Thank you very much for your positive comments and constructive feedback, you addressed some important points. Your clarifications helped to make the manuscript clearer for the reader. Our responses are provided in green (changes made in the manuscript are written in ***bold***) together with your original comments in black.

We really appreciate your time and insight in reviewing our manuscript!

Kind regards,
Susanna (on behalf of all co-authors)

Reviewer #1:

Overarching comment:

Heat transport cannot be calculated for cross-sections that do not conserve mass because the results are sensitive to the choice or reference temperature. Either the authors need to show that the lines are mass-conserving (or at least volume-conserving), or use the term “temperature transport” rather than “heat transport” (as in Johns et al. 2011). To this end, it would be good to show the net volume transport for each of the reanalyses. This is shown to some degree in Fig. A3, but a time series of volume transport each month for each product would be enlightening.

We agree that strictly unambiguous heat transports would require closed (or at least volume-conserving) sections and that otherwise the transport depends on the choice of reference temperature. As discussed by Schauer and Beszczynska-Möller (2009) and related studies, that condition is generally not fulfilled for partial sections such as OSNAP East and West.

In line with common practice in the literature (e.g. Tsubouchi et al., 2012, 2018; Muilwijk et al., 2018; Shu et al., 2022; Heuzé et al., 2023), we therefore compute heat transports relative to a fixed reference temperature (θ_{ref} = 0 °C).
We use the term “heat transport” in this conventional, reference-dependent sense rather than “temperature transport”, as the transported quantity **represents the enthalpy of seawater rather than temperature itself**, and to maintain consistency with previous studies (Winkelbauer et al., 2024a; Winkelbauer et al., 2024b).

We state this more clearly in Section 2.1 now (L128 and following):
Additionally, the potential temperature θ and a reference temperature θ_{ref} are needed for estimating heat transports. ***Strictly speaking, unambiguous heat transports would require closed mass transports across the examined section, which is generally not the case for partial sections such as OSNAP East and West and only approximately satisfied for the total oceanic transport (Schauer and Beszczynska-Möller, 2009). Therefore, heat transports depend on the choice of reference temperature. To minimize the ambiguity arising from this choice, θ_{ref} should be chosen to represent the section-mean temperature of the flow across the considered section (e.g. Bacon et al., 2015). However, to ensure internal consistency across products***

we calculate all heat transports relative to a constant reference temperature of θ_ref = 0 °C, following common practice (e.g., Tsubouchi et al., 2012,* Muilwijk et al., 2018; Shu et al., 2022; Heuzé et al., 2023*). Throughout this study, the term "heat transport" is therefore used in this conventional, reference-dependent sense and describes the transport of heat referenced to water at 0°C.

To further clarify the degree of volume imbalance, we now include time series of net volume transport for each reanalysis product in the Appendix:

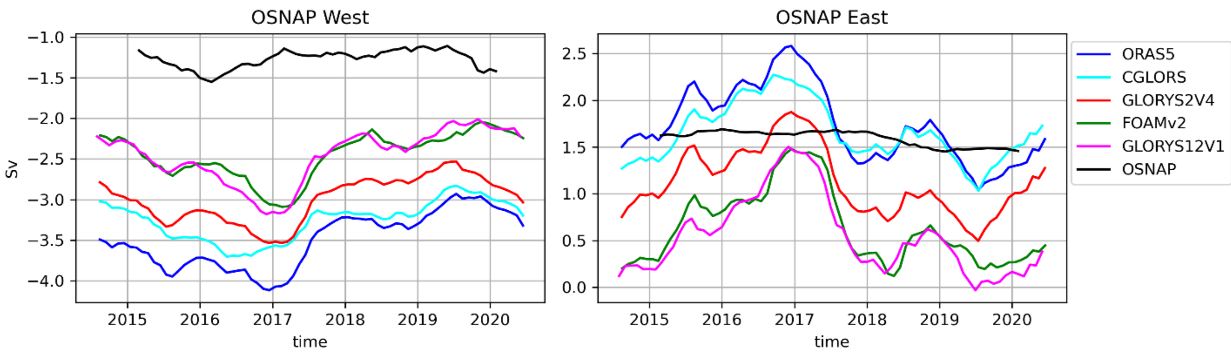

Figure A3. Net volume transport across the eastern and western OSNAP section derived from OSNAP observations (computed from the gridded sections) and for each reanalysis.

