Validating the spatial variability of the semidiurnal internal tide in a realistic global ocean simulation with Argo and mooring data

Gaspard Geoffroy[1], Jonas Nycander[1], Maarten C. Buijsman[2], Jay F. Shriver[3], and Brian K. Arbic[4]

[1]Dept. of Meteorology, Stockholm University, Stockholm, Sweden
[2]School of Ocean Science and Engineering, University of Southern Mississippi, Stennis Space Center, MS, USA
[3]Naval Research Laboratory, Stennis Space Center, MS, USA
[4]Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA

Correspondence: Gaspard Geoffroy (gaspard.geoffroy@misu.su.se)

Abstract. The total variance and decorrelation of the semidiurnal internal tide (IT) are examined in a 32-day segment of a global run of the HYbrid Coordinate Ocean Model (HYCOM). This numerical simulation, with 41 vertical layers and 1/25° horizontal resolution, includes tidal and atmospheric forcing allowing for the generation and propagation of IT to take place within a realistic eddying general circulation. The HYCOM data are in turn compared with global observations of the IT around

5 1,000 dbar, from Argo float park phase data and mooring records. HYCOM is found to be globally biased low in terms of total variance and decorrelation of the semidiurnal IT over timescales shorter than 32 days. Except in the Southern Ocean, where limitations in the model causes the discrepancy with in situ measurements to grow poleward, the spatial correlation between the Argo and HYCOM inferred total variance suggests that the generation of low-mode semidiurnal IT is globally well captured by the model.

10 1 Introduction

Internal tides (IT) are internal waves generated by the interaction of tidal currents with rough bathymetry. The radiated beams can travel thousands of kilometers (e.g., Zhao et al., 2016; Buijsman et al., 2020) over which they undergo dissipative processes as they interact with the eddying ocean and bottom topography, to eventually break (MacKinnon et al., 2017). The dissipation of IT represents a major source of vertical mixing in the ocean interior (de Lavergne et al., 2020). As such, IT have a key

15 influence on the ocean state (Melet et al., 2016), and therefore on the global climate system (see Melet et al., 2022, for a review on the subject).

At any given position in a stationary medium, the waves would have a constant phase difference to the astronomical forcing at their generation site. However, since they propagate within the time varying ocean circulation, IT are subject to a variety of mechanisms that causes their phase difference to the tidal forcing to shift with time (Rainville and Pinkel, 2006; Zaron and

20 Egbert, 2014; Buijsman et al., 2017). In other words, IT lose *coherence* by interacting with the eddying ocean. They decorrelate: the autocovariance of a time series representing the internal tide variability at a fixed position (away from the source) inevitably decays with time lag (Caspar-Cohen et al., 2022; Geoffroy and Nycander, 2022).

When analyzing such an autocovariance, it is important to understand that only the *coherent* fraction of the IT energy decays with time lag, that is the energy carried by waves with a constant phase difference to the astronomical forcing. Conversely, the
25 *incoherent* fraction of the energy grows, so that the total IT energy at a given location (the sum of the coherent and incoherent components) is unaffected by the decorrelation. For very long time lags, the coherent and the complementary incoherent asymptotic limits are often called *stationary* and *nonstationary*, respectively. Generally, the longer the waves propagate, the more energy has likely been shifted out of phase by decorrelating processes. Hence, the stationary fraction decreases with increasing distance from the generation site (Shriver et al., 2014).

30 Some current high-resolution global ocean circulation model enable the estimation of the barotropic and internal tides concurrently with the ocean circulation (Arbic et al., 2010; Shriver et al., 2012; Buijsman et al., 2020). Such fully nonlinear numerical simulations also incorporate the interactions of the generated IT with the eddying ocean and its boundaries. The model utilized in this study is the HYbrid Coordinate Ocean Model (HYCOM; Chassignet et al., 2006) with 41 vertical layers and 1/25° horizontal resolution. The literature on the validation of HYCOM with observations is already vast (see
35 Buijsman et al., 2020; Arbic, 2022, for recent accounts). In the latest developments, surface drifters have been used to validate the geographical variability of the kinetic energy in various frequency bands at the global scale (Arbic et al., 2022). While the global coverage of drifter data is comparable to that of satellite altimetry, the contribution from the baroclinic tides to the kinetic energy observed by drifters has not been determined yet. Hence, until recently, only altimetry could unveil the geographical variability of the IT at the global scale.

40 The empirical mapping of IT from altimetry remains challenging from various aspects (Egbert and Ray, 2017). Most notably, the long sampling intervals of altimeters and the low signal-to-noise ratio preclude any direct estimation in the time domain of the total IT variance at a single location. Notwithstanding, Zaron (2017) analysed along-track wavenumber spectra of the sea surface height to map the total and nonstationary semidiurnal IT variance (for the baroclinic mode 1 only). The author found a global mean ratio of nonstationary to total semidiurnal IT variance of 44%. He also outlined the spatial inhomogeneity of the
45 tidal variability, with much of the equatorial Pacific and Indian Oceans affected by nonstationary tides larger than stationary tides (Zaron, 2017). In a comparison with data from HYCOM, Nelson et al. (2019) showed this 'k-space' methodology to miss a large fraction of both the nonstationary and total variance. Nevertheless, the spatial correlation of the nonstationary fraction (i.e., the ratio of the nonstationary to total variance) between the model and altimetry suggests that the model at least qualitatively captures the generation of IT and their interactions with the background circulation (Nelson et al., 2019).

50 Lately, Geoffroy and Nycander (2022) used observations from Argo park phase data (Argo, 2000) to empirically map the variance of the semidiurnal IT at 1,000 dbar. The high sampling rate of the floats allows to capture the total variance of the IT, i.e. the autocovariance at short time lags, before any loss of coherence occurs. On the other hand, the Lagrangian sampling of the drifting floats results in decorrelating effects on top of the intrinsic decorrelation of the IT (Zaron and Elipot, 2021; Caspar-Cohen et al., 2022; Geoffroy and Nycander, 2022). The latter decorrelation cannot be disentangled from the former
55 using Argo data only. Thus, to gain insights on the intrinsic decorrelation of IT around 1,000 dbar, one can instead apply the methodology of Geoffroy and Nycander (2022) to Eulerian observations from moorings. In the present work we will compare

observations of the total variance and decorrelation of the semidiurnal IT around 1,000 dbar from Argo floats and moorings, respectively, to data from a global HYCOM run.

The remainder of this work is organized as follows: In Sect. 2, the in situ datasets and the numerical simulation are presented. In Sect. 3, we review the methodology of Geoffroy and Nycander (2022) in an example location from the Lagrangian and Eulerian perspectives. In Sect. 4, the main results are presented. We compare the geographical variability of the total variance and decorrelation of the semidiurnal IT obtained from the in situ data and numerical simulation. Then, the model data are used to quantify the strength of the decorrelation induced by the Lagrangian sampling. Finally, we outline the potential biases affecting the datasets. We conclude in Sect. 5, by discussing the results and giving a summary.

2 Data

2.1 Argo Floats With Iridium Communication

We use data from a global collection of Argo Iridium floats deployed by the University of Washington as part of the National Ocean Partnership Program during the period 2004–2022. In between the descending and ascending profiling phases, these floats also record temperature and pressure with an hourly resolution while adrift at 1,000 dbar. This so-called park phase typically lasts 10 days. Stitching together data from successive cycles, by filling the time between park phases (typically 6 hr) with NaNs, one can construct longer time series. In this study, we use segments of 32 days of data (i.e., the duration of the segment of numerical simulation introduced in Sect. 2.3). Note that, in practice, the sampling frequency of the park phase can vary. To ensure evenly spaced time series, we linearly interpolate each concatenated record of 32 days onto a time axis with constant 1 h step. Any interpolated value lying between two original records that are more than 1.5 h apart is replaced by NaN.

The position of the floats can only be fixed when they reach the surface. We assume straight trajectories in between two successive surfacings (typically 10 days apart). This assumption has been shown reasonable by Geoffroy and Nycander (2022), especially since we discard segments of data for which the mean speed of the float is larger than 0.1 m s^{-1}. This criterion is primarily intended to avoid contamination by lee waves: a flow with speed $U \sim \mathcal{O}(0.1)$ m s^{-1} passing over a bathymetric feature with horizontal length scale $\lambda \sim \mathcal{O}(10)$ km will generate lee waves with angular frequency $\omega \sim (2\pi/\lambda)U$ rad s^{-1} of the order of the diurnal frequency. A more detailed description of the Argo data and quality controls we employed is available from Hennon et al. (2014) and Geoffroy and Nycander (2022).

2.2 Global Multi-Archive Current Meter Database (GMACMD)

The Global Multi-Archive Current Meter Database (GMACMD; Scott et al., 2011) compiles tens of thousands of oceanographic time series from moorings. Previous model-data validation efforts involving both HYCOM and the GMACMD notably include Luecke et al. (2020), where the authors compared the temperature variance and kinetic energy, over various frequency bands, in realistic global ocean simulations to more than 3,800 instrumental records.

Here, we extracted 331 temperature time series spanning 1972–2010 and meeting the following criteria:

1. The mooring lies in water deeper than 2,000 m

90 2. The record is longer than 64 days with a sampling interval more frequent than 3 hr, for adequate resolution of the semidiurnal tidal signals

3. The instrument depth is within ± 200 m from 1,000 m

One mooring in particular is used as an example in Sect. 3. It recorded during 366 days in the years 1982-83 at the position 28.00° N, 151.95° W (north of the Hawaiian Ridge). As previously documented by Alford and Zhao (2007), we refer to the instrument at 1,119 m depth as mooring No. 2. This instrument sampled temperature with a 0.25 h resolution.

95 2.3 HYbrid Coordinate Ocean Model (HYCOM)

This study mainly uses 32 days, covering all of June, of hourly output at 1,000 m depth from a global run of HYCOM, with 41 vertical layers and 1/25° horizontal resolution. The non data-assimilative simulation, designated 'GLBy190.04', includes realistic tidal and atmospheric forcing allowing the generation and propagation of IT within the eddying general circulation. The high resolution 2D fields are complemented by monthly-mean 3D fields of temperature and salinity subsampled to 1°.

100   A Lagrangian analysis of the simulation is also used for a direct model-data validation with Argo floats. The Argo quasi-Lagrangian sampling is mimicked by releasing 41644 particles randomly across the world oceans. We let the particles be advected by the 2D velocity field at 1,000 m for 32 days while sampling temperature with an hourly resolution. This Lagrangian sampling of HYCOM is achieved using the software Parcels (Van Sebille et al., 2021). As for the Argo data, we discard particles with a mean speed larger than 0.1 m s^{-1}.

105 3   Methods

We start by comparing the Argo observations to the Lagrangian sampling of the HYCOM data. Next, we investigate the effects of the drift by comparing Lagrangian and Eulerian samplings of the numerical simulation. We end the section by introducing the comparison between the Eulerian HYCOM time series and mooring observations.

3.1   Lagrangian sampling of the isopycnal displacement at 1,000 dbar

110 As in Geoffroy and Nycander (2022), we define the vertical isotherm displacement observed by a Lagrangian particle η_{1000}^{L} as

$$\eta_{1000}^{L}(t) = \frac{T(t) - \overline{T}}{(\mathrm{d}T/\mathrm{d}P)_{1000}(t)}, \tag{1}$$

where $\overline{T}$ is the time average of the temperature measurements $T(t)$ over a particle trajectory, and $(\mathrm{d}T/\mathrm{d}P)_{1000}(t)$ is the temperature gradient at 1,000 dbar, interpolated at the successive particle positions from the monthly-mean 3D temperature field. In the case of Argo floats, the temperature gradient at 1,000 dbar is obtained from the average of the two neighboring 115 profile measurements within 100 dbar of the parking pressure. The low frequency background activity is filtered out from

η_{1000}^L using a fourth order Butterworth filter with a cut-off frequency of 0.3 cpd. The inertial peak is not removed by this filter poleward of about $\pm 10°$ of latitude.

