

Authors' response to anonymous referee #2

Authors: Luca G. Severino, Evelyn Mühlhofer, Nishadh Kalladath, and David N. Bresch
Date: 20 October 2026
Journal: *Natural Hazards and Earth System Sciences*
Revised manuscript: *Beyond the flood map: an open-source framework for forecasting cascading impacts of natural hazards on people and essential services*

1 General comments

Referee comment: I enjoyed reading the article by Severino et al. which presents an open-source framework for forecasting cascading impacts from riverine floods, with a focus on impacts on healthcare access using case studies in East Africa. The topic is of great interest, given the increasing trend to focus early warning systems on disaster impacts, rather than just hazard magnitude. Overall the article is well written, though the grammar and English wording would benefit from a review by an English native speaker. Models and data used are mostly open source, which is appreciated. I think that the proposed framework has strong potential to improve the state of the art in flood impact forecasting, provided that some improvements are made on the current version. Main concerns about this work are detailed below, followed by a list of more specific comments.

Authors' response: We thank the reviewer for the kind and constructive words. We have taken into account the different suggestions to improve our manuscript. In particular, the revised manuscript contains the following changes with regard to the reviewer's comments:

- The section "Model evaluation against satellite observations" has been moved to the appendix following the suggestion of reducing the importance given to the validation of the flood modelling data. In addition, the discussion dedicated to this aspect has been reduced.
- We complemented the manuscript to clarify some of our modelling choices and remind the readers that the purpose of the publication is to illustrate the use of the model, and that some of the parameter choices have been made subjectively.
- We have added a small complementary analysis, using flood satellite data instead of modelled flood data, following the suggestion to include this in our analysis.
- We have updated some of the figures to improve their readability and interpretability.

Below you will find our detailed responses to the different comments and suggestions. The changes refer to the revised version of the manuscript that we are planning to resubmit. We are grateful for the time and attention that anonymous referee #2 and the editor gave to our submission and

hope that the revisions will improve the quality of the manuscript.

2 Major comments

2.1 Major comment 2.1

Referee comment: First, the original contribution of the article, that is the estimation of cascading impacts from riverine flooding, is somewhat diluted within other analyses that are less prioritary for this work. I'm not sure why the authors put so much focus on validating GloFAS forecasts versus satellite based estimates, which is certainly of interest for GloFAS developers but less for this work. The meteorological and hydrological model uncertainty plays a big role in the evaluation and these can be excluded if the evaluation is made directly using satellite derived inundation extent as input and compared it with observed data for the events.

Authors' response: Our original contribution is not only the estimation of cascading impacts from riverine flooding, but also the evaluation of its application to early warning and anticipatory action. In order for this evaluation to be complete, we insist on keeping the analysis of errors originating from the hazard forecasting in the analysis, as this plays a big role in the evaluation. However, we agree that this aspect might have taken too much space in the paper and in the discussion. Hence, we revised the manuscript to make our original contribution clearer and put less emphasis on the evaluation of the flood forecast. In the revised manuscript, the section "Model evaluation against satellite observations" has been moved to the appendix, and discussion relating to the performance of the flood forecast model has been shortened. In addition, results obtained by running the model using the satellite-detected instead of the modelled flood extents are shown in the section "Model evaluation against survey data and uncertainty analysis" and discussed in the discussion section. The paragraph amended to the section "Model evaluation against survey data and uncertainty analysis" and the updated figure can be found below:

We test whether the model can better reproduce the surveyed post-flood healthcare accessibility rates by replacing forecast flood extents with satellite-detected flood extents obtained from UNOSAT (2023) to compute the post-flood healthcare access. We use as a default model configuration the parameter values defined in Tab. 1, WorldPop as the population data source, OSM as the road data source, Maina et al. (2019) as the healthcare data source and the impact functions with a 50cm threshold for all three types of exposures. The model runs obtained using the default model configuration but replacing the flood model with satellite data are shown as blue dots in Fig. 7. Model runs obtained using the default configuration and the median forecasted flood extent, without including flood protection, are shown as yellow dots, for comparison. The comparison reveals that replacing the flood model with satellite observations only helps significantly reduce the the modelling error in Hirshabelle for the healthcare access loss, where it reduces the modelling error from -0.25 to -0.08, and for the healthcare access undisrupted in Jubaland where it reduces the modelling error

from -0.35 to 0.01 . In the remaining cases, replacing the flood model with satellite data does not help significantly reduce the modelling error with respect to the survey data. This suggests that significant sources of error originate from the other components of the risk modelling framework, exposures, vulnerability, and the modelling of service accessibility.

