https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Estimates of Atlantic meridional heat transport from spatiotemporal fusion of Argo, altimetry and gravimetry data

Francisco M. Calafat[1,2], Parvathi Vallivattathillam[2], Eleanor Frajka-Williams[3]

[1]Physics Department, University of the Balearic Islands, Palma, 07122, Spain
[2]Marine Physics and Ocean Climate, National Oceanography Centre, Liverpool, L3 5DA, UK
[3]Institute of Oceanography, University of Hamburg, 20146, Germany

Correspondence to: Francisco M. Calafat (francisco.mcalafat@uib.es)

Abstract. Understanding how changes in Atlantic meridional heat transport (MHT) and the Earth's climate relate to one another is crucial to our ability to predict the future climate response to anthropogenic forcing. Attaining this understanding requires continuous and accurate records of MHT across the whole Atlantic. While such records can be obtained through direct ocean observing systems, these systems are expensive to install and maintain and thus, in practice, records of MHT derived in this way are restricted to a few latitudes. An alternative approach, based on hydrographic and satellite components of the global ocean observing system, consists of inferring heat transport convergence as a residual from the difference between ocean heat content (OHC) changes and surface heat flux. In its simplest form, this approach derives the OHC from hydrographic observations alone, however these observations are spatially sparse and unevenly distributed, which can introduce significant errors and biases into the MHT estimates. Here, we combine data from hydrography, satellite altimetry and satellite gravimetry through joint spatiotemporal modelling to generate probabilistic estimates of MHT for the period 2004-2020 at 3-month resolution across 12 latitudinal sections of the Atlantic Ocean between 65º N and 35º S. Our approach leverages the higher spatial sampling of the satellite observations to compensate for the sparseness and irregular distribution of the hydrographic data, leading to significantly improved estimates of MHT compared to those derived from hydrographic data alone. The fusion of the various data sets is done using rigorous Bayesian statistical methods that account for the spatial resolution mismatch between data sets and ensure an adequate representation and propagation of uncertainty. Our estimates of MHT at 26º N agree remarkably well with estimates based on direct ocean observations from the RAPID array, in terms of both the magnitude and phase of the variability, with a correlation of 0.77 for quarterly (3-monthly) time series and 0.93 after applying a 5-quarter running mean. The time-mean MHT at 26º N is also captured by our approach, with a value of 1.17 PW [1.04,1.30] (5-95% credible interval). Estimates of MHT at other latitudes are also consistent with what we expect based on earlier estimates as well as on our current understanding of MHT in the Atlantic Ocean.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

1 Introduction

Changes of the Earth's climate since the Industrial Revolution are primarily a result of the excess heat trapped in the climate system by the accumulation of greenhouse gases, leading to global warming. A substantial portion of this extra heat, over 90%, has been absorbed by the world's oceans (Meyssignac et al., 2019; von Schuckmann et al., 2020), temporarily slowing the warming of the atmosphere, albeit at the cost of higher sea levels, accelerated ice sheet melting, and harm to marine ecosystems. However, the ocean's role extends beyond being merely a passive thermal buffer against global warming. It also plays a central role in mediating the climate system response to greenhouse gas emissions by helping to shape – through heat redistribution – the pattern of sea-surface warming, on which climate feedback processes depend (Andrews et al., 2018; Dong et al., 2020). Ocean currents also play a crucial part in regulating the regional climate by transporting heat poleward from the tropics and then releasing it into the atmosphere at higher latitudes (Woollings et al., 2012; Buckley and Marshall, 2016; Zhang et al., 2019; Yin and Zhao, 2021). In the Atlantic Ocean, heat is transported northward throughout the basin by a vast system of ocean currents – the Atlantic meridional overturning circulation (AMOC) (Frajka-Williams et al., 2019) – carrying warm surface waters northward and cold waters southward at deeper levels. Importantly, the AMOC is capable of storing heat (and carbon dioxide) deep in the ocean where it can remain sequestered for centuries before resurfacing thousands of kilometers away. This unique capability endows the AMOC with the potential to affect the global climate over long time scales.

Extensive research efforts in recent decades have been dedicated to monitoring the AMOC and meridional heat transport (MHT) through various multi-observational approaches (Frajka-Williams et al., 2019; Li et al., 2021), leading to significant progress in our understanding of these two critical climate-relevant factors (Srokosz et al., 2021). Nevertheless, despite this progress, significant knowledge gaps persist, such as questions about the latitudinal coherence of the AMOC (and MHT) and whether it is weakening (Jackson et al., 2022; Piecuch & Beal, 2023; Volkov et al., 2024), among other major concerns. Filling these remaining gaps is crucial to advancing our understanding of future climate change. However, ongoing efforts to achieve this are faced with challenges related to limitations in observing capability as well as in the methods currently being used for combining noisy and sparse data from multiple sources. This study is particularly motivated by those limitations and focuses, specifically in the context of quantifying MHT.

Past changes in Atlantic MHT have been estimated mainly through two different approaches. The first approach, employed by both the RAPID/Meridional Overturning Circulation and Heat-flux Array/Western Boundary Time Series (hereafter RAPID) programme (Cunningham et al., 2007; Johns et al., 2011; McCarthy et al., 2015, Johns

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

et al., 2023) and the Overturning in the Subpolar North Atlantic Program (OSNAP) (Lozier et al., 2019; Li et al., 2021), calculates MHT directly by integrating the product of the temperature and cross-sectional velocity across designated transbasin sections (26ºN in RAPID and 50ºN-60ºN in OSNAP). Temperatures and velocities are estimated based on hydrographic data from transport mooring arrays, Argo profiling floats and, in the case of the ONSAP section, also ocean gliders. Heat transport through the Florida Straits in the RAPID section is estimated based on measurements from a submarine cable (Volkov et al. 2024). This approach is widely regarded as the gold standard for monitoring both MHT and the AMOC, but it still comes with limitations. In particular, such an observing array system is time-consuming and expensive to install and maintain, making it impractical for ocean-wide monitoring. Consequently, estimates of MHT based on this approach are restricted to these two latitudes, and thus they are not sufficiently latitudinally dense to characterize the spatiotemporal structure of the MHT.

The second approach, which is the focus of this study, attempts to fill these existing latitudinal gaps in observing. It consists of inferring ocean heat transport convergence (HTC) as a residual from the imbalance between changes in ocean heat content (OHC) and surface heat flux (HF). Here, we shall concern ourselves mainly with the estimation of OHC changes and will rely on estimates of HF derived elsewhere in the literature. However, it is important to emphasize that both components of the energy balance are crucial to the success of this approach in their own right, requiring their own thorough consideration. Most past studies derive OHC changes from gridded temperature (T) and salinity (S) data sets produced through objective analysis of hydrographic profiles from various instruments (e.g., Argo floats, bathythermographs, bottles, etc.) (Roberts et al, 2017; Cheng et al., 2020; von Schuckmann et al., 2020). While such profiles have almost ocean-wide coverage, they are spatially sparse (including the Argo era), almost non-existent below 2000 m, irregularly spaced, noisy and highly heterogenous across instruments in terms of accuracy. These data issues can introduce significant biases and uncertainties into the gridded T/S products and, by propagation, into the estimates of OHC changes, especially on regional scales, restricting the accuracy with which we can estimate MHT through this heat-budget approach. Some studies use T and S data from ocean reanalyses as an alternative to observations (Trenberth and Fasullo, 2017; Trenberth et al., 2019), but such reanalyses have biases and uncertainties of their own and crucially depend on the availability of hydrographic data for assimilation, thus they face similar issues to in-situ observations.

A promising solution to the issues discussed above, with a view to improving the accuracy of OHC estimates, is to combine hydrography-derived thermosteric (TS) and halosteric (HS) heights with sea level (SL) from satellite altimetry and ocean mass (OM) from satellite gravimetry, leveraging the relatively good spatial sampling of the satellite observations. The key idea is to exploit the fact that TS is directly linked to OHC and that SL is related to

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

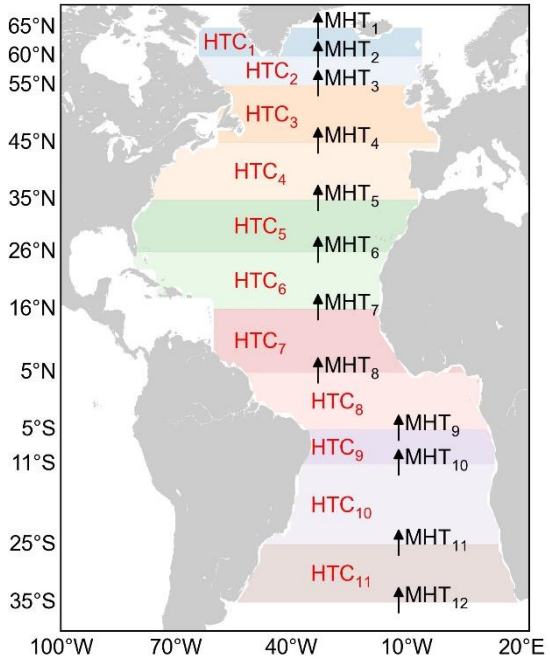
TS through SL=TS+HS+OM. A special case of this approach arises when the focus is on estimating global average OHC because, on global scales, halosteric effects are negligible (Lowe and Gregory, 2006; Gregory et al., 2019) and thus TS (and OHC) can be derived from satellite data alone as the residual of SL and OM (Dieng et al., 2015; Meyssignac et al., 2019; Hakuba et al., 2021; Marti et al., 2022). However, the estimation of MHT requires knowledge of regional OHC changes and, on such scales, halosteric effects can no longer be ignored (Maes, 1998; Wang et al., 2017) and need to be estimated from hydrographic data. In this case, a straightforward combination of the satellite and hydrographic data is direct pointwise subtraction, where the operation TS=SL-OM-HS is performed independently at each grid point. This simple data-merging approach has already shown improvements over estimates of MHT based solely on hydrographic data (Meyssignac et al.; 2024), but it lacks a formal statistical framework to address the complexities of the three data sets. These complexities include data error structures, spatial dependencies and resolution mismatches, among others. For example, the three data sets differ in spatial resolutions, making them inherently incompatible without adjustments – a challenge known as the change of support problem (Gelfand et al., 2001; Gotway and Young, 2002). Moreover, each data set has unique and complex error structures in both time and space. Successfully integrating the three data sets requires accounting for both the resolution mismatch and the error structures within a statistically coherent framework that models all data sets and their associated uncertainties simultaneously and comprehensively.

One of the first attempts to quantify MHT through joint modelling of hydrographic and satellite data was presented in the work of Kelly et al. (2014, 2016). They estimated HTCs by evaluating the heat budget over latitudinally-bounded regions of the Atlantic Ocean based on data from hydrography, altimetry and gravimetry together with reanalysis-derived surface heat fluxes. They used a two-step procedure wherein they first calculated spatial averages over each of the regions independently for each data set and then assessed the heat budget through joint modelling of the spatially averaged time series. While conceptually simple, this procedure has several limitations. By first calculating spatial averages separately for each variable, the procedure ignores any spatial dependencies between the variables and loses the opportunity to leverage cross-variable spatial information, both of which can lead to suboptimal estimates of spatially averaged values. Also, such a modelling choice makes the estimation of uncertainties in the spatially averaged values challenging, often requiring ad-hoc or approximate methods.

Here, we present a Bayesian hierarchical framework (see Cressie & Wikle (2011) for a general description of spatiotemporal hierarchical models) for estimating MHT that combines data from hydrography, altimetry and gravimetry in a statistically rigorous way. Our approach extends that of Kelly et al. (2016) by accounting for spatiotemporal dependencies between processes (i.e., TS, HS, and OM) and enabling information sharing across

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

the various data sets. This is achieved by simultaneous spatiotemporal modelling of the observational fields and their error structures, in contrast to time series modelling of spatially averaged values as done in Kelly et al. (2016). The idea of combining multi-source climatic data through spatiotemporal Bayesian modelling has been successfully used before, for example to assess Antarctic ice mass changes (Zammit-Mangion et al., 2014; Zammit-Mangion et al., 2015) and sea-level trends (Piecuch et al., 2018; Calafat et al., 2022), but to our knowledge has not been applied to MHT and merging hydrography, altimetry and gravimetry. Our approach overcomes the limitations of hydrography-only based analyses and addresses the issues associated with combining data from disparate sources, leading to more robust and accurate estimates of both OHC and MHT. Importantly, by considering error structures jointly, the hierarchical approach provides a coherent way to propagate uncertainty in the data, model and parameters through the analysis to the estimates of the MHT. We use our Bayesian hierarchical model (BHM) to produce observation-based probabilistic estimates of non-seasonal quarterly (3-month-averaged) MHT for the period 2004-2020 across 12 latitudinal sections over the Atlantic Ocean between 65° N and 35° S. The sections are shown in Fig. (1) and have been chosen either because other MHT or AMOC volume transport estimates are available, or because they are interesting oceanographically.



https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Figure 1. Latitude lines across which MHT is estimated using a spatiotemporal BHM. The coloured areas denote the regions over which the heat budgets are evaluated. Both the MHT across each latitude and the HTC in each budget region have been labelled with numbers and this is the notation that we follow in Section 5.1.

2 Data

2.1. Hydrography-derived TS and HS heights

TS and HS heights are calculated using monthly gridded fields of T and S from the ISAS20 product (available at https://www.seanoe.org/data/00412/52367/) (Gaillard et al., 2016), which provides data on a 1/2° × 1/2° grid for the period January 2002 to December 2020 and is based solely on Argo profiles. ISAS20 also provides uncertainty estimates for the objectively analyzed T and S fields. While Argo floats can go as deep as 2000 m, only about 58% of the Atlantic profiles have data below 1900 m on average over the period 2004-2020 (lower than that in the early years of that period and higher in the most recent years). In contrast, about 72% of the profiles reach, on average, a depth of at least 1500 m (see also Wong et al., 2020). For this reason, we decide to use only T and S data from the surface to 1500 m in the calculation of TS and HS (the contribution from below 1500 m is accounted for by inflating the uncertainty in the TS and HS data, as explained later). It is also important to mention that we exclude the Gulf of Mexico and the Caribbean Sea in the evaluation of the heat budgets (Fig. 1) as there appears to be a problem with the hydrographic data from ISAS20 in those regions.