3.2 Averaging sample autocovariance series

The HYCOM derived Lagrangian time series η_{1000}^L can be analyzed in the same way as in Geoffroy and Nycander (2022). We illustrate below the mapping process. The geographical patch presented here is also described in Geoffroy and Nycander (2022). For this example area, Fig. 1 shows the location of the 8 32-day segments of Argo data and 13 HYCOM particles selected using their median position. The circular patch is centered on the mooring No. 2.
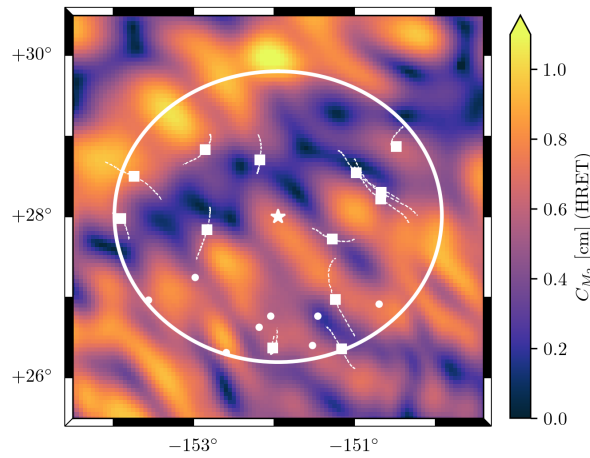
Figure 1. Example circular geographical patch of radius 200 km (the white star denotes mooring No. 2 at the center). The segments of Argo data and the HYCOM particles are shown by white filled circles and white squares, respectively. The dashed white curves represent the trajectories of the HYCOM particles over the 32 days of numerical simulation. In the background we show the amplitude of the M_2 baroclinic sea level anomaly from the High-Resolution Empirical Tide model (HRET; Zaron, 2019).

From a finite time series $\eta(t)$ one can calculate the sample autocovariance $\hat{R}(\tau)$, which is an unbiased estimate of the true autocovariance:

$$\hat{R}(\tau) = \frac{1}{N-\tau} \sum_{t=1}^{N-\tau} (\eta_t - \hat{\bar{\eta}})(\eta_{t+\tau} - \hat{\bar{\eta}}), \tag{2}$$

where N is the total number of observations, and $\hat{\bar{\eta}}$ is the sample mean of the series. Here, N is taken as the number of non-missing observations, accounting for gaps in the Argo time series (during the descent, main profiling, and surface transmission phases of the float cycle, or because of failed quality controls). Note that $\hat{R}(0)$ is the sample variance of the series. In the remainder of this article we only compute the variance and autocovariance of vertical isotherm displacement time series. For the sake of brevity, we simply refer to them as the variance and autocovariance.

We compute a sample autocovariance for each binned HYCOM particle and average them over the subset to obtain a local mean autocovariance $\overline{R}_{\mathrm{HYCOM}}^{L}$. The local mean autocovariance from the Argo data, $\overline{R}_{\mathrm{Argo}}$, is computed in the same way. The sampling error affecting a single autocovariance series is estimated from the standard deviation (STD) of the autocovariance series over the subset. The standard error of the local mean autocovariance is computed as

$$\mathrm{SEM}(\tau) = \frac{\mathrm{STD}(\tau)}{\sqrt{N_p}}, \tag{3}$$

where N_p is the number of particles in the subset. Fig. 2a demonstrates the good agreement of the two datasets until $\tau \simeq 200$ h. Past this time lag limit, $\overline{R}_{\mathrm{Argo}}$ falls under its 95% confidence interval (95% C.I. $\sim \pm 2\,\mathrm{SEM}$) and thus cannot be considered significantly different from white noise.

A convenient tool for monitoring the evolution of the autocovariance of the semidiurnal IT is complex demodulation. Here, it consists in the least squares fitting of $C\cos(\omega_{\mathrm{SD}}\tau) + S\sin(\omega_{\mathrm{SD}}\tau)$, where $\omega_{\mathrm{SD}} = (\omega_{M_2} + \omega_{S_2})/2$ is the semidiurnal frequency, to the sample autocovariance in 48-h windows with 75% overlap. This is equivalent to fitting $A\cos(\omega_{\mathrm{SD}}\tau + \Phi)$, where A and Φ are the amplitude and phase, respectively. We then define the complex demodulate as $A = \sqrt{C^2 + S^2}$. This positive definite amplitude follows the sinusoidal envelope of the modulated sinusoidal with frequency ω_{SD}. Note that this definition differs from the complex demodulation used in Geoffroy and Nycander (2022), where the authors fitted $A\cos(\omega_{\mathrm{SD}}\tau)$, i.e. with $\Phi = 0$. In practice, we compute the complex demodulate of the autocovariance following the harmonic analysis method of Cherniawsky et al. (2001). The error affecting the complex demodulate of a mean autocovariance series over a 48-h window can be obtained from the standard error of the mean autocovariance, as defined in Eq. (3), and the covariance matrix of the fitted parameters. In the remainder of this article we refer to it as the total error:

$$\mathrm{TE} = \sqrt{\widetilde{\mathrm{SEM}}^2 + \mathrm{Var}A}, \tag{4}$$

where $\widetilde{\mathrm{SEM}}$ is the median of the standard error of the mean autocovariance over the 48-h window, and $\mathrm{Var}A$ is the variance of the complex demodulate A in the same 48-h window. We computed $\mathrm{Var}A$ from the covariance matrix of the fitted parameters C and S, taking into account potential correlations between C and S, using the Uncertainties python package (Lebigot, 2022).

The result of the complex demodulation at the semidiurnal frequency applied to our two mean autocovariance series is presented in Fig. 2b. The first 48-h complex demodulate is a direct estimate of the total semidiurnal IT variance. From the Argo data plotted in Fig. 2b we obtain 23.5 m² with a 95% C.I. of [10.8, 36.2] m² (here, 95% C.I. $\sim \pm 2\,\mathrm{TE}$). For this local example, the first demodulate of $\overline{R}_{\mathrm{HYCOM}}^{L}$ is almost identical to the first demodulate of $\overline{R}_{\mathrm{Argo}}$. Taking into account the errorbars, the two demodulate series are not significantly different, even at longer time lags.

3.3 Eulerian perspectives

The decay of the demodulates with time lag observable in Fig. 2b mirrors the decorrelation captured by the Lagrangian sampling of the Argo floats. The motion of the instruments results in additional decorrelating effects that cannot be isolated from the loss of coherence of the IT. However, by analyzing the HYCOM data within a Eulerian framework, one can directly monitor the intrinsic decorrelation of the IT. The Eulerian HYCOM time series can then be compared with observations from moorings.
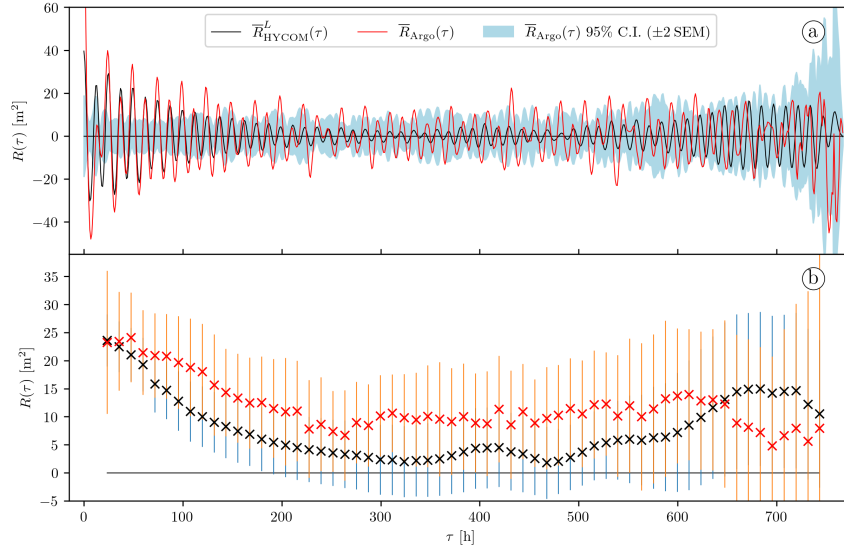
Figure 2. (a) Local mean autocovariance $\overline{R}^L_{\mathrm{HYCOM}}$ computed from the HYCOM particles (solid black curve), and local mean autocovariance $\overline{R}_{\mathrm{Argo}}$ computed from the 32-day segments of Argo data (solid red curve) and 95% confidence interval of $\overline{R}_{\mathrm{Argo}}$ (light blue shading). The data are from the geographical patch shown in Fig. 1. The Argo variance lies above the figure scale, here $\overline{R}_{\mathrm{Argo}}(0) \simeq 121 \ \mathrm{m}^2$. (b) Complex demodulates at the semidiurnal frequency of the autocovariance series shown in (a) and their 95% C.I.

In the Eulerian framework, our methods remain practically unchanged. We now define the vertical isotherm displacement at a given location as

$$\eta^E_{1000}(x,y,t) = \frac{T(x,y,t) - \overline{T}(x,y)}{(\mathrm{d}T/\mathrm{d}P)_{1000}(x,y)} \tag{5}$$

where $\overline{T}(x,y)$ is the time average of the temperature field $T(x,y,t)$, and $(\mathrm{d}T/\mathrm{d}P)_{1000}(x,y)$ is the temperature gradient at 1,000 dbar, interpolated from the monthly-mean 3D temperature field. As for the Lagrangian time series, we apply a fourth order Butterworth filter with a cut-off frequency of 0.3 cpd on η^E_{1000}.

For each HYCOM particle, we compute 32-day long time series of η^E_{1000} at the successive positions of the particle subsampled at a 12-h rate. We then calculate the sample autocovariance from each of these time series and average the results over the particle's trajectory. As in the Lagrangian framework, the resulting averages can be averaged further to compute local and global mean autocovariance series.

Eq. (5) is also used to compute vertical displacement time series from the mooring temperature records. In this case, the temperature gradient at 1,000 dbar is computed from the annual mean climatology (WOA18; Boyer et al., 2018) as an average within 100 dbar of the instrument depth. Each mooring time series is split into successive 32-day segments. We then compute the sample autocovariance for each high pass filtered segment and average them to obtain a mean autocovariance.

The Eulerian sampling serves two main purposes: (*i*) validating the total variance of the IT measured by the Lagrangian particles, and (*ii*) comparing the decorrelation of the IT in the HYCOM data to mooring observations. We illustrate these two aspects for the local example introduced in section 3.2 in Fig. 3, and 4, respectively.
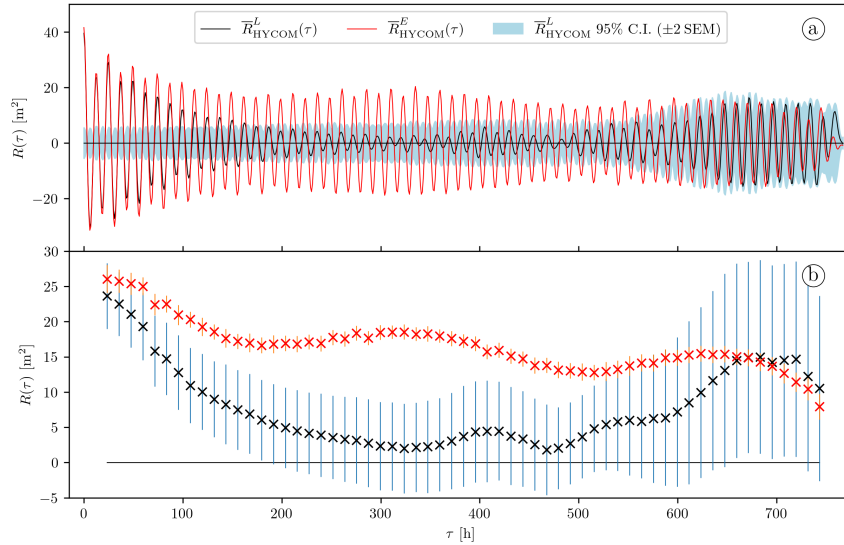
Figure 3. (a) Mean autocovariance computed from η_{1000}^{L} (solid black curve) and 95% confidence interval (light blue shading), and mean autocovariance computed from η_{1000}^{E} (solid red curve). The data are from the geographical patch shown in Fig. 1. (b) Complex demodulates at the semidiurnal frequency of the autocovariance series shown in (a) and their 95% C.I.

