Dear Reviewer,

Thanks very much for your positive feedback and very useful comments. We fix the problems and address your comments as listed below. Your original comments are copied for your reference. Following the journal's publishing guideline, the revised manuscript is not uploaded at this time, so all line numbers and figure numbers are based on the original manuscript.

<Reviewer's Comment #1>

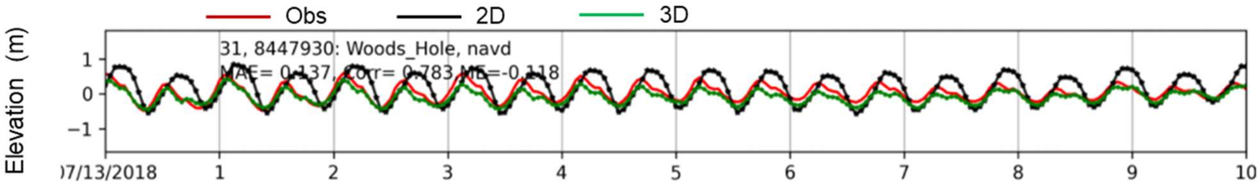
The authors suggest that a 3D baroclinic model is necessary to accurately capture the water level response. However, I suspect that utilizing such a complex model (plus the very large mesh size) would require a very large computational expense. The authors do not present results from using a more simplified approach (2D depth-averaged for example), so we don't know whether the additional complexity is actually needed to capture the water level response. The manuscript would benefit from some discussion about the trade offs between model complexity and computational burden, as well as some details about how long each simulation takes to run.

<Response>

We agree with the reviewer and add more texts after Line 105:

"The trade-off between 2D and 3D must be carefully weighed, especially for operationalization. In short, the advantage of 2D is the speed (about 80 times faster than 3D baroclinic) and the simplicity of the set up. The disadvantage of 2D is that it misses the effects from baroclinic processes that may become important at certain time and locations. The baroclinic effects during the adjustment phase after Hurricane Irene (2011) are discussed in details in Ye et al. (2020), using a similar model setup as the one used here. Even though different setups (2D, 3D barotropic, and 3D baroclinic) were tuned to the best possible skill, the 3D baroclinic setup was shown to better capture the post-storm adjustment phase. In addition, during the ongoing effort to operationalize the model, we found that including 3D processes greatly reduces the need for bottom friction parameterization at some coastal locations."

The figure below shows the model-data comparison for 2D and 3D with a similar model setup as the current paper. Although it is possible to improve the 2D skill by locally tuning bottom friction, the 3D model achieved good skill without any local tuning. We are looking into the momentum budget at this location and will report the findings in a future publication.



We also add the information on computational cost after Line 164 in "Section 3. Model description" of the original manuscript:

"Although the model covers a large domain, most of the elements are quasi-2D, making it efficient enough for operational forecast. For the baseline run, the real time to simulation time ratio is 80 with 1440 cores on TACC's Stampede2 and 30 with 480 cores on W&M's Sciclone. Intel Skylake cores with a nominal clock speed of 2.1 GHz were used on both clusters. A 3-day simulation will take about 0.9 hours using 1440 cores or 2.4 hours using 480 cores."

<Reviewer's Comment #2>

In my experience, the NCEI CUDEM does not capture the bathymetry of coastal streams well, and tends to significantly underestimate the channel depth. Did the authors make any corrections/modifications to the raw DEM to better represent coastal streams?

<Response>

No, we did not do any modifications on the NCEI CUDEM. We also noticed some underestimation of channel depth, especially in the South Carolina watersheds, which may have contributed to the errors on HWMs in Fig. 11c.

Our collaborators at NOAA (including those in the author list) is working on introducing improved DEMs to our model, including some DEMs not yet open to the public. Sometimes we manually correct DEMs to make channels continuous based on navigation charts, imagery, and our best judgement; but we have not done so in the current study.

We will add this information on Line 137 of the original manuscript, so that the revised manuscript reads:

"Customary of all SCHISM applications, no manipulation or smoothing of bathymetry was done in the computational grid after interpolation of the depths from DEMs (including steep slopes in the Caribbean and all shipping channels). From our experience, CUDEM may underestimate the depth of coastal streams (e.g., in the South Carolina watersheds), which is a potential error source of our model."