TS and HS changes reflect the expansion and contraction of the water column induced by T and S variations, respectively. Assuming that such variations in T and S are small relative to the time-mean value, TS and HS anomalies can be calculated at each horizontal grid point (latitude – longitude) and for each month as follows (Gill and Niiler, 1973):

$$TS = \int_{-1500}^{0} \alpha T' dz \quad (1)$$

$$HS = -\int_{-1500}^{0} \beta S' dz \quad (2)$$

where α and β are the coefficients of thermal expansion and haline contraction, respectively, and the prime denotes deviations from the time-mean fields (i.e., anomalies). The integration is carried out over the vertical coordinate z.

To obtain uncertainty estimates for TS and HS, we use Monte Carlo simulation to propagate uncertainties in T and S through Eqs. (1) and (2). This procedure involves first generating random profiles of T and S at each horizontal grid point and for each month under the assumption that the errors provided for the gridded data are normally

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

distributed, where we allow for vertically correlated errors. Mathematically, the procedure would be expressed as follows:

$$T^i \sim N(\mu_T, \Sigma_T) \quad (3)$$

$$S^i \sim N(\mu_S, \Sigma_S) \quad (4)$$

where the superscript 'i' denotes a random profile ($i = 1, \ldots, 100$), μ_T and μ_S represent the means of the normal distributions which are set equal to the original gridded T and S values, and Σ_T and Σ_S are covariance matrices that encode the magnitude of the errors and their dependency in the vertical direction. The covariance matrices are specified by assuming an exponential model of the form $\Sigma_{nm} = \sigma_n \sigma_m e^{-d_{nm}/l}$, where d_{nm} is the distance between the n-th and m-th vertical levels, σ_n and σ_m are the error standard deviations provided by the gridded products at the corresponding levels, and l is a vertical decorrelation length scale. We set l equal to 100 m, based on the decorrelation length scales typically used in the objective analysis of Argo data (Good et al., 2013).

Then, for each random profile T^i and S^i we calculate the corresponding TS and HS values by evaluating Eqs. (1) and (2). This yields 100 random samples of TS and HS at each grid point and for each month. The standard deviations of the random samples provide an estimate of the uncertainties associated with the TS and HS fields. To account for the contribution of waters below 1500 m to TS and HS we inflate the estimated uncertainty by 20%.

2.2. SL from satellite altimetry

The altimetry sea-level data are from the gridded sea surface height product (based on a stable two-satellite constellation) (SEALEVEL_GLO_PHY_CLIMATE_L4_MY_008_057) produced and distributed by the Copernicus Climate Change Service (C3S). These data are available at http://marine.copernicus.eu/ and are provided as daily fields on a 1/4° × 1/4° near global grid. For this study, C3S refers to the period from January 2004 to December 2020. The C3S data are provided with all standard corrections applied, including corrections for tropospheric (wet and dry) and ionospheric path delays, sea state bias, tides (solid earth, ocean, loading, and pole), and barotropic atmospheric effects (wind and atmospheric pressure for periods<20 days and inverse barometer effects for longer periods). We also adjust the sea-level fields for glacial isostatic adjustment (GIA) using the estimates derived by Frederikse et al. (2020) as well as for deformation effects on the sea floor (+0.1 mm yr^{-1}, spatially uniform) due to contemporary mass changes of the Greenland and Antarctic ice sheets, glaciers, and terrestrial water storage (Frederikse et al., 2017).

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

The altimetry data are affected by several sources of uncertainty (Prandi et al., 2021), including mapping errors, high-frequency errors arising from both orbit determination and any of the geophysical corrections mentioned above, low-frequency errors associated with the wet tropospheric correction, drift errors from orbit determination, inter-mission biases (from Jason-1 to Jason-2 and from Jason-2 to Jason-3), and errors from the GIA adjustment. We account for all these error source contributions. The mapping errors are provided by C3S as part of the gridded product whereas estimates for the other error sources are from Prandi et al. (2021) (available at https://www.seanoe.org/data/00637/74862/). Details of how these uncertainties are taken into account in the data fusion analysis are given in Section 3.

2.3. OM from satellite gravimetry

The OM data are based on measurements collected by the Gravity Recovery and Climate Experiment (GRACE). Here, we use the global time-variable gravity mascon solution (RL06v2.0) from the NASA Goddard Space Flight Center (GSFC) (Loomis et al., 2019), available at https://earth.gsfc.nasa.gov/geo/data/grace-mascons. The data span the period from April 2003 to December 2021 and are provided in the form of monthly mascons on an equal-area 1° × 1° (at the equator) grid. Non-tidal ocean bottom pressure variations (GAD product) with their global ocean mean removed have been restored and a GIA correction has been applied, ensuring that the resulting OM data are comparable to the residual of altimetric SL and hydrography-derived TS+HS. The product supplies uncertainty estimates at each mascon, accounting for both serially uncorrelated errors due to leakage and stochastic noise and for leakage trends.

2.4. Surface heat flux

We use surface HF data from three products: 1) the Diagnosing Earth's Energy Pathways in the Climate system (DEEP-C, version 5) project (https://researchdata.reading.ac.uk/347/) (Liu et al., 2015; Liu and Allan, 2022), which provides monthly mean data on a $0.7° \times 0.7°$ global grid for the period; 2) the ERA5 reanalysis (https://cds.climate.copernicus.eu) (Hersbach et al., 2023), which provides monthly means on a $1/4° \times 1/4°$ grid; and 3) the NCEP/NCAR reanalysis (https://psl.noaa.gov/data/reanalysis/reanalysis.shtml) (Kalnay, 1996), which provides monthly mean data on a $2° \times 2°$ global grid. DEEP-C is an observation-based and dedicated product that has been developed specifically for a robust estimation of the surface HF and thus, in principle, it is preferable to the reanalysis-based products, however, it only provides data up to November 2017. By including surface HF data from reanalyses, we can obtain estimates of MHT for the period 2004-2020. The reanalysis-based HF across the air-sea interface is calculated as the sum of radiative and turbulent fluxes from the reanalysis. None of the HF

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

products provides uncertainty estimates, thus we derive a measure of uncertainty from the spread across the three products as follows. In comparing the three HF products we note that the variability across the various products is very similar, but differences in the time-mean HF can be larger. For this reason, we obtain uncertainty estimates separately for the HF variability and the time-mean HF. The uncertainties associated with the HF variability are estimated as the standard deviation (SD) of the HF anomalies (with respect to the time mean) over the three products at each time step whereas those for the time-mean HF are computed as the SD of the time means across the three products. Note that for the period beyond November 2017, uncertainties are based solely on data from ERA5 and NCEP.

2.5. Resolving the spatial resolution mismatch

Here we aim to combine three data sets that are incompatible in terms of spatial resolution. Both the hydrographic data and the GRACE data are provided on relatively fine grids, but their effective spatial resolution is much lower than what such grids imply. In particular, monthly GRACE data can be regarded as spatial averages over a ~300 km footprint (Tapley et al., 2019). The gridded hydrographic product is based on Argo profiles. Although such profiles represent point-level measurements, they are sparsely collected over the ocean with an average spatial separation of about 3 degrees (in any given month). This separation sets a practical limit to the smallest features that can be resolved (on average) by the Argo data. Additionally, the Argo-derived gridded data incorporates some degree of spatial smoothing due to the interpolation process, which is largely determined by the decorrelation lengths scales used by the gridding methods. Such length scales are typically on the order of 300 km (Good et al., 2013), except within a few degrees of the equator where they are significantly larger. The effective resolution of the gridded altimetry data is higher than that of the GRACE and hydrographic data, although it varies with latitude. The decorrelation length scales used in the mapping of the altimetry data, which we take as a rough measure of spatial resolution, range from ~350 km in a low-latitude band (±15^{o}N) to ~150 km poleward of this band (Pujol et al., 2016). In considering ways of resolving the resolution mismatch problem, it is important to remember that, here, we aim to assess heat budgets over large ocean regions and this only requires spatial averages of the relevant variables over such regions (i.e., we do not need point-level estimates). In this context, the simplest solution to the spatial support problem is to aggregate each of the three data sets into areal units (i.e., grid cells) of a size similar to the resolution of the coarsest resolution data. In this way, the aggregated data will all have the same level of spatial resolution and can then be combined using areal spatiotemporal modelling (Banerjee et al., 2014). In this study, we adopt this approach.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

We set the size of the target areal units to $3° \times 3°$ as this aligns with the decorrelation length scales used in the mapping of the hydrographic data as well as with the resolution of the GRACE data. For all three data sets, we aggregate the data by averaging the original grid cells over the target areal units through proportional area weighting (i.e., weights are computed as the proportion of the target areal unit that lies within each source grid cell).

The uncertainties in the aggregated data are computed by propagating the source data uncertainties through the spatial averaging process using standard error propagation formulae. In particular, the error variance of the aggregated data at any given target cell is calculated as:

$$\sigma^2 = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j \neq i}^{n} w_i w_j \rho_{ij} \sigma_i \sigma_j \quad (5)$$

where w_i is the weight assigned to the i-th source grid cell in the aggregation process, σ_i is the standard error of the source data at the i-th grid cell, and ρ_{ij} is the error correlation between the i-th and j-th source grid cells. Note that the correlations ρ_{ij} are not known exactly and thus they need to be approximated somehow. Here, we assume a spatial correlation function of the form $\rho_{ij} = e^{-d_{ij}/l}$, where d_{ij} is the geodesic distance between the centroids of the i-th and j-th source grid cells, and l denotes a decorrelation length scale. For the GRACE data, we assume that errors are perfectly correlated in space. For the TS and HS data, we set l=300 km at latitudes poleward of 5° and l=600 km otherwise (based on the decorrelation length scales used in the mapping process). Finally, for the altimetry data, we use different length scales depending on the type of error (see Section 2.2). For the interpolation errors, we set l=150 km at latitudes poleward of 15° and l=350 km otherwise (based on the length scales used in the original gridding of the data). The high-frequency errors are assumed to be spatially uncorrelated. For the errors associated with the wet tropospheric correction and the GIA adjustment, we set l=300 km. The drift errors and inter-mission biases are assumed to be perfectly correlated in space (over the length scales on which the spatial aggregation takes place).

2.6. Considerations on data temporal resolution

Although the data described above are available at monthly resolution, their effective temporal resolution is somewhat lower due to the relatively low sampling rate of the observations on which the gridded products are based. For example, most satellite altimeters have a repeat cycle that ranges from 10 days (e.g., Jason satellites) to 35 days (e.g., Envisat satellite), whereas GRACE has a repeat cycle of 30 days. Argo floats provide a vertical profile once every 10 days. In practice, this means that the month-to-month variability in the gridded products exhibits significant levels of noise. Such noise at high frequencies can be greatly amplified by the time derivatives involved in the calculation of MHT (the amplification factor is proportional to frequency) and, in turn, can corrupt

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

our estimates of MHT. To minimize this issue, we trade off some temporal resolution for a significant reduction in noise by converting the monthly data to 3-month averages (i.e., to quarterly data: Jan-Feb-Mar, and so on).

The uncertainties associated with the quarterly data are calculated by propagating the uncertainties in the monthly data using Eq. (5), where now the sub-indices i and j refer to the i-th and j-th months, respectively. For the uncertainties in TS and HS as well as for the high-frequency errors in the GRACE and altimetry data (and the errors in the wet tropospheric correction) we set $w_i = 1/3$ and assume serially uncorrelated errors ($\rho_{ij} = 0$). In contrast, the leakage trends in GRACE are, by definition, perfectly correlated in time and thus such errors are not reduced by the temporal averaging (their value remains the same). Similarly, for altimetry, the drift errors, inter-mission biases and GIA errors are all modelled as a linear trend error and thus they are also perfectly correlated in time.