180   Fig. 3 shows the local mean autocovariance series at 1,000 dbar computed from both the Lagrangian η_{1000}^{L} (solid black curve) and Eulerian η_{1000}^{E} (solid red curve). As expected, the two autocovariance series demonstrate a close agreement at short time lags, before the motion of the particles causes $\overline{R}_{\mathrm{HYCOM}}^{L}$ to decay faster. In contrast, $\overline{R}_{\mathrm{HYCOM}}^{E}$ remains close to the mean autocovariance computed from the mooring No. 2 time series $\overline{R}_{\mathrm{Moor}}$ (see Fig. 4). In conclusion, for this local example, the model agrees very well with the in situ observations, both in terms of total variance and decorrelation of the semidiurnal IT at 185   1,000 dbar.

4   Results

4.1   The total semidiurnal IT at 1,000 dbar

We bin the global collection of HYCOM particles based on their median position, using circular geographical patches of radius 200 km centered on a regular $2.5° \times 2.5°$ grid. Here, we use only particles for which (*i*) the mean speed is lower than $0.1\ \mathrm{m\ s^{-1}}$ 190   (to avoid contamination by lee waves), and (*ii*) the variance of η_{1000} is lower than $2 \times 10^{4}\ \mathrm{m^{2}}$ (outliers). As in sections 3.2 and 3.3, we compute an autocovariance series for each particle, in the Lagrangian and Eulerian framework, separately. Then,
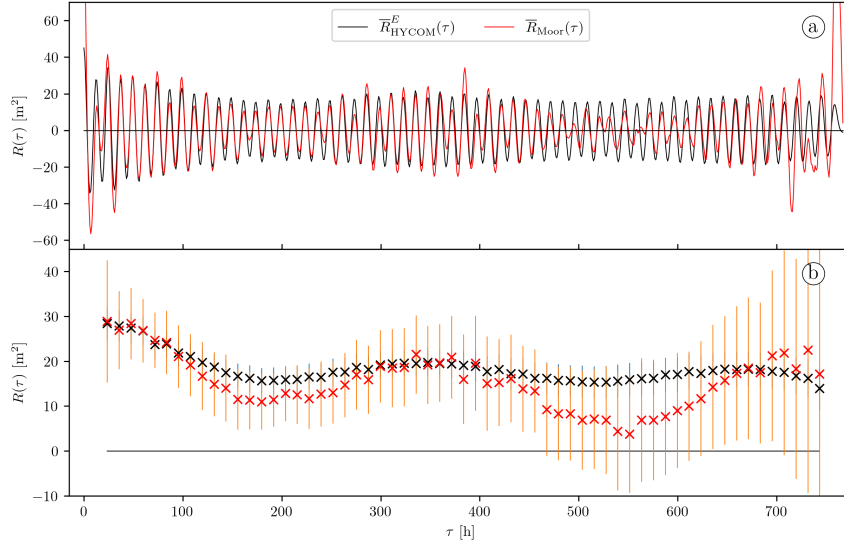
Figure 4. (a) Mean autocovariance computed from η^E_{1000} (solid black curve), and mean autocovariance computed from eleven 32-day segments of the mooring No. 2 time series (solid red curve). (b) Complex demodulates at the semidiurnal frequency of the autocovariance series shown in (a) and their 95% C.I.

we average these autocovariances over the corresponding patches to obtain local mean autocovariance series $\overline{R}^L_{\mathrm{HYCOM}}$ and $\overline{R}^E_{\mathrm{HYCOM}}$.

We start by checking how the Lagrangian sampling affects the total semidiurnal IT variance. Fig. 5 shows the scatter of the first 48-h complex demodulate of $\overline{R}^E_{\mathrm{HYCOM}}$ as a function of the first demodulate of $\overline{R}^L_{\mathrm{HYCOM}}$ for our collection of geographical patches. The agreement is close to perfect (Pearson's $R^2 \simeq 1.0$ in log-log domain). Thus, the motion of the Lagrangian particles, and therefore of the Argo floats, have no impact on the measured total variance of the IT.

We can then map the total semidiurnal IT variance (here taken as the first 48-h complex demodulate of $\overline{R}^L_{\mathrm{HYCOM}}$) and the associated total error, as defined in Eq. (4), on a $2.5° \times 2.5°$ grid (see Fig. 6a and 6b, respectively). In each figure we show only the bins which yield a total IT variance larger than one total error. Fig. 6a can be directly compared with the global map of the total semidiurnal IT variance computed from Argo data, (see Fig. 6c and 6d, updated from Geoffroy and Nycander, 2022). As documented in Buijsman et al. (2020), HYCOM is known to be subject to a thermobaric instability (TBI) in the north Pacific (dashed black rectangle in Fig. 6a). Since only a few patches of Argo data are available at the edge of this TBI area, we do not exclude it from the subsequent analysis.

The main patterns visible in Fig. 6a and Fig. 6c broadly agree, with energy radiating away from low-mode IT generation hotspots, namely near Madagascar, Hawaii, and the tropical south and southwest Pacific. In Fig. 7a we show a scatter plot of the Argo derived semidiurnal IT variance as a function of the simulated one. The two datasets are well correlated (Pearson's $R^2 = 0.52$ in the log domain), but one can easily identify a systematic bias.
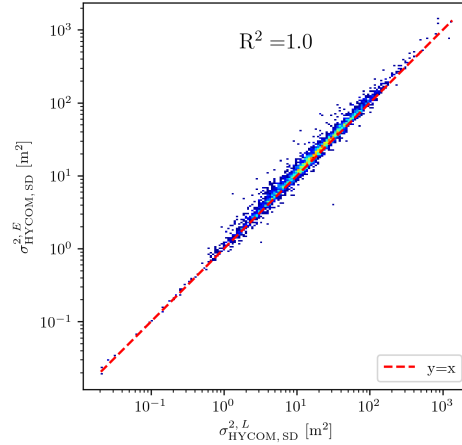
Figure 5. First 48-h complex demodulate of $\overline{R}_{\mathrm{HYCOM}}^{E}$ as a function of the first demodulate of $\overline{R}_{\mathrm{HYCOM}}^{L}$ for the unmasked bins in Fig. 6a. A warmer color indicates a denser scatter.
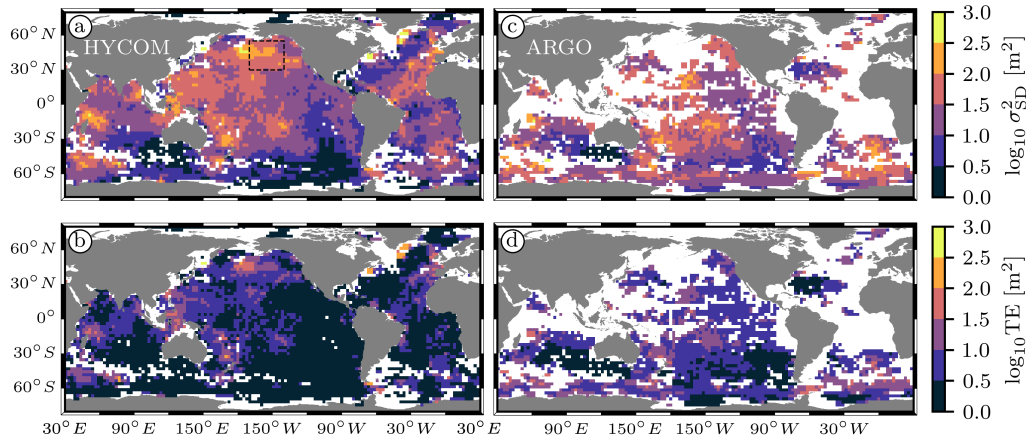
Figure 6. (a) Atlas of the total semidiurnal IT variance computed as the first 48-h complex demodulate of the local mean autocovariance series at 1,000 dbar ($\overline{R}_{\mathrm{HYCOM}}^{L}$), and (b) corresponding total error computed from Eq. (4). (c) and (d) Same as (a) and (b), respectively, but computed from the Argo data ($\overline{R}_{\mathrm{Argo}}$). (c) and (d) are updated from Geoffroy and Nycander (2022). The area where the simulation is affected by the TBI is shown by the dashed black rectangle in (a).

We investigate this bias by plotting the geographical distribution of the HYCOM to Argo semidiurnal IT variance ratio (see Fig. 7b). The ratio is relatively homogeneous globally (in the range $[0.5, 2]$), except over the Southern Ocean where the Argo inferred total variance is significantly larger. Furthermore, we show the zonal mean of the Argo and HYCOM derived semidiurnal IT variance as a function of latitude in Fig. 7c. The zonal mean variance from Argo is generally larger, except around 40° N where the HYCOM inferred zonal mean variance peaks at 1.7 times the Argo one. Discarding the data in the TBI

area has virtually no effect on this peak (not shown). More noteworthy than this localized feature, the ratio of the zonal mean

215 variances increases approaching the poles (poleward of $\pm 50°$ of latitude). In contrast, equatorward of $\pm 50°$, the ratio remains

fairly constant, oscillating around a mean value of 0.74 with a standard deviation of 0.27. The global mean ratio is 0.60 with a

standard deviation of 0.34.
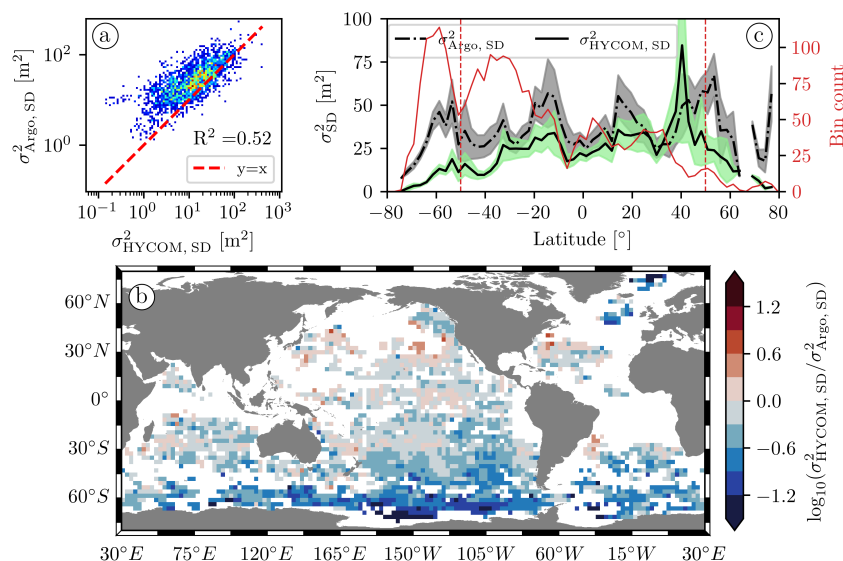
Figure 7. (a) Scatter plot of the Argo derived semidiurnal IT variance as a function of the simulated one for the collection of geographical bins shown in Fig. 7b. A warmer color indicates a denser scatter. (b) Atlas of the HYCOM to Argo semidiurnal IT variance ratio at 1,000 dbar. Note that the colorbar is in log of the ratio, for reference $\log_{10}^{-1}(0.3) = 2$. (c) Zonal mean of the semidiurnal IT variance from Argo (dash-dotted black curve) and HYCOM (solid black curve) as a function of latitude and their respective 95% C.I. (gray and light green shading, respectively). The red curve represents the number of geographical bins used to compute the zonal mean at a given latitude (red axis on the right). The vertical dashed red lines are placed at $50°$ S and $50°$ N.

