSL and GRACE-based ocean mass), and (2) a long-wavelength mapping for reducing uncertainty due to meso-scale eddies not well observed by Argo.

Such a new data set could be of great value for some studies that investigate the long-term upper ocean warming. It could be also considered as a useful new indicator of the contribution of the upper ocean to the Earth energy imbalance. However (and this is not mentioned in the manuscript), I would be more cautious about using it for assessing global and regional sea level budget closure, as explained below.

While such a data set may be worth to be published, the manuscript needs substantial revision and clarification.

- *The authors do not explain well the novelty of their results neither the specific applications of their data set. To me, such a data set would be useful as an indicator of upper ocean warming but should not be used for sea level budget closure assessments. In effect, the use altimetry-based GMSL corrected for the ocean mass contribution (equivalent to the steric sea level) as first guess (or reference) will have a strong weight during the first 5 years of the record in any sea level budget assessment because of the poor Argo data coverage, hence the limited information brought by Argo data in the optimal interpolation process. Over this time span, the computed steric sea level will thus be dominated by the 'Altimetry minus GRACE' signal. As a result, one may expect that the GMSL budget of the early part of the record will be artificially closed because the estimated steric sea level from this study will be - by construction- essentially equal to the 'Altimetry minus GRACE' time series.*

Response: This dataset will be as useful in sea level budget studies as any other Argo-based T/S gridded product. Such Argo T/S products for the upper 2000m are routinely used in sea level budget analyses. While our grids will default back to the altimetry-GRACE reference in early years (strongest before 2005, reducing from there to about 2008 as demonstrated in the analysis), the Argo data controls the mapping post-2008 (as we discuss).

We have revised the manuscript to make it clearer that these products are intended to provide SSL, TSL, and OHC anomalies above 2000m. For example, in the discussion section, we now state:

"There are limitations to the data. These maps are intended to represent SSL, TSL, and OHC for the upper 2000m of the ocean. They do not account for any steric variations below 2000m.

However, this is true of any Argo-based mapped product of T/S or OHC above 2000m. New data from the deep Argo array indicate the possibility of relatively large steric anomalies below 2000m in at least one small area of the tropical Atlantic (Zilberman et al., 2025), but complete knowledge of deep steric signals is still very much an open research area. The maps we have produced will be as useful as any other current Argo-based product for global and basin-scale sea level and heat budgets that are available at the present, provided the studies are for areas larger than about 1000km by 1000km."

- *The term 'long-wavelength' is unexplained (500 km?, 1000km? more?...).*

Response: Apologies. We assumed this would be clear from the equations used for the OI. We have added text after Equation (7) to be explicit about what we mean by "longwave" and "shortwave."

"Throughout the remainder of this paper, when we say "longwave" or "long wavelength" we mean the portion of SSL that has a covariance function equal to $25e^{-\left(\frac{r}{1675\ km}\right)}$, i.e., an exponential decay with a roll-off of 1675 km and max covariance of 25 cm^2. By "shortwave" or "short wavelength" we mean the portion of SSL that has a covariance function equal to $82e^{-\left(\frac{r}{100\ km}\right)^2}$, i.e., a Gaussian with a roll-off of 100 km and max covariance of 82 cm^2. By "eddy resolution" we mean the full covariance structure of Equation (7). A spectral analysis of the longwave maps compared to the original, unsmoothed data indicates the longwave maps keep nearly 100% of the power for wavelengths longer than 1500 km."