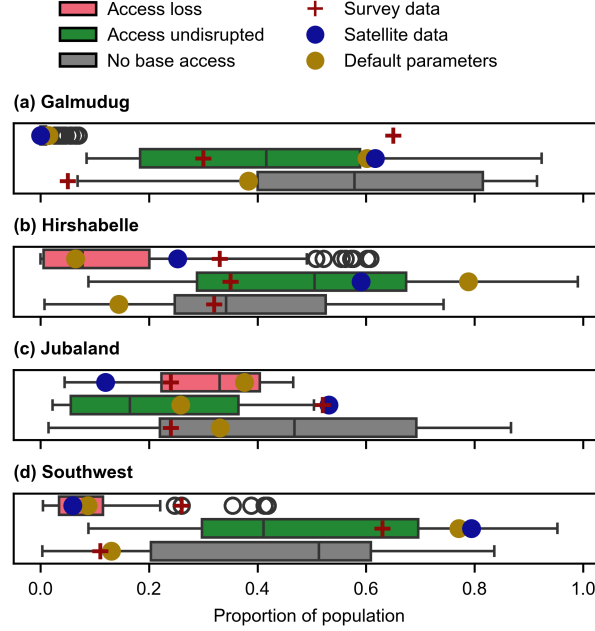


Figure 7: Comparison of modelled versus surveyed flood-induced healthcare access loss for the 2023 Deyr floods in Somalia. Model spread is quantified following the uncertainty analysis described in section “Uncertainty and sensitivity analysis”. Panel (a) shows results for the region of Galmudug, (b) Hirshabelle, (c) Jubaland, and (d) Southwest. Red crosses represent survey data obtained from [Federal Government of Somalia et al. \(2024\)](#), yellow dots represent model runs with the default configuration and using the median flood forecast from the 50-members ensemble without protection. Blue dots represent model runs with the default configuration but replacing the flood forecasts with satellite-detected flood extents from [UNOSAT \(2023\)](#). The default model configuration is defined with the default parameter values defined in Tab. 1, WorldPop as the population data source, OSM as the road data source, [Maina et al. \(2019\)](#) as the healthcare data source, and the impact functions with a 50cm threshold for all three types of exposures.

2.2 Major comment 2.2

Referee comment: Second, there are a number of assumptions made that have a considerable influence on the model results and weaken the approach. While I appreciate the inclusion of a sensitivity analysis to assess the different assumptions made, some choices should be made more subjectively. For example, by default I think that the median (50%) of the ensemble should be

taken, rather than the 90th percentile. Anyways, this would not be needed in case the satellite based inundation map is taken as input.

Authors’ response: Our intention with this paper is to illustrate a modelling pipeline for a certain application. To this end, we selected some modelling parameters, and made some assumptions, which we considered to be the best suited to illustrate certain aspects of this modelling pipeline and its application. We have revised the manuscript to clarify that these assumptions are modelling choices that have been made mainly for illustrating purposes and that they should not be interpreted as recommended parameters or settings. We recommend users to revise these modelling assumptions for operational purposes. Regarding the choice of the 90th percentile over the 50th percentile, we chose a more extreme percentile to illustrate a “worst-case scenario”, as we explain in the manuscript. However, users are free to use any percentile for their application, and we do not claim that using this specific percentile is better suited than any other arbitrarily chosen percentile. Maps obtained using other percentiles are available in the appendix, for comparison (cf Fig. A1 and B1 in the first submission of the manuscript).

Added in the section “Limitations, improvements and future research”:

Similarly, our analysis focused on demonstrating the approach, and we did not perform a thorough calibration of the model. The different modelling and parameter choices taken in this study are illustrative and do not represent ideal parameters or settings that should be adopted for operational use of the framework for early warning or anticipatory action. We recommend users to revise these modelling assumptions for operational purposes. For instance, using data at higher resolution might help reduce the modelling errors.