The representativity of the zonal mean variances is smaller north of $40°$ N, because of the scarcity of available data (solid red curve in Fig. 7c). We therefore focus on the pronounced discrepancy affecting the Southern Ocean. We gather the data

220 available over the unmasked bins in Fig. 7b into two groups, north and south of $50°$ S. For each group and the global collection of bins, we compute a mean autocovariance by averaging the corresponding Argo and HYCOM local mean autocovariance series (see Fig. 8). A significant fraction of the divergence visible in the global mean autocovariance at short time lags (see Fig. 8a and 8b) can be explained by the larger discrepancy south of $50°$ S (see Fig. 8e and 8f). In this region, the Argo derived total semidiurnal IT variance is close to 4 times larger than the simulated one. The agreement is better north of $50°$ S, where

225 the corresponding factor is 1.5 (see Fig. 8c and 8d).
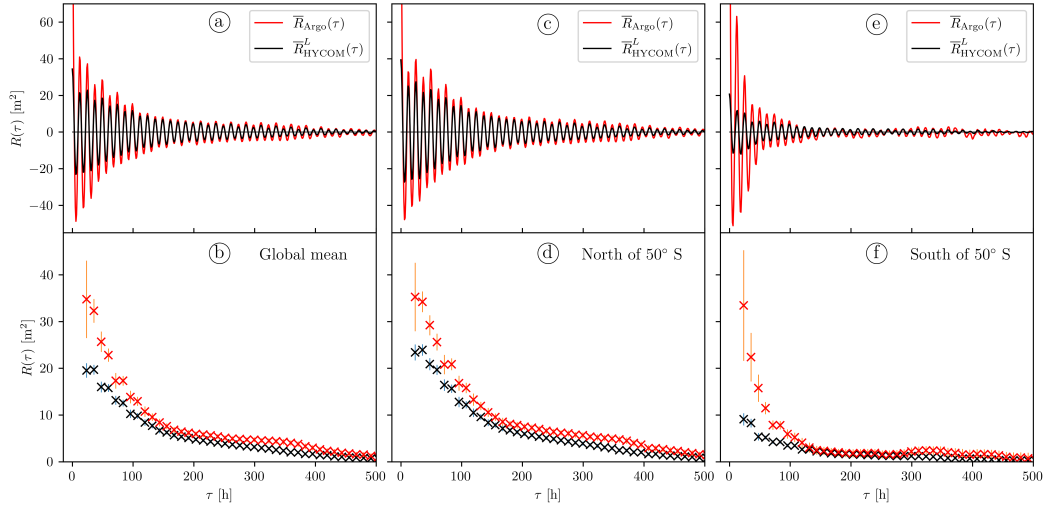
Figure 8. (a) Global mean autocovariance at 1,000 dbar computed from Argo data (solid red curve) and Lagrangian HYCOM particles (solid black curve), and (b) corresponding complex demodulates at the semidiurnal frequency and their 95% C.I. (c) and (d) Same as (a) and (b), respectively, but using only data north of $50°$ S. (e) and (f) Same as (a) and (b), respectively, but using only data south of $50°$ S. In this figure we truncated τ at 500 h, since past this limit the autocovariance series are undistinguishable from 0.

4.2 The decorrelation of the semidiurnal IT at 1,000 dbar

In contrast to Argo floats, the Eulerian sampling of moorings allows us to directly monitor the intrinsic decorrelation of the IT. In a procedure similar to that in Sect. 4.1, we bin the global collection of HYCOM particles based on their median position, but this time using geographical patches (of radius 200 km) centered on the available moorings. We compute a

230   sample autocovariance for each particle in the Eulerian framework, and average them within the corresponding patches to obtain local mean autocovariance series. Again, we use only particles for which (*i*) the mean speed is lower than 0.1 m s^{-1} (to avoid contamination by lee waves), and (*ii*) the variance of η_{1000} is less than 2×10^4 m^2 (outliers). After some additional quality controls, mainly discarding bins where the complex demodulates fall under one total error within 15 days of time lag, we are left with 167 moored instruments and the corresponding patches of simulated data.

235   For all the datasets used in this study, we found the probability distribution of the global collection of local mean autocovariance (and their demodulates) at any given timelag to be right skewed (not shown). This is not an issue when computing average statistics from the Argo and the corresponding simulated data, since the number of samples (i.e., the number of geographical bins) is very large and the sample mean is therefore expected to be normally distributed, by virtue of the central limit theorem. In contrast, geographical bins where mooring data are available are fewer. Thus, the influence of the tail of the distribution

240   on the sample mean is larger when analyzing the relatively small collection of bins where mooring data are available than when considering the global collections of Argo and HYCOM data. This precludes the use of statistics that assume a normal distribution when describing regional or even global averages of the autocovariances computed from moorings. To limit the

effects of skewness, we discard bins for which either the mooring or the simulated first 48-h demodulate is above the 95th percentile of its observed distribution (here $P_{\text{Moor}}^{95} \sim 720 \text{ m}^2$ and $P_{\text{HYCOM}}^{95} \sim 270 \text{ m}^2$, respectively). The latter criterion accounts

245 for 9 additional discarded bins, all located in moderate to strong mesoscale activity areas. Note that even after discarding these extreme samples, the data remain highly skewed.

As a measure of the strength of the decorrelation affecting the Eulerian mean autocovariances, we define the semidiurnal coherent variance fraction (SCVF$_{15}$) as the ratio between the 48-h demodulate at $\tau \simeq 15$ days (i.e., half the spring-neap period) and the first demodulate. Note that, due to the limiting duration of our time series, we cannot distinguish a strong but short from

250 a weak but long decorrelating process. Fig. 9a shows the scatter plot of the SCVF$_{15}$ computed from the moored instruments as a function of the SCVF$_{15}$ from the corresponding HYCOM local mean autocovariance series. Although the correlation is small (Pearson's $R^2 = 0.13$ in the log domain), the distribution is well centered on the $y = x$ line. In Fig. 9b and 9c we show the geographical location of the patches along with the ratio of the SCVF$_{15}$ computed from the HYCOM and mooring data, and the histogram of this ratio, respectively.
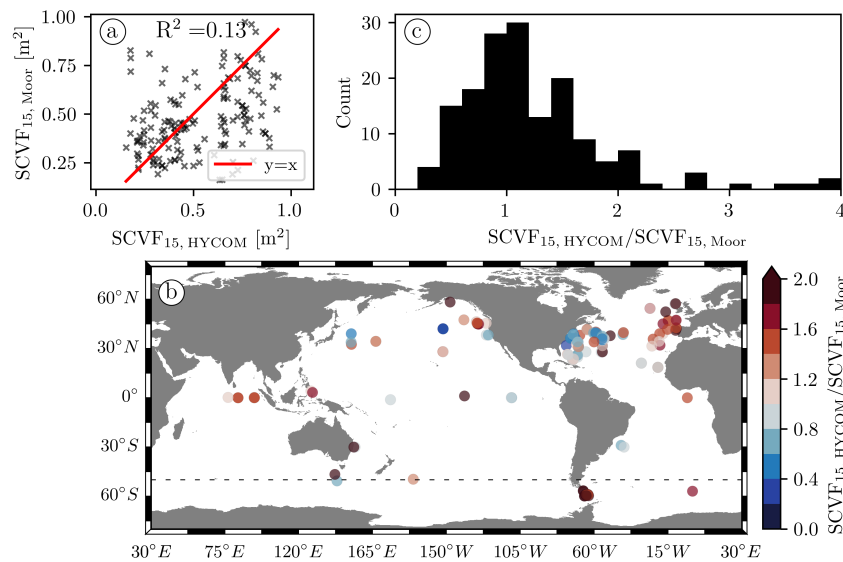
Figure 9. (a) Scatter of the SCVF$_{15}$ computed from the moored instruments as a function of the SCVF$_{15}$ computed from the HYCOM local mean autocovariance series. (b) Map of the 158 geographical bins presented in (a) along with the ratio of the SCVF$_{15}$ computed from the HYCOM and mooring data, and (c) the histogram of this ratio. The dashed black line in (b) indicates $50°$ S.

255 Fig. 10a displays the global mean autocovariances calculated from the mooring ($\overline{R}_{\text{Moor}}$, solid red curve) and HYCOM data ($\overline{R}_{\text{HYCOM}}$, solid black curve), as the average of the local mean autocovariance series over the collection of geographical bins presented in Fig. 9b. In Fig. 10b we show different statistics of the observed distribution of the demodulates of the local mean autocovariances in the form of boxplots as a function of time lag. Fig. 10b suggests that the mooring records exhibit both a larger total semidiurnal IT variance and a stronger decorrelation on a global average: the mean of the first demodulates and the

260 mean SCVF$_{15}$ are 83 m^2 and 0.50, and 40 m^2 and 0.64, for the moorings and HYCOM data, respectively. Using median values

instead leads to the same conclusions (not shown).

As in Sect. 4.1, we investigate this discrepancy by plotting the mean autocovariance series computed separately from the ge-

ographical bins lying north (144 instruments) and south (14 instruments) of 50° S (see Fig. 10c and 10d, and Fig. 10e and 10f,

respectively). The total mean variance and the mean SCVF$_{15}$ for each group are shown in Table 1. Again, the divergence be-

265 tween the two datasets appears enhanced south of 50° S. Moreover, contrarily to moorings, the HYCOM SCVF^{15} is noticeably

larger in the Southern Ocean. This indicates a weaker, or slower, decorrelation of the IT in HYCOM.

The spatial distribution of the moorings is sparse and tends to be denser in particular areas (e.g. the Gulf Stream region).

This is all the more true south of 50° S , where much fewer moorings are available than in the northern region, and where they

are mostly located in the Drake Passage. For both these reasons, the mean autocovariances computed from moorings north and

270 especially south of 50° S cannot be considered truely representative of the IT in these vast regions. Nonetheless, they remain

representative of the collection of geographical bins used to construct them.

Table 1. Summarizing numerics of the autocovariance series plotted in Fig. 10.

		Mooring	HYCOM	HYCOM/Mooring
Global	σ^2_{SD}	83 m^2	40 m^2	0.48
	SCVF$_{15}$	0.50	0.64	1.28
North of 50° S	σ^2_{SD}	71 m^2	36 m^2	0.51
	SCVF$_{15}$	0.55	0.62	1.13
South of 50° S	σ^2_{SD}	213 m^2	84 m^2	0.40
	SCVF$_{15}$	0.35	0.75	2.16

4.3 Apparent decorrelation

In Sect. 4.1 we demonstrated that the motion of the floats has no impact on the measured total variance of the IT. However, the

Doppler effect and spatial decorrelation induced by the Lagrangian sampling of the floats both act as decorrelating processes

275 causing the autocovariance to decay at longer time lags (Geoffroy and Nycander, 2022). Following Caspar-Cohen et al. (2022)

we call this mecanism apparent decorrelation, as it is unrelated to the intrinsic decorrelation of the propagating IT. Geoffroy

and Nycander (2022) estimated the apparent decorrelation timescale to be longer than that of the observed Lagrangian auto-

covariance function, on a global average. Thus, they concluded that the decay of the global mean autocovariance observed by

Argo floats is primarily due to the intrinsic decorrelation of the IT.