3. Bayesian hierarchical framework

Here, we develop a Bayesian hierarchical framework that integrates observations from altimetry, GRACE, and hydrography together with surface HF data to estimate HTC from ocean heat budgets over a set of latitudinally-bounded regions (see Fig.1). A schematic of the Bayesian model architecture is shown in Figure 2. The model is composed of two distinct but interconnected parts: one that models the spatiotemporal evolution of the sea-level components (TS, HS, and OM) based on the 3° × 3° areal units of quarterly (3-month-averaged) observations from hydrography, altimetry and GRACE; and another part that determines non-seasonal quarterly HTC as the residual of OHC tendency (derived using estimates of TS from the first component of the model) less surface HF over selected regions bounded by latitude lines (Fig. 1). The heat budgets are evaluated over all regions simultaneously, allowing for, but not enforcing, correlation in HTC between regions. It is important to note that the two parts described above are components of the same Bayesian model, i.e., we compute a single posterior distribution from which we derive expectation values (means and quantiles) for all the quantities of interest. We perform three different fits based on the three surface HF products (DEEP-C, ERA5 and NCEP), which hereafter we refer to, respectively, as BHM1, BHM2, BHM3. The solution BHM1 spans the period from the first quarter of 2004 to the second quarter of 2017, whereas the other two solutions cover the period from the first quarter of 2004 to the last quarter of December 2020.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

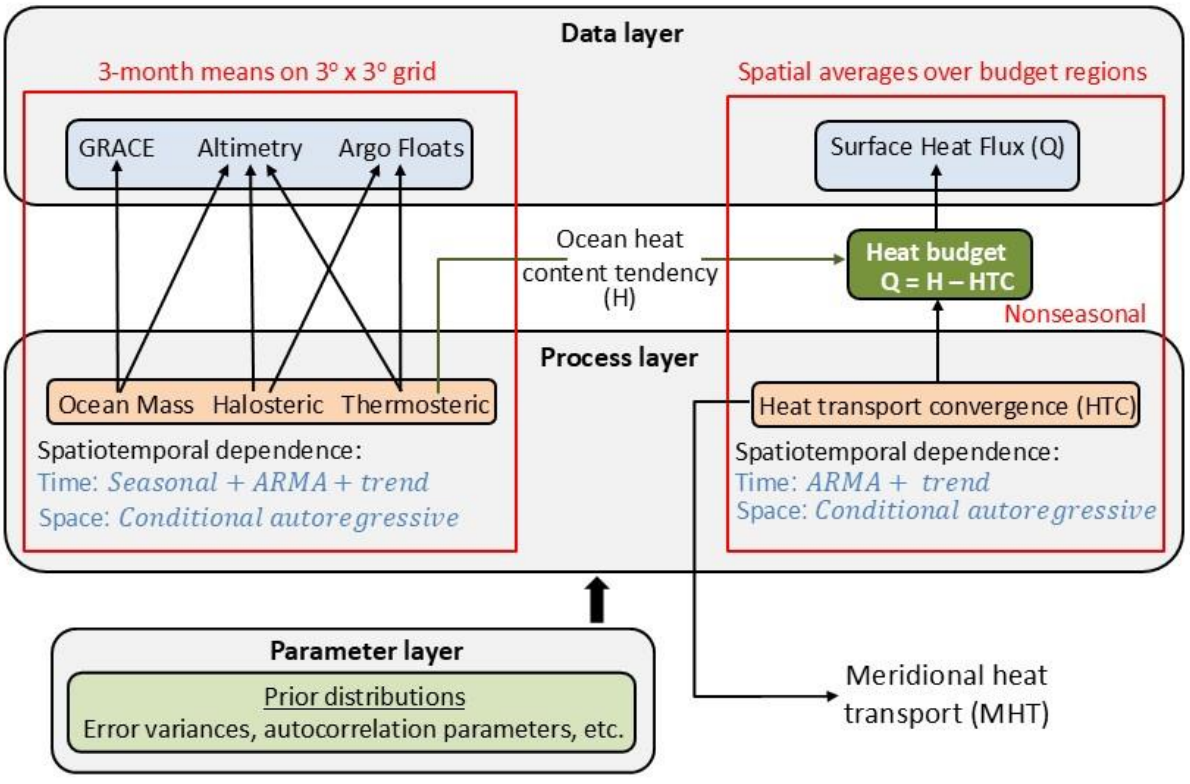


Figure 2. Schematic of the Bayesian hierarchical framework used to estimate MHT.

Our approach centers on the notion that the data, being observed with error, can be regarded as the sum of a true latent spatiotemporal process plus observation error. For example, the GRACE data can be viewed as noisy observations of the OM latent process, and similarly for all the other data sets. Here, the term process refers to a stochastic process (i.e., a collection of random variables indexed by either time, space or both), whereas the term latent process means that the process is not directly measured but inferred from noisy observations. The aim is then to separate the various latent processes (TS, HS, OM, and HTC) from the observational noise through joint statistical modelling of all the available data (i.e., hydrography, altimetry, GRACE and surface HF). By conducting joint modelling and accounting for spatial dependencies, we allow for information exchange both between observational data sets and across locations, improving the estimation of the true underlying process. The success of this approach relies on the careful representation of uncertainty, not least because the model uses uncertainty to decide which observations should have more influence. Furthermore, we need to consider not only uncertainty associated with measurement noise, but also uncertainty arising from limited knowledge of the latent process as

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

well as uncertainty in unknown parameters of the BHM such as decorrelation length scales and error variances. This problem can be conveniently formulated within a Bayesian hierarchical framework.

Here, we develop a BHM with three layers (each layer takes the form of a probability model): 1) a data model that describes the distribution of the observations (SL, TS, HS, OM and HF) conditional on the latent processes; 2) a process model that describes the spatiotemporal evolution of the latent processes conditional on a set of parameters; and 3) a parameter model that describes the uncertainty in the model parameters and encodes any prior information that we have about the data and the processes. In our BHM, each component of sea level (i.e., TS, HS and OM) is modelled as the sum of three contributions, namely seasonal changes, variability, and a linear trend. The non-seasonal HTC is modelled as the combination of variability, a linear trend and an intercept. Next, we describe each layer of the BHM, starting with the data layer.

3.1. Observation layer.

Let $B = \{B_i \mid i = 1, \dots, N\}$ denote the set of $3° \times 3°$ cells on which the observations have been aggregated, and $z_{p,t}(B_i)$ denote an observation on the *i*-th cell at time step *t*, for $p \in P$ where $P = \{SL, TS, HS, OM\}$. The data layer of the BHM for the sea-level observations can then be written as:

$$z_{p,t}(B_i) = y_{p,t}(B_i) + a_p(B_i) + b_p(B_i)t\mathbf{1}_{\{SL,OM\}}(p) + \hat{b}_p(B_i)\hat{t}\mathbf{1}_{\{OM\}}(p) + e_{p,t}(B_i), \quad i = 1, \dots, N \quad t = 1, \dots, T \tag{6}$$

where $y_{p,t}(B_i)$ denote the true latent process of interest. The term $a_p(B_i)$ represents data offsets, modelled as independent and identically distributed (i.i.d.) normal random variables with variance of 1 m^2, $a_p(B_i) \overset{iid}{\sim} N(0,1)$. Such offsets are introduced to cater for the fact that the sea-level height data are referenced to different vertical datums. The variables $b_p(B_i)$ denote data error trends, modelled as i.i.d. normal random variables, $b_p(B_i) \overset{iid}{\sim} N(0, \gamma_p^2(B_i))$. These include orbit errors, inter-mission biases and GIA uncertainties in the altimetry data, and leakage trends in GRACE. The factor $\mathbf{1}_A(x)$ is an indicator function such that $\mathbf{1}_A(x) = 1$ if $x \in A$, and $\mathbf{1}_A(x) = 0$ otherwise. The variables $\hat{b}_{OM}(B_i)$ are introduced to account for a gradual increase in instrumental noise during the final years of the GRACE mission and are modelled in the process layer. The variable $\hat{t}$ is defined such that $\hat{t} = t$ if t corresponds to the years between (and including) 2012 and 2017, and $\hat{t} = 0$ otherwise. Finally, $e_{p,t}(B_i)$ are assumed to be serially and spatially uncorrelated observation errors, $e_{p,t}(B_i) \overset{ind}{\sim} N(0, \sigma_{p,t}^2(B_i))$. The variances $\gamma_p^2(B_i)$ and $\sigma_{p,t}^2(B_i)$ are specified based on the data uncertainties calculated as described in Section 2.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Next, we define the data model for the surface HF data. Let $R = \{R_j \mid j = 1, \dots, M\}$ denote the set of regions over which the heat budgets are evaluated, and $Q_t(R_j)$ be the non-seasonal surface HF into the ocean at time t spatially averaged over R_j. We express $Q_t(R_j)$ as the sum of its time-mean value, $\langle Q_t(R_j)\rangle$, plus a fluctuating part, $Q'_t(R_j)$, as this allows us to model uncertainty in both components explicitly. With this, the observation layer for the HF data is written as a heat budget:

$$Q'_t(R_j) = H'_t(R_j) - U'_t(R_j) + v_t(R_j), \qquad j = 1, \dots, M \tag{7}$$

$$\langle Q_t(R_j)\rangle = \langle H_t(R_j) - U_t(R_j)\rangle + q(R_j), \qquad j = 1, \dots, M \tag{8}$$

where $H_t(R_j)$ is the non-seasonal OHC tendency spatially averaged over R_j, $U_t(R_j)$ is the HTC through the lateral boundaries of R_j, and $v_t(R_j)$ and $q(R_j)$ are serially and spatially uncorrelated observation errors, $v_t(R_j) \overset{ind}{\sim} N(0, \eta_t^2(R_j))$ and $q(R_j) \overset{ind}{\sim} N(0, \delta^2(R_j))$. The SDs $\eta_t(R_j)$ and $\delta(R_j)$ are set equal to the SD of the HF over the three HF products (DEEP-C, ERA5 and NCEP) for the anomalies (at each time step) and the time mean, respectively. The prime $'$ denotes deviation with respect to the time mean.

3.2. Process layer.

Here, we describe the process layer of the BHM. The degree to which the noise increases during the final years of the GRACE mission (see Equation 6) appears to be similar for cells that are close to one another. To capture this spatial dependency, the variables $\hat{b}_{OM}(B_i)$ are assumed to follow a spatial conditional autoregressive (CAR) process (Cressie and Wikle, 2011):

$$\hat{b}_{OM}(B_i) \sim MVN(0, (I_N - \alpha_b K)^{-1}\tau_b^2 D) \tag{9}$$

where MVN denotes a multivariate normal distribution, I_N is the identity matrix of size N, α_b is a parameter that controls the degree of spatial autocorrelation (to be estimated), K is the adjacency matrix ($k_{ii} = 0$, $k_{ij} = 1$ if B_i and B_j are neighbors, and $k_{ij} = 0$ otherwise), τ_b is a SD parameter (to be estimated), and $D = diag(1/n_i)$ is an $N \times N$ diagonal matrix with n_i equal to the number of neighbors of the *i*-th grid cell. Two cells are considered to be neighbors if the distance between their centroids is no larger than two times the size of the cells (i.e., 6 degrees). CAR models are classes of Markov random fields commonly used to describe spatial autocorrelation in areally-aggregated data.

Seasonal changes in the sea-level components are modelled as:

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

$$x_{p,t}^{\text{Seas}}(B_i) = a_{1,p}(B_i)\cos(\omega t) + a_{2,p}(B_i)\sin(\omega t) \qquad p \in P\backslash\{SL\} \tag{10}$$

where ω is the angular velocity of the annual cycle, and $a_{k,p}(B_i)$, for $k = 1,2$, are CAR processes:

$$a_{k,p}(B_i) \sim MVN\left(0, \left(I_N - \alpha_{a,p}K\right)^{-1}\tau_{a,p}^2 D\right) \tag{11}$$

The non-seasonal variability in TS, HS and OM is modelled as spatial fields that evolve through time according to a first-order autoregressive moving average (ARMA) process. Here, we note that the TS and HS components of sea level tend to covary inversely. To capture this spatiotemporal interaction between TS and HS, we use the method of coregionalization (Gelfand et al., 2004), which assumes that the interaction is local. With this, the model for the non-seasonal variability can be written as:

$$x_{p,t}^{\text{Var}}(B_i) = \rho_p x_{p,t-1}^{\text{Var}}(B_i) + \theta_p m_{p,t-1}(B_i) + m_{p,t}(B_i) - \psi m_{TS,t}(B_i)\mathbf{1}_{\{HS\}}(p) \qquad p \in P\backslash\{SL\} \tag{12}$$

where ρ_p and θ_p, for $p \in P\backslash\{SL\}$, are spatially constant coefficients that control, respectively, the degree of temporal autocorrelation and past-noise dependence, ψ is a positive parameter (to be estimated) that determines the strength of the interaction between the TS and HS fields, and $m_{p,t}(B_i)$ are CAR processes, $m_{p,t}(B_i) \sim MVN\left(0, \left(I_N - \alpha_{m,p}K\right)^{-1}\tau_{m,p}^2 D\right)$.

The linear trends in the latent sea-level processes are modelled as spatial fields, where again we capture the interaction between the TS and HS fields through coregionalization:

$$x_{p,t}^{\text{Trend}}(B_i) = \left(g_p(B_i) - \phi g_{TS}(B_i)\mathbf{1}_{\{HS\}}(p)\right)t, \qquad p \in P\backslash\{SL\} \tag{13}$$

where ϕ is an interaction parameter (to be estimated), and $g_p(B_i)$ are CAR processes, $g_p(B_i) \sim MVN(\mu_{g,p}, (I_N - \alpha_{g,p}K)^{-1}\tau_{g,p}^2 D)$, for $p \in P\backslash\{SL\}$. The mean of the CAR process, $\mu_{g,p}$, is set to 0 for $p \in \{TS, HS\}$ and to 2 mm yr^{-1} for $p \in \{OM\}$ based on the spatially averaged trends from the observational data (computed using ordinary least squares).

With all the contributions now defined, the true latent process for each sea-level component is given by:

$$y_{p,t}(B_i) = x_{p,t}^{\text{Seas}}(B_i) + x_{p,t}^{\text{Var}}(B_i) + x_{p,t}^{\text{Trend}}(B_i), \qquad p \in P\backslash\{SL\} \tag{14}$$

$$y_{SL,t}(B_i) = y_{TS,t}(B_i) + y_{HS,t}(B_i) + y_{OM,t}(B_i) \tag{15}$$

The non-seasonal OHC tendency spatially averaged over R_j is computed as follows. Let $y_{TS,t}^{Des}(R_j)$ be the deseasonalized TS at time t spatially averaged over R_j, computed as:

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

$$y_{TS,t}^{Des}(R_j) = \frac{\sum_{i=1}^{N} w_{ij}\left(x_{TS,t}^{\mathrm{Var}}(B_i) + x_{TS,t}^{\mathrm{Trend}}(B_i)\right)}{\sum_{i=1}^{N} w_{ij}}, \qquad j = 1, \dots, M \tag{16}$$

where the weights w_{ij} are computed as the proportion of the region R_j area that lies within the grid cell B_i. The non-seasonal OHC tendency is then calculated using central differences as:

$$H_t(R_j) = \frac{\rho_0 c(R_j)}{\alpha(R_j)}\left(\frac{y_{TS,t+1}^{Des}(R_j) - y_{TS,t-1}^{Des}(R_j)}{2}\right), \qquad j = 1, \dots, M \tag{17}$$

where ρ_0 is a reference seawater density, $c(R_j)$ is the time-mean heat capacity of seawater spatially averaged over R_j, and $\alpha(R_j)$ is the time-mean coefficient of thermal expansion spatially averaged over R_j and vertically averaged over the top 1500 m of the ocean (averaging the coefficient of thermal expansion over different water depths has almost no effect on the phase of the estimated HTC variability and only a small effect on its amplitude).

