Dear Reviewer,

Thank you again for taking the time to read through and comment on our manuscript. We responded to your greatest criticism in a previous reply. In this comment, we give a point-by-point response to all your comments and detail our proposed changes to the manuscript.

R: Reviewer's comment
A: Author's response
C: Proposed changes to the manuscript; text changes in blue

[All references that we cite herein can be found in the reference list of the modified manuscript.]

R: This paper by Pringle et al. presents recent developments of the circulation model ADCIRC that allow simulating efficiently tides and storm surges at global scale. The paper is well-written and organized, the figures are clear and the topic addressed fits well the scope of the journal. However, while storm surge predictions are rather good for a global model, tidal predictions are locally weak compared to other well-established global tidal models. Thus, in the Bay of Biscay, the RMSE on M2 reaches 0.12-0.15 m, that is more than 10% once normalized by the amplitude of this constituent. Over the Patagonian Shelf, RMSE on M2 reaches 0.25 m, which again represents errors over 10 %. In these regions, other global models have errors of a few % in these areas, see for instance a paper describing the hydrodynamic version of FES2014 (i.e. without assimilation) under discussion in Ocean Science (Lyard et al., 2020). For this reason, I think that the paper cannot be considered further for publication until the authors explain why the model is locally not reproducing tides correctly or better, improve their results. Indeed, only discussing the improvements compared to the previous version of global ADCIRC is not sufficient as tidal predictions from this version of the model were really bad (i.e. errors on M2 locally > 20%).

A: Thank you for your positive comments regarding the general organization and presentation of this manuscript. As per our previous response, we highlight that the major point of our paper is not to present a model with the lowest tidal errors possible. Instead, it is to; 1) highlight improvements to the treatment of the governing equations and implicit time-integration in the new version of ADCIRC (v55), and 2) explore the effects of unstructured mesh design on storm tide solutions. In the previous response we also highlighted that modeled tidal solutions are dictated by the three major mechanisms: bathymetry, internal tide wave drag, and bottom friction/bed stress. Analysis of these solution controlling mechanisms have been detailed in previous studies (Lyard et al., 2020; Pringle et al., 2018), and in this study we specifically avoided the excessive tuning of the model through these three controls. The previous response also provides an example figure of the M2 tidal solution errors (errors are generally smaller than presented in this manuscript) of a more tuned version of the model used in this study. In the next paragraph we detail our proposed changes to the manuscript to "explain why the model is locally not reproducing tides correctly" and state our aims and decisions more clearly.

C:
1) We explicitly state the aims of the study at the end of "Section 1: Introduction", Lines 56-58:
Section 3.3 summarizes the timing results with ADCIRC v55, highlighting its computational efficiency using a semi-implicit time-integration scheme. In summary, this study aims to: 1) highlight improvements to the treatment of the governing equations and implicit time-integration in the new version of ADCIRC (v55), and 2) explore the effects of unstructured mesh design on storm tide solutions.

2) In "Section 2.4: Datasets and Model Setup". At the end of the first paragraph which specified the bathymetric data used as well as other data use in the model setup, we propose to add the following sentences that outline how we specified the bottom friction and the internal wave drag coefficients and the reasons for this (Lines 193-201):
We note here that the accuracy of global tidal solutions strongly depends on the quality of the bathymetric data, the internal wave drag tensor, and the bottom stress term which can all be tuned to minimize tidal errors (Pringle et al., 2018a, Lyard et al., 2020). Since this study is focused on the effects of mesh design and the improvements to the governing

equations in the new version of the ADCIRC model, we deliberately avoided excessive tuning of the model with the aim to minimize tidal solution errors. Instead we chose to use a global constant value of Cit which gives the same available potential tidal energy as compared to the TPXO9-Atlas, and employ a global constant Cf of 0.0025 except in the Indian Ocean and Western Pacific Ocean where it is spatially varying per the specifications by a previous study of ours (Pringle et al., 2018a) (see Sect. S2 for additional details of model specifications).

3) We included additional details on the model properties that affect tide solutions into "Section 4: Discussion":

Lines 378-380:

Indeed, a recent study conducts a 432-member ensemble of perturbations to bathymetric depths, and bottom friction and internal wave drag coefficients to obtain smaller tidal errors than this study, particularly in shallow water (Lyard et al, 2020).

Lines 402-407:

Last, it is widely recognized that sensitivities to local high resolution bathymetry datasets, internal tide wave drag, and spatially varying bottom friction and surface ice friction are important (Lefevre et al., 2000; Le Bars et al., 2010; Zaron, 2017; Pringle et al., 2018a; Zaron, 2019; Lyard et al., 2020) likely more so than the mesh resolution effects that we concentrate on here. We aim to develop a unified framework for globally calibrating spatially varying internal tide wave drag and bottom friction coefficients with improved local high resolution bathymetric datasets in future work. Doing so should result in smaller storm tide elevation discrepancies especially in shallow water (e.g., Lyard et al., 2020).