280 The HYCOM data allow to study this by directly comparing the global mean autocovariance computed in the Lagrangian and

Eulerian frameworks (see Fig. 11a and 11b). As expected, the Lagrangian ($\overline{R}^L_{\text{HYCOM}}$, solid black curve) and Eulerian ($\overline{R}^E_{\text{HYCOM}}$,

solid red curve) autocovariances are virtually identical until $\tau \simeq 50$ h. Past this limit, the more rapid decay of $\overline{R}^L_{\text{HYCOM}}$ can
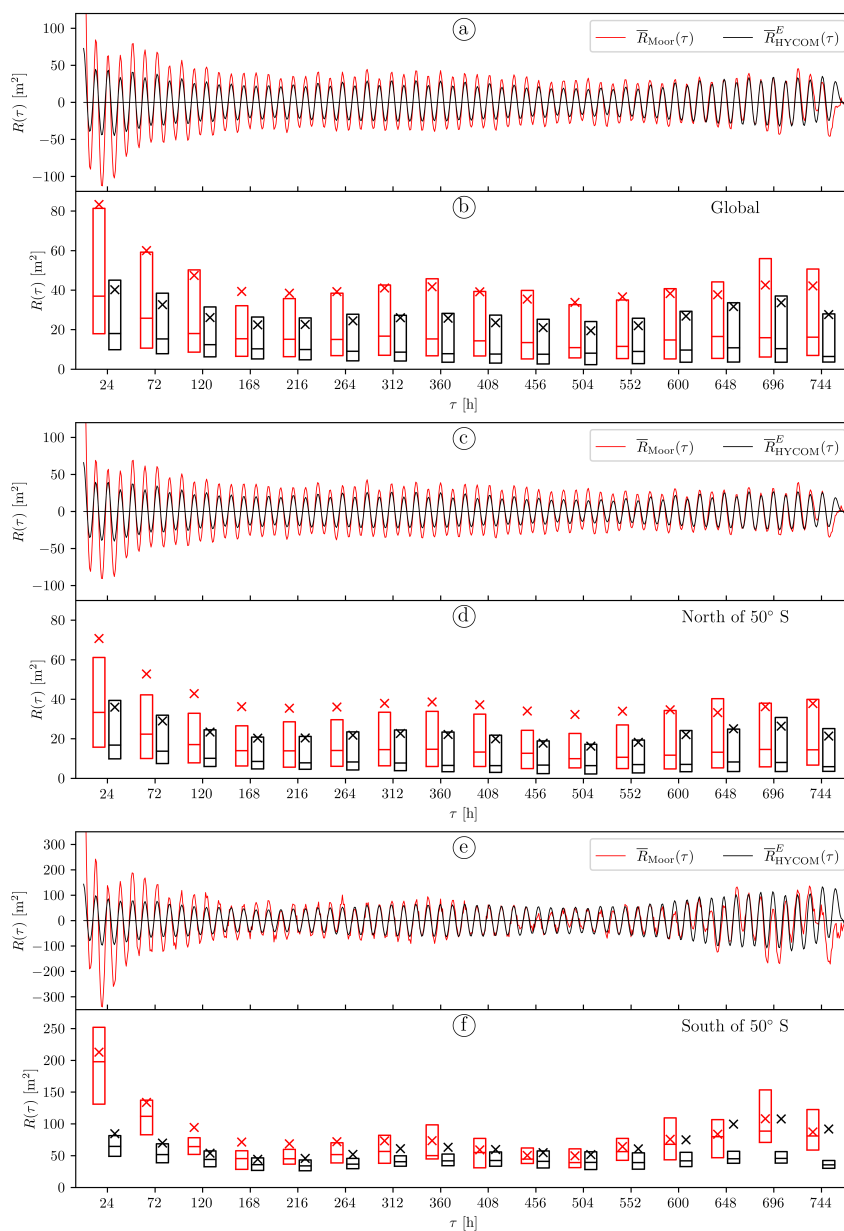
Figure 10. (a) Global mean autocovariance computed as the average of the local mean autocovariance series from the HYCOM particles over the geographical bins presented in Fig. 9b (solid black curve), and global mean autocovariance from the moored instruments computed in the same way (solid red curve); (b) Boxplots of the observed distribution of the complex demodulates at the semidiurnal frequency of the local mean autocovariances for the collection of geographical bins presented in Fig. 9b as a function of time lag. Each boxplot is constituted by a rectangle extanding from the first quartile to the third quartile of the data, with a line at the median and a cross at the mean. For a given time lag, the red and black boxplots, offset on either side of the time lag value, represent the distribution of the demodulates computed over the 48-h window centered on that time lag value from moorings and HYCOM, respectively. (c) and (d) Same as (a) and (b), respectively, but using only the 144 geographical bins north of $50°$ S. (e) and (f) Same as (a) and (b), respectively, but using only the 14 geographical bins south of $50°$ S.

only be caused by the apparent decorrelation, while $\overline{R}_{\mathrm{HYCOM}}^{E}$ continues to solely reflect the intrinsic decorrelation of the IT. Our aim in the present section is to find estimates for the characteristic timescales of the apparent and intrinsic decorrelation in

285   HYCOM and compare it with observations.
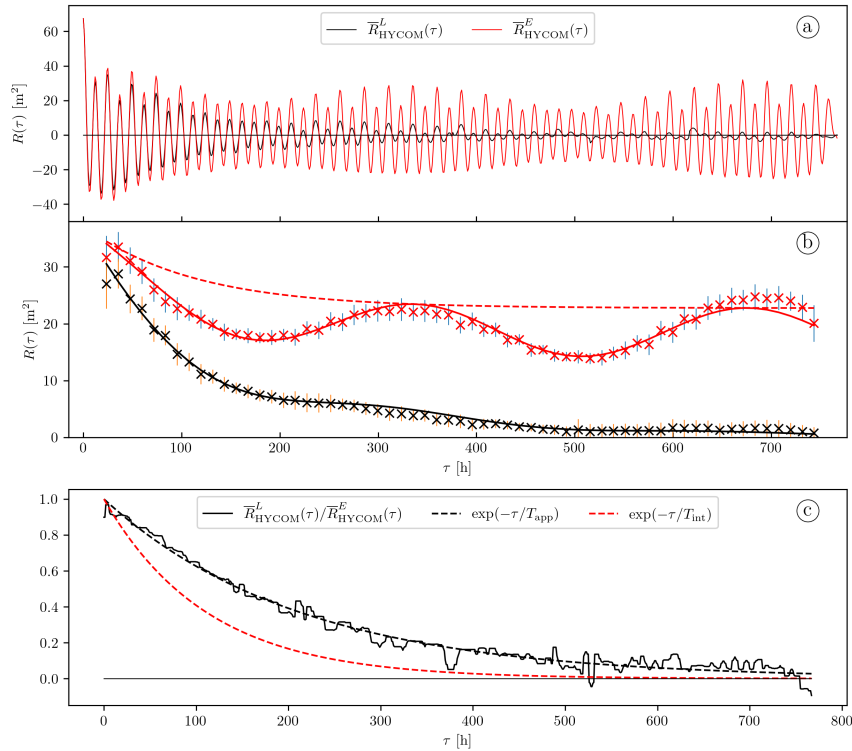
Figure 11. (a) Global mean autocovariance at 1,000 dbar computed from the global collection of HYCOM derived η^{L} (solid black curve) and η^{E} (solid red curve) time series, and (b) corresponding complex demodulates at the semidiurnal frequency and their 95% C.I. The solid and dashed red curves represent the result of fitting the model Eq. (6) to the Eulerian demodulates and the underlying exponential decay due to the intrinsic decorrelation, respectively. The solid black curve corresponds to the model Eq. (7) with the different parameters set as described in the text. (c) Median filtered ratio of $\overline{R}_{\mathrm{HYCOM}}^{L}$ to $\overline{R}_{\mathrm{HYCOM}}^{E}$ (solid black curve) and fitted decaying exponential with the apparent decorrelation timescale (dashed black curve). For reference, we overlay a decaying exponential with the intrinsic decorrelation timescale obtained by fitting the model Eq. (6) to the Eulerian demodulates (dashed red curve).

Taking inspiration from Geoffroy and Nycander (2022) and Caspar-Cohen et al. (2022), we try a simple model for the complex demodulate at the semidiurnal frequency of the Eulerian autocovariance computed over a time lag window centered on τ':

$$\mathcal{C}_{\mathrm{SD}}^{E}(\tau') = \sigma_{\mathrm{SD}}^{2}\big(\alpha + (1-\alpha)\exp(-\tau'/T_{\mathrm{int}})\big) + \sigma_{\mathrm{AM}}^{2}\big(\cos(\omega_{\mathrm{AM}}\tau') - 1\big). \tag{6}$$

290   Here, σ_{SD}^{2} is the total semidiurnal internal tide variance, α is the stationary fraction, T_{int} is the intrinsic decorrelation timescale, and σ_{AM}^{2} and ω_{AM} are the variance and angular frequency of an amplitude modulating sinusoidal, respectively. A heuristic

model for the demodulate of the corresponding Lagrangian autocovariance is then

$$\mathcal{C}_{\mathrm{SD}}^L(\tau') = \mathcal{C}_{\mathrm{SD}}^E(\tau') \exp(-\tau'/T_{\mathrm{app}}), \tag{7}$$

where T_{app} is the apparent decorrelation timescale.

295   A constrained least squares fit of the model Eq. (6) to the complex demodulate series computed from $\overline{R}_{\mathrm{HYCOM}}^E$ (red crosses in Fig. 11b) yields $T_{\mathrm{int}} = 112$ h. The fitted model as well as the exponential decay due to the intrinsic decorrelation (i.e., only the term proportional to σ_{SD}^2) are represented in Fig. 11b as the solid and dashed red curves, respectively. On a global average, the complete decorrelation of the nonstationary IT in HYCOM is therefore achieved within $3T_{\mathrm{int}} \sim \mathcal{O}(400)$ h, i.e. well under the 32 days of data.

300   By dividing $\overline{R}_{\mathrm{HYCOM}}^L$ by $\overline{R}_{\mathrm{HYCOM}}^E$, we can isolate the effects of the apparent decorrelation on $\overline{R}_{\mathrm{HYCOM}}^L$. In Fig. 11c we plot this ratio after applying a median filter with a window of 18 h (solid black curve). A least squares fit of a simple decaying exponential to the obtained curve yields an apparent decorrelation timescale $T_{\mathrm{app}} = 213$ h (dashed black curve). For verification, we compute the Lagrangian $\mathcal{C}_{\mathrm{SD}}^L$ from Eq. (7), by multiplying the fitted model of $\mathcal{C}_{\mathrm{SD}}^E$ (solid red curve Fig. 11b) by the fitted exponential decay due to the apparent decorrelation (dashed black curve in Fig. 11c). The result (solid black curve in Fig. 11b)
305   closely follows the demodulated Lagrangian autocovariance (black crosses in Fig. 11b). The different parameters obtained from the Eulerian and Lagrangian simulated data are summarized in Table 2. Note that the amplitude modulating sinusoide is closely related to the spring-neap cycle, here $\omega_{\mathrm{AM}} \sim |\omega_{\mathrm{M}_2} - \omega_{\mathrm{S}_2}| = 0.0177\,\mathrm{h}^{-1}$.

Table 2. Summary of the parameters estimated from the simulated Eulerian and Lagrangian data. These values are used to compute $\mathcal{C}_{\mathrm{SD}}^L$ from the model Eq. (7) (solid black curve in Fig. 10b).

σ_{SD}^2	α	T_{int}	σ_{AM}^2	ω_{AM}	T_{app}
$37\,\mathrm{m}^2$	0.61	112 h	$4\,\mathrm{m}^2$	$0.0186\,\mathrm{h}^{-1}$	213 h

Lastly, we fit the model Eq. (7) and (6) to the demodulates of the global mean autocovariance computed from all the Argo floats available (see Fig. 12, updated from Geoffroy and Nycander, 2022). We emphasize that the geographical coverage of
310   the Argo data is less than that of HYCOM, but we are simply interested in a qualitative comparison of the two. The fitted parameters are gathered in Table 3. The value of T_{app} and T_{int} obtained here, from Argo data only, are similar to that reported by Geoffroy and Nycander (2022), where the authors estimated it from a comparable collection of Argo floats and HRET (in particular, the stationary limit was determined from HRET instead of by fitting). As from the simulated data, we recover a T_{int} shorter than T_{app}. While the Argo T_{app} is almost identical to the HYCOM value, T_{int} is about twice smaller.

Table 3. Parameters from the fitting of the model Eq. (7) to the demodulates of the global mean autocovariance computed from Argo data (see Fig. 12).