Finally, the HTC through the lateral boundaries of R_j, $U_t(R_j)$, is modelled as the sum of three contributions, namely variability, a linear trend, and an intercept. The variability in $U_t(R_j)$ is allowed to be correlated across regions, and thus we modeled it as a spatial CAR process that evolves through time according to a first-order ARMA process:

$$u_t^{\mathrm{Var}}(R_j) = \rho_U u_{t-1}^{\mathrm{Var}}(R_j) + \theta_U f_{t-1}(R_j) + f_t(R_j) \tag{18}$$

where ρ_U and θ_U are spatially constant coefficients that control, respectively, the degree of autocorrelation and past-noise dependence, and $f_t(R_j)$ is a spatial CAR process $f_t(R_j) \sim MVN(0, (I_M - \alpha_U L)^{-1}\tau_U^2 D_U)$. Here, L is the $M \times M$ adjacency matrix, and $D_U = diag(1/n_j)$ is an $M \times M$ diagonal matrix with n_j equal to the number of neighbors of the j-th region. In this case, two regions are considered to be neighbors if they share a common border.

Both the linear trend and the intercept are modelled as i.i.d. normal random variables with SD of 1.3 W m^{-2} yr^{-1} and 127 W m^{-2}, $u^{\mathrm{Trend}}(R_j) \overset{iid}{\sim} N(0, 1.3^2)$ and $u^{\mathrm{Intc}}(R_j) \overset{iid}{\sim} N(0, 127^2)$.

With this, the HTC is calculated as:

$$U_t(R_j) = u_t^{\mathrm{Var}}(R_j) + t u^{\mathrm{Trend}}(R_j) + u^{\mathrm{Intc}}(R_j)\,. \tag{19}$$

3.3. Parameter layer

The BHM is completed by defining the parameter layer. The prior distributions ascribed to the hyperparameter of the BHM are summarized in Table 1. Here, we provide justification for the informative the priors. First, the priors assigned to the ARMA parameters (ρ_* and θ_*) ensure that the ARMA processes are stationary and invertible by enforcing the following conditions: $0 < \rho_* < 1$ and $|\theta_*| < 1$. The constraint applied to the autocorrelation

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

parameters of the CAR processes, $\alpha_* < \lambda_{max}^{-1}$, is necessary to ensure that the covariance matrix of the CAR processes is positive definite. The choice of the location parameter for the truncated normal distributions assigned to the SDs of the CAR trend processes (i.e., $\tau_{g,TS}$, $\tau_{g,HS}$ and $\tau_{g,OM}$) is based on the SD of the trend fields derived from the observations. Finally, the interaction parameters (ψ and ϕ) are assumed to be positive since the TS and HS components tend to be anticorrelated, with values likely ranging from 0 to 1.

Table 1. Prior distributions ascribed to the hyperparameters of the BHM. The constant λ_{max} is the maximum eigenvalue of the corresponding adjacency matrix (K for $\alpha_{a,p}$, $\alpha_{m,p}$, $\alpha_{g,p}$ and α_b, and L for α_U). The notation half-$\mathcal{N}$ denotes a half normal distribution whereas trunc-$\mathcal{N}$(,)[a,b] represents a truncated normal distribution with support in the interval [a,b].

Parameter	Units	Description	Prior distribution
ρ_p $p \in \{TS, HS, OM, U\}$	-	ARMA autocorrelation	Uniform(0,1)
θ_p $p \in \{TS, HS, OM, U\}$	-	ARMA past noise	Uniform$(-1,1)$
$\alpha_{a,p}, \alpha_{m,p}, \alpha_{g,p}, \alpha_b, \alpha_U$ $p \in \{TS, HS, OM\}$	-	CAR autocorrelation	Uniform$(0,\lambda_{max}^{-1})$
$\tau_{a,p}, \tau_{m,p}$ $p \in \{TS, HS, OM\}$	m	CAR standard deviation	half-$\mathcal{N}(0,1)$
τ_b	mm yr^{-1}	CAR standard deviation	half-$\mathcal{N}(0,10^2)$
$\tau_{g,TS}$	mm yr^{-1}	CAR standard deviation	trunc-$\mathcal{N}(3.5,1)[0,\infty]$
$\tau_{g,HS}$	mm yr^{-1}	CAR standard deviation	trunc-$\mathcal{N}(2.0,0.3^2)[0,\infty]$
$\tau_{g,OM}$	mm yr^{-1}	CAR standard deviation	trunc-$\mathcal{N}(1.3,0.3^2)[0,\infty]$
τ_U	W m^{-2}	CAR standard deviation	half-$\mathcal{N}(0,127^2)$
ψ, ϕ	-	Interaction TS-HS	half-$\mathcal{N}(0,1)$

3.4. Inference

Inference in the BHM is accomplished by numerically sampling from the posterior distribution of the processes and parameters given the observational data using the No-U-Turn Sampler (NUTS) as implemented by the Numpyro probabilistic programming language (Phan et al., 2019). NUTS is a Markov chain Monte Carlo (MCMC) method that uses Hamiltonian dynamics to enable rapid exploration of the posterior distribution. We run the

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

sampler with four chains of 1000 iterations each (warm-up=1000) for a total of 4000 post-warm-up draws. Such draws represent samples from the posterior distribution.

4. BHM evaluation

4.1. MCMC diagnostics

We begin the evaluation of the model by assessing how accurately the MCMC samples characterize the posterior distribution. To this aim, we use a number of MCMC diagnostics that are designed to diagnose problems with the sampler and assess convergence and mixing. In this context, convergence means that the Markov chains have reached a stationary distribution that coincides with the true posterior distribution, whereas mixing refers to the number of iterations required for a Markov chain to approximate the posterior distribution adequately. While there are no definitive tests of convergence, we can use various diagnostics to determine whether Markov chains appear to have converged. One of such diagnostics is the potential scale reduction statistic (Gelman and Rubin, 1992), $\hat{R}$, which checks whether Markov chains initialized from different values have the same distribution (a necessary, although insufficient, condition for convergence). Mathematically, $\hat{R}$ compares the sample variances both within and between Markov chains. When all the Markov chains have converged to the same distribution, values of $\hat{R}$ should be close to 1 for all model parameters. While there is no universally accepted cut-off point for convergence based on $\hat{R}$, values of $\hat{R}$ >1.2 are typically considered to be suggestive of non-convergence. In addition to convergence, evaluating the mixing of the Markov chains is also important as this can be poor in complex models due to high autocorrelation of the MCMC samples. The higher the autocorrelation the larger the MCMC standard error (given a fixed number of iterations), and thus the larger the error of the estimates derived from the posterior samples. As a measure of mixing and autocorrelation, we use estimates of the effective sample size, n_{eff}, for each hyperparameter (Gelman et al., 2014). In general, a value of n_{eff} per iteration <0.001 is indicative of poorly mixing chains and suggestive of possible biased estimates.

We find $\hat{R}$ to be <1.2 for all hyperparameters in the BHM, suggesting that the Markov chains have converged to the equilibrium distribution and are providing a good approximation to the posterior distribution. Additionally, the n_{eff} per iteration is >0.001 for all hyperparameters with an average value ranging from 0.21 to 0.31 (depending on the BHM fit), indicating low autocorrelation and good mixing.

We use additional diagnostic tools, specific to Hamiltonian Monte Carlo, that offer information about the ability of the NUTS sampler to explore the posterior distribution. These tools include divergent transitions and tree-depth saturation. The presence of divergences would indicate that the sampler has run into regions of challenging

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

posterior geometry that it is unable to explore well, whereas the appearance of tree-depth saturations would indicate that the sampler is terminating prematurely to avoid excessively long execution time thus leaving regions of the posterior distribution unexplored. We confirm that none of the iterations in the three model fits (BHM1, BHM2 and BHM3) showed any divergent transitions or tree-depth saturations, giving us additional confidence in the reliability of our Bayesian solutions.

4.2. Goodness of fit

Here, we evaluate the performance of our Bayesian model by examining its ability to accurately fit the observational data. We begin by examining the residuals from the observation model (Equation 6), focusing on the SL process as this integrates all the sea-level components (i.e., $z_{SL,t} - y_{SL,t} - a_{SL} - b_{SL}t$). The observation model assumes that the residuals are normally distributed, and thus gross violations of this assumption would signal the inadequacy of the model to describe the underlying structure of the data. It is, therefore, important to confirm the appropriateness of the normality assumption. Here, we do this by testing the null hypothesis that the residuals conform to a normal distribution using the Anderson-Darling test (Anderson & Darling, 1954). We apply the test separately to the time series of residuals at each grid cell and for each iteration. The test fails to reject the null hypothesis (5% significance level) in about 95% of the cases (4000 iterations x 641 cells) for all three BHM fits, confirming the validity of the normality assumption. We also verify that there are no systematic departures between the Bayesian solutions and the observations. In particular, we find that the time-mean of the residuals across all the grid cells are distributed symmetrically around zero, with a mean value <0.4 mm and a SD of about 4 mm.

Both the MCMC diagnostics and the residual analysis presented above indicate a good fit of the Bayesian model to the observational data. Despite this, it is still important to check that the posterior inferences from the model look plausible when compared to the observational data. We do this by plotting estimates of SL, TS, HS and OM against observational time series at a randomly selected grid cell (Fig. 3). We show results only for the solution based on the surface HF from DEEP-C (i.e., BHM1) but estimates from the other three Bayesian solutions show a similar performance. All the sea-level components show considerable inter-annual variability, although such variability is significantly larger in the SL and the TS with peak-to-peak fluctuations that can reach more than 20 cm as compared to 8 cm for the HS and 2 cm for the OM. While the inter-annual variability is similar between the observations and the Bayesian solutions, as we would expect, there are also significant differences between the two, especially for the TS and HS. These differences reflect the relatively large uncertainties associated with the TS and HS observations and also demonstrate a crucial point: the ability of the BHM to constrain the TS and HS estimates based on information from the SL and OM data. This ability leads to more accurate estimates of TS and

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

HS, in turn, allowing us to obtain more reliable estimates of MHT and other quantities of interest that depend on TS and HS.

For completeness, we have also plotted the linear trend estimates derived from the BHM on top of the time series (Fig. 3). In assessing the trends, it is important to note that the Bayesian trends are not directly comparable to least squares trends calculated from individual observed time series. Bayesian trends aim to capture the true underlying trend, free from the influence of AR1 variability and noise, whereas least squares trends will be affected by these factors. In practice, this means that the Bayesian trends will not necessarily be entirely aligned with what visually appears to be the tendency of the time series, although in general we would expect some degree of alignment, especially if the trend is large relative to the variability. In the example of the figure, the trends do agree to a large extent with the long-term tendency displayed by the time series, including in the OM time series for which the trend is the dominant signal.

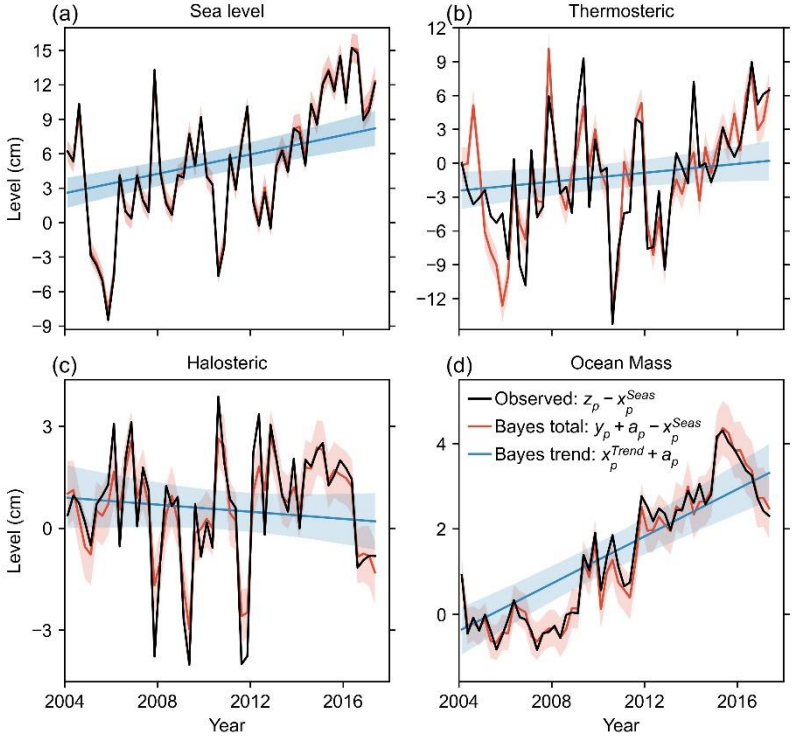


Figure 3. Quarterly (3-month-averaged) time series of (a) sea level, (b) thermosteric, (c) halosteric and (d) ocean mass as derived from BHM1 at a randomly selected grid cell (red) together with the corresponding observed time series (black). Bayesian estimates of the linear trend at the same grid cell (blue) are also shown for the SL and its components. The shading around the red and blue lines represents the 5-95% credible interval.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

