We would like first to thank both referees for their valuable inputs. Many of their remarks proved pertinent, and overall contributed to make the manuscript better.

General comments:

This is a very valuable and comprehensive study that compares the semidiurnal tidal variance and autocovariance in a global numerical model and in in situ observations. This paper is well written and mostly logically organized. I would recommend publication after the main concern below is addressed as well as the smaller technical comments.

I have some serious concerns about the method employed to compare autocovariance function estimates. After calculating average autocovariance functions, the authors essentially estimate the autocovariance function amplitudes and their associated confidence intervals. Yet, the method employed for calculating these confidence intervals appear to be flawed because it assumes that the amplitude estimates are normally distributed, which is not the case. This is clearly illustrated as an example in Figure 2b that shows confidence intervals crossing the zero value: true amplitude values cannot be less than zero. I suggest for the authors to properly derive error estimates and confidence intervals for the autocovariance amplitude before pursuing the rest of this study. I provide some potential ways of doing so in my detailed comments below. In fact, in Figure 10, as an example, the authors take a better approach by displaying quantiles of the distributions: why not taking that approach from the beginning?

Ans.: We wrongly assumed Gaussian distributed complex demodulates, hence our confidence intervals were incorrect. Referee #1 also pointed at the questionable definition of the total error in equation (4). For both these reasons, we now use a Monte Carlo method to estimate the confidence interval of the complex demodulates.

I also suggest to replace the term "demodulate" by something more meaningful: perhaps autocovariance envelope or amplitude? As noted below, the method to obtain the "complex demodulate" could be improved by simply computing the analytic transform of the autocovariance functions.

Ans.: From our understanding, the analytic transform would not capture the amplitude at the M2 frequency without band-pass filtering the \eta_1000 time series a priori.

For the validating part of the study, section 3, the current organization of the material does not make sense to me and the conclusions are not clearly laid out. First, I would like to know how the model does in an Eulerian framework, then I would like to know if the Lagrangian framework or method is valid, and third I would like to know the result of comparing Argo and Lagrangian HYCOM particles. As such, I suggest the following reorganization of the material of section 3:
- comparison of HYCOM Eulerian results and mooring (Eulerian) results to assess the model: what is the conclusion?
- comparison of HYCOM Eulerian and Lagrangian results to assess the method of using Lagragian data: what is the conclusion on the potential Lagrangian bias?
- Comparison of HYCOM Lagrangian results and Argo (Lagrangian) results: what is the conclusion?

Ans.: Section 3 in itself is not about validating HYCOM, but introducing, using a local example, the methods used to further validate HYCOM (section 4). There are other papers describing strict Eulerian point-to-point comparisons between HYCOM and moorings (c.f., cited papers Ansong, 2017 and Luecke, 2020). Rather, the Eulerian component of our analysis is designed to bolster and extend the main Lagrangian component (added clarification at the end of section 1). Therefore, the *logical beginning of section 3 is the methodology developed by Geoffroy and Nycander (2022) to estimate the variance of the semidiurnal IT using Lagrangian data. We then developed a Eulerian framework using the HYCOM data primarily to validate our Lagrangian methodology. Incidentally, it also enables the analysis of the decorrelation of the IT. This mirrors the organization of section 4.*

Comments:

Abstract:

line 1: In the abstract, unless you explain there what you mean by "decorrelate" as you do in the main text, I think that instead of "correlation" you should write "auto-correlation" or "auto-covariance" which are established statistical terms. An abstract should be able to stand alone. It may already be in the title but perhaps you could rephrase the abstract to provide a summary statement of what you are doing: validating a model by comparing it to in situ observations.
Ans.: Reworked abstract. Suppressed "decorrelate".

l46: It is not obvious (to me) what the k-space methodology is. I suggest that you rephrase or explain. Does this refer to the method of Zaron (2017)?
Ans.: Yes. Rephrased.

l71-72: It is not good practice to refer to a section ahead. Simply explain that the duration was chosen to match the numerical output you are using/comparing?
Ans.: Corrected

Section 2.3:

Perhaps a reference for HYCOM and that specific simulation is needed. Should you add more details about the use of the Parcels software that would allow readers to replicate your experiment?
Ans.: The reference for HYCOM (Chassignet et al., 2006) was given in the introduction, the first time we used the HYCOM acronym. We do not have any other reference for this particular simulation (apart from `GLBy190.04'). Added information on the Lagrangian simulation.