R: -L35: I would indicate somewhere that all these studies neglected the contribution of short waves, although this process can drive a "regional setup" (i.e. a storm surge extending outside surf zone) reaching 0.5 m (e.g. Fortunato et al., 2017).

A: We added the following sentence following the citation to these previous studies on extreme sea levels.

C: Line 35-37: Note that these previous studies neglected the contributions to extreme sea levels by short waves that can drive a significant regional setup (e.g., Fortunato et al., 2017).

R: -L73: as the model is used to compute storm surges, you should explain how Cd is computed/which bulk formula is used.

A: We define Cd on Line 84 (old L83) so we added the drag law formulation information to that line.

C: Line 87 (old L83): … computed using the Garratt (1977) drag law.

R: -L100: please explain how much larger

A: This information (Courant number = 5-22 with 120 s time step) is contained within "Section 3.3 Computational Performance" so we modified the sentence to refer the reader to this section for details.

C: Lines 105-107 (old L100): With a semi-implicit time integration scheme, the computational time step permitted is larger than the CFL constraint and as a result facilitates computationally efficient global simulations (see Sect. 3.3 for details).

R: -L106: "obtain" rather than facilitate?

C: Line 111 (old L106): changed to obtain

R: -L157: Gulf of Mexico rather than Western North Atlantic?

A: The Western North Atlantic here refers to one of the basins where tropical cyclones form and we are including the Gulf of Mexico within that definition of Western North Atlantic. Therefore, we decided not to edit this.

R: -Table 3: please compare with Figure 12 in Lyard et al. (2020), where FES2014 yields errors on M2 < 0.5 cm in deep water and <4 cm on the shelf, that is about one order of magnitude smaller than here.

A: We agree to compare with hydrodynamic FES2014 (Lyard et al., 2020) here but we take numbers from Table 1 in Lyard et al. (2020) which gives the overall RMS of the vector difference which we can compare to our numbers shown in the Table 3 (the errors are greater than the reviewer states). We also realize that the Ngodock et al. (2016); Schindelegger et al. (2018); Lyard et al. (2020) are computing errors only for latitudes equatorward of ±66°, so we included our result for within these latitudes as well (results are not that different). We also add some two sentences to "Section 3.1.1: Validation of the Reference Mesh" commenting on the comparison to the FES2014 results.

C:

1) New Table 3:

Table 3. $\overline{\mathrm{RMSE}}_t$ [cm] (c.f., Appendix A) values for simulated tidal results using ADCIRC v55 (upgrade) and ADCIRC v54 in deep ($h > 1$ km) and shallow ($h < 1$ km) waters on the Ref mesh. Results from other forward barotropic tidal models (Stammer et al., 2014; Ngodock et al., 2016; Schindelegger et al., 2018; Lyard et al., 2020) are included for comparison where known.

		M_2 $\overline{\mathrm{RMSE}}_t$ [cm]		$\overline{\mathrm{RMSE}}_{t\|tot}$ [cm]	
Model	Latitudes	Deep water	Shallow water	Deep water	Shallow water
ADCIRC v54	All	6.5	18.5	7.92	22.1
ADCIRC v55	All	2.87	13.9	3.89	17.2
ADCIRC v55	$\leq \pm 66°$	2.85	14.7	3.81	18.2
Stammer et al. (2014)*	All	5.25-7.76	18.6-27.9	-	-
Ngodock et al. (2016)*#	$\leq \pm 66°$	2.6-3.2	-	-	-
Schindelegger et al. (2018)*	$\leq \pm 66°$	4.4	14.6	-	-
Lyard et al. (2020)**	$\leq \pm 66°$	1.53	6.44	-	-

*: $\overline{\mathrm{RMSE}}_t$ is computed against TPXO8-Atlas rather than TPXO9-Atlas.

#: Uses state ensemble Kalman Filter (perturbed data assimilation).

**: $\overline{\mathrm{RMSE}}_t$ results for hydrodynamic FES2014 computed against satellite cross-over points.

2) Modified Lines 225-234 (end of Section 3.1.1): The deep ocean M2 RMSEt = 2.87 cm (Table 3) is smaller than for the majority of previously non-assimilated barotropic tidal models (Stammer et al., 2014; Schindelegger et al., 2018), and within the range of errors computed for solutions obtained by embedding a state ensemble Kalman Filter (perturbed data assimilation) into a forward ocean circulation model (Ngodock et al., 2016). The recent study by Lyard et al. (2020) carefully tunes local bathymetric data and dissipation parameters (Cf and Cit) to obtain smaller errors (M2 RMSEt = 1.53 cm) than presented here. As noted in Sect. 2.4, in this study we deliberately avoided excessive tuning of the model with the aim to minimize tidal solution errors. Nevertheless, the 5-constituent total tidal error, RMSEt|tot, is less than 4 cm in the deep ocean. In shallow regions, the M2 RMSEt is 13.9 cm, which is essentially the same as presented in Schindelegger et al. (2018), but significantly greater than in Lyard et al. (2020). The total tidal error in shallow water, RMSEt|tot is 17.2 cm, but note that the area-weighted median value of shallow water RMSEt|tot (c.f. Appendix A) is just 6.63 cm.

R: -L249: as shown by several studies (e.g. Townend and Pethick, 2002) and synthetized in Idier et al., (2019), representing flooding in storm surge models results in lower water levels seaward compared to simulations where the flooding is not represented. Therefore, I expect that water levels in the present simulations are biased high due to this process, possibly by 0.5 to 1.0 m considering previous studies on the topic.

A: We agree with the point that you raised regarding that including inundation in the simulation would result in lower water levels seaward. In other words, the maximum coastal water levels shown in Figure 9 are likely biased high. However, here we are comparing to high water marks (HWM) measured on land using the closest modeled wet point. Runup onto the land can amplify the water levels beyond those recorded seaward, but since our simulated results at the coast are likely biased high there is a degree of cancellation involved. Nevertheless, we noted that our closest wet point results might only follow the lower envelope of HWMs as noted by Mori et al. (2014) since the amplification could be greater than the low bias due to not simulating inundation especially in the presence of steep topography.

C: Added to the end of Lines 267-268 (old L249): ... (although ignoring inundation in our simulations is expected to overestimate the seaward maximum storm tide heights (Idier et al., 2019) that likely cancels out some of the otherwise low bias when compared to HWMs).

R: -L256: please refer to Bricker and Roeber (2015) who showed that Hayan also drove very large infragravity waves, which could explain the large scatter on HWMs observed.
A: Thank you, it is a good idea to point out the large scatter in the HWM measurements and this potentially being related to infragravity wave generation over reefs.
C: Added to end of paragraph on Lines 274-276 (after old L256): The large scatter present in the HWM measurements (SD ≈ 1.3 m for all MinEle) could be related to the generation of infragravity waves over fringing reefs in the region leading to amplified coastal runup (Roeber and Bricker, 2015).

R: -Figure 11: for Katrina, the model displays a 0.5 m negative bias before the surge peak, could the authors comment on the possible causes? Could it be related to the 2DH approach which only allows for a crude representation of Ekman transport?
A: On lines 261-264 of the original manuscript we noted that this negative bias could have been due to the neglect of the regional wave setup since a previous ADCIRC-based study that coupled to short waves better matched the time series before the surge peak (Roberts and Cobell, 2017). However, after subsequent simulations by our group on separate but related research, we do not think that this bias is mostly attributable to the insufficient generation of the surge forerunner (e.g., Kennedy et al., 2011). This fact indeed arises from the crude representation of Ekman transport by the 2DH approach as the reviewer surmises, but the negative effect can be mitigated by setting the bottom friction coefficient to a very small value on the shelf. The previous studies by Bunya et al. (2010) and Roberts and Cobell (2017) used a Manning's formulation for the bottom friction coefficient (where n ~ 0.02 in the ocean) which leads to very small values of the Cf on the continental shelf (~50-200 m deep).
C: Changed the old lines 261-264 to the following (new Lines 283-289): We think that this negative bias is mostly attributable to the insufficient generation of the surge forerunner and partly also to the omission of regional wave setup. The surge forerunner is generated through the Ekman setup process (Kennedy et al., 2011) which is crudely represented by the depth-averaged model used here. Previous depth-averaged ADCIRC-based studies that used a Manning's bottom friction formulation so that Cf becomes very small on the continental shelf appear to be better able to generate the surge forerunner, as well as employing wind wave-coupling that generates wave setup, indeed show better agreement with the time series prior to the peak storm tide (Bunya et al., 2010; Roberts and Cobell, 2017).

R: -L357: please correct "are be able"
C: Line 398 (old L357): … are able to …

R: -L376: I'm not sure that this conclusion is very robust based on a model that does not represent flooding (see my previous comment).
A: We agree with your comment that when including inundation in the simulation the seaward maximum storm tide heights would be decreased. So, we modified parts of paragraph in paragraph 4 of "Section 4: Discussion" in addition to this line in "Section 5: Conclusions" to comment on this potential effect, noting that the coarser models would have a greater coastal flooding potential.

C:

1) Lines 389-393: In practice, higher peak storm tide heights in coarser models translates to greater coastal flooding potential. Including inundation in the model would decrease the storm tide elevations along the coast (Idier et al., 2019) perhaps leading to more similar coastal storm tide elevations between the different mesh resolutions since more flooding may occur in the coarser model. Overall, the impacts of mesh resolution on the HWM errors were relatively small, especially for Super Typhoon Haiyan. However, the …

2) Line 419 (Old L376): We found that in general, peak storm tide elevations along the open coast are decreased (therefore the coastal flooding potential is decreased) …