We agree that the seen differences in volume transport variability can in principle affect reference-dependent heat transports. To quantify the possible impact of this effect we perform a back-of-the-envelope estimate based on the full volume-transport time series:

$$\Delta OHT_{max}(t) = \rho c_p \Delta V(t) \Delta\theta(t)$$

where $\Delta V(t)$ is the net volume transport across the section and $\Delta\theta$ is the mean temperature bias between OSNAP and the respective reanalysis.

The resulting time series of $\Delta \mathrm{OHT}_{\mathrm{max}}$ is shown in the figure (right Plot) below. The mean value across all reanalyses is approximately **1.2 TW**, more than two orders of magnitude smaller than the mean heat transport (~400 TW) and more than an order of magnitude smaller than the observed 2015 anomaly (~60 TW).

This demonstrates that net volume-transport imbalances and reference-temperature ambiguity cannot explain the 2015 discrepancy, which must instead be dominated by differences in the thermal structure and spatial distribution of the flow.

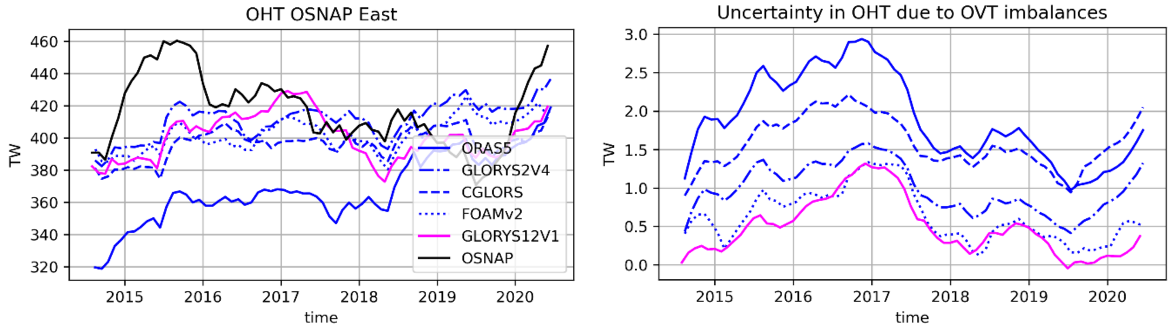

Fig.: *left*: mean OHT at OSNAP East; *right*: uncertainty in OHT due to OVT imbalances estimated via a back-of-envelope approach

We added the following paragraph to the manuscript:

Volume transport time series are shown in Fig. A1. OSNAP-derived net volume transport (computed from the gridded sections) across OSNAP East has a mean comparable to the reanalyses but shows substantially reduced variability. This is expected because OSNAP's velocity reconstruction combines time-varying geostrophic shear with a constrained barotropic component (transport closure/compensation), which dampens section integrated volume-transport variability. In contrast, ocean reanalysis permit time-varying transports in response to atmospheric forcing and freshwater fluxes, leading to larger variability in net volume transport across the open OSNAP section. While the realism of this variability cannot be independently assessed here, its magnitude remains small compared to what would be required to explain the pronounced heat-transport anomaly in 2015. A conservative upper-bound estimate shows that the associated reference-temperature-dependent contribution to heat transport at OSNAP East is of order 1–2 TW on average, which is negligible compared to the mean transport (~400 TW), indicating that differences in the thermal structure and distribution of the flow play a more dominant role in the 2015 anomaly.

Specific comments:

l. 54-55: "As reanalyses generally do not assimilate direct observations of ocean currents, their transport estimates depend largely on model dynamics and parameterizations rather than observational constraints" – this is not entirely true. Ocean reanalyses assimilate SSH and T/S, which together constrain the geostrophic circulation. Most of the AMOC (and resulting MHT) is in geostrophic balance, thus the components of the velocity field that are important to this paper are indeed assimilated. The one exception to this would be the boundary currents, where direct velocity measurements from ADCPs and current meters are indeed not assimilated by the reanalyses. This sentence should be rewritten to convey this information.