2.3 Major comment 2.3

Referee comment: Third point is related to the model resolution, where for flood impact assessment is very important to use the highest available. Why did the authors downscale the resolution of the inundation maps and of the population grids? Both data are available from their respective source at 90m resolution, hence it would be important to run the assessment at the same resolution, and given the limited size of the two case studies it should not be a problem from the computational side.

Authors’ response: Thank you for highlighting this. We agree that using high-resolution data is important and can help to obtain more precise and accurate modeling results. However, we focused our analysis on demonstrating a modelling framework, and we did not seek to necessarily obtain the most precise results possible. Both flood data and population data were up-scaled for computation purposes. Using the flood data at its original resolution would have resulted in memory requirements that exceed typical capacities of commercial laptops to process the full ensemble of 50 members with 10 separate time steps. For instance, for a country like Sudan and the 103 days considered in our analysis this would have resulted in a requirement of approximately 150GB of RAM and 10TB of disk storage. For these reasons, we decided to run the analysis in lower

resolution. This is further motivated by the fact that we believe that using lower resolution data does not qualitatively change our results, which are mainly about illustrating the application of the model. We up-scaled the population data for similar computation reasons, as the uncertainty and sensitivity analysis required generating an ensemble of 5632 simulations for each of the four regions assessed. We tested the influence of the resolution of the population data in the appendix (see supplementary section “Sensitivity to the resolution of the population data” below). The results show that the resolution of the population data does not change our results qualitatively, and only changes them quantitatively to a limited extent. Maps obtained using high-resolution population data can also be found in the supplementary (cf Fig. C1 and D1 in the first submission of the manuscript). In addition, we added in the discussion that using data at higher resolution might contribute to reduce the modeling errors (cf. our response to the comment 2.2).

Supplementary section “Sensitivity to the resolution of the population data”:

We assess the sensitivity to the resolution of the population data by computing the difference between modelling results obtained with WorldPop at its native resolution of 1km, and with the up-scaled resolution of 2km. Table B1 shows the difference between the two resolutions as:

$$\Delta S(\text{resolution}) = S(\text{WorldPop}_{2\text{km}}) - S(\text{WorldPop}_{1\text{km}}) \quad (1)$$

where S represents one of the access statistics (healthcare access loss, healthcare access undisrupted, healthcare no base access), expressed in terms of proportion of population affected. The comparison reveals that the resolution of the population data only limitedly affects the accessibility statistics, with an average difference below 0.05 of population affected for each statistic across the four regions. Comparing the influence of population resolution across the different regions, the strongest influence of the population resolution is in Jubaland, with an average difference across the three metrics of 0.08. In Galmudug, Southwest, and Hirshabelle, the average difference is below 0.02. This suggests that the resolution of the population data does not affect our results qualitatively, and that results obtained using a higher resolution for the population data are likely to be contained in the uncertainty intervals displayed in Fig. 7. However, we are not able to rule out the effects from parameter interaction with this complementary analysis, which might increase the total influence of the population resolution on the results.

Table B1: Comparison between population resolution

Region	Access loss	Access undisrupted	No base access
Galmudug	−0.0	−0.03	0.03
Southwest	0.0	−0.02	0.02
Hirshabelle	0.0	−0.01	0.01
Jubaland	−0.04	−0.08	0.11

2.4 Major comment 2.4

Referee comment: Fourth, the choice of the figures and their content. Visually they are all nice looking, but from the practical aspects of presenting results, some of them (Figure 4, 5, 7, A1, B1, C1, D1, H1, I1, J1) are a bit too crowded with information, especially color shades, and difficult to understand in all their aspects. Some thinking should be made to try to improve the presentation of results.

Authors' response: Thank you for these suggestions. We have revised the figures to improve their readability. You can find a revised version of Fig. 4 below. Same changes have been applied to Fig. 5, A1, B1, C1, D1. Fig. 7 (Fig. A1 and A2 in this document), has also been revised for improved readability (cf. specific comment on Figure 7 below).