We have also added another line to equation (7) to make this separation abundantly clear in the equation:

$$C^{sat}(r) = 82e^{-\left(\frac{r}{100\ km}\right)^2} + 25e^{-\left(\frac{r}{1675\ km}\right)}\ , \qquad (7)$$
$$C^{sat}(r) = C^{sat}_{short}(r) + C^{sat}_{long}(r).$$

We also now state the definition of long-wavelength in the abstract, and verify that averages over a box that is 1500km by 1500km in an eddy-rich region is consistent with the average over the raw data in the same region:

New test on lines 439-445 of the revised text: "We performed an additional test over a smaller area (1500 km by 1500 km) east of Japan where there is significant eddy activity. Correlations remain 0.99 or higher (even after removing a trend and annual variations) and standard deviation of the differences only increases to 2 mm, with trend errors < 0.2 mm yr^{-1}. However, if smaller areas are considered (for example 500 km by 500 km in the same area), non-seasonal correlation drops to 0.8 and the standard deviation of differences increases to 4 cm, with significant trend differences (1.7 mm yr^{-1}). Thus, we conclude the longwave mapping is sufficiently accurate to resolve average monthly SSL, TSL, OHC anomalies for the upper 2000m over areas greater than 1500 x 1500 km. At smaller areas, especially within eddy regions, the longwave maps cannot fully resolve the average and should not be used for studies over regions smaller than this.”

- *Above all, the results need thorough validation in order to convince users about the interest of using this new data set. Only a single comparison with one Argo product is proposed in the manuscript. More is needed. For example, the derived OHC data set could be compared with CERES data.*

Response: We would argue that the Altimeter-GRACE SSL global curves are essentially independent post 2008 when our mapping does not appreciably change under different reference grids (including with no reference) – see Figure 7 in the new version (Figure 6 in old version). This comparison is already included in the analysis.

In response to comments by **Reviewer # 1**, we have eliminated a discussion of OHU, as dOHC/dt for the upper 2000m is only one (albeit, a major) component of it. Instead, we have focused only on OHC and dOHC/dt for the upper 2000m. In this case, we feel it is perfectly fair to compare with other Argo-based products to show where our results agree and disagree. We have, however, added another dataset, from the Lyman and Johnson (2023) analysis.

Because the CERES-based estimates of OHU (such as those calculated by Mayer et al) would include portions our maps do not measure (such as OHU below 2000m or in the Artic Ocean). While we could add estimates from other studies (which is done in any Argo-based estimate of OHU), we feel showing the estimates of OHC and dOHC/dt for the upper 2000m against two other Argo-based products is sufficient.

- *It is a pity that all presented steric sea level curves include the seasonal cycle while one is likely more interested in the interannual variability. The discussion about trend differences is quite superficial.*

Response: The annual/semiannual signals have been removed and interannual variations (unsmoothed) are shown for all time-series now as a separate panel. We believe it is also important to show that the seasonal variation is captured as well, which is why it remains.

Statistics for the basin-averages in Table 1 are now computed for the full series and the interannual-only ones. Note that removing the seasonal variation has no impact on correlation and the standard deviation of residuals becomes smaller, reflecting that much of the difference arises from small changes in the seasonal variation, not the interannual variations.

Additional analysis on the interannual variations has been added throughout the revised manuscript.

- *Finally, while the authors are well aware of the Argo-based salinity measurement errors as of 2015, there is no discussion on the reliability of their halosteric data set. This is an important limitation of this study.*

Response: We stated that there are clearly residual errors in the SSL data that are most likely from salinity, even though we use the latest version of the Argo profile data with new flagging and adjustments (e.g., Wong et al., 2023). We have revised this section to make this clearer. In addition, in the Methods section, we have explicitly stated that the flagging we use should eliminate floats with detected salinity drift using the methods of Wong et al. (2023).

New text: (in Section 2.1, **new text is bold**): "Additionally, profiles were retained only if the QC flags for position, date, pressure, salinity, and temperature were all set to "1," indicating the highest quality measurements (Wong et al., 2023). **In particular, this flag indicates no sign of apparent salinity drift in the float**."