5. Atlantic meridional heat transport

5.1. Calculation of the MHT

In this section, we describe how estimates of MHT over the Atlantic Ocean are derived from our Bayesian estimates of HTC. The BHM does not provide MHT directly as an output, but this can be calculated by meridionally integrating the HTCs. The integration can be started from any of the twelve latitude lines (see Fig. 1), but here we choose (for reasons that will become obvious later) to start from the northernmost latitude (i.e., 65°N). In this case, the northward heat transport across the *i*-th latitudinal section at time t, $\mathrm{MHT}_{i,t}$ (using the same labeling as shown in Fig. 1), can be calculated as:

$$\mathrm{MHT}_{i,t} = \mathrm{MHT}_{1,t} + \sum_{j=1}^{i-1} \mathrm{HTC}_{j,t} \,, \qquad i = 1, \dots, M+1 \tag{20}$$

where $\mathrm{MHT}_{1,t}$ is the northward heat transport at 65°N and $\mathrm{HTC}_{j,t}$ is the heat transport convergence in region R_j at time t (i.e., $U_t(R_j)$). Clearly, to evaluate Eq. (20) at any latitude line we need an estimate of $\mathrm{MHT}_{1,t}$, but such an estimate is not available. Let's now suppose that we set $\mathrm{MHT}_{1,t} = 0$ and then evaluate Eq. (20). This will produce estimates of MHT that will be accurate at any latitude where the true transport, $\mathrm{MHT}_{i,t}^{\mathrm{True}}$, is large relative to the true transport at 65°N, i.e. where the following condition is met: $\mathrm{MHT}_{i,t}^{\mathrm{True}} \gg \mathrm{MHT}_{1,t}^{\mathrm{True}}$. While the validity of this approximation cannot be tested at all latitudes due to the lack of MHT estimates, it can be assessed at 26°N by comparing estimates of MHT from the RAPID array (at 26°N) with estimates from the OSNAP array (at 50°N-60°N). Such comparison shows that the SD of the MHT time series (3-month means; overlapping period) is over four times larger at RAPID than at OSNAP, whereas the time-mean MHT is more comparable in magnitude with values of 1.19 PW (RAPID) and 0.51 PW (OSNAP). This indicates that ignoring the variability of MHT_1 in Eq. (20) will provide a very good approximation to the MHT variability at 26°N, assuming that the MHT variability at 65°N is similar to (or smaller than) that at the latitude of OSNAP. Ignoring the time-mean value of MHT_1 will, however, incur a larger error. Considering this, we approximate Eq. (20) as follows (we drop the temporal subscript t to simplify notation):

$$\mathrm{MHT}_i = \langle \mathrm{MHT}^{\mathrm{OSNAP}} \rangle - \langle \mathrm{HTC}_1 \rangle + \sum_{j=1}^{i-1} \mathrm{HTC}_j \,, \qquad i = 2, \dots, M+1 \tag{21}$$

where the angle brackets denote the time mean and $\langle \mathrm{MHT}^{\mathrm{OSNAP}} \rangle$ is the time-mean MHT from the OSNAP array over the period 2014-2018 (i.e., it is set equal to 0.506 PW). Hence, in Eq. (21) we are essentially setting $\mathrm{MHT}_1 = 0$ and the time-mean value of MHT at 60°N equal to $\langle \mathrm{MHT}^{\mathrm{OSNAP}} \rangle$. In making this approximation, we have implicitly

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

assumed that the time-mean MHT at 60ºN is constant over the period of our analysis, which appears to be a reasonable assumption based on the lack of a significant trend in the OSNAP MHT time series.

At this point, one might wonder why not start the integration of the HTCs at 26ºN instead of 65ºN and then just use the MHT from the RAPID array to fix the constant of integration, as this would circumvent the need to make any approximations. This, of course, is straightforward to do but it raises its own issues. First, this would make our estimates dependent on the RAPID-derived MHT and, thus, would leave us with almost no observation-based estimates with which to validate our results. Second, the magnitude of the MHT variability at 26ºN appears to be significantly larger than that at middle and high latitudes of the North Atlantic, and thus any discrepancies between our HTCs and those that would be consistent with the MHT from RAPID would manifest as large errors at higher latitudes where the MHT variability is relatively small. Of course, the approximation used in Eq. (21) will also introduce errors at latitudes where the MHT variability is small, but the advantage in this case is that our estimates of MHT are entirely independent from the RAPID-derived MHT.

5.2. MHT at 26ºN

Here, we compare our Bayesian estimates of MHT at 26º N calculated using Eq. (21) with the MHT from the RAPID array. The comparison is done for two different periods: 2004-2017 for BHM1, and 2004-2020 for BHM2 and BHM3. In the following, any estimates derived from the BHM solutions will be summarized by the posterior mean and the 5-95% credible interval (CI), where the CIs will be denoted by square brackets.

Before proceeding with the comparison, we note a recent study by Volkov et al. (2024) reporting on a spurious drift in the submarine cable measurements of the Florida Current that are used in the calculation of the RAPID-derived AMOC and MHT. In this study, they show that, after correcting for this spurious drift, the negative trend that is detectable in the uncorrected RAPID AMOC time series becomes barely statistically significant. This drift is also certain to affect the RAPID-derived MHT time series, but no correction for the MHT is publicly available at the time of writing this article. In view of this, we choose to remove a linear trend from the time series of MHT before the comparison to ensure that the comparison metrics reflect the true degree of concordance between the Bayesian and the RAPID estimates. That said and given the current debate on whether or not the AMOC is weakening, it is worth mentioning that of the three Bayesian estimates of MHT at 26º N, only BHM3 shows a statistically significant trend, and even then only marginally. More precisely, the MHT trends derived from BHM1, BHM2, BHM3 for the overlapping period 2004-2017 are, respectively, 0.06 PW decade^{-1} [-0.12,0.25], 0.09 PW decade^{-1} [-0.09,0.26], and 0.21 PW decade^{-1} [0.02,0.39]. Given the strong correlation between the AMOC and the

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

MHT at 26ºN (Johns et al., 2023), the lack of a statistically significant trend in two of the three Bayesian solutions appears to support the findings of Volkov et al. (2024).

Moving now to the comparison, we begin by showing the quarterly MHT time series as derived from the BHM together with the MHT from RAPID (Fig. 4a-c). Note that these time series represent 3-month-averaged values without any smoothing or scaling applied to them (i.e., they are the direct solutions from the BHM, except for the removal of a linear trend). Visually, it is already obvious that the variability in the MHT from RAPID is very well captured by the three Bayesian solutions, in terms of both the phase and the magnitude of the variability. This includes the prominent drop in MHT around 2010, the magnitude and timing of which are both captured with remarkable precision by the Bayesian solutions. More quantitatively, the correlation between the RAPID and Bayesian time series for the common period 2004-2017 is very significant with values of 0.77, 0.72 and 0.74 for BHM1, BHM2 and BHM3 respectively. For the period 2004-2020, the correlation is 0.63 for both BHM2 and BHM3. The lower correlation observed during the longer period of 2004-2020 is primarily due to a discrepancy in 2020. During this year, the Bayesian estimates indicate a decline in MHT, while the RAPID time series shows a pronounced spike. This discrepancy is the most significant inconsistency between the Bayesian and RAPID time series throughout the entire period. The SD of the MHT time series, which provides a measure of the magnitude of the variability, is also completely consistent between the Bayesian solutions and the MHT from RAPID, with values of (common period 2004-2017) 0.19±0.02 PW (RAPID), 0.19 PW [0.16,022] (BHM1), 0.18 PW [0.15,0.21] (BHM2) and 0.19 PW [0.16,0.22] (BHM3) (the uncertainty in the RAPID SD represents the frequentist 90% confidence interval).

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

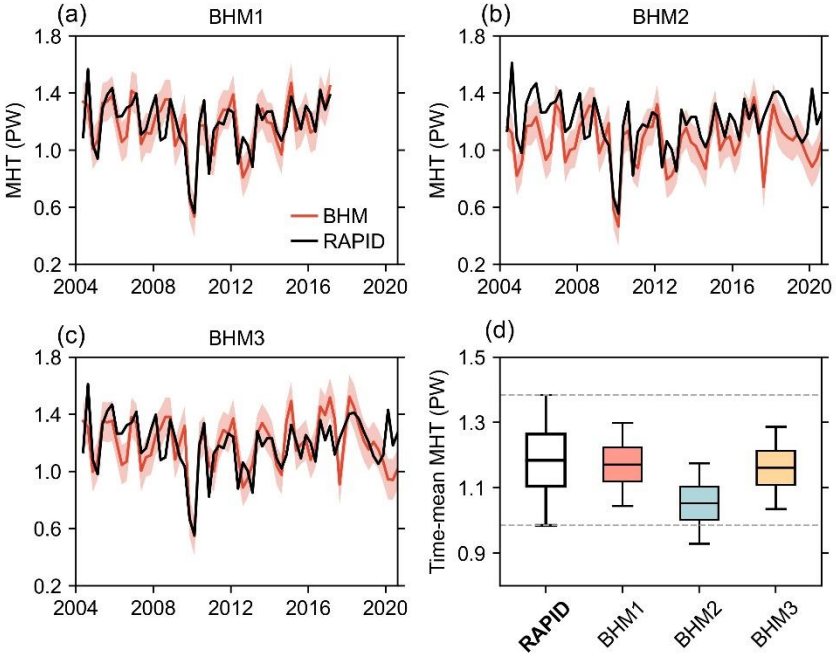


Figure 4. Quarterly (3-month-averaged) time series of MHT at 26º N (red line) as derived from (a) BHM1, (b) BHM2, and (c) BHM3 together with the RAPID-derived MHT (black line). The shading around the Bayesian time series represents the 5-95% CI. (d) Time-mean MHT from solutions BHM1, BHM2 and BHM3, summarized by the posterior mean (central mark), interquartile range (box) and 5–95% CI (whiskers). The CIs associated with the Bayesian estimates of the time-mean MHT reflect both statistical uncertainty due to the variability of the MHT time series as well as uncertainty from all sources considered in the BHM. For comparison, box and whiskers for the time-mean MHT from RAPID are also shown (in this case the quantiles have been computed based on the uncertainty provided by Johns et al. (2023), assuming a normal distribution).

Focusing now on the time-mean MHT, we can already sense from the comparison of the time series (Fig. 4a-c) that BHM1 and BHM3 match the time-mean MHT from RAPID very well, whereas BHM2 displays a slight, almost imperceptible, downward offset. This visual intuition is confirmed by calculating the time-mean value of the time series and its uncertainty for the common period 2004-2017 (Fig. 4d). The time-mean MHT from RAPID is 1.18±0.20 PW (90% confidence interval), whereas for BHM1, BHM2 and BHM3 we obtain values of 1.17 PW [1.04,1.30], 1.05 PW [0.93,1.18] and 1.16 PW [1.03,1.29]. In short, the time-mean MHT from BHM1 and BHM3 is in perfect agreement with that from RAPID, while that from BHM2 is slightly lower but still consistent with RAPID when considering the uncertainty.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

To complete our comparison, we evaluate the agreement between the Bayesian solutions and the RAPID-derived MHT at lower frequencies by applying a 5-quarter running mean to the MHT time series (Fig. 5). For clarity, and since we have already discussed the time-mean MHT, we now present the anomalies with respect to the time mean. We find that, at such lower frequencies, the agreement with RAPID is also remarkably good, with correlations for the common period 2004-2017 of 0.93, 0.81 and 0.86 for BHM1, BHM2 and BHM3, respectively. For the longer period of 2004-2020, the correlation for BHM2 and BHM3 is slightly lower with a value of 0.76. The best agreement with the MHT from RAPID is observed again for BHM1, which shows almost a perfect match (Fig. 5a), capturing both the timing and magnitude of nearly all the prominent features in the RAPID-derived time series, including the big drop around 2010 and the recovery afterwards. One of the most significant discrepancies between the Bayesian and RAPID time series occurs during the period 2005–2007, when the latter consistently exhibits higher MHT values. However, the elevated MHT values in the RAPID time series around 2006 are expected to decrease following the application of the drift correction proposed by Volkov et al. (2024). This adjustment would likely improve the agreement between the RAPID estimates and the Bayesian solutions.

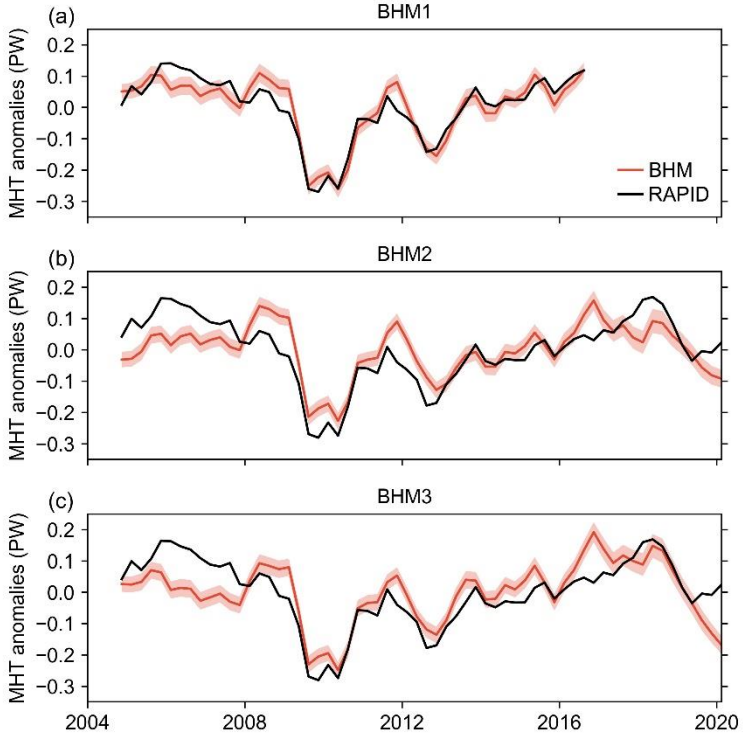


Figure 5. Same as Figs. 4a-c but with a 5-quarter running mean applied to the time series and the time-mean MHT removed.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

5.3. MHT over the Atlantic Ocean

Although the excellent agreement of the Bayesian estimates with the MHT from RAPID at 26º N does not guarantee a similar performance at other latitudes, it does give us confidence that the estimates are robust and likely to reflect true changes. In this section we show our estimates of MHT across eleven latitude lines over the Atlantic (we do not include 65º N as our estimates are set to zero at this latitude), both in their quarterly (3-month-averaged) resolution (Fig. 6a) and after applying a 5-quarter running mean (Fig. 6b). Note that, unlike for the comparison with the MHT from RAPID, here we do not remove a linear trend from any of the time series. While long records of MHT from direct ocean observations are restricted to only RAPID at 26º N, several earlier studies have produced estimates of MHT with good spatial coverage indirectly as the residual from energy budgets. Here, for comparison and as an additional check on our approach, we plot estimates from one of such studies (Trenberth et al., 2019) on top of ours, although this comparison is only possible for low-pass filtered time series because estimates from Trenberth et al. (2019) are only available as 12-month running means.