l96: Why "mainly"? And can you simply state why you used only 32 days of the model? Data/space constrains?
Ans.: Suppressed "mainly". Added paragraph at the beginning of the section regarding the data we use. The model wasn't run specifically for this study, we used the data that were available and suitable for our methodology. Hence, there are no real technical constraints to mention.

l101: Why 41644? Does this correspond to a mean geographical density?
Ans.: Yes, it roughly corresponds to a mean density of 15 particles in our final 200 km radius circular patches. We feel this is unnecessary to be added, moreover it would not be easy to motivate clearly at this stage of the paper.

l106: The effects of the drift? Do you mean potential Lagrangian biases?
Ans.: As pointed at by referee #1, we prefer to call these effects "Lagrangian decorrelation".

Section 3.1:

eq. 1: Could we get here an explanation of this quantity and why it represents the vertical displacement of an isotherm? Perhaps cite Hennon et al. 2014 as you did in Geoffroy and Nycander 2022? Are you correcting for the float displacement as you did in that paper?
Ans.: Added explanations. We do correct for the float displacement for the Argo data .

l113-115: which monthly-mean 3D temperature field? Is it from a product for the case of Argo? Please provide more details; I do not understand how you get that gradient for the in situ data.
Ans.: Added information. It is the modeled monthly-mean 3D temperature field introduced in section 2.3. For the Argo data, we compute the temperature gradient at 1000 dbar for a given park phase using the temperature profiles recorded by the float immediately before and after that park phase (now made clearer in section 3).

Figure 1: The Argo segments are shown as dots? How are these segments? Could you plot the assumed rectilinear trajectories of the Argo floats?
Ans.: Figure 1 shows the median position of the Argo segments as dots (now made clear in the text). Added Argo trajectories.

l 131: I don't get this: what is a "binned HYCOM particle"? Do you mean that you average the individual autocovariance estimates in Eulerian bins?
Ans.: Rephrased.

l136: Don't you think that in that figure the R_{argo} falls below its CI at ~100h rather than at ~200h?
Ans.: Here R_{argo} is the red curve. We do see it fall below its CI at ~200h.

Eq4 and after: I am not sure that this is the right way to compute the confidence interval for A: what you call the complex demodulate, or rather its square value (A^2), should be distributed like a chi-square variable with 2 degrees of freedom (like a spectral estimate), and not distributed like a Gaussian variable. Thus, confidence intervals as plus or minus two standard errors are likely incorrect. Consider your figure 2b: the CIs suggest that A can take negative values whereas it is clearly a positive quantity. I suggest you revise the derivation of the CI for A and reassess your overall results.
Ans.: Correct. Referee #1 also pointed at the questionable definition of the total error in equation (4). For both these reasons, we now use a Monte Carlo method to estimate the confidence interval of the complex demodulates.

Figure 2b: the CIs for the two curves are superimposed and thus cannot be distinguished; please modify the figure so that the reader can see both.
Ans.: Figure modified.

l156: Considering my remark above that your CIs are likely incorrect, I think you should revisit that statement.
Ans.: Also discussed by referee #1, modified the sentence.

I believe that what you are trying to plot in Figure 2b is the envelope of the autocovariance function. Your method is probably fine but the envelope can be easily obtained by computing the amplitude of the analytic signal of the autocovariance, see Lilly and Gascard (2006) as an example (The analytic signal can be calculated using the Hilbert transform in python with the scipy package or the anatrans.m function of the jLab toolbox for Matlab). One way to get a confidence interval for the amplitude of the analytic transform would be to look at the distribution of all the individual transform amplitudes, lag value by lag value (as you do in Figure 10 later).
Ans.: Our complex demodulate method specifically selects the amplitude at the semidiurnal frequency. Our understanding is that unless band-pass filtering the time series prior to computing the autocovariance, the analytic transform cannot be used to isolate the amplitude of the oscillations at the semidiurnal frequency.

l190: "outliers": please use sentence to explain what you mean.
Ans.: Also pointed at by referee #1. Added sentence. \eta_1000 values can be unstable when facing small temperature gradients in the denominator of Eq. (1) and (5), leading to unrealistically large variance. For consistency with the Argo results, we use the same quality checks on the variance of \eta_1000 computed using HYCOM data. Added sentence.