New text: lines 595-603 in revision: "Moreover, when global thermosteric sea level (GTSL) is compared instead (**Fig. 10**), there is no change in behavior during 2015, but USF and SIO GTSL agree well for all other periods. This suggests that the difference between USF and SIO GSSL in 2018-2020.5 is due to unresolved salinity errors that are not flagged (and so get into the USF processing) but that SIO has edited out. Unfortunately, there is no documentation that we can find on recent SIO processing standards that can explain how this was accomplished. Moreover, the SIO and USF GSSL curves both depart significantly from the altimetry-GRACE curve after 2020, whereas the SIO and USF TSL curves differ by a smaller amount. This further suggests unresolved salinity errors in a large number of profiles after 2018, even though we have utilized a release of Argo T/S profiles that has added new flags and adjustments to identify such problem floats (Wong et al., 2023). We removed such flagged/adjusted floats from our processing stream (see Section 2.1). However, this does not prevent large and unrealistic global halosteric signals post 2018, so we must conclude that unidentified salinity errors remain in enough profiles to affect the global halosteric signal."

Detailed comments

- *Abstract, line 9: indicate the wavelength range. What means 'long-wavelength'?*

Response: Done.

- *Abstract, line 11: explain what is a 'mean climatology'*

Response: Changed to read: "…the satellite estimate is used instead of a mean monthly climatology or zero, as other analysis centers use."

- *Introduction, line 23: I suppose that you mean 'open ocean' rather than 'deep ocean'?*

Response: Yes. Changed.

- *Introduction, lines 49-62: indicate here that your altimetry data set does not account for the Jason-3 radiometer drift (a problem that may impact your results, especially at the end of the record)*

Response: We feel this is more relevant in the data section than the Introduction, so have added it there. However, we note that not using the radiometer drift correction will not affect the final grids, as this only affects data after 2016 and that in this period the mapping is controlled by Argo, not altimetry-GRACE (see Figure 7). Also, we do use the radiometer drift correction for Figures 9, 10, and 11 (and we already note that in the text).

New text in Section 2 reads: "We note that these data do not include a correction for the recently discovered small drift of the Jason-3 radiometer (Brown et al., 2023), which will affect the altimeter after January 2016. However, during this time, Argo sampling is of sufficient density (Fig. S1 and S2) that they should correct for any small drift errors in the reference surface used. We will demonstrate this is so in Section 3 by utilizing different initial reference grids, including zero and a mean monthly climatology, neither of which are affected by any radiometer drift."

We considered re-doing our analysis when the newest Beckley et al analysis was released which included the J3 radiometer drift correction. But a quick test with a couple of months indicated no change in the final grids.

- *Introduction, lines 57-58: deep ocean warming may no more be negligible (e.g., Johnson and Purkey, 2024).*

Response: We have deleted that sentence and rephrased it to better emphasize this is still very much an open research question and that groups have to add in an estimate of SSL contributions below 2000m to study the sea level budget.

New text: "Fully quantifying deep ocean contributions to steric sea level is still very much an open area of research (e.g., Purkey and Johnson, 2010; Desbruyères et al., 2016; Johnson and Purkey, 2024). For example, a recent study suggests that short-term changes in small areas may be of the order of ± 2 cm for periods of up to a year (Zilberman et al., 2025), Because of this, studies that study the sea level budget have to add an estimate of deep steric changes to any Argo-based product that can only resolve changes above 2000m. With that caveat, numerous studies have compared the two methods on regional and global scales, finding good agreement in ocean heat content (von Schuckmann et al., 2014; Hakuba et al., 2021; 2024; Marti et al., 2022; Meyssignac et al., 2019;) and global thermosteric sea level (Blaquez et al., 2018; Barnoud et al., 2021), but with substantial differences in steric sea level after 2015 (Blaquez et al., 2018; Chen et al., 2020; Barnoud et al., 2021), likely due to salinity problems in the gridded Argo T/S products."

Again, this is not a unique problem to the data we have created, but is true of ANY Argo-based SSL estimate.

- *Section 2.1, lines 116-118: the sentence 'we kept only profiles deeper than 750 dba' is unclear. Which depth? Why?*

Response: 750 dbar is the pressure level (so ~750 m depth). This is actually a typo – we used an initial selection to 750 dbar, but further restricted the max to 1000 dbar in the gridding step. We have revised the sentence to be clearer on this, and added a note on a test where we found it made little difference on the mapping using some floats to 1000m and some to 2000m, or all to 2000m (really, anything deeper than 1000m).