Looking at the quarterly MHT time series (Fig. 6a), we note several interesting features. First, MHT estimates from the three Bayesian solutions are very similar to one another at all latitudes, both in terms of phase and magnitude of the variability. Indeed, the correlation over the overlapping period 2004-2017 for all pairs of (detrended) time series from BHM1, BHM2 and BHM3 is > 0.9 in most cases, whereas differences in the SD of the time series are always smaller than 25% of the SD value. In comparing the three solutions, we also note that BHM3 shows slightly larger MHT values than BHM2 in the last years at most latitudes. This is primarily due to a small difference in the long-term trend between the two solutions rather than differences in variability. Comparing now the MHT time series from different latitudes, we note that the amplitude of the MHT fluctuations tends to increase gradually from north to south. For example, taking the SD of the time series as a measure of variability, we find that SD values for BHM1 range from 0.02 PW [0.02,0.03] at 60º N to a maximum value of 0.29 PW [0.24,0.34] at 35º S. Similar SD values are found for BHM2 and BHM3. Peak-to-peak fluctuations can be as large as 0.7 PW at 26º N and 1.5 PW at both 35º S and 25º S. The MHT fluctuations show a strong latitudinal coherence within two bands of latitudes, namely between 35º N and 16º N and from 5º N to 35º S. Such bands are not entirely decoupled from one another, but the MHT coherence within a band is much stronger than between bands. In particular, the event of 2010, which is so prominent in the MHT time series at 26º N, is also clearly visible at 16º N but much less so at 5º N. The lack of coherence south of 16°N during this event can be attributed to its main underlying cause − a southward shift of the North Atlantic subtropical gyre driven by a southward displacement of the mid-latitude westerlies (Evans et al., 2017) – the effects of which were largely confined to the mid-latitude regions with limited influence further south.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

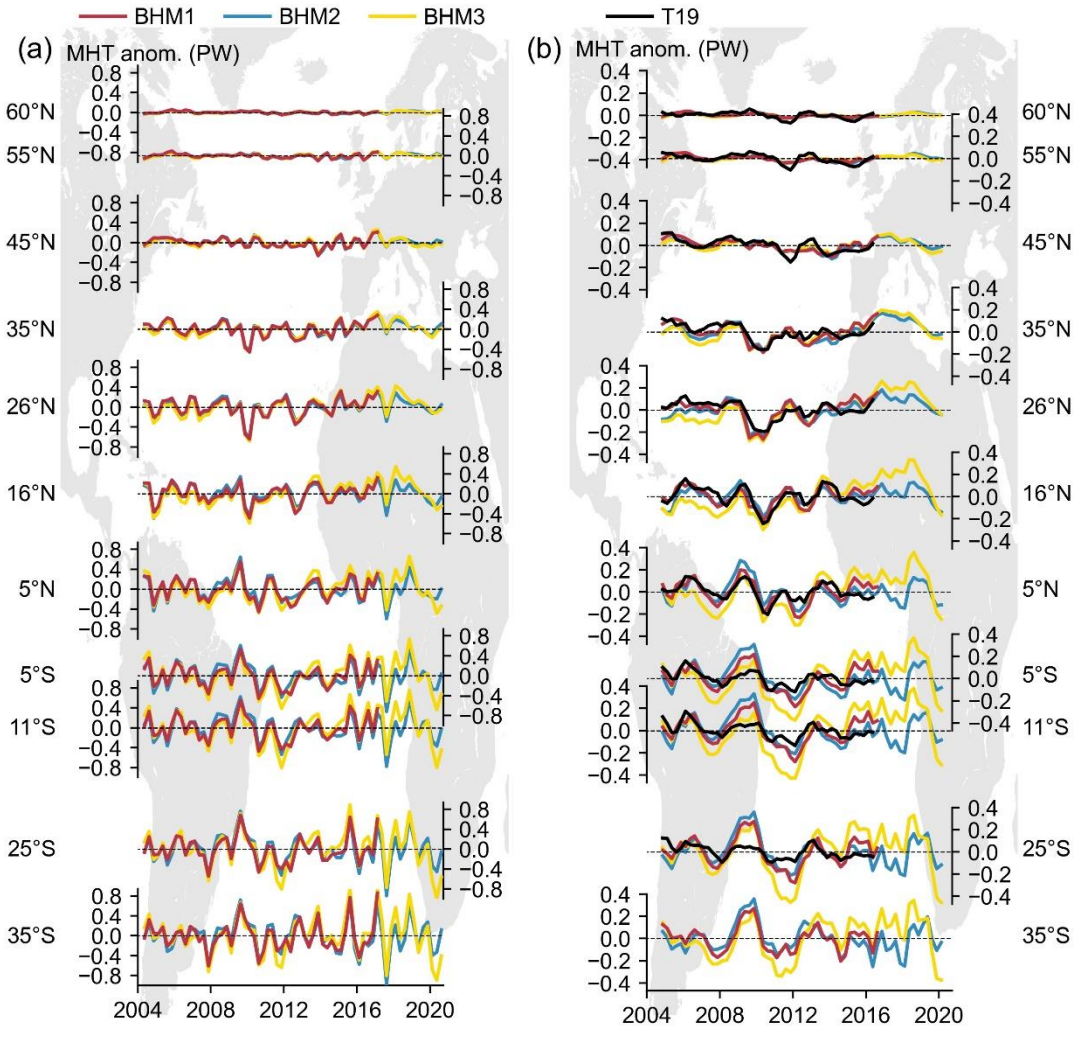


Figure 6. (a) Quarterly (3-month-averaged) time series of MHT anomalies across the latitudinal sections denoted on the vertical axis as estimated by BHM1 (red), BHM2 (blue) and BHM3 (yellow). (b) Same as panel (a) but with a 5-quarter running mean applied to the time series. For comparison, estimates of MHT from Trenberth et al. (2019) across the same sections are also shown in black.

Turning now our attention to the low-pass filtered time series (Fig. 6b), we note that again the three Bayesian solutions are in good agreement. The tendency of BHM3 to show higher MHT values in the last years compared to BHM2 is more noticeable in the smoothed time series since the short-term variability has been filtered out, but it is still mostly due to differences in the long-term trend between the two solutions. Just like for the quarterly

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

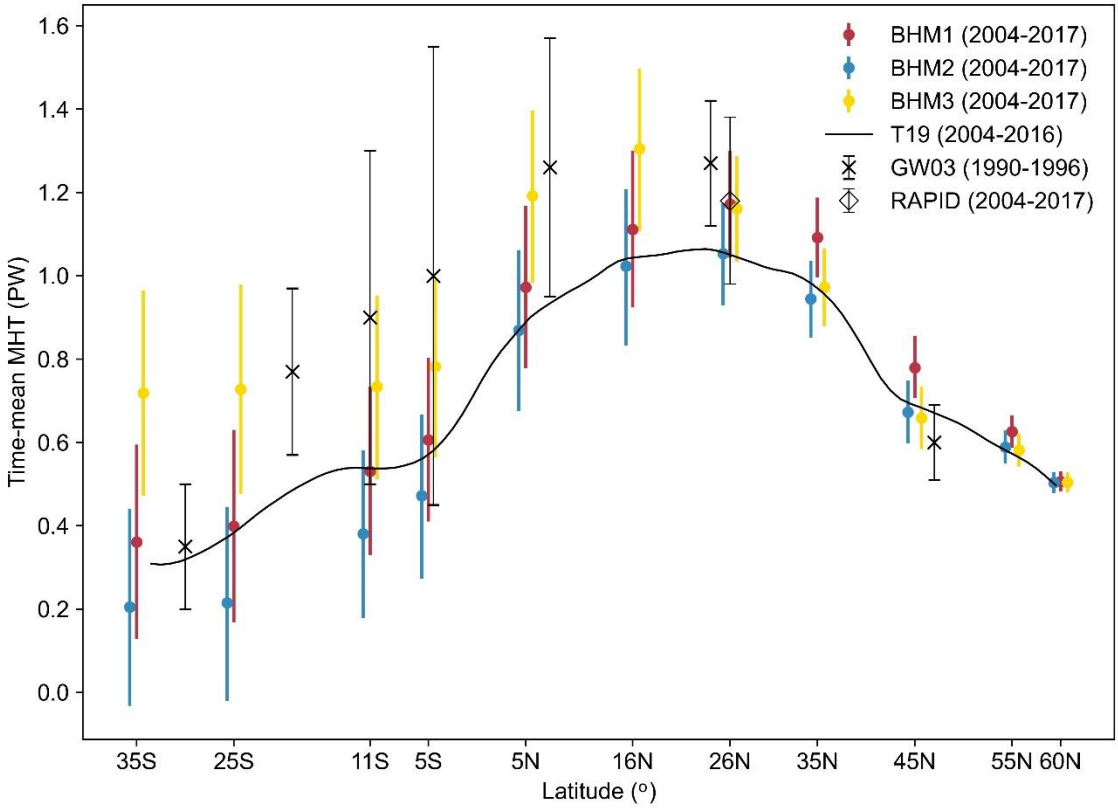
values, the magnitude of the fluctuations increases southward, although differences in magnitude across latitudes are smaller. The bands of latitudinal coherence are essentially the same as noted for the quarterly values. Overall, the Bayesian estimates are in good agreement with the time series of Trenberth et al., (2019) over the period 2004-2016, with correlation values of (maximum across BHM1, BHM2 and BHM3; detrended time series) 0.49 (60° N), 0.69 (55° N), 0.67 (45° N), 0.79 (35° N), 0.81 (26° N), 0.76 (16° N), 0.71 (5° N), 0.71 (5° S), 0.72 (11° S), and 0.63 (25° S). There are, however, noticeable differences in the strength of the variability between the two estimates at some latitudes. In particular, the size of the fluctuations is significantly larger in our estimates than in those of Trenberth et al. (2019) at all latitudes south of the equator.

Finally, it is also interesting to plot the time-mean MHT from the three Bayesian solutions across all the latitudinal sections of the Atlantic (Fig. 7). For comparison, observation-based estimates from the RAPID array (at 26° N), Trenberth et al. (2019) as denoted by T19, and Ganachaud & Wunsch (2003) as denoted by GW03 are also shown. The time-mean MHT is computed over the period 2004-2017 for the three Bayesian solutions and RAPID, 2004-2016 for T19, and 1990-1996 for GW03. Focusing first on the Bayesian estimates, we note that the time-mean MHT is positive, and thus northward, at all latitudes, achieving its maximum value between 26° N and 16° N depending on the solution. Overall, the latitudinal structure of the time-mean MHT is very similar across the three Bayesian solutions but there can be differences in the actual values (i.e., the posterior means). More precisely, the three solutions are in very good agreement in the 60° N−26° N latitude band, but south of 26° N the spread becomes wider, with BHM3 showing considerably higher positive values than BHM1 and BHM2. The largest difference between BHM3 and the other two Bayesian solutions is found at 35° S where BHM3 gives a value of 0.72 PW [0.47,0.97] compared to 0.36 PW [0.13,0.59] and 0.20 PW [-0.03,0.44] for BHM1 and BHM2, respectively. Note that the Bayesian CIs tend to become wider as we move southwards. This reflects both increased MHT variability as well as larger differences between surface HF products (recall that the uncertainty in the surface HF data has been calculated as the spread over the three HF products).

Taking BHM1 as our best solution − given its reliance on an observation-based HF product and its superior agreement with RAPID − we find that the time-mean MHT has a value of 0.51 PW [0.50,0.52] at 60° N (recall this value is set to match the time-mean MHT from the OSNAP array), which gradually increases southward reaching a maximum value of 1.17 PW [1.04,1.30] at 26° N, before decreasing again to a minimum value of 0.36 PW [0.13,0.59] at 35° S. This latitudinal structure of the time-mean MHT aligns well with the other observation-based estimates (RAPID, T19 and GW03), although noticeable differences exist in the values. Among all the estimates at 26° N, BHM1 shows the closest agreement with RAPID, while GW03 displays higher values (1.27 PW) and T19 lower values (1.05 PW). Overall, the Bayesian estimates align more closely with T19 than with GW03. In fact,

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

GW03 consistently shows higher values than all the other estimates across most latitude, including RAPID, although this might be because the GW03 estimates refer to a different period.



690 **Figure 7.** Time-mean MHT derived from BHM1 (red), BHM2 (blue) and BHM3 (yellow) along with the associated 5-95% CIs (whiskers) at multiple latitudes of the Atlantic Ocean. The CIs reflect both statistical uncertainty due to the variability of the MHT time series as well as uncertainty from all sources considered in the BHM. Observation-based estimates from the RAPID array, Trenberth et al. (2019) as denoted by T19, and Ganachaud & Wunsch (2003) as denoted by GW03 are also shown. The Bayesian solutions have been slightly shifted in latitude for better
695 visualisation, but they all correspond to the same latitude. The confidence intervals associated with the RAPID (90% interval) and GW03 (1-sigma interval) estimates are derived from the values provided in the corresponding studies.