We thank the reviewer for this clarification and agree that the original wording was imprecise. While ocean reanalyses do not assimilate direct current measurements, they do assimilate sea level anomalies and temperature/salinity profiles, which together strongly constrain the large-scale geostrophic circulation. We have revised the text to reflect this:

As reanalyses generally do not assimilate direct observations of ocean currents, their transport estimates depend on a combination of model dynamics, parametrizations ***and observational constraints provided indirectly through the assimilation of sea level anomalies and temperature/salinity profiles. Since much of the heat transport associated with the AMOC is in geostrophic balance, these components of the velocity field are indirectly constrained by observations, while limitations remain particularly for boundary currents and narrow passages, where direct velocity measurements are not assimilated.***

l. 83-85: are the vertical cross-sections from GLORYS12V1 re-mapped onto a ¼° grid to be comparable to the other reanalyses? If not, the mean RMSE shown in Figs. 3 and 5 could be aliased by the different spatial resolution.

The GLORYS12V1 sections were not remapped first to a ¼° grid. Instead, all reanalyses are interpolated directly onto the same OSNAP gridded section prior to the calculation of biases and RMSE. This effectively evaluates all products at the OSNAP resolution and smooths higher-resolution features in GLORYS12V1. As the OSNAP grid is similar coarse as the ¼° grid we do not expect aliasing problems. Nevertheless, to assess whether the different native resolutions could bias the RMSE, we additionally tested to remap GLORYS12V1 first to the ¼° grid and then interpolated onto the OSNAP grid. The resulting RMSE values are virtually unchanged, suggesting that the differences in native model resolution do not alias the RMSE values shown in Figs. 3 and 5.

l. 86: "…they differ in their data assimilation methods…" it would be good to clarify what these differences are. A table would be a good way to organize this information.

We added the following Table in section 2.1

Reanalysis	Resolution	Assimilated data	Assimilation scheme
ORAS5 (Zuo et al., 2019)	1/4°	T/S profiles, SLA (50°S-50°N), SST, SIC	NEMOVAR (3D-Var) (Mogensen and Balmaseda, 2012)
CGLORSv7 (Storto and Masina, 2016)	1/4°	T/S profiles, SLA (global, ice-free), SST, SIC	OceanVar (3D-Var) (Dobricic and Pinardi, 2008)
GLORYS2V4 (Garric and L.Parent, 2016)	1/4°	T/S profiles, SLA (global, ice-free), SST, SIC	SAM2 (SEEK, multivariate)
FOAMv2/GloRanV14 (MacLachlan et al., 2015)	1/4°	T/S profiles, SLA (global, ice-free), SST, SIC	NEMOVAR (3D-Var)
GLORYS12V1 (Lellouche et al., 2018)	1/12°	T/S profiles, SLA (global, ice-free), SST, SIC	SAM (SEEK, multivariate)

Table 1. List of used ocean reanalyses. All reanalyses assimilate temperature and salinity profiles and sea level anomalies, which constrain the large-scale geostrophic circulation. None of the products assimilate direct velocity observations from current meters or ADCPs, including those from the OSNAP array.

l. 90: "they can be considered independent of OSNAP in that regard". As mentioned above, though the velocities are not assimilated, much of the OSNAP velocity field is determined from SLA and geostrophy so the only place there is any independence is in the boundary currents. This should be specified.

We rephrased that part to:

All reanalyses assimilate in situ temperature and salinity profiles and SLA, which constrain the large-scale geostrophic circulation, ***but none assimilate direct velocity observations from current meters or ADCPs. In contrast, OSNAP transport estimates are derived from direct, full-depth observations of velocity, temperature, and salinity obtained from moorings, gliders and hydrographic measurements. As a result, OSNAP and the reanalyses differ fundamentally in how ocean velocities, in particular boundary currents, are constrained.***

Fig. 1: what is the mooring in the center of the Labrador Sea?

We removed that mooring from the figure as it is not used in the OSNAP calculation.

l. 133: the reanalyses used in this paper are not volume (or mass) conserving so to which 'conservation properties' are the authors referring?

We clarify that the ocean reanalyses used in this study are indeed volume conserving, even though the net volume transport across an open section such as OSNAP East or West is not required to be zero at monthly timescales.