"Since our goal is to integrate vertically to obtain a value for upper ocean ΔSSL, etc, we kept only profiles that had a maximum pressure/depth greater than 1000 dbar, had a minimum pressure ≤ 50dbar, and that had more than 50 observations in each profile. Although we allowed profiles that only sampled the upper 1000m and not the full 2000m of the upper ocean, this was a relatively small number and tests indicated it did not significantly alter recovered maps over using a restriction to profiles that observed down to > 1900 dbar."

- *Section 2, line 145, I suppose that 'who which' should read 'who wish'…*

Response: Yes. Changed.

- *Section 2.2: line 152: indicate which corrections are applied (and mention that the Jason-3 radiometer drift is not accounted for)*

Response: We note first that we don't make corrections. These are done by Beckley et al. and their final product is fully corrected SSH anomalies. It is beyond the scope to provide a full description of their processing (or specific models used), but we have added a note about some of the main ones and referred the reader to the appropriate document for more info. We have also added the note about the radiometer.

New text: "These data have consistent geophysical corrections and orbits and have had all standard geophysical corrections applied (e.g., inverted barometer, ocean tides, wet and dry troposphere, ionosphere, sea state bias) as described in the data documentation (Beckley et al., 2010; 2022). We note that these data do not include a correction for the recently discovered small drift of the Jason-3 radiometer (Brown et al., 2023), which will affect the altimeter after January 2016. However, during this time, Argo sampling is of sufficient density (Fig. S1 and S2) that they should correct for any small drift errors in the reference surface used. We will demonstrate this is so in Section 3 by utilizing different initial reference grids, including zero and a mean monthly climatology, neither of which are affected by any radiometer drift."

- *Section 2.2: lines 173-174: mention (for readers not familiar with the sea level budget) that 'altimetry-based sea level minus GRACE-based ocean mass' represents the steric sea level*

Response: We stated this multiple times throughout the Introduction and do not feel it is necessary to repeat it again here in the data section. We feel it is well established in our manuscript that we are using altimetry-GRACE for the satellite estimate of SSL.

- *Section 2.3, Optimal interpolation: this section is not easy to read. The numerical values mentioned in the equations seem to have been pulled out of a hat…*

Response: Numbers are not "pulled out of a hat." We explained the process to compute the rolloff and amplitudes for the Gaussan and exponential functions of the covariance right before Equation (7):

"A single autocovariance was computed for all months between 2003 and 2023 – we tested computing month-specific functions, but the differences were minimal, so we used the single covariance function to reduce complexity. We then approximated the covariance values with a continuous function comprised of a Gaussian for short wavelength signals and exponential decay functions for the longwave portion, along with a random component to match the full variance of the data, similar to that done in previous studies (e.g., Willis et al., 2008; Roemmich and Gilson, 2009). Optimal parameters for the roll-off and amplitude parameters were estimated using non-linear least squares based on iterating values of the roll-off parameters over a range of expected values. "

- *Section 2.3, line 215: what means '4-10x'?*

Response: Sorry. "x" in this context is often used for "times" and that is what we meant: 4-10 times slower. We have changed it to "times."

- *Section 2.3, line 272: what means 'mean surface' for GRACE data? (and Argo); line 289: 'corrected' for what?*

Response: All 3 datasets (GRACE, altimetry, Argo) are anomalies relative to the mean state for some period. If that state were fixed and not evolving in time, this would not be a problem. But, for example, the mean ocean surface in 2000 is (on average) than the mean in 2010, or 2012. Averaging of climate patterns like El Niño can also cause regional biases differences. We have revised this section to emphasize these data are all anomalies computed relative to different mean periods.