3 Specific comments

Referee comment: L95: Perhaps here you could motivate the regridding from 3 to 20 arc seconds

Authors' response: The aggregation has been made to reduce the computational and memory burden. This is clarified in the revised manuscript as follows: *For computational purposes, these maps are up-scaled using an aggregation factor of 7.*

Referee comment: L104: 2 km resolution is rather coarse for impact analysis, given the high resolution needed to represent flood scenarios. How this assumption can increase the uncertainty of the model output?

Authors' response: We added a small comparison of the results obtained at different resolutions in the appendix of the revised manuscript. Results show that increasing the resolution only changes our results to a limited extent (cf. our response to the major comment 2.3). Hence, we believe that this assumption does not significantly increases the uncertainty of the model output.

Referee comment: L107-109: The choice of a 50 cm threshold would benefit from some supporting reasoning. Is it also used to bias correct the model output? I find it quite high, as even just 30-40 cm seems enough to cause disruption to the population and services.

Authors' response: The supporting reasoning is first to use this a priori as it has been used in other studies (e.g. [Kam et al., 2021](#); [Riedel et al., 2024](#)). The 50 cm threshold is not used to bias correct the model output, it is only used as a vulnerability threshold to compute the impacts. However, we have masked flood extents where flood heights are below this value on the plots to highlight where the exposure interacts with a flood high enough to produce impacts. Secondly, the uncertainty and sensitivity analysis shows that this parameter has almost no influence on the results, so we believe that the choice of the threshold value is not critical for this particular application. This aspect is also discussed in the results and discussion of the manuscript, see below.

Discussion in section “Sensitivity analysis”

We also note that our modelling results do not show a significant sensitivity to vulnerability thresh-

olds. This may be because the forecasted flood depths are systematically higher or lower than the vulnerability thresholds.

Discussion in section “Limitations, improvements and future research”

Similarly, the uncertainty associated with other modelling components such as the vulnerability might have been underestimated due to their simplistic representation in our modelling framework.

Referee comment: L151-153: FLOPROS is a dataset of flood protection standards. Please clarify the meaning of “Uncertainty associated with flood adaptation is approximated using the FLOPROS dataset”

Authors’ response: This has been clarified by adding a discussion of how [Riedel et al. \(2024\)](#) incorporates FLOPROS to their flood modelling pipeline at the end of section “Flood forecast” and by updating the sentence *Uncertainty associated with flood adaptation is approximated using the FLOPROS dataset* with *We quantify the uncertainty associated with flood protection measures by considering the FLOPROS flood layer from the model of [Riedel et al. \(2024\)](#).*

Amended to section “Flood forecast”:

The model of [Riedel et al. \(2024\)](#) also proposes an additional flood layer that uses the FLOPROS dataset to model the effects of flood protection measures on inundation ([Scussolini et al., 2016](#)). FLOPROS is a globally consistent database of flood protection standards which has been derived by combining information from different sources such as policy documents, technical reports or modelling results. [Riedel et al. \(2024\)](#) assigns flood depths of 0 to locations where the modelled flood return period is lower than the flood protection standards of the FLOPROS merged layer at the same location.

Referee comment: L187: International Organization for Migration IOM appears twice

Authors’ response: Thank you. This has been noted and removed.

Referee comment: Figure 2: The figure caption should be expanded with more details. Also, there should be a reference to each figure in the text (before the actual figure).

Authors’ response: We have added a reference to the figure in the revised manuscript and added a more detailed figure caption.

Referee comment: Figure 3: The caption should clearly state that these time series are model output. Are these figures based on reanalysis or forecasts?

Authors’ response: The time series are model outputs, based on the reanalysis. We have updated the caption accordingly, thank you for the suggestion.

Referee comment: L233-236: The part of the triggers should be clarified. I have the impression that these triggers were set a posteriori, looking at when affected population experienced sharp increases. However no forecast data was mentioned so far, so the process leading to “response

funds payout based on predefined triggers to be released earlier” is not clear at this point.

Authors’ response: The triggers are not set a posteriori. One of the triggers is an operational trigger that was used during this event by the IFRC and Sudanese Red Crescent Society. The second trigger is an hypothetical one; we use the operational trigger value to look up the time at which our simulations reach this operational trigger value. From this we obtain our “Alternative trigger date”, which only represents when our model including indirect impact would have reached this trigger value. In the discussion, we have clarified how this may be used for rapid impact assessment, which does not necessarily require forecast but can be done with reanalysis or observations.