<Reviewer's Comment #3>

There is no mention of infiltration in this manuscript, which is an important factor controlling the pluvial flood dynamics. Did the authors completely neglect infiltration in this work? And if so, there should be some discussion/justification about why infiltration was not accounted for. Since much of the underlying soil in the NC region is sand, I expect infiltration losses would be non-negligible and neglecting infiltration could cause over-estimation of the rainfall-induced flooding.

<Response>

Yes, we neglected infiltration in this work. We agree that infiltration may be an important process. But in the case of Hurricane Florence induced flooding, we expect the effect of infiltration to be minor.

The revised manuscript reads:

[Line 184] As a model limitation, infiltration is neglected in this work. In the case of Hurricane Florence induced flooding, we expect the effect of infiltration to be minor. According to the NOAA's weather map (https://www.wpc.ncep.noaa.gov/dailywxmap/), there was continuous rainfall along the US east coast from Sep 11, 2018 to the date of Florence's landfall (Sep 14, 2018), so the infiltration capacity of the soil was already reduced. Moreover, "wet" storms like Hurricane Florence (2018) and Hurricane Harvey (2017) tend to dump large amount of rain fall at a location for days because of the slow movement of the storm, so most of the rainfall is on saturated soil. The drainage in urban settings is not included in our model. This may lead to some big errors in the prediction of elevations on high water marks; for example, the one large error in the urban area of Fig. 13d. We do have a plan of explicitly accounting for infiltration as volume sinks based on NWM (or other hydrologic models). However, considering the additional uncertainty this would bring, for now we choose to continue improving more important aspects of the model (especially the quality of model grid, which is likely responsible for most of the large errors in Fig. 11c) for operational use.

< Reviewer's Specific Comment #1>

Line 106: What is meant by "bona fide" in this context? Please clarify.

<Response>

We delete these two words and clearly identify the merit of the current model:

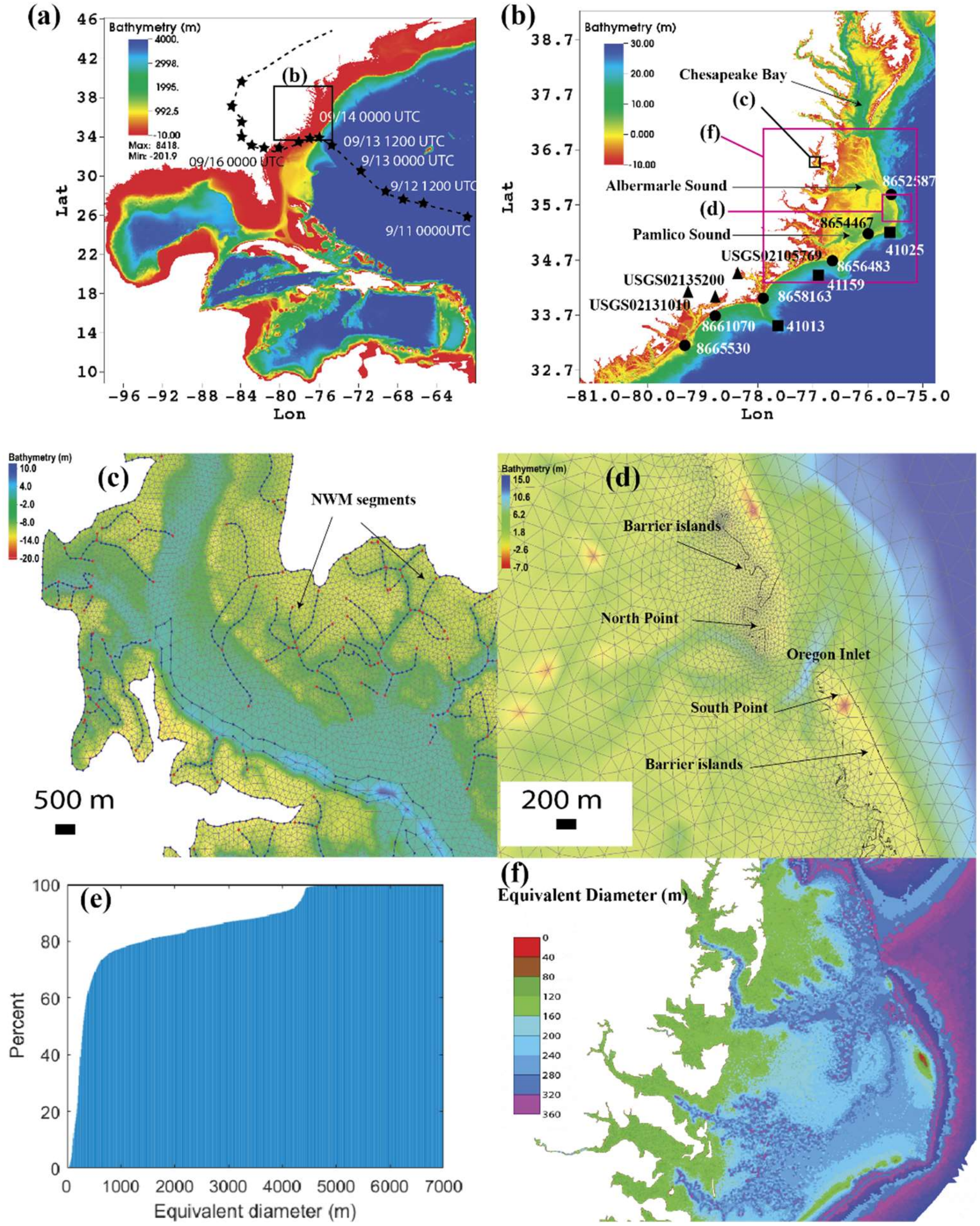
As we shall see, this model solves the physical processes from the watershed to the ocean with the same set of governing equations, qualifying for Santiago-Collazo et al. (2019)'s definition of a fully-coupled compound surge and flood model.

< Reviewer's Specific Comment #2>

Figure 3: I am a bit confused by figures 3b and 3c. The box showing the location of figure 3c indicates that 3c is north of Albermale Sound. However, when I look at Figure 3c it appears to be depicting the Cape Fear River Estuary, which is at the southern tip of NC. Please confirm/update the actual location of Fig 3c.

<Response>

Thanks for pointing this out. We have corrected the location of Fig. 3c:



< Reviewer's Specific Comment #3>

Table 1: The difference between the No_NWM_precip and Ocean scenario is not clear to me. It seems they both do not include river or precip forcing?

<Response>

"No_NWM_precip" and "Ocean scenario" are identical. We will only use "Ocean scenario" in the revised manuscript. Thanks for catching this issue.

< Reviewer's Specific Comment #4>

Fig 9: It is difficult to distinguish the differences between the model runs, especially for figures 9a-d. I suggest the authors shorten the depicted time window to only show 1-2 days before landfall through 4-5 days after landfall for a-d. This will allow a reader to see more clearly the differences between the scenarios and the comparison to the observed water levels.

<Response>

Another reviewer also asked for changes on this figure. To accommodate both reviewers' request, we will shorten the time window and also add sub-tidal comparisons in Fig 9b (see AC3).

<References>

Santiago-Collazo, F.L., Bilskie, M.V., Hagen, S.C.: A comprehensive review of compound inundation models in low-gradient coastal watersheds. Environ. Model. Software 119:166–181. https://doi.org/10.1016/j.envsoft.2019.06.002, 2019.

Ye, F., Zhang, Y.J., Yu, H., Sun, W., Moghimi, S., Myers, E., Nunez, K., Zhang, R., Wang, H.V., Roland, A. and Martins, K., 2020. Simulating storm surge and compound flooding events with a creek-to-ocean model: Importance of baroclinic effects. *Ocean Modelling*, *145*, p.101526.

Zhang, Y.J., Ye, F., Yu, H., Sun, W., Moghimi, S., Myers, E., Nunez, K., Zhang, R., Wang, H., Roland, A. and Du, J., 2020. Simulating compound flooding events in a hurricane. Ocean Dynamics, 70(5), pp.621-640.