By "conservation properties" we refer to the numerical consistency of fluxes as represented on the native model grids. Interpolating velocity fields prior to transport calculations can introduce spurious signals and compromise the model-internal conservation of fluxes. We have clarified this point in the revised manuscript:

*To avoid interpolation **and preserve the numerical consistency of fluxes on the native model grids**, net integrated transports from reanalyses are calculated using StraitFlux's line-integration method.*

l. 136: when the heat transports are calculated at monthly time scales, is it calculated from the monthly mean of the heat transport or calculated from the monthly mean velocity and temperature fields? The former accounts for the v'T' term, while the other does not.

In general, heat transports in this study are calculated from monthly mean velocity and temperature fields (for data availability reasons) and therefore do not explicitly include sub-monthly covariance terms (v'T' term). However, the comparison to daily GLORYS12V1 data discussed at line 136 is based on daily vT products and therefore does include the v'T' term. As discussed, the resulting differences between transports computed from daily fields and those derived from monthly mean fields are comparatively small, with negligible impact on temporal variability. This demonstrates that neglecting sub-monthly covariance terms does not substantially affect the results presented in this study. We adapted the wording slightly:

*All transport calculations in this study are based on monthly mean output from the ocean reanalyses. To evaluate the potential influence of temporal resolution, we additionally tested calculations based on daily velocity and temperature fields for GLORYS12V1, **which include sub-monthly covariance terms** **(v'T'**) **that are not explicitly resolved when using monthly mean fields. The resulting differences in integrated heat transport across the OSNAP section amount to about 2 TW on average over the analysis period (corresponding to approximately 0.5% of the mean transport, see Fig. A2), with negligible impact on variability. This indicates that monthly output provides a sufficiently accurate representation for the purposes of this study.***

l. 136: the authors refer to a 0.5% error… is this a percentage of PW? Heat transport has very small variability compared to its mean value. So it would be more clear if the authors just reported a value of heat transport in PW rather than a %.

We have revised the text to report the difference in heat transport in TW, with the percentage given only for reference. See comment above.

l. 157: What is meant by "Mass-consistent heat transport estimates"?

By "mass-consistent heat transport estimates" we refer to heat transports inferred from atmospheric energy budgets that are explicitly constrained to satisfy mass continuity. We have clarified this wording in the revised manuscript:

Heat transport estimates inferred from mass-consistent atmospheric energy budgets (see, e.g., Mayer et al., 2021, Mayer et al., 2024) *are used at two different choke-points: the Greenland–Scotland Ridge (GSR), and the combination of Fram Strait (FS) and the Barents Sea Opening (BSO).*

Table 1: this is an impressive list of data sets. Why was JRA-55 used rather than the updated version (JRA-3Q)?

We have already started producing mass consistent inferred surface heat fluxes based on the newer JRA-3Q reanalysis. However, these JRA-3Q energy budgets are still under consolidation and have not yet been formally published or fully documented. Therefore, we rely on the JRA55-based budgets in this work.

We added the following to the manuscript:

Fs is estimated indirectly from atmospheric budgets, so these are much better constrained by independent observations than parameterized surface fluxes, which typically are more uncertain and depend on the sea state (Mayer et al., 2023; Trenberth et al., 2019). Therefore, divergences and tendencies from atmospheric reanalyses ERA5 (mass-consistent energy budgets, Mayer et al., 2021a), MERRA2 (Gelaro et al., 2017) and JRA55 (Kobayashi et al., 2015) are combined with top-of-atmosphere (TOA) fluxes from CERES-EBAF TOA version 4.2 (Scott et al., 2022; NASA/LARC/SD/ASDC, 2025). ***An updated implementation based on the newer JRA-3Q (Kosaka et al., 2024) reanalysis is currently under development and will be addressed in future work once the corresponding energy budgets are fully consolidated.***

Fig. 3: Consider using a different colorbar to depict RMSE – at first look, this appears as a consistent high bias in the reanalyses compared to OSNAP.

We have revised Fig. 3 and Fig. 5 to use a more neutral colormap for RMSE, which more clearly represents error magnitude without implying a systematic bias.

l. 315 and 404: it is unclear to me whether this 2015 event was captured by OSNAP because there was a glider in that year (and not afterwards), or if this was truly an anomalous event. It would be interesting to analyze an OSNAP gridded section that does not include the glider in 2014-2016. Does the event appear if the glider is not included? Determining whether the event is real or an artifact of changing observational structure would go a long way toward understanding the authors' thoughts in the conclusions about the importance of a consistent set of observations.