Amended to the section “Application to anticipatory action”:

Our results also suggest that including indirect impacts, such as disruption in access to essential services, amplify the estimated magnitude of the impacts of a disaster, in agreement with the literature (Mühlhofer et al., 2023). Including indirect impacts in damage assessments can thus help better estimate the scale of a disaster and the needs of the affected population. Incorporating indirect impacts to rapid post-disaster damage assessments could be used to help adjust the disaster response, for instance by leading pre-defined triggers for disaster fund release to be reached earlier.

Referee comment: Section 3.1: Section names read as if they were article titles. I suggest turning them to more standard and concise versions.

Authors’ response: We have updated the section names according to your suggestion in the revised manuscript.

Referee comment: Figure 6: “Fourth row shows the exceedance probabilities”: there are only 3 rows, please correct the caption. Also, the 5% threshold is not shown, the caption must be changed to reflect it. I suggest explicitly naming the 2 columns at the top to make it clearer. On the conceptual side, I find the panels with the probability of return period exceedance not easy to read and of little value for the presented work, considering that the FLOPROS dataset is already used to include the information on the level of flood protection.

Authors’ response: We have incorporated your suggestions in the revised manuscript regarding the figures. The panels with the probability of return period exceedance are meant to represent current practice for early action protocols for river flooding and they can be compared with our impact-based thresholds. FLOPROS does not have an influence on the predicted discharge levels in our modelling set-up. We have thus left this part of the plot as is.

Referee comment: L331: F1-score, precision and recall are widely used performance scores. However, a brief description including how to interpret them should be added in the methods section.

Authors’ response: This section has been moved to the appendix in the revised manuscript. We have thus added an explanation regarding their interpretation there.

Referee comment: Figure 7: I find the overlaying of 3 different flood extents not easy to read.

Perhaps I'd split each of the two left panels in one panel with only modeled flood extent (forecast and reanalysis) and another panel for satellite flood extent.

Authors' response: We have changed the figures to have separate panels for flood forecast and reanalysis, see below.