6. Data availability

The Bayesian estimates of MHT and HTC are available at Zenodo via https://doi.org/10.5281/zenodo.14931921. (At the
700 moment, access to files is restricted but will be made open once the paper is accepted for publication. In the meantime, files

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

can be accessed via the link: https://zenodo.org/records/14931921?token=eyJhbGciOiJIUzUxMiJ9.eyJpZCI6IjY0MGEyOWVhLWQwYjYtNDIxNi1hNGI0LTE1YmUwYjZmMjgwZCIsImRhdGEiOnt9LCJyYW5kb20iOiJhMzdmMzA5NTgzOGQxNDhlM2QxZDZlMzcxODI3YjFkNiJ9.qIn7PtJ9B9wBa-a3jlU3py-6lrCnEKRp60OTjJq5mc3uGRL8cbabbAt0CeyXXZUurqRV7sH8HsfwugDR7ISu9g)

7. Conclusions

Here, we have generated observation-based probabilistic estimates of MHT for the period 2004-2020 at 3-month-mean resolution across 12 latitudinal sections of the Atlantic Ocean between 65º N and 35º S. The MHT has been calculated based on estimates of HTC which, in turn, have been inferred as a residual from the difference between OHC tendency and surface HF through joint spatiotemporal modelling of observations from hydrography, satellite altimetry and satellite gravimetry. We have produced three different Bayesian solutions based on three different surface HF products, namely DEEP-C, ERA5 and NCEP.

Our estimates of MHT agree remarkably well with estimates based on direct ocean observations from the RAPID array at 26º N, capturing both the magnitude and phase of the MHT variability in the RAPID time series with high fidelity. The time-mean MHT at 26º N is also very well captured by the three Bayesian solutions. The solution based on DEEP-C (BHM1) shows the closest agreement to RAPID-derived MHT, with a correlation of 0.77 for the quarterly time series and 0.93 for low-pass filtered (5-quarter running mean) time series. Such correlation (0.93) is noticeably higher than the correlations reported by previous studies for annual or 12-month low-pass filtered data, such as 0.54 by Meyssignac et al. (2024) for the period 2005-2018, 0.66 by Liu et al. (2020) for 2008-2016, or 0.72 by Mayers et al. (2022) for 2004-2016. This demonstrates the importance and benefit of accounting for spatiotemporal dependencies when combining hydrographic and satellite measurements.

The MHT variability is highly consistent across BHM1, BHM2 and BHM3 at all latitudes, both in terms of phase and magnitude, indicating that estimates of the MHT variability are robust to the choice of the surface HF product. In contrast to the variability, differences in the time-mean MHT across the three Bayesian solutions can be relatively large, especially in the South Atlantic and particularly between BHM1/BHM2 and BHM3. Given that the only difference among the three Bayesian solutions is the surface HF product, these results demonstrate that the surface HF makes an important contribution to the time-mean MHT and highlight the importance to our approach of having accurate, unbiased estimates of such a quantity. We have found that the magnitude of the variability tends to grow gradually from north to south, with peak-to-peak fluctuations that can reach 1.5 PW at the southernmost latitudinal sections. The MHT variability is not coherent across the whole Atlantic Ocean but instead covaries within, although not between, two distinct bands of latitudes, namely between 35º N and 16º N and from 5º N to 35º S. Regarding

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

the time-mean MHT, all three Bayesian solutions show northward MHT across the entire Atlantic Ocean, with a latitudinal profile that is characterized by relatively small transport in the South Atlantic Ocean, particularly south of 11° S, and maximum transport between 26° N and 16° N.

Our results are exceptionally good if we accept the agreement with RAPID-derived MHT as a measure of performance, but there are some limitations that the users of these estimates need to be aware of. The first one is that our estimates of MHT have been derived by assuming no variability in the MHT at 65° N. As discussed earlier, this assumption will produce very accurate estimates of MHT variability over most of the Atlantic Ocean, but the incurred errors will be larger at high latitudes of the North Atlantic where the MHT variability is smaller in magnitude, particularly north of 45° N. Additionally, while estimates of the time-mean MHT are not affected by this assumption, it is important to remember that the time-mean MHT at 60° N has been set equal to that derived from the OSNAP array over the period 2014-2018. This assumes a constant time-mean MHT at 60° N over the analysis period (2004-2020), and thus deviations from this premise will introduce an error into the time-mean MHT at all the other latitudes. That said, such an error is expected to be very small based on the absence of a trend in the MHT time series from OSNAP and the excellent agreement with the RAPID-derived MHT at 26° N. Finally, it is also important to note that our estimates are derived from observational data spatially aggregated into 3° × 3° areal units, hence our approach is unlikely to resolve variations in MHT between latitude lines spaced by less than 3°.

References

Anderson, T. W., and Darling, D. A.: A test of goodness of fit, J. Am. Stat. Assoc., **49**(268), 765–769, https://doi.org/10.2307/2281537, 1954.

Andrews, T., Gregory, J. M., Paynter, D., Silvers, L. G., Zhou, C., Mauritsen, T., Webb, M. J., Armour, K. C., Forster, P. M., and Titchner, H.: Accounting for Changing Temperature Patterns Increases Historical Estimates of Climate Sensitivity, Geophys. Res. Lett., 45, 8490–8499, https://doi.org/10.1029/2018GL078887, 2018.

Banerjee S, Carlin BP, and Gelfand AE.: Hierarchical modeling and analysis for spatial data (2nd edn), London, UK: Chapman & Hall, 2014.

Buckley, M. W., and Marshall, J.: Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review, Rev. Geophys., 54, 5–63, https://doi.org/10.1002/2015RG000493, 2016.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Calafat, F. M., Frederikse, T., and Horsburgh, K.: The sources of sea-level changes in the Mediterranean Sea since 1960, J. Geophys. Res. Oceans, 127, e2022JC019061, https://doi.org/10.1029/2022JC019061, 2022.

Cheng, L., Abraham, J., Trenberth, K. E., Fasullo, J., Boyer, T., Locarnini, R., Zhang, B., Yu, F., Wan, L., Chen, X., Song, X., Liu, Y., Mann, M. E., Reseghetti, F., Simoncelli, S., Gouretski, V., Chen, G., Mishonov, A., Reagan, J., and Zhu, J.: Upper Ocean Temperatures Hit Record High in 2020, Adv. Atmos. Sci., 38, 523–530, https://doi.org/10.1007/s00376-021-0447-x, 2021.

Cressie, N. and Wikle, C. K.: Statistics for Spatio-Temporal Data, John Wiley & Sons, 2011.

Cunningham, S. A., Kanzow, T., Rayner, D., Baringer, M. O., Johns, W. E., Marotzke, J., Longworth, H. R., Grant, E. M., Hirschi, J. J.-M., Beal, L. M., Meinen, C. S., and Bryden, H. L.: Temporal variability of the Atlantic Meridional Overturning Circulation at 26◦ N, Science, 317, 935–938, doi:10.1126/science.1141304, 2007.

Dieng, H. B., Palanisamy, H., Cazenave, A., Meyssignac, B., and von Schuckmann, K.: The sea level budget since 2003: Inference on the Deep Ocean Heat content, Surv. Geophys., 36(2), 209–229. https://doi.org/10.1007/s10712-015-9314-6, 2015.

Dong, Y., Proistosescu, C., Armour, K. C., and Battisti, D. S.: Attributing historical and future evolution of radiative feedbacks to regional warming patterns using a Green's Function approach: The preeminence of the western Pacific, J. Clim., 32(17), 5471–5491, https://doi.org/10.1175/JCLI-D-18-0843.1, 2019.

Evans, D. G., Toole, J., Forget, G., Zika, J. D., Naveira Garabato, A. C., Nurser, A. J. G., and Yu, L.: Recent wind-driven variability in Atlantic water mass distribution and Meridional Overturning Circulation, J. Phys. Oceanogr., 47(3), 633–647. https://doi.org/10.1175/JPO-D-16-0089.1, 2017.

Frajka-Williams, E., Ansorge, I. J., Baehr, J., Bryden, H. L., Paz Chidichimo, M., Cunningham, S. A., Danabasoglu, G., Dong, S., Donohue, K. A., Elipot, S., Heimbach, P., Holliday, N. P., Hummels, R., Jackson, L. C., Karstensen, J., Lankhorst, M., Le Bras, I. A., Lozier, M. S., McDonagh, E. L., Meinen, C. S., Mercier, H., Moat, B. I., Perez, R. C., Piecuch, C. G., Rhein, M., Srokosz, M. A., Trenberth, K. E., Bacon, S., Forget, G., Goni, G., Kieke, D., Koelling, J., Lamont, T., McCarthy, G. D., Mertens, C., Send, U., Smeed, D. A., Speich, S., van den Berg, M., Volkov, D., and Wilson, C.: Atlantic meridional overturning circulation: Observed transport and variability, Front. mar. sci., 6, 260, https://doi.org/10.3389/fmars.2019.00260, 2019.

Frederikse, T., Riva, R. E. M., and King, M. A.: Ocean Bottom Deformation Due To Present-Day Mass Redistribution and Its Impact on Sea Level Observations, Geophys. Res. Lett., 44, 12,306–12,314, https://doi.org/https://doi.org/10.1002/2017GL075419, 2017.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Frederikse, T., Landerer, F., Caron, L., Adhikari, S., Parkes, D., Humphrey, V. W., Dangendorf, S., Hogarth, P., Zanna, L., Cheng, L., and Wu, Y.-H.: The causes of sea-level rise since 1900, Nature, 584, 393–397, https://doi.org/10.1038/s41586-020-2591-3, 2020.

Gaillard, F., Reynaud, T., Thierry, V., Kolodziejczyk, N., von Schuckmann, K.: In-situ based reanalysis of the global ocean temperature and salinity with ISAS: variability of the heat content and steric height, J. Clim., 29(4), 1305-1323. https://doi.org/10.1175/JCLI-D-15-0028.1, 2016.

Ganachaud, A., and C. Wunsch: Large-scale ocean heat and freshwater transport during the World Ocean Circulation Experiment, J. Climate, 16, 696–705, https://doi.org/10.1175/1520-0442(2003)016%3C0696:LSOHAF%3E2.0.CO;2, 2003.

Gelfand, A. E., Zhu, L., Carlin, B. P.: On the change of support problem for spatio-temporal data, Biostatistics, 2(1), 31–45. https://doi.org/10.1093/biostatistics/2.1.31, 2001.

Gelfand, A. E., Schmidt, A. M., Banerjee, S., and Sirmans, C.: Nonstationary multivariate process modeling through spatially varying coregionalization, Test, 13(2), 263–312, https://doi.org/10.1007/BF02595775, 2004.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B.: Bayesian Data Analysis, 3rd ed., Chapman and Hall/CRC 398 Press, Boca Raton, Fla, 2014.

Gelman, A., and Rubin, D. B.: Inference from Iterative Simulation Using Multiple Sequences, Stat. Sci., 7(4), 457–72, https://doi.org/10.1214/ss/1177011136, 1992.

Gill, A. E., and Niiler, P. P.: The theory of the seasonal variability of the ocean, Deep Sea Res., 20, 141– 177, https://doi.org/10.1016/0011-7471(73)90049-1, 1973.

Good, S. A., Martin, M. J., and Rayner, N. A.: EN4: quality controlled ocean temperature and salinity profiles and monthly objective analyses with uncertainty estimates, J. Geophys. Res. Oceans, doi:10.1002/2013JC009067, 2013.

Gotway, C. A., and Young, L. J.: Combining incompatible spatial data, J. Am. Stat. Assoc., 97, 632–48. https://doi.org/10.1198/016214502760047140, 2002.

Gregory, J. M., Griffies, S. M., Hughes, C. W., Lowe, J. A., Church, J. A., Fukimori, I., Gomez, N., Kopp, R. E., Landerer, F., Cozannet, G. L., Ponte, R. M., Stammer, D., Tamisiea, M. E., and van de Wal, R. S. W.: Concepts and Terminology for Sea Level: Mean, Variability and Change, Both Local and Global, Surv. Geophys., 40, 1251–1289, https://doi.org/10.1007/s10712-019- 09525-z, 2019.

Hakuba, M. Z., Frederikse, T., and Landerer, F. W.: Earth's energy imbalance from the ocean perspective (2005–2019), Geophys. Res. Lett., 48, e2021GL093624. https://doi.org/10.1029/2021GL093624, 2021.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J-N.: ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.adbb2d47, 2023.

Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., and Robson, J.: The evolution of the North Atlantic meridional overturning circulation since 1980, Nat. Rev. Earth Environ., 3(4), 241–254. https://doi.org/10.1038/s43017-022-00263-2, 2022.

Johns, W. E., Baringer, M. O., Beal, L. M., Cunningham, S. A., Kanzow, T., Bryden, H. L., Hirschi, J. J. M., Marotzke, J., Meinen, C. S., Shaw, B., and Curry, R.: Continuous, array-based estimates of Atlantic Ocean heat transport at 26.5°N, J. Clim., 24(10), 2429–2449, doi:10.1175/2010JCLI3997.1, 2011.

Johns, W., Elipot, S., Smeed, D., Moat, B., King, B., Volkov, D., and Smith, R.: Towards two decades of Atlantic Ocean mass and heat transports at 26.5°N, Philos. Transact. A Math. Phys. Eng. Sci., 381(2262), 20220188. https://doi.org/10.1098/rsta.2022.0188, 2023.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-Year Reanalysis Project, Bull. Amer. Meteor. Soc., 77, 437–471, https://doi.org/10.1175/1520-0477(1996)077%3C0437:TNYRP%3E2.0.CO;2, 1996.

Kelly, K. A., Drushka, K., Thompson, L., Le Bars, D., and McDonagh, E. L.: Impact of slowdown of Atlantic overturning circulation on heat and freshwater transports, Geophys. Res. Lett., 43, 7625–7631, doi:10.1002/2016GL069789, 2016.

