Reply on RC1

General Comments:

This article presents a new flood modelling framework, focusing on compound flood hazard from fluvial, pluvial and coastal flooding. It is well written and provides appropriate detail around the methodology. The model, which uses global inputs to build a local scale flood model will be a valuable tool to investigate compound flooding.

The models appear to have skill on par with other models built on global data, although additional case studies and/or validations would be nice to see how the framework performs in a range of different situations. I would recommend the article for publication with minor modifications:

We would like to thank the reviewer for the thorough review and comments, which we believe have led to an improvement in the manuscript. We are pleased to read that the reviewer considers the manuscript to be well written and the presented framework a valuable tool for compound flood analysis. Based on the suggestions of both reviewers we have made some changes to the manuscript.

To improve the robustness of the analysis we have also extended the validation dataset with two more flood extent images, one for each event. To obtain this extended dataset in a consistent manner, we have processed the Sentinel-1 images with a slightly different algorithm as this algorithm was available to our team (see section 3.3.1). The new algorithm provides very similar results compared to the RAPID product used in the original submission. Furthermore, we have made some small improvements to the model setup by including more small rivers as boundary conditions to the SFINCS model and focus the analysis on the areas connected to the Buzi and Pungwe floodplains.

Our response to the specific comments can be found in the paragraphs below.

Specific Comments:

1. As above, case studies for a different location would strengthen the validation.

For this manuscript we have deliberately selected only one case study which allows us to provide more context on the studied events and to demonstrate how the framework can be used to analyze compound events. We have clarified the goal of the paper at the end of the introduction. However, we agree that additional validation would be useful in further research, as mentioned in the discussion.

L70: The goal of this study is to present the framework, to test its ability to simulate compound floods in data-sparse coastal deltas, and to demonstrate how it can be used for compound flood analysis

L518: To further increase the credibility of the model it should be validated against a larger set of flood events, for instance using the recently published Global Flood Database based on MODIS data (Tellman et al., 2021) or the RAPID sentinel-1 database over the continental United States (Yang et al., 2021).

2. Validation observations: The observations used in the validation are based on single snapshots for each event.

- Firstly it would be useful to add some context at where these snapshots fit into the event (how close are they to the fluvial/ coastal flooding peaks for example).
- Similarly, can you comment on how the timing of the observations influences the validation? E.g., Would the results in Table 4 -3 showing reduced sensitivity to coastal drivers be connected to the time of observations being before/after the coastal flooding had receded?
- Are there any alternative observations of the flood hazard for the same events you could use (e.g. extents from Copernicus EMS)?

Thanks for these suggestions. We have expanded the analysis by processing additional Sentinel-1 observations using an unsupervised histogram-based surface water mapping algorithm (Markert et al., 2020). We did not use additional images from the Copernicus EMS as these only cover a part of our area of interest and are mostly based on the same Sentinel-1 images. For each event we now have two observations on subsequent days. We compare the individual observed flood extents with the maximum simulated extent from the same day and the maximum observed extent per event with the maximum simulated extent during all days with observations. The observations during Idai are around the flood peak and for Eloise just before the flood peak. We describe the algorithm and the temporal context of the observations in section 3.3.1 (L339 & L452)

The sensitivity analysis is based on the maximum extents per event. We added a sentence in section 4.1 (L424) to clarify that the sensitivity to the drivers might be different if compared for a snapshot after the surge peak.

L339: High-resolution (10 m) flood extent data are derived from Sentinel-1 Synthetic Aperture Radar (SAR) images. We use VV-polarized ground range detected level data, provided by Google Earth Engine (GEE), which has undergone geometric terrain correction and provides radar backscatter in decibel (dB) units. These data are processed using the GEE with an unsupervised histogram-based surface water mapping algorithm that consists of three steps (Markert et al., 2020). First, noise is reduced using the Refined Lee speckle filter (Lee, 1981). Second, a threshold to distinguish water and dry cells is detected using the Edge Otsu thresholding algorithm (Donchyts et al., 2016). Third, cells *with a relative elevation of more than 50 m above the nearest stream are excluded from the water class to avoid false positives. We process each image individually and combine flood extents from ascending and descending orbits during the same day. In total we obtain flood extents for four days based on eight images: on the 19 and 20 March 2019 for Tropical Cyclone Idai which is around the peak of the flood event, and on 25 and 26 January 2021 for Tropical Cyclone Eloise which is just before the peak of the flood event. Finally, the flood extents are reprojected to the SFINCS model grid.*

L352: We compare the individual observed flood extents with the maximum simulated extent from the same day and the maximum observed extent per event with the maximum simulated extent during all days with observations.