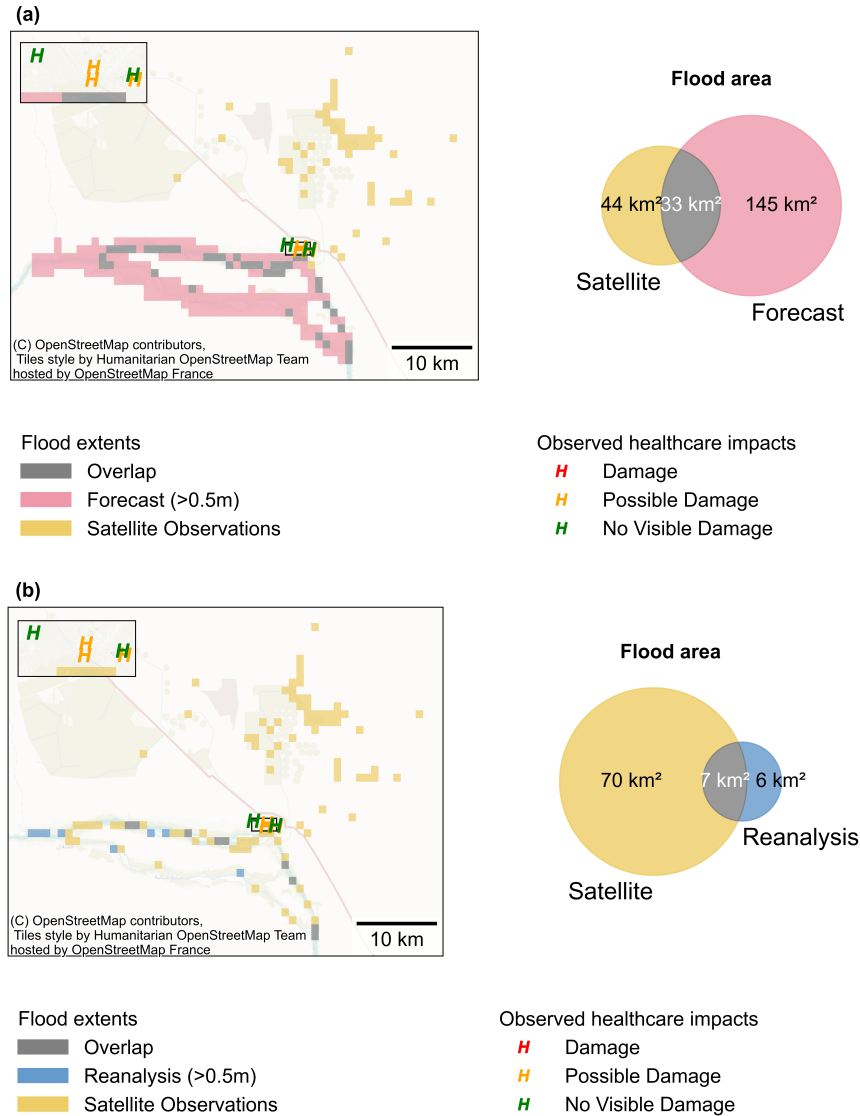


Figure A1: Comparison between modelled and satellite-detected flood extents for the city of Abu Hamad. Panel (a) compares the maximum flood extent forecasted with a lead time of 10 days between 10 July and 30 October (red shading). Panel (b) compares the maximum flood extent from the reanalysis over the same period (blue shading). Yellow shading indicates satellite-detected flood extent acquired by UNOSAT on 29 August 2024 (both panels). Colored H symbols represent possible damages to health facilities detected from the satellite.