Kelly, K. A., Thompson, L., and Lyman, J.: The coherence and impact of meridional heat transport anomalies in the Atlantic ocean inferred from observations, J. Clim., 27(4), 1469–1487. https://doi.org/10.1175/JCLI-D-12-00131.1, 2014.

Li, F., Lozier, M. S., Holliday, N. P., Johns, W. E., Le Bras, I. A., Moat, B. I., Cunningham, S. A., and de Jong, M. F.: Observation-based estimates of heat and freshwater exchanges from the subtropical North Atlantic to the Arctic, Prog. Oceanogr., 197, 102640, https://doi.org/10.1016/j.pocean.2021.102640, 2021.

Liu, C., Allan, R. P., Berrisford, P., Mayer, M., Hyder, P., Loeb, N., Smith, D., Vidale, P.-L., and Edwards, J. M.: Combining satellite observations and reanalysis energy transports to estimate global net surface energy fluxes 1985–2012, J. Geophys. Res. Atmos., 120, 9374–9389, doi:10.1002/2015JD023264, 2015.

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Liu, C., Allan, R.P., Mayer, M. *et al*. Variability in the global energy budget and transports 1985–2017. *Clim Dyn* **55**, 3381–3396 (2020). https://doi.org/10.1007/s00382-020-05451-8.

Liu, C., Allan, R. P., Mayer, M., Hyder, P., Desbruyères, D., Cheng, L., Xu, J., Xu, F., and Zhang, Y.: Variability in the global energy budget and transports 1985–2017, Clim. Dynam., 55, 3381– 3396, https://doi.org/10.1007/s00382-020-05451-8, 2020.

Loomis, B. D., Luthcke, S.B., and Sabaka, T.J.: Regularization and error characterization of GRACE mascons, J. Geod., 93, 1381–1398, https://doi.org/10.1007/s00190-019-01252-y, 2019.

Lowe, J. A., and Gregory, J. M.: Understanding projections of sea level rise in a Hadley Centre coupled climate model, J. Geophys. Res., 111, C11014, doi:10.1029/2005JC003421, 2006.

Lozier, M. S., Li, F., Bacon, S., Bahr, F., Bower, A. S., Cunningham, S. A., de Jong, M. F., de Steur, L., deYoung, B., Fischer, J., Gary, S. F., Greenan, B. J. W., Holliday, N. P., Houk, A., Houpert, L., Inall, M. E., Johns, W. E., Johnson, H. L., Johnson, C., Karstensen, J., Koman, G., le Bras, I. A., Lin, X., Mackay, N., Marshall, D. P., Mercier, H., Oltmanns, M., Pickart, R. S., Ramsey, A. L., Rayner, D., Straneo, F., Thierry, V., Torres, D. J., Williams, R. G., Wilson, C., Yang, J., Yashayaev, I., and Zhao, J.: A sea change in our view of overturning in the subpolar North Atlantic, Science, 363(6426), 516–521. https://doi.org/10.1126/science.aau6592, 2019.

Maes, C.: Estimating the influence of salinity on sea level anomaly in the ocean, Geophys. Res. Lett., 25, 3551–3554. https://doi.org/10.1029/98GL02758, 1998.

Marti, F., Blazquez, A., Meyssignac, B., Ablain, M., Barnoud, A., Fraudeau, R., Jugier, R., Chenal, J., Larnicol, G., Pfeffer, J., Restano, M., and Benveniste, J.: Monitoring the ocean heat content change and the Earth energy imbalance from space altimetry and space gravimetry, Earth Syst. Sci. Data, 14, 229–249, https://doi.org/10.5194/essd-14-229-2022, 2022.

Mayer J., Mayer M., Haimberger L., Liu C.: Comparison of surface energy fluxes from global to local scale, J. Clim., 35(14), 4551-4569, https://doi.org/10.1175/JCLI-D-21-0598.1, 2022.

McCarthy, G. D., Smeed, D. A., Johns, W. E., Frajka-Williams, E., Moat, B. I., Rayner, D., Baringer, M. O., Meinen, C. S., Collins, J., and Bryden, H. L.: Measuring the Atlantic Meridional Overturning Circulation at 26°N, Prog. Oceanogr., 130, 91–111, doi:10.1016/j.pocean.2014.10.006, 2015.

Meyssignac, B., Boyer, T., Zhao, Z., Hakuba, M. Z., Landerer, F. W., Stammer, D., Köhl, A., Kato, S., L'Ecuyer, T., Ablain, M., Abraham, J. P., Blazquez, A., Cazenave, A., Church, J. A., Cowley, R., Cheng, L., Domingues, C. M., Giglio, D., Gouretski, V., Ishii, M., Johnson, G. C., Killick, R. E., Legler, D., Llovel, W., Lyman, J., Palmer, M. D., Piotrowicz, S., Purkey, S. G., Roemmich, D., Roca, R., Savita, A., von

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Schuckmann, K., Speich, S., Stephens, G., Wang, G., Wijffels, S. E., and Zilberman, N.: Measuring Global Ocean Heat Content to Estimate the Earth Energy Imbalance, Front. Mar. Sci., 6:432, doi: 10.3389/fmars.2019.00432, 2019.

Meyssignac, B., Fourest, S., Mayer, M., Johnson, G. C., Calafat, F. M., Ablain, M., Boyer, T., Cheng. L., Desbruyères, D., Forget, G., Giglio, D., Kuusela, M., Locarnini, R., Lyman, J. M., Llovel, W., Mishonov, A., Reagan, J., Rousseau, V., and Benveniste, J.: North Atlantic Heat Transport Convergence Derived from a Regional Energy Budget Using Different Ocean Heat Content Estimates, Surv. Geophys., https://doi.org/10.1007/s10712-024-09865-5, 2024.

Phan, D., Pradhan, N., and Jankowiak, M.: Composable effects for flexible and accelerated probabilistic programming in NumPyro, ArXiv preprint, arXiv:1912.11554, 2019.

Piecuch, C. G., Huybers, P., Hay, C. C., Kemp, A. C., Little, C. M., Mitrovica, J. X., Ponte, R. M., and Tingley, M. P.: Origin of spatial variation in US East Coast sea-level trends during 1900– 2017, Nature, 564, 400–404, https://doi.org/10.1038/s41586- 018-0787-6, 2018.

Piecuch, C. G., and Beal, L. M.: Robust weakening of the Gulf Stream during the past four decades observed in the Florida Straits, Geophys. Res. Lett., 50, e2023GL105170. https://doi.org/10.1029/2023GL105170, 2023.

Prandi, P., Meyssignac, B., Ablain, M., Spada, G., Ribes, A., and Benveniste, J.: Local sea level trends, accelerations and uncertainties over 1993–2019, Sci. Data, 8, 1, https://doi.org/10.1038/s41597-020-00786-7, 2021.

Pujol, M.-I., Faugère, Y., Taburet, G., Dupuy, S., Pelloquin, C., Ablain, M., and Picot, N.: DUACS DT2014: the new multi-mission altimeter data set reprocessed over 20 years, Ocean Sci., 12, 1067–1090, https://doi.org/10.5194/os-12-1067-2016, 2016.

Roberts, C. D., Palmer, M. D., Allan, R. P., Desbruyeres, D. G., Hyder, P., Liu, C., and Smith, D.: Surface flux and ocean heat transport convergence contributions to seasonal and interannual variations of ocean heat content, J. Geophys. Res. Oceans, 122, 726–744, doi:10.1002/2016JC012278, 2017.

Srokosz, M., Danabasoglu, G., and Patterson, M.: Atlantic meridional overturning circulation: Reviews of observational and modeling advances-An introduction, J. Geophys. Res. Oceans, 126, e2020JC016745. https://doi.org/10.1029/2020JC016745, 2021.

Tapley, B. D., Watkins, M. M., Flechtner, F., Reigber, C., Bettadpur, S., Rodell, M., Sasgen, I., Famiglietti, J. S., Landerer, F. W., Chambers, D. P., Reager, J. T., Gardner, A. S., Save, H., Ivins, E. R., Swenson, S. C., Boening, C., Dahle, C., Wiese, D. N., Dobslaw, H., Tamisiea, M. E., and Velicogna, I.: Contributions of

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

GRACE to understanding climate change, Nat. Clim. Change, 5, 358–369, https://doi.org/10.1038/s41558-019-0456-2, 2019.

Trenberth, K. E., and Fasullo, J. T.: Atlantic meridional heat transports computed from balancing Earth's energy locally, Geophys. Res. Lett., 44, 1919–1927, doi:10.1002/2016GL072475, 2017.

Trenberth, K. E., Zhang, Y., Fasullo, J. T., and Cheng, L. (2019) Observation-based estimate of global and basin ocean meridional heat transport time series, J. Clim., 32, 4567–4583, https://doi.org/10.1175/JCLI-D-18-0872.1, 2019.

Volkov, D. L., Smith, R. H., Garcia, R. F. Smeed, D. A., Moat, B. I., Johns, W. E., and Baringer, M. O.: Florida Current transport observations reveal four decades of steady state, Nat Commun, 15, 7780, https://doi.org/10.1038/s41467-024-51879-5, (2024).

von Schuckmann, K., Cheng, L., Palmer, M. D., Hansen, J., Tassone, C., Aich, V., Adusumilli, S., Beltrami, H., Boyer, T., Cuesta-Valero, F. J., Desbruyères, D., Domingues, C., García-García, A., Gentine, P., Gilson, J., Gorfer, M., Haimberger, L., Ishii, M., Johnson, G. C., Killick, R., King, B. A., Kirchengast, G., Kolodziejczyk, N., Lyman, J., Marzeion, B., Mayer, M., Monier, M., Monselesan, D. P., Purkey, S., Roemmich, D., Schweiger, A., Seneviratne, S. I., Shepherd, A., Slater, D. A., Steiner, A. K., Straneo, F., Timmermans, M.-L., and Wijffels, S. E.: Heat stored in the Earth system: where does the energy go?, Earth Syst. Sci. Data, 12, 2013–2041, https://doi.org/10.5194/essd-12-2013-2020, 2020.

Wang, G., Cheng, L., Boyer, T., and Li, C.: Halosteric Sea Level Changes during the Argo Era, Water, *9*, 484. https://doi.org/10.3390/w9070484, 2017.

Wong, A. P., Wijffels, S. E., Riser, S. C., Pouliquen, S., Hosoda, S., Roemmich, D., Gilson, J., Johnson, G. C., Martini, K., Murphy, D. J., Scanderbeg, M., Bhaskar, T. V., Buck, J. J., Merceur, F., Carval, T., Maze, G., Cabanes, C., André, X., Poffa, N., Yashayaev, I., Barker, P. M., Guinehut, S., Belbéoch, M., Ignaszewski, M., Baringer, M. O., Schmid, C., Lyman, J. M., McTaggart, Kr. E., Purkey, S. G., Zilberman, N., Alkire, M. B., Swift, D., Owens, W. B., Hersh, C., Robbins, P., West-Mack, D., Bahr, F., Yoshida, S., Sutton, P. J., Cancouët, R., Coatanoan, C., Dobbler, D., Juan, Andrea G., Gourrion, J., Kolodziejczyk, N., Bernard, V., Bourlès, B., Claustre, H., D'Ortenzio, F., Le Reste, S., Le Traon, P. Y., Rannou, J. P., Saout-Grit, C., Speich, S., Thierry, V., Verbrugge, N., Angel-Benavides, I. M., Klein, B., Notarstefano, G., Poulain, P. M., Vélez-Belchí, P., Suga, T., Ando, K., Iwasaska, N., Kobayashi, T., Masuda, S., Oka, E., Sato, K., Nakamura, T., Takatsuki, Y., Yoshida, T., Cowley, R., Lovell, J. L., Oke, P. R., van Wijk, E. M., Carse, F., Donnelly, M., Gould, W. J., Gowers, K., King, B. A., Loch, S. G., Mowat, M., Turton, J., Rama Rao , E. P., Ravichandran, M., Freeland, H. J., Gaboury, I., Gilbert, D., Greenan, B. J., Ouellet, M., Ross, T.,

https://doi.org/10.5194/egusphere-2025-1216
Preprint. Discussion started: 21 March 2025

Tran, A., Dong, M., Liu, Z., Xu, Jia., Kang, Ki., Jo, H., Kim, S., and Park, H.: Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations From a Global Array of Profiling Floats, Front. Mar. Sci., 7, 700,https://doi.org/10.3389/fmars.2020.00700, 2020.

Woollings, T., Gregory, J. M., Pinto, J. G., Reyers, M., and Brayshaw, D. J.: Response of the North Atlantic storm track to climate change shaped by ocean–atmosphere coupling, Nat. Geosci., 5, 313–317, https://doi.org/10.1038/ngeo1438, 2012.

Yin, J., and Zhao, M: Influence of the Atlantic meridional overturning circulation on the U.S. extreme cold weather, Commun. Earth. Environ., 2, 218, https://doi.org/10.1038/s43247-021-00290-9, 2021.

Zammit-Mangion, A., Rougier, J., Bamber, J., and Schön, N.: Resolving the Antarctic contribution to sea-level rise: a hierarchical modelling framework, Environmetrics, 25, 245–264, https://doi.org/10.1002/env.2247, 2014.

Zammit-Mangion, A., Rougier, J., Schön, N., Lindgren, F., and Bamber, J.: Multivariate spatio-temporal modelling for assessing Antarctica's present-day contribution to sea-level rise, Environmetrics, 26, 159-177, https://doi.org/10.1002/env.2323, 2015.

Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.-O., Marsh, R., Yeager, S. G., Amrhein, D. E., and Little, C. M.: A Review of the Role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and Associated Climate Impacts, Rev. Geophys., 57, 316–375, https://doi.org/10.1029/2019RG000644, 2019.