Summary

This study presents an uncertainty analysis for process-based flood models during a compound flood event (Hurricane Harvey). The authors verify different sources of uncertainty and determine which contributes more to the overall error using a machine-learning approach.

Major Comments

I would like to congratulate the authors for a wonderful piece of sciene the have in this manuscript. The study is worth of publication given its nature and novelty around compound flood (CF). However, the manuscript needs some improvements to be publishable in this journal. First, the introduction section needs improvements to help the reader follow the manuscript story. First, I would suggest breaking up the model uncertainty sources for process-based models into individual paragraphs instead of grouping them into pairs in two very long paragraphs. Second, I was impressed not to find the study objective/aims within this section. For example, the section ends with a literature review paragraph about previous uncertainty studies instead of the study goals/objectives and the manuscript layout of the following sections. Furthermore, the authors fail to lay out the research gaps this study tries to bridge and thus are unclear why and what they are trying to do. If it was not for the abstract that I read first, I will be completely lost about the study goals based only in the introduction.

Second, the order of some sections could be improved to enhance the story-telling flow of the manuscript. For example, I should know first which hydrodynamic model (Section 2.3) you used instead of the data (Section 2.2) and case study (Section 2.1). This is highlighted in the caption of Figure 1 that mentioned Delft-FM and 2D HEC-RAS but I was unaware that the authors were planning to do two different modeling approaches. This could also be improved if mentioned in the introduction, as suggested above. I recommend talking about the models selected first, then the case study and lastly the data for this section. If not, you should at least create a brief paragraph before 2.1 and give the reader a broader vision of the methods, such as the models to be used, the study area, and any other relevant details, like a summary of this section. Third, the calibration section (Section 2.3.2) seems quite long and confusing regarding the different calibrating scenarios. I would divide this subsection into three different ones, one for tides only calibration, another for hurricane calibration, and another for hurricanes validation. Also, this could be considered as its section outside of the methods since they include numerical model results. Fourth, a limitation section within the manuscript should be added. Currently, limitations are all over the manuscript, and it would help the reader if they are summarized in a single location.

Lastly, the result and discussion section is very long. I would recommend splitting this into a result and discussion section separately. This will follow the traditional journal articles more, and we can better differentiate the discussion from the results.

Minor Comments

- L28: the use of "(56%) 49%" is not clear in the provided context and could confuse the reader. I suggest to revise this statement.
- L32: I will remove the "i.e." and just put it in parenthesis next to the "gross domestic product"., similarly to the below statement of the population percent.
- L35: you should define what is "in the past five years" since it could be from 2023 (probably when you wrote the statement) but the manuscript could be published in 2024.
- L38: I would add "flood" between "coastal drivers" to emphasize the flood hazard. Also, the compound flood definition should be referenced to some of the first publications that studied and defined this in detail, like Bilskie and Hagen (2018).
- L47-65: On this paragraph you mention the three main ways to compute CF. However, you only explain two of them in detail. You should add a couple of sentences describing the hybrid approach since the reader may not be familiar with that term.
- L83-68: while the statement about CERA is truth, the authors should comment that is only of the modes that CERA operates, since it also have a compound flood tools for LA. I will suggest rewriting the statement to highlight the above.
- L90: another source of model uncertainty within the model parameters is the soil moisture (antecedent conditions), and should be briefly discussed in this paragraph, especially if you are talking about compound floods.
- L102: I was expecting that the authors would also include the coupling approach as part of the model structure uncertainty. There is vast literature comparing the different coupling approaches (one-way, two-way, tightly and, fully coupled) for CF and how that affects the results. Regardless, if the authors did not test this option, I would still include it in this paragraph to highlight the potential of an additional uncertainty source.
- Figure 1: what is the purpose of having panel b and c? They look very similar (regarding the topobathy) and there is no discussion about this on the text. Also, the figure caption says that Ike was in 2009 but it should 2008. I would prefer to see the numerical mesh of both models side-by-side than the topobathy.
- L148: the authors should comment why they did selected these two events as case studies. These two events are hurricane and we can classify them like CF events. However, sometimes their impacts does not reflect a CF event. For example, Ike was an event mainly dominated by coastal process flooding, whereas Harvey was the opposite and dominated by the hydrologic process. I would like to see this type of statement in this section.

- L186: it should say “Forcing or boundary conditions”, right? Also, WL is already defined, why defined it again?
- L194: I do not see in Figure 1a the HWM from Ike. Is that what the authors are referring to? Please rephase the sentence since if my interpretation was incorrect.
- L191-194: Why do the authors only mention the HWM marks from Ike and not Harvey? I would assume there are multiple reports of flood levels for Harvey that could be used. I would also mention them here.
- L197: why did the authors not use a higher resolution precipitation source, such as the Stage IV dataset from NCEP (https://data.eol.ucar.edu/dataset/21.093), which is at a 4km spatial resolution and available for the US? Please justify your selection since we typically use the ERA5 rainfall data for remote locations that does not have these high resolution datasets. There are even studies in this journal that talk about the inaccurate performance of this dataset (ERA5) for compound flood (https://nhess.copernicus.org/articles/23/3379/2023/). Also, gridded rainfall have proven to be more accurate for flood estimation that rain gauges due to their limited coverage and are mostly use to correct the gridded rainfall products.
- L206-208: I strongly suggest the authors remove the governing equations the models are solving, especially if they are not modifying them directly. They seem to be unnecessary and the authors can reference to other publications that introduce the model and its governing equations.
- L223: you should be consistent with your use regarding G-Bay.
- L232: this statement is repeated, regarding hec-ras capabilities. Same on L256.
- L281-282: The authors should provide the equations used to evaluate the model performance, at least in the appendix or cited from another source that used the same equations.
- L296-297: This seems to be a very vague justification of why they did not run the tidal simulations. There is currently a linux version of HEC-RAS (v 6.1) that can be run in HPC systems and is easy to install. Probably, the justification could be that the Delft can run in parallel while HEC-RAS will run in series within HPC and would take significantly more computer resources and time. Please address this.
- Table 1: Why does the optimal only tides row not have a roughness value for all except open water? It is my understanding that tides were only run in Delft; thus, we should see values here, right?
- Figure 3: I will rearrange the order of the panels in this figure. It would be better for the reader if the panels are group by storm event in each column and by gauge in each row. I would also remove the NOAA tide gauge id and replace it with the location name for easy recognition. The authors have also to shown where in Figure 1a are each of these gauges located since they only have with a start and not the specific gauge name/ID.
- L351-352: rephrase the sentence to mention that Delft is better than HEC-RAS for G-Bay. As it is now, it seems like Delft is the best model in the modeling community for G-Bay.

- Figure 4: move the figure earlier in the text, it should be near line 353. Also, it is not clear the model that was used to create panel c and d, please specify on legend. Why not include both max depth flood maps for each model and hurricane? Why not consider presenting a different plot map based on the flood depths as raster (thus, eliminating the different mesh configurations)?
- L355: why the uncertainty assessment with 5 scenarios are only performed with Harvey? I would expect both so the reader can see if the results are associated with the dominating flood. For example, Ike was mainly driven by surge, while Harvey was hydrologic. Furthermore, why on table 2 you mention that would use Harvey but are using calibrated values from Ike. Similarly, why not consider those 5 scenarios both all of them for HEC-RAS and repeat for all of them Delft, instead than a combination? The text on this section explain well all this, but the table does not, thus potentially confusing the reader.
- Figure 6: those this figure intended to have eight panels? I see reference on the caption to panels (g,h) but only see up to d. I prefer to see all scenarios within the manuscript instead of going back and forth to the supplement figures. I would suggest keeping all the zoom-out maps for all scenarios in the manuscript and moving to supplement the zoom-in maps for all scenarios.