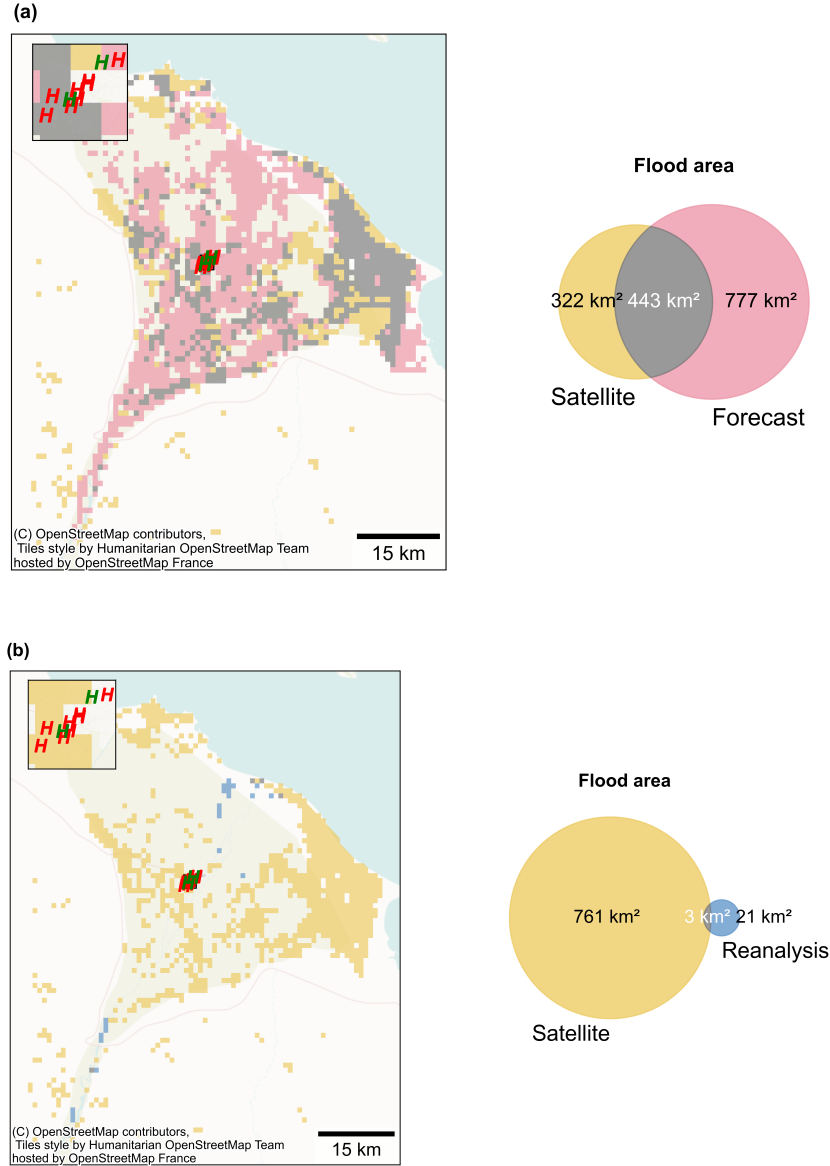


Figure A2: Comparison between modelled and satellite-detected flood extents for the city of Abu Hamad. Panel (a) compares the maximum flood extent forecasted with a lead time of 10 days between 10 July and 30 October (red shading). Panel (b) compares the maximum flood extent from the reanalysis over the same period (blue shading). Yellow shading indicates satellite-detected flood extent acquired by UNOSAT on 2 September 2024 (both panels). Colored H symbols represent possible damages to health facilities detected from the satellite.

Referee comment: L366-367: from figure 8, indirectly affected people seem similar to those directly affected. I don't see the ratio 500k to 1M mentioned in the text.

Authors’ response: We believe that there was a confusion between the “total affected” and “directly affected” curves. If this is something else that we are missing, we apologize and hope that this will be clarified through the next round of revision.

Referee comment: L398: for or four? Double check

Authors’ response: Thank you, this has been double checked.

Referee comment: L421-422: This sentence reads like a general model result. It should be clarified that this result is specific of the (2) events shown. Also, I found it a bit misleading as this reflects purely the performance of the chosen hydrological forecasting model, rather than that of the proposed framework especially for indirect impact assessment.

Authors’ response: We have clarified in the revised manuscript that this is a result obtained from an analysis done on two events. We also discuss the implications in terms of robustness in the section “Limitations, improvements and future research”:

The lack of impact data on post-disaster healthcare access and infrastructure damage restricted our analysis to a couple of case studies. We expect an analysis based on a more extensive number of case studies, spanning longer time frames and expanding over different regions to provide more robust results, yielding a clearer picture of the actual benefits of impact forecasting and the inclusion of indirect impacts. This is particularly true for our analysis of how the impact forecast can be used to trigger anticipatory action, which is based on only two locations for a single event and is therefore not robust. An extended analysis based on multiple events and locations requires structured data documenting impacts on service accessibility, at specific levels of spatial and temporal disaggregation. While an established impact database with the desired qualities is currently lacking, initiatives for systematic and structured collection of detailed impact data such as the DesInventar Sendai monitoring framework ([UNDRR](#)), and data mining projects leveraging artificial intelligence such as [Li et al. \(2026\)](#) or [Hasbini et al. \(2026\)](#) can help address these data limitations.

Referee comment: L433: Suggested “...in Sudan by 40%”

Authors’ response: This has been implemented in the revised manuscript.

References

- Federal Government of Somalia, United Nations, World Bank, and European Union: Somalia 2023 Deyr Floods Rapid-Post Disaster Needs Assessment, Tech. rep., Mogadishu, Somalia, 2024.
- Hasbini, L., Severino, L. G., Brito, M. M. d., Gesualdo, G. C., Rotaru, A. M., Bresch, D. N., Mühlhofer, E., Wang, J., and Carvalho, T. M. N.: ROUGE : A database of disaster impacts in the Global South using Red Cross reports and Large Language Models, <https://doi.org/10.21203/rs.3.rs-8778674/v2>, iSSN: 2693-5015, 2026.
- Kam, P. M., Aznar-Siguan, G., Schewe, J., Milano, L., Ginnetti, J., Willner, S., McCaughey, J. W.,

- and Bresch, D. N.: Global warming and population change both heighten future risk of human displacement due to river floods, *Environmental Research Letters*, 16, 044 026, <https://doi.org/10.1088/1748-9326/abd26c>, 2021.
- Li, N., Thiery, W., Zahra, S., Madruga de Brito, M., Worou, K., Kurfali, M., Lampe, S., Muñoz, P., Flynn, C., Trigos, C., Nivre, J., Zscheischler, J., and Messori, G.: Wikim-pacts 1.0: a new global climate impact database based on automated information extraction from Wikