## **Response to Referee #1:**

Dear C. K. Shum,

Thank you for your interest in this manuscript and for the comments and suggestions you make. I reply to all your comments, corrections and suggestions to change hereafter in blue. Best regards,

Loren Carrere

## **General comments:**

This was a very detailed assessment and good comparison of available IT models supported by qualitative, quantitative and spectral analyses. The manuscript is in overall well structured, clearly written and organized. Starting from the abstract, all sections in the manuscript were captured concisely. Next, the introduction paragraph provided a summarized background on satellite altimetry and stated its limitations for tidal analyses, supported by relevant references. Also, it explains the need for a validation process for available IT models and delineates what to expect in subsequent sections but omits some information on the resolution of altimetry data used. Next, in the presentation of the participating IT models' section, all the participating models were well described in terms of their methodology, except for Ray model. Next, the qualitative and quantitative comparison sections are well presented, requiring minor grammar corrections. All the plots are accurately described, but some plots were difficult to read due to their font size. Also, mathematical support or presentation of equations used (as cited in the literature), are rather necessary in the quantitative comparison section. Next, the presentation of the altimeter database section is well written and clearly indicate the reasons for selecting the J2 and C2 missions for the comparison. Next, the method of comparison section was well explained and follows a logical order but needs for a clearer or possibly an enumerated sequence. In the next section, variance reduction analysis using satellite altimeter data, SSH and SLA variances plots were produced for both the M2 and K1 constituents. The results were properly described, and variances of the corrections were stated for each participating IT model. Some plots were difficult to read due to their font size. Next, the wavelength analysis of M2 section corroborates previous results by estimating the amount of energy removed at from participating IT models. Results for this section are well described; only need to further elaborate on EGBERT's model performance in the Gulf of Guinea. Next, the discussion session clearly summarizes previous sections and offers conclusive explanations on the performance of each model based on final results. Finally, the appendices section supplemented the manuscript, with results of remaining constituents. This section was clearly described and presented similarly to the variance reduction analysis of M2 and K1 section. As previous sections, the font size on some plots were not legible

## Specific comments:

1. Line 83: Remove "Altika". The mission is not used for the quantitative comparison or anywhere in the manuscript. LC: OK removed.

2. Line 84: The section omits the resolution of the altimetry databases used in the validation process e.g. HRM/LRM - LC : LRM added

3. Lines 135-144: RAY model needs to be better described in the presentation of participating internal tide models. The model methodology should be included as for other participating models or more information is rather necessary. LC: added reference to paper Ray and Zaron 2016 section3 + a few sentences on methodology

4. Lines 240-243: Consider mentioning why from all the seven regions of interest, NPAC and Luzon regions were selected for the comparison e.g. more energetic regions of all seven. LC : they are more energetic regions + all tested models are available on NPAC region and Luzon area is characterized by strong semi-diurnal and diurnal baroclinic tides. Information added in the text.

5. Line 294: Include Stammer, 2014 equations for the calculation of STD of IT models for the reader convenience. LC: done

6. Line 325: Here the resolution of 1-hz is indicated (LRM). Consider adding it to line 84. : LC: OK

7. Line 351: For future research: consider adding seasonal barotropic tide correction in best performing models. LC: I agree that this point could be an interesting point to notice, but as this correction is not yet available and used in the present dataset, I think that mentioning it might make the definition of the SSH a bit confusing.

8. Lines 371-408: This section explains well the methodology for the analysis. But would be convenient to enumerate each step to follow a sequential order. LC: OK modified in the text.

9. Lines 416-425: Reiterate or remind that the quantification and regional impact of the M2 IT correction were performed using all participating IT models but not the same case for K1. LC: OK added in the text in sections M2 and K1.

10. Line 450: Independent results from C2 shows similar patterns as J2 mission, albeit J2 bias in empirical IT models. From all IT models listed in table 1, only ZARON incorporates Altika mission. Consider adding Altika to support C2 independent results and further corroborate J2 bias towards empirical model. LC: Indeed the tests with Altika mission have also been performed and presented at some conferences (OSTST), but for the clarity of the paper we prefered not to include it. Moreover analysis with Altika mission gives close results to C2 and J2 tests.

11. Line 419-460: Speculate or provide a possible explanation of why Dushaw and HYCOM models rise SSH or SLA variances in some locations as supported by conclusive evidence in the variance computation e.g. areas of strong currents, others? LC: some comments have been added in the text:

The DUSHAW model raises SLA variance in several mesoscale regions (Gulf Stream, Agulhas current, Malvinas region and Kuroshio currents), indicating that the model does not properly separate IT and other ocean signals in these strong current areas.

HYCOM raises the variance over wider regions in the three oceans than the empirical and assimilative models do, likely due to its intrinsic characteristic of free hydrodynamic model which may induce more phase errors compared to constrained/empirical models + due to the short HYCOM time series duration used to extract the IT atlas and that induces stronger IT amplitudes (see Ansong et al. 2015 and Buijsman 2020)

12. Line 559: Further elaborate on EGBERT model's performance in the Gulf of Guinea. EGBERT model appears to reduce energy in shorter modes for this region, compared to other models. LC: comment added

## **Technical corrections:**

LC : All technical corrections proposed have been taken into account and the font size has been enlarged on the plots to be more legible.