σ_{SD}^2	α	T_{int}	σ_{AM}^2	ω_{AM}	T_{app}
$48\,\mathrm{m}^2$	0.44	54 h	$3\,\mathrm{m}^2$	$0.0181\,\mathrm{h}^{-1}$	220 h
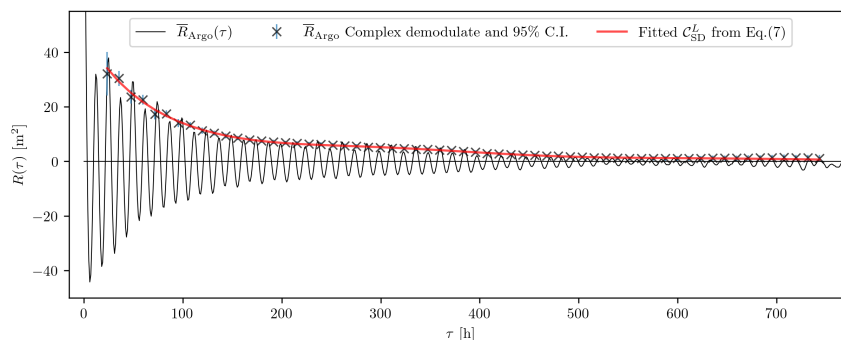
Figure 12. Global mean autocovariance at 1,000 dbar computed from all the Argo floats available (solid black curve), and corresponding complex demodulates at the semidiurnal frequency and their 95% C.I. (black crosses). The solid red curve represents the result of fitting the model Eq. (7) and (6) to the demodulates.

315   Similarly to Geoffroy and Nycander (2022), we conclude that the intrinsic decorrelation is more rapid than the apparent decorrelation on a global average. Yet, according to the simulated data, it is only more rapid by a factor of 2 (a factor of 4 according to the Argo data). Hence, the intrinsic decorrelation typically is at least as important as the apparent one over the first few days of time lag. This result might not hold true everywhere, since the geographical variability of T_{int} and α is expected to be large.

320   **4.4   Potential sources of bias**

Why is the IT variance lower and IT intrinsic decorrelation weaker in HYCOM than in the observations, particularly in the Southern Ocean ? At the time of writing we cannot think of a particular reason for either the Argo or the mooring derived IT variance to be biased high globally. We did investigate whether the bias in the Southern Ocean could be related to the contamination of the first 48-h demodulate at ω_{SD} by near-inertial waves as we approach the M_2 critical latitude (where

325   $f = \omega_{\mathrm{M}_2}$, at about 74° S). To check this, we can map the semidiurnal IT variance from the Argo data anew (as shown in Fig. 6c), this time fitting an additional $\cos(f\tau)$, where f is the local Coriolis frequency, to the local mean autocovariances. The result of the fit becomes unstable at 74° S, but the zonal mean of the demodulates at f does reach a maximum around 60° S while the zonal mean of the demodulates at ω_{SD} remains unaffected (not shown). We conclude that the first 48-h demodulate at the semidiurnal frequency is not affected by near-inertial waves, at any latitude.

330   As far to the model, equatorward of $\pm 50°$ latitude and for seafloor depths deeper than 1250 m, the horizontal grid spacing allows the correct resolution of vertical modes 1 to 5 (Buijsman et al., 2020). Approaching the poles, the number of layers below the mixed layer, and hence the vertical resolution, decreases. This further limits the number of resolved modes in HYCOM south of 50° S: roughly, modes 3 and 2 are only partially resolved poleward of 60° S and 65° S, respectively. Although the bulk of the IT variance at 1,000 dbar is likely related to mode-1 waves on a global average (Geoffroy and Nycander, 2022),

335   the contribution from higher modes can become significant locally. In principle, both Argo floats, moorings, and HYCOM data

include the effect of higher modes, but these are less well resolved in HYCOM, particularly in the Southern Ocean. Together with the general idea that higher-mode IT are less coherent (Egbert and Ray, 2017), this may explain the lower IT variance and weaker IT intrinsic decorrelation in HYCOM than in the observations, both on a global average and more specifically in the Southern Ocean.

340   The mode-m vertical structure of the isopycnal displacement $\Phi_m(z)$ is obtained by solving the Sturm-Liouville problem

$$\frac{\mathrm{d}^2\Phi_m(z)}{\mathrm{d}z^2} + \frac{N^2(z)}{c_m^2}\Phi_m(z) = 0, \tag{8}$$

with the boundary conditions $\Phi_m(0) = \Phi_m(-H) = 0$, where H is the ocean depth, $N(z)$ is the buoyancy frequency profile, and $-1/c_m^2$ is the eigenvalue corresponding to the eigenfunction $\Phi_m(z)$ for mode-m (Gill, 1982). The modal partitioning of the IT energy at a given location is mainly determined by the conversion rate (both local and remote) and lifetime of each mode

345   (de Lavergne et al., 2019). Additionally, depending on the local stratification and ocean depth, the parking depth at 1,000 dbar can be more or less close to the anti-node (point of maximal displacement) and node (point of no displacement) of the different vertical modes. Therefore, the normalized contribution of mode-m relative to mode-1 waves to the variance recorded at this depth is weighted by a coefficient γ_{m1}. In Appendix B of Geoffroy and Nycander (2022), the authors derived an expression for this coefficient:

350    $$\gamma_{m1}(z) = \frac{\Phi_m^2(z)\int_{-H}^{0} N^2\Phi_1^2\mathrm{d}z}{\Phi_1^2(z)\int_{-H}^{0} N^2\Phi_m^2\mathrm{d}z}. \tag{9}$$

In Fig. 13a and 13b we plot the global maps of γ_{21} and γ_{31} at 1,000 dbar, respectively, computed from Eq. (9) after solving the eigenvalue problem Eq. (8) for the $1/4°$ WOA18 summer climatology. Here, we extended the climatology down to the reference bathymetry from the General Bathymetric Chart of the Oceans 2019 (GEBCO, 2019), wherever deeper, by appending the deepest valid record. A visual comparison with Fig. 7b in the Southern Ocean suggests that low HYCOM to Argo IT

355   variance ratios (darker blue pixels in Fig. 7b) spatially coincide with large γ_{31}, mostly, or γ_{21} to some extent (especially in the Weddell Sea). South of $60°$ S, either γ_{31} or γ_{21} is larger than unity, hence the contribution from mode-3 or mode-2 waves, respectively, to the IT variance recorded at 1,000 dbar is magnified compared to the contribution from mode-1 waves. However, modes 2 and up are not well resolved in HYCOM at these latitudes.

The magnification of the contributions from modes 2 and 3 to the variance at 1,000 dbar in the Southern Ocean only affects

360   how propagating IT are perceived at the Argo parking depth. This has no connection with the generation and dissipation processes that set the underlying modal partitioning of the IT energy. Additional explanations might be found by examining whether the main parameters affecting the generation of IT (namely the bottom topography, barotropic tidal forcing, and stratification) are less accurate in this region than in the rest of the globe.

The pattern of the enhanced discrepancies between Argo and HYCOM in the Southern Ocean (darker blue in Fig. 7) could

365   be correlated with the distribution of bathymetric features. For instance, the large values around $45°$ S, $105°$ W in the South Pacific are centered on the Chile Rise, a known IT generation area. To date, only about 19% of the global ocean seafloor has been mapped using shipborne techniques. Therefore, global bathymetric products widely rely on depth prediction from satellite gravimetry (Smith and Sandwell, 1994). Small scale features such as abyssal hills are not resolved by this technique.
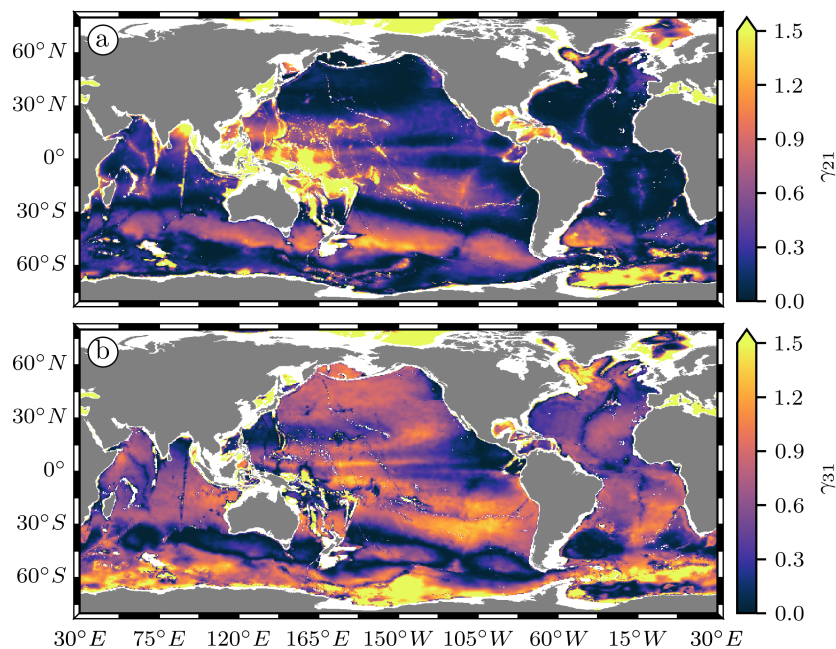
Figure 13. (a) Weight of the normalized contribution of mode-2 relative to mode-1 waves to the IT variance recorded at 1,000 dbar, computed from the WOA18 stratification. (b) Same as (a) but for the relative contribution of mode-3 waves.

In the recent SRTM15+ bathymetry (Tozer et al., 2019), gravity predicted depths were estimated to have root-mean-square
370 (RMS) uncertainties and maximum error of the order of ± 150 m and 1800 m, respectively (based on 50 cruises distributed globally). With about 22% seafloor coverage by high quality direct measurements, the International Bathymetric Chart of the Southern Ocean v2 (Dorschel et al., 2022, IBCSO) is the state of the art bathymetric chart south of $50°$ S. A cell by cell comparison between IBSCO v2 and SRTM15+ v2.2 showed marked disparities, in particular for water depths between -4000 m and -1000 m with differences reaching up to 1700 m (Dorschel et al., 2022). Most of the bathymetric errors only
375 affect the generation of high-mode IT which are bound to dissipate locally. Note that HYCOM may not resolve these high-mode IT anyway, as discussed in the present section. Still, the less frequent errors of the order of thousands of meters are likely to impact the generation of low-mode IT.

Efforts have been made to improve the accuracy of the M_2 barotropic tides embedded in HYCOM. The simulation used in this study incorporates the framework presented in Ngodock et al. (2016) aiming at minimizing the tidal elevation RMS
380 errors with respect to TPXO, a state of the art data-assimilative tide model (Egbert and Erofeeva, 2002). However, the tidal elevations in TPXO itself also have errors. Both Stammer et al. (2014) and Zaron and Elipot (2021) point at the imperfection of the modeled tidal elevations near Antarctica (using data from GRACE and CryoSat-2, respectively). On the other hand, it is not so much the sea surface height that matters here, but the tidal currents. Zaron and Elipot (2021) used surface drifter

observations for evaluating TPXO predicted tidal currents throughout much of the global oceans. Unfortunately, the spatial
385 density of observations is too poor to evaluate the model performance around much of Antarctica.

Lastly, to assess the stratification in HYCOM, we compare the phase speed of a mode-1 gravity wave in the model with the
phase speed determined from climatology. The phase speed of a mode-m gravity wave c_m is directly related to the eigenvalue
$-1/c_m^2$ obtained by solving the Sturm-Liouville problem Eq. (8). Having already solved Eq. (8) for the climatology, we now
solve it for our simulated monthly-mean 3D fields of temperature and salinity. As for the climatology, the HYCOM data were
390 extended down to the reference bathymetry from GEBCO 2019, wherever deeper, by appending the stored bottom value. We
then linearly interpolated the climatological phase speed of a mode-1 gravity wave c_1^{WOA} onto the coarser HYCOM grid.