L424: The skill is likely more sensitive to the coastal water level forcing if assessed for a snapshot around the surge peaks of both events instead of the multi-day maximum flood extent.

3. It seems like there are only river inputs at the boundaries of the domain, is this correct? Does that mean in the coupled framework, the pluvial rain/runoff on-grid implementation is relied on to convey water in tributaries completely within the domain? Is there likely to be an underestimation of river flows due to this (for example if this framework is used to simulate larger domains)?

For this case study we indeed only applied discharge boundaries at the boundaries of the model domain where a river enters the model domain, where rivers are defined based on a minimal contribution area of 100 km2. Within the domain we apply the ERA5 runoff (i.e. net rainfall) directly to all the SFINCS grid cells and thus rely on SFINCS to convey the water. To prevent under/overestimating river flow in large domain models it is important to correctly schematize the conveyance capacity of the rivers and smaller streams (see also section 4.3).

Technical comments/ corrections:

1. Abstract: You refer to 'local scale' models (L13) to refer to models built with high quality local information, but then refer to the SFINCS model as a 'local model' in L22 and L24. Could you reword this to make it clear these are referring to different things?

Good suggestion. We now refer to the SFINCS model within the presented framework as the "globally-applicable model" throughout the manuscript..

2. L17: 'loosely coupled' is not clear, consider rewording.

We have removed loosely in the abstract and introduction, but explain it in more detail in the methods:

L282: SFINCS is forced based on output from global models, which is automatically transformed to the input data format that SFINCS requires. This is also referred to as a loose coupling between models (Santiago-Collazo et al., 2019).

3. L135: This is a little confusing. It would be useful to know that the total water height consists of two components (L138-140) first. Also that the ERA5 data comes from the ECMWF Ocean Wave Model.

We have clarified that the ERA5 significant wave height data originates from the ECMWF OCean Wave Model.

L129: Hourly time series of significant height of wind waves (Hs) are extracted at GTSM output locations from the 30 arcmin ERA5 dataset and based on the ECMWF Ocean Wave Model (Bidlot, 2012; Hersbach et al., 2020).

4. L144: It's not too clear where the gauge locations are here.

We have clarified the location (port of Beira) in the text.

L149: In comparison with the tidal constituents of International Hydrographic Organization (IHO) station at the Port of Beira as retrieved using the Delft Dashboard (van Ormondt et al., 2020), […]

5. Table 1: Are you using MERIT DEM or the MERIT-hydro hydrologically adjusted elevations?

We use the MERIT-hydro dataset as these data provide a better estimate of river elevation, which we use to estimate the river bathymetry (see section 3.2.1).

6. L255: should this final sentence have a separate bullet point?

Yes, we have changed the text accordingly.

7. L302: is the 'gridded discharge dataset' referring to CaMa-flood?

Yes, that's correct. For step 7-9 we first discuss the general implementation, followed by the specific data and choices made for the case study. In L304 we explain that for this case study we use discharge from the gridded CaMa-Flood output.

8. L307: add: maximum relative error in upstream area

Good suggestion. We have changed the text accordingly.

L304: [..] based on a maximum relative error of 5% or absolute error of 100 km2 in upstream area.

9. Figure 6: Panels are mislabelled

In this case the caption of Figure 6 was incorrect (not the labels). We have changed the caption to match the labels.

10. L552: The zenodo link doesn't work (but does if the final slash is removed).

We have fixed the URL.

References

Markert, K. N., Markert, A. M., Mayer, T., Nauman, C., Haag, A., Poortinga, A., Bhandari, B., Thwal, N. S., Kunlamai, T., Chishtie, F., Kwant, M., Phongsapan, K., Clinton, N., Towashiraporn, P., and Saah, D.: Comparing Sentinel-1 Surface Water Mapping Algorithms and Radiometric Terrain Correction Processing in Southeast Asia Utilizing Google Earth Engine, Remote Sensing, 12, 2469, https://doi.org/10.3390/rs12152469, 2020.