l194: Figure 5 does not look like a scatter plot but a 2D density plot. Is the R^2 exactly 1 as written in the plot or is it approximately 1 as stated in the text? I am surprised that it is so close to one. What is it in a domain that is not logarithmic? What is your R2 anyway? The adjective "Pearson" is usually used for the correlation coefficient while the coefficient of determination is the correlation squared for linear regression.
Ans.: Deleted "scatter". There was an error in the script calculating R^2. The correct value in log domain is 0.98, in non-log domain it is 0.74. Our R^2 is (Pearson's R)^2. Changed to r^2 to avoid any confusion with the autocovariance (denoted R).

l198: "taken as ..." : state this earlier to remind the reader.
Ans.: Added statement earlier in the text.

l208: a bias which means that HYCOM underestimate Argo, correct?
Ans.: Correct. Rephrased.

Figure 7b: try the ratio x/(x+y) instead of log10(x/y) as in Arbic, Elipot et al. 2022. In this way you will not have to use log10 and truncate the scale. The results will look the same but it is a better statistic that is robust to outliers.
Ans.: Replaced log10(x/y) by x/(x+y).

l214: The ratio increases approaching the poles? Where is this seen?
Ans.: The ratio itself is not shown, added comment.

l225: Should you conclude the section with some statement?
Ans.: Added sentence referring to the discussion on potential sources of biases.

Section 4.2:

l227-228: I do not understand what you mean by that. Please explain what is the intrinsic decorrelation. Do you mean Eulerian? In fixed space?
Ans.: Also noted by referee #1. Everywhere replaced "intrinsic decorrelation" by "decorrelation of the IT", or "decorrelation", and further "apparent decorrelation" by "Lagrangian decorrelation".

l229-230: "particle in the Eulerian framework": what do you mean? You average in Eulerian/geographical bins? I think you should use "Eulerian framework" for computing autocovariance from Eulerian time series (model grid and moorings) and "Lagrangian framework" for computing autocovariance from Lagrangian time series (model particles and Argo).
Ans.: Changed the proposition to "We compute a sample autocovariance in the Eulerian framework for each particle". This refers to the sample autocovariance computed in our Eulerian framework as explained in section 3.3. The sample autocovariance in the Eulerian framework is computed along the particle's trajectory using Eulerian data.

l235: yes indeed because the autocovariance and its amplitude are probably not gaussian distributed!

Ans.: Agreed, by definition our demodulates are Rice distributed. Deleted "(and their demodulates)". However, we emphasize that the sample mean autocovariance at a given time lag can be considered Gaussian distributed in the two following cases:
- In a local geographical patch: we assume the particles are randomly sampling a wave field with uniform statistics.
- When computing the sample mean autocovariance from a very large population of particles (global or regional scales), by virtue of the central limit theorem.

l251: "the distribution is well centered on the y = x": I strongly suggest you revise this assessment. Figure 9a suggests no linear relation between the mooring results and the model results.
Ans.: As noted by referee #1, as a ratio of sample statistics, SCVF_15 is expected to be noisy. We now plot the IT variance instead.

l253: "log domain" : this figure appears to be on a linear scale?
Ans.: Plot changed.

l270: truely -> truly
Ans.: Corrected

Figure 11b: a legend for the various fitted curve would be very helpful.
Ans.: Added legend

l293: Why is it 3 times T_{int}?
Ans.: For 95% of the exponential decay is achieved within 3 time constants (exp(-3)~0.05). Added comment.

l306: "Note that ..." : this should be moved earlier just after your eq 6.
Ans.: This remark explains the results of the fitting. We did not assume this when defining our model.

l315-319: What are the implications of this comparison for HYCOM? Could you expand? I understand you address this next but a transition sentence at the end of a section would be useful.
Ans.: Expanded on the implications for HYCOM.

Section 4.4:

l323: If your method holds, should you not rather say that the model is biased low?
Ans.: Since we do not see any reasons for the in situ data/processing to be biased high, indeed our conclusion is that HYCOM is biased low.

Data availability: A statement on the HYCOM data availability is missing.
Ans.: Added statement