We thank the reviewer for raising this important point. We first clarify that the discussion at line 315 refers to a mesoscale eddy observed in the western glider region, which we explicitly state is unrelated to the 2015 heat transport peak. The 2015 anomaly discussed at line 404, by contrast, refers to the basin-scale heat transport maximum at OSNAP East, which we trace primarily to an intensified NAC inflow in the eastern glider region.

We agree that assessing the sensitivity of the OSNAP heat transport estimates to changes in the observing system would be valuable. However, recomputing OSNAP transports without glider data is beyond the scope of this study and not feasible with the publicly available OSNAP gridded product.
To acknowledge this uncertainty, we have added a brief statement to the Discussion:

We cannot exclude that changes in observational coverage, including the use of gliders early in the OSNAP record, may contribute to the observed amplitude of the 2015 heat transport peak. Future sensitivity studies assessing OSNAP transport estimates with and without specific observing components, such as gliders, could help further quantify the impact of observational heterogeneity.

Fig. 11: the units on the x-axis are a bit strange… 2.5-3.3 x 10^6 m… I suggest using km and specifying that this refers to the along-section distance from the OSNAP western boundary.

We changed the units to km.

l. 412-418: In this paragraph, the authors express more confidence in reanalyses products than is justified from the results of this paper. While it is true that the discrepancy in OHT between OSNAP and the reanalyses in 2015 is interesting and raises questions about the coverage and consistency of OSNAP, the authors have not presented any independent evidence that reanalyses can provide error estimates for the observing system (OSNAP in this case). I agree that this is a possible use of ocean reanalyses once they are validated, but the authors would need to present independent data that justify this usage. Given how much the reanalyses disagree with one another (in this paper and in others, e.g. Jackson et al. 2019), I would proceed down this path with a lot of caution – and much more caution than interpreting the direct observations from OSNAP.

We agree that our original wording overstated the degree of confidence that can currently be placed in ocean reanalyses as tools to diagnose errors in observational systems. Our intention was not to suggest that reanalyses can provide quantitative error estimates for OSNAP, but rather that systematic and coherent discrepancies across multiple reanalyses may help highlight regions where both observing systems and models are challenged. We have revised the paragraph to:

While well-maintained mooring lines provide the gold standard for MHT variability estimation, systematic discrepancies between OSNAP and reanalyses ***may be used to find regions of increased uncertainty arising from limitations in both observing systems and models.*** *There are many examples in atmospheric sciences where reanalyses could be used to find and even estimate biases in global observing systems (Hollingsworth et al., 1986; Haimberger et al., 2012).* ***In this sense, our results suggest that present ocean reanalyses can serve as a valuable complementary tool for diagnosing uncertainty.***

Conclusions: the authors could also mention the use of reanalyses to replace the use of moorings in regions where lower frequency variability is dominant. This would save costs and is currently being pursued by the RAPID team (Petit et al., (in review)).

We added a paragraph in the conclusion section:

While well-maintained mooring lines provide the gold standard for MHT variability estimation, systematic discrepancies between OSNAP and reanalyses may be used to find regions of heightened uncertainty arising from limitations in both observing systems and models. There are many examples in atmospheric sciences where reanalyses could be used to find and even estimate biases in global observing systems (Hollingsworth et al., 1986; Haimberger et al., 2012). In this sense, our results suggest that present ocean reanalyses can serve as a valuable complementary tool for diagnosing uncertainty. ***At the same time, in regions where low-frequency variability dominates and where reanalyses demonstrate robust skill, reanalysis products may offer complementary means to extend or support observational estimates (see e.g., Mayer et al., 2023, Fritz et al., 2023). Such approaches are currently being explored within the RAPID program as part of efforts to develop more sustainable and cost-effective long-term observing strategies (Petit et al., 2025).***

References:

Johns et al. (2011): https://doi.org/10.1175/2010JCLI3997.1

Jackson et al. (2019): https://doi.org/10.1029/2019JC015210