New text: "However, before combining the satellite and Argo data in the OI scheme, one must account for differences in the reference period used for the satellite altimetry, GRACE/FO, and Argo anomalies. For example, GRACE/FO anomalies are referenced to a 2005-2010 mean period, the altimetry SSH anomalies to a mean period of approximately 1993-2018, and Argo to T/S means for 2004-2018."

- *Section 3, Table 1: since you do not remove the seasonal cycle, the correlation essentially refers to this dominant seasonal cycle. What about the correlations at interannual time scale?*

Response: Correlation of interannual (seasonal removed) and standard deviation of interannual residuals has been added to Table 1. In all cases, there is no significant difference in interannual versus full comparisons.

- *Section 3, Table 1: trend differences around 0.3 mm/yr are not negligible at all (same order of magnitude as the GIA signal in the altimetry-based GMSL)*

Response: These were typos. In recomputing the data for the new Figure 1, we discovered we had transcribed differences in the bias term, not the trend term, from the calculation to the Table. These have all been corrected. New maximum values are ± 0.18 mm yr^{-1}, which are well within the uncertainty of the GMSL trends from satellite altimetry.

- ***Section 3, Figures 5 and 6: Show difference time series!!! Remove the seasonal cycle***

Response: The SSL and TSL figures now also include a panel showing the interannual variations after removing the best seasonal fit. We choose not so show differences as well, as these can be deduced from the plots we show.

- *Section 3, lines 59-362: The trend difference over 2005-2024 (1.04 versus 1.17 mm/yr) is not significantly larger than the one over 2011-2024 (0.08 mm/yr), so that the use of the geodetic reference has limited impact.*

Response: Possibly, but it does have an impact and the difference is outside the uncertainty, so does show the impact on GMSL of the switch between a climatology and the satellite reference.

- *Section 3, results: Why not show maps of regional trend differences? I suppose that the regional trend differences are not uniform but may depend on the regional Argo data sampling*

Response: Comparing regional trend differences between our grids and those from, for example, SIO, would mainly emphasize mesoscale eddies, so are not relevant. And comparisons with altimetry-GRACE would show general differences between the satellite and Argo SSL products, which has been described by several other groups. This is also more of a scientific investigation, rather than a data description paper, which is the purpose of this manuscript, and why it was submitted to Earth System Science Data. We have, however, added a plot of standard deviation of differences between the altimetry-GRACE reference and final grids (as well as with an altimetry reference) (new Figure 3) at the request of Reviewer #1, which provides some information on the agreement.

- *Section 3, line 380: 'The satellite data tend to dominate the mapping from 2003 to 2008…': see my comment above*

Response: This sentence was discussing where the monthly maps are based on the satellite reference or Argo data, not the impact of the reference on the calculation of the global average, and will affect areas with significant interannual variability. We have added a sentence commenting on this:

New text: "This should result in an improved estimate of SSL, TSL, and OHC in these areas over using a mean monthly climatology if there are significant interannual fluctuations there – e.g., El Niño/La Niño signals in the eastern Pacific, which did not have complete Argo coverage in every month until 2008 **(Fig. S1)**."

- *Section 3, lines 400-404: the discussion about the corrections should appear much earlier in the manuscript (see comment above)*

Response: We have added the note on the radiometer correction earlier and why it was not applied in the original processing but why it is here.

- *Section 3, figures 8 and 9:* ***show difference time series!!! Remove the seasonal cycle***

Response: The figures now also include a panel showing the interannual variations after removing the best seasonal fit. We choose not so show differences as well, as these can be deduced from the plots we show.

- *Section 3 figures 8 and 9:* ***what is the impact of salinity measurement errors on the USF SSL curves?*** *Please discuss*

Response: we felt we did discuss this, but have added some more discussion in the revision:

New text: "This suggests that the difference between USF and SIO GSSL in 2018-2020.5 is due to unresolved salinity errors that are not flagged (and so get into the USF processing) but that SIO has edited out. Unfortunately, there is no documentation that we can find on recent SIO processing standards that can explain how this was accomplished. Moreover, the SIO and USF GSSL curves both depart significantly from the altimetry-GRACE curve after 2020, whereas the SIO and USF TSL curves differ by a smaller amount. This further suggests unresolved salinity errors in a large number of profiles after 2018, even though we have utilized a release of Argo T/S profiles that has added new flags and adjustments to identify such problem floats (Wong et al., 2023). We removed such flagged/adjusted floats from our processing stream (see Section 2.1). However, this does not prevent large and unrealistic global halosteric signals post 2018, so we must conclude that unidentified salinity errors remain in enough profiles to affect the global halosteric signal."

- *Section 3, figure 10: there is no explanation for the 2015 difference between USF and SIO OHC. Please discuss*

Response: We did discuss the 2015 difference in OHC, tying it to the observed difference in TSL (since TSL is proportional to OHC). We have added a new OHC and dOHC/dt curve from the PMEL analysis and have added even more discussion than in the original draft:

New text: "Here, we compare the global OHC anomalies above 2000m computed from our OHC grids with those from computed from the SIO T/S grids as well as those computed by Lyman and Johnson (2023), which use altimetry and satellite sea surface temperature measurements as a reference before combining with Argo (**Fig. 11a**). Before 2015, the three datasets agree well, showing similar overall trends and interannual variability. While the

Lyman and Johnson (2023) grids are specifically formulated to resolve eddy signals, it is clear this has little effect on the global average and that the longwave mapping we use is sufficient.   within estimated errors, except for the periods that include 2015 (**Figure 11**). During 2015 to early 2016, there is disagreement as noted previously (e.g., **Figs. 9 & 10** and associated discussion). The USF OHC appears to be the outlier in 2015, but by early 2016, it agrees with the PMEL OHC, whereas the SIO OHC has a significant drop throughout 2016 until 2017. This further supports the idea that there may be unresolved issues with Argo temperatures in some floats in 2015 and 2016 that warrants further investigation.

After 2020, there is a significant change in behavior between the USF/SIO OHC curves and that from the Lyman and Johnson (2023) analysis. The PMEL analysis shows a steady rise in OHC after 2020, whereas USF and SIO grids indicate more interannual variability, with a drop from 2020-2021, followed by a subsequent rise. It is interesting that the PMEL curve follows the general trend in satellite altimetry over this time (e.g., **Fig. 10**), which suggests that the global OHC from the PMEL analysis may be more dependent on the altimetry reference than our analysis, as we find little to no impact of different references in the global average (**Fig. 7**). Notably, the USF curve post-2020 follows that of SIO, which uses only Argo data (and an Argo-based climatology) in the mapping.

These subtle differences in OHC are reflected in the time-derivative (**Fig. 11b**), which we calculated using running 2-year trends (along with annual and semiannual sinusoids) from the global average OHC (and converting J m^{-2} yr^{-1} to W m^{-2}) – the time stamp used is the middle of each 2-year window and a one-month step was used. Values agree reasonably well before 2020 (noting the small differences in 2015-2017 noted earlier). The mean values for 2005 to 2023 are similar for USF and SIO (USF: 0.58 ± 0.18 W m^{-2}; SIO: 0.54 ± 0.21 W m^{-2}) but are significantly higher for the PMEL OHC derivative (0.96 ± 0.19 W m^{-2}). This is primarily caused by higher values post 2020. In the first half or of the record (2005-2015), the mean PMEL values of dOHC/dt are 0.63 W ± 0.24 W m^{-2}, whereas in the second half (2015-2024), the mean of the PMEL series is 1.36 ± 0.20 W m^{-2}."

- *Section3, figure 10: the OHC time series show clear quasi periodic variation. Could you compute the periodogram and discuss the observed quasi periodicity?*

Response: We feel it is beyond the scope of this manuscript to explain the source of these periodic fluctuations in dOHC/dt. They are seen in all three time series (and in previous studies looking at OHU). This is more of a scientific investigation than a data description, which is the intent of this manuscript and why it was submitted to *Earth System Science Data*.