In Fig. 14a we plot the climatology to HYCOM phase speed ratio. Most of the visible differences fall in the range of
expected interanual variability of less than 10% (Chelton et al., 1998). However, in a few areas around Antarctica there are
larger departures of the ratio from unity. Fig. 14b shows the zonal mean of the mode-1 phase speed from Argo and HYCOM as a
395 function of latitude. Equatorward of $\pm 60°$, the two curves agree almost perfectly, and they also visually agree with the results of
Chelton et al. (1998) (not shown). Poleward of $60°$ S and $60°$ N, however, differences steadily grow. In the Southern Ocean, the
zonal mean climatology to HYCOM phase speed ratio peaks at 1.7 (see Fig. 14c). In terms of wavelength (roughly $\propto 1/c$) the
latter ratio is reversed, with HYCOM being biased high in the Southern Ocean, on a zonal mean. The energy conversion from
the barotropic tide to the mode-m baroclinic tide mostly occurs where the bathymetry horizontal wavelength is comparable
400 to the mode-m wavelength. Hence, a biased mode-m wavelength in HYCOM results in a biased energy conversion from the
barotropic tide to the mode-m IT.

5 Conclusions

In this work we compare a 32-day segment of a global run with the HYCOM model, including realistic tidal and atmospheric
forcing, to in situ observations of the semidiurnal IT around 1,000 dbar. First, a Lagrangian sampling of the simulation is
405 compared to Argo floats park phase data to validate the geographical variability of the total variance of the semidiurnal IT
in HYCOM (see Sect. 4.1). Then, the Eulerian simulation outputs are directly compared to geographically sparser mooring
records, in terms of total variance and intrinsic decorrelation of the IT (see Sect. 4.2).

The main spatial patterns of the simulated IT variance at 1,000 dbar broadly agree with Argo observations, with energy
radiating away from low-mode IT generation hotspots (see Fig. 6). Nonetheless, on a global average, the HYCOM data exhibit
410 a smaller total semidiurnal IT variance than observed by Argo floats, by a factor of 0.60. This is in line with the results of
Ansong et al. (2017) and Luecke et al. (2020), who found HYCOM values to be biased low by a similar factor when comparing
the simulated M_2 IT energy flux and temperature variance to historical mooring observations.

While the difference between the model and Argo data appears reasonably homogeneous across most of the world ocean, it
steadily increases towards the poles (see Fig. 7). Because of the scarcity of Argo floats available in the northmost region, we
415 focus on the Southern Ocean. On average, south of $50°$ S, we find that the Argo derived total semidiurnal IT variance is close
to 4 times larger than the simulated one. North of $50°$ S, this factor is 1.5.
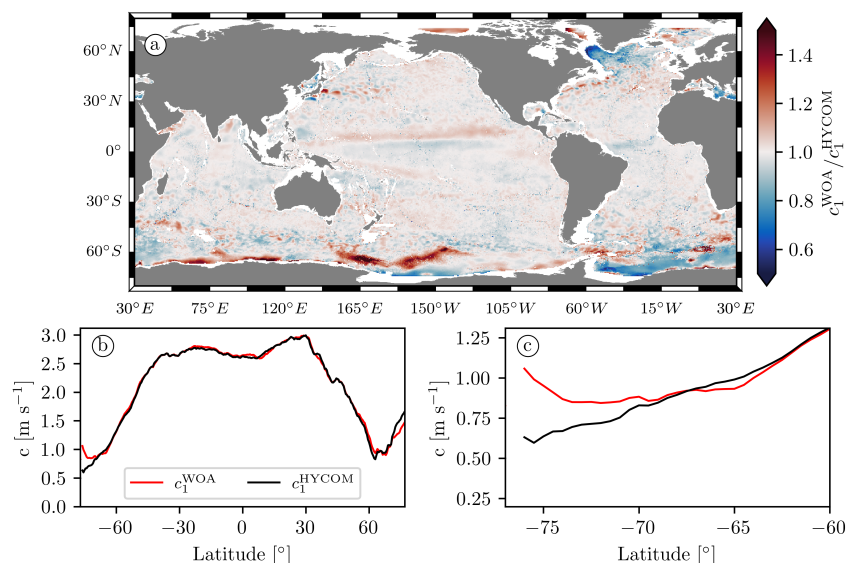
Figure 14. (a) Ratio of the phase speeds of a mode-1 gravity wave computed from the WOA18 and HYCOM stratification profiles. (b) Zonal mean of the phase speed of a mode-1 gravity wave from WOA18 (solid red curve) and HYCOM (solid black curve) as a function of latitude. (c) Same as (b) but zoomed in between $77°$ S and $60°$ S.

The mooring data support the above results for the total variance. Additionally, we find that the decorrelation affecting the semidiurnal IT in HYCOM over a 32-day window is weaker than observed in the mooring records, on average (see Fig. 10 and Table 1). In other words, over timescales shorter than 32 days, the IT in HYCOM are more coherent than in observations. We

420 emphasize that, depending on the location, the complete decorrelation of the nonstationary IT is not systematically observable in a 32-day duration. Longer time series are thus needed to accurately describe the intrinsic decorrelation of IT. Nonetheless, we find that the semidiurnal IT in HYCOM actually become stationnary within $\mathcal{O}(400)$ h on a global average, i.e. well under our 32 days of simulated data (see Fig. 11). Put together, these results support the conclusions of Buijsman et al. (2020), who found that the stationary (i.e., the longterm coherent) M_2 IT solution from HYCOM is too energetic compared with altimetry.

425 Note that their comparison was based on one-year simulated time series corrected for the duration difference with 17-year long altimetry records.

We also investigate the effects of the Lagrangian sampling inherent to the Argo floats. When comparing autocovariances computed from the HYCOM data sampled in the Lagrangian and Eulerian frameworks, respectively, we find the total variance to be unaffected (see Fig. 5). Moreover, the simulated apparent decorrelation (the decorrelation due to the motion of a

430 Lagrangian particle) is seen to agree very well with the apparent decorrelation experienced by Argo floats, on a global average (see Sect. 4.3). The intrinsic decorrelation of IT in the simulated data, on the other hand, typically is half as rapid as observed by Argo floats. This would make the intrinsic decorrelation at least as important as the apparent one over the first few days of

time lag. Note, however, that the geographical variability of the duration and strength of the intrinsic decorrelation is expected
to be large.

435   Finally, we discuss the potential sources of bias. We cannot think of a particular reason for the IT variance obtained from
either the Argo or the mooring data to be biased high, particularly in the Southern Ocean. However, HYCOM is subject to
various limitations. First and foremost, the model can only correctly resolve vertical modes up to 5 in most of the global
oceans. Approaching the poles, the reduced number of layers limits further the number of resolved modes. While mode-1 IT
supposedly account for most of the tidal variability at 1,000 dbar on a global average (Geoffroy and Nycander, 2022), our in
440 situ instruments also record the contribution from higher modes that can become significant locally. This may explain why the
IT variance is larger in the in situ data than in HYCOM, particularly in the Southern Ocean. Additionally, we discuss limitations
in the modeled bathymetry, barotropic tidal forcing, and stratification in the Southern Ocean. These could perhaps play a role
for the discrepancy with observations we find south of $50°$ S.

Code and data availability. Argo data were obtained from U.S. GDAC (ftp://usgodae.org/pub/outgoing/argo). These data were collected
445 and made freely available by the International Argo Program and the national programs that contribute to it. (https://argo.ucsd.edu, https:
//www.ocean-ops.org). The Argo Program is part of the Global Ocean Observing System. An Argo Iridium float list is maintained by
Stephen C. Riser (http://runt.ocean.washington.edu/argo/heterographs/rollcall.html). The code used to download and process the Argo data
is available at https://doi.org/10.57669/geoffroy-2022-argoit-1.0.0 (Geoffroy, 2022a). A global map of the total semidiurnal internal tide
variance at 1,000 dbar produced using the latter code is available at https://doi.org/10.17043/geoffroy-2022-argoit-1 (Geoffroy, 2022b). Cli-
450 matological data were obtained from the World Ocean Atlas 2018 (https://accession.nodc.noaa.gov/NCEI-WOA18). Bathymetric data were
obtained from the GEBCO Compilation Group (https://www.gebco.net/data_and_products/gridded_bathymetry_data/). The Global Multi-
Archive Current Meter Database is is not publicly available but can be obtained through request (http://stockage.univ-brest.fr/\protect$\relax\
sim$scott/GMACMD/gmacmd.html). Netcdf versions of the baroclinic tidal harmonic constants from the High Resolution Empirical Tide
model are made available by Edward D. Zaron (https://ezaron.hopto.org/~ezaron/downloads.html).

455 *Author contributions.* Gaspard Geoffroy: Conceptualization, Formal analysis, Project administration, Software, Visualization, Writing - orig-
inal draft, Writing - review & editing. Jonas Nycander: Conceptualization, Funding acquisition, Writing - original draft, Writing - review
& editing. Maarten C. Buijsman: Project administration, Writing - review & editing. Jay F. Shriver: Formal analysis, Resources, Software.
Brian K. Arbic: Writing - review & editing.

Competing interests. The authors declare that they have no conflict of interest.

460 *Acknowledgements.* This work was supported by grant 2017-04623 from the Swedish Research Council. The computations and data handling were enabled by resources provided by the Swedish National Infrastructure for Computing (SNIC) at the National Supercomputer Centre (NSC) partially funded by the Swedish Research Council through grant agreement no. 2018-05973. B.K.A. acknowledges support from Office of Naval Research (ONR) grant N00014-19-1-2712.

References

465   Alford, M. H. and Zhao, Z.: Global Patterns of Low-Mode Internal-Wave Propagation. Part I: Energy and Energy Flux, Journal of Physical
        Oceanography, 37, 1829 – 1848, https://doi.org/10.1175/JPO3085.1, 2007.

Ansong, J. K., Arbic, B. K., Alford, M. H., Buijsman, M. C., Shriver, J. F., Zhao, Z., Richman, J. G., Simmons, H. L., Timko, P. G., Wallcraft,
        A. J., and Zamudio, L.: Semidiurnal internal tide energy fluxes and their variability in a Global Ocean Model and moored observations,
        Journal of Geophysical Research: Oceans, 122, 1882–1900, https://doi.org/https://doi.org/10.1002/2016JC012184, 2017.

470   Arbic, B. K.: Incorporating tides and internal gravity waves within global ocean general circulation models: A review, Progress in Oceanog-
        raphy, 206, 102 824, https://doi.org/https://doi.org/10.1016/j.pocean.2022.102824, 2022.

Arbic, B. K., Wallcraft, A. J., and Metzger, E. J.: Concurrent simulation of the eddying general circulation and tides in a global ocean model,
        Ocean Modelling, 32, 175–187, https://doi.org/https://doi.org/10.1016/j.ocemod.2010.01.007, the magic of modelling: A special volume
        commemorating the contributions of Peter D. Killworth – Part 2, 2010.

475   Arbic, B. K., Elipot, S., Brasch, J. M., Menemenlis, D., Ponte, A. L., Shriver, J. F., Yu, X., Zaron, E. D., Alford, M. H., Buijsman, M. C.,
        Abernathey, R., Garcia, D., Guan, L., Martin, P. E., and Nelson, A. D.: Frequency dependence of near-surface oceanic kinetic energy from
        drifter observations and global high-resolution models, https://doi.org/10.48550/ARXIV.2202.08877, 2022.

Argo: https://doi.org/10.17882/42182, Argo float data and metadata from Global Data Assembly Centre (Argo GDAC), 2000.

Boyer, T. P., Garcia, H. E., Locarnini, Ricardo A.and Zweng, M. M., Mishonov, A. V., Reagan, J. R., Weathers, K. A., Baranova, O. K., Seidov,
480   D., and Smolyar, I. V.: World Ocean Atlas 2018. Temperature, and Salinity, Dataset, https://accession.nodc.noaa.gov/NCEI-WOA18,
        Accessed 11-10-2021, 2018.

Buijsman, M. C., Arbic, B. K., Richman, J. G., Shriver, J. F., Wallcraft, A. J., and Zamudio, L.: Semidiurnal internal tide incoherence in the
        equatorial Pacific, Journal of Geophysical Research: Oceans, 122, 5286–5305, https://doi.org/10.1002/2016JC012590, 2017.

Buijsman, M. C., Stephenson, G. R., Ansong, J. K., Arbic, B. K., Green, J. M., Richman, J. G., Shriver, J. F., Vic, C., Wallcraft, A. J., and
485   Zhao, Z.: On the interplay between horizontal resolution and wave drag and their effect on tidal baroclinic mode waves in realistic global
        ocean simulations, Ocean Modelling, 152, 101 656, https://doi.org/https://doi.org/10.1016/j.ocemod.2020.101656, 2020.

Caspar-Cohen, Z., Ponte, A., Lahaye, N., Carton, X., Yu, X., and Gentil, S. L.: Characterization of Internal Tide Incoherence: Eulerian versus
        Lagrangian Perspectives, Journal of Physical Oceanography, 52, 1245 – 1259, https://doi.org/10.1175/JPO-D-21-0088.1, 2022.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Wallcraft, A. J., Metzger, E. J., Blanton, B. O., Lozano, C., Rao, D. B.,
490   and Hogan, P. J.: Generalized vertical coordinates for eddy-resolving global and coastal ocean forecasts, Oceanography, 19, 20, 2006.

Chelton, D. B., deSzoeke, R. A., Schlax, M. G., El Naggar, K., and Siwertz, N.: Geographical Variability of the First
        Baroclinic Rossby Radius of Deformation, Journal of Physical Oceanography, 28, 433–460, https://doi.org/10.1175/1520-
        0485(1998)028<0433:GVOTFB>2.0.CO;2, 1998.

Cherniawsky, J. Y., Foreman, M. G. G., Crawford, W. R., and Henry, R. F.: Ocean Tides from TOPEX/Poseidon Sea Level Data, Journal of
495   Atmospheric and Oceanic Technology, 18, 649 – 664, https://doi.org/10.1175/1520-0426(2001)018<0649:OTFTPS>2.0.CO;2, 2001.

de Lavergne, C., Falahat, S., Madec, G., Roquet, F., Nycander, J., and Vic, C.: Toward global maps of internal tide energy sinks, Ocean
        Modelling, 137, 52–75, https://doi.org/https://doi.org/10.1016/j.ocemod.2019.03.010, 2019.

de Lavergne, C., Vic, C., Madec, G., Roquet, F., Waterhouse, A. F., Whalen, C. B., Cuypers, Y., Bouruet-Aubertot, P., Ferron, B., and Hibiya,
        T.: A Parameterization of Local and Remote Tidal Mixing, Journal of Advances in Modeling Earth Systems, 12, e2020MS002 065,
500   https://doi.org/10.1029/2020MS002065, 2020.

Dorschel, B., Hehemann, L., Viquerat, S., Warnke, F., Dreutter, S., Tenberge, Y. S., Accettella, D., An, L., Barrios, F., Bazhenova, E., et al.: The International Bathymetric Chart of the Southern Ocean Version 2, Scientific Data, 9, 275, https://doi.org/https://doi.org/10.1038/s41597-022-01366-7, 2022.

Egbert, G. D. and Erofeeva, S. Y.: Efficient Inverse Modeling of Barotropic Ocean Tides, Journal of Atmospheric and Oceanic Technology, 19, 183 – 204, https://doi.org/10.1175/1520-0426(2002)019<0183:EIMOBO>2.0.CO;2, 2002.

Egbert, G. D. and Ray, R. D.: Tidal Prediction, Journal of Marine Research, 75, 189–237, https://doi.org/10.1357/002224017821836761, 2017.

GEBCO: Dataset, https://doi.org/10.5285/836f016a-33be-6ddc-e053-6c86abc0788e, GEBCO 2019 Grid, 2019.

Geoffroy, G.: Scripts for mapping the total semidiurnal internal tide variance at 1,000 dbar using Argo data, https://doi.org/10.57669/geoffroy-2022-argoit-1.0.0, 2022a.

Geoffroy, G.: Global map of the total semidiurnal internal tide variance at 1,000 dbar from Argo data, Dataset, https://doi.org/10.17043/geoffroy-2022-argoit-1, 2022b.

Geoffroy, G. and Nycander, J.: Global Mapping of the Nonstationary Semidiurnal Internal Tide Using Argo Data, Journal of Geophysical Research: Oceans, 127, e2021JC018 283, https://doi.org/https://doi.org/10.1029/2021JC018283, 2022.

Gill, A. E.: Atmosphere-Ocean Dynamics, Elsevier, academic press edn., 1982.

Hennon, T. D., Riser, S. C., and Alford, M. H.: Observations of Internal Gravity Waves by Argo Floats, Journal of Physical Oceanography, 44, 2370–2386, https://doi.org/10.1175/JPO-D-13-0222.1, 2014.

Lebigot, E. O.: Uncertainties: a Python package for calculations with uncertainties, http://pythonhosted.org/uncertainties/, 2022.

Luecke, C. A., Arbic, B. K., Richman, J. G., Shriver, J. F., Alford, M. H., Ansong, J. K., Bassette, S. L., Buijsman, M. C., Menemenlis, D., Scott, R. B., Timko, P. G., Voet, G., Wallcraft, A. J., and Zamudio, L.: Statistical Comparisons of Temperature Variance and Kinetic Energy in Global Ocean Models and Observations: Results From Mesoscale to Internal Wave Frequencies, Journal of Geophysical Research: Oceans, 125, e2019JC015 306, https://doi.org/https://doi.org/10.1029/2019JC015306, 2020.

MacKinnon, J. A., Zhao, Z., Whalen, C. B., Waterhouse, A. F., Trossman, D. S., Sun, O. M., Laurent, L. C. S., Simmons, H. L., Polzin, K., Pinkel, R., Pickering, A., Norton, N. J., Nash, J. D., Musgrave, R., Merchant, L. M., Melet, A. V., Mater, B., Legg, S., Large, W. G., Kunze, E., Klymak, J. M., Jochum, M., Jayne, S. R., Hallberg, R. W., Griffies, S. M., Diggs, S., Danabasoglu, G., Chassignet, E. P., Buijsman, M. C., Bryan, F. O., Briegleb, B. P., Barna, A., Arbic, B. K., Ansong, J. K., and Alford, M. H.: Climate Process Team on Internal Wave–Driven Ocean Mixing, Bulletin of the American Meteorological Society, 98, 2429 – 2454, https://doi.org/10.1175/BAMS-D-16-0030.1, 2017.

Melet, A., Legg, S., and Hallberg, R.: Climatic Impacts of Parameterized Local and Remote Tidal Mixing, Journal of Climate, 29, 3473 – 3500, https://doi.org/10.1175/JCLI-D-15-0153.1, 2016.

Melet, A., Hallberg, R., and Marshall, D. P.: Chapter 2 - The role of ocean mixing in the climate system, in: Ocean Mixing, edited by Meredith, M. and Naveira Garabato, A., pp. 5–34, Elsevier, https://doi.org/10.1016/B978-0-12-821512-8.00009-8, 2022.

Nelson, A. D., Arbic, B. K., Zaron, E. D., Savage, A. C., Richman, J. G., Buijsman, M. C., and Shriver, J. F.: Toward Realistic Nonstationarity of Semidiurnal Baroclinic Tides in a Hydrodynamic Model, Journal of Geophysical Research: Oceans, 124, 6632–6642, https://doi.org/10.1029/2018JC014737, 2019.

Ngodock, H. E., Souopgui, I., Wallcraft, A. J., Richman, J. G., Shriver, J. F., and Arbic, B. K.: On improving the accuracy of the M2 barotropic tides embedded in a high-resolution global ocean circulation model, Ocean Modelling, 97, 16–26, https://doi.org/https://doi.org/10.1016/j.ocemod.2015.10.011, 2016.

Rainville, L. and Pinkel, R.: Propagation of Low-Mode Internal Waves through the Ocean, Journal of Physical Oceanography, 36, 1220 – 1236, https://doi.org/10.1175/JPO2889.1, 2006.

Scott, R. B., Goff, J. A., Naveira Garabato, A. C., and Nurser, A. J. G.: Global rate and spectral characteristics of internal gravity wave generation by geostrophic flow over topography, Journal of Geophysical Research: Oceans, 116, https://doi.org/10.1029/2011JC007005, 2011.

Shriver, J. F., Arbic, B. K., Richman, J. G., Ray, R. D., Metzger, E. J., Wallcraft, A. J., and Timko, P. G.: An evaluation of the barotropic and internal tides in a high-resolution global ocean circulation model, Journal of Geophysical Research: Oceans, 117, https://doi.org/https://doi.org/10.1029/2012JC008170, 2012.

Shriver, J. F., Richman, J. G., and Arbic, B. K.: How stationary are the internal tides in a high-resolution global ocean circulation model?, Journal of Geophysical Research: Oceans, 119, 2769–2787, https://doi.org/10.1002/2013JC009423, 2014.

Smith, W. H. F. and Sandwell, D. T.: Bathymetric prediction from dense satellite altimetry and sparse shipboard bathymetry, Journal of Geophysical Research: Solid Earth, 99, 21 803–21 824, https://doi.org/https://doi.org/10.1029/94JB00988, 1994.

Stammer, D., Ray, R. D., Andersen, O. B., Arbic, B. K., Bosch, W., Carrère, L., Cheng, Y., Chinn, D. S., Dushaw, B. D., Egbert, G. D., Erofeeva, S. Y., Fok, H. S., Green, J. A. M., Griffiths, S., King, M. A., Lapin, V., Lemoine, F. G., Luthcke, S. B., Lyard, F., Morison, J., Müller, M., Padman, L., Richman, J. G., Shriver, J. F., Shum, C. K., Taguchi, E., and Yi, Y.: Accuracy assessment of global barotropic ocean tide models, Reviews of Geophysics, 52, 243–282, https://doi.org/https://doi.org/10.1002/2014RG000450, 2014.

Tozer, B., Sandwell, D. T., Smith, W. H. F., Olson, C., Beale, J. R., and Wessel, P.: Global Bathymetry and Topography at 15 Arc Sec: SRTM15+, Earth and Space Science, 6, 1847–1864, https://doi.org/https://doi.org/10.1029/2019EA000658, 2019.

Van Sebille, E., Kehl, C., Lange, M., Delandmeter, P., and contributors, T. P.: Parcels, https://doi.org/10.5281/zenodo.7035503, 2021.

Zaron, E. D.: Mapping the nonstationary internal tide with satellite altimetry, Journal of Geophysical Research: Oceans, 122, 539–554, https://doi.org/10.1002/2016JC012487, 2017.

Zaron, E. D.: Baroclinic Tidal Sea Level from Exact-Repeat Mission Altimetry, Journal of Physical Oceanography, 49, 193–210, https://doi.org/10.1175/JPO-D-18-0127.1, 2019.

Zaron, E. D. and Egbert, G. D.: Time-Variable Refraction of the Internal Tide at the Hawaiian Ridge, Journal of Physical Oceanography, 44, 538 – 557, https://doi.org/10.1175/JPO-D-12-0238.1, 2014.

Zaron, E. D. and Elipot, S.: An Assessment of Global Ocean Barotropic Tide Models Using Geodetic Mission Altimetry and Surface Drifters, Journal of Physical Oceanography, 51, 63 – 82, https://doi.org/10.1175/JPO-D-20-0089.1, 2021.

Zhao, Z., Alford, M. H., Girton, J. B., Rainville, L., and Simmons, H. L.: Global Observations of Open-Ocean Mode-1 M2 Internal Tides, Journal of Physical Oceanography, 46, 1657–1684, https://doi.org/10.1175/JPO-D-15-0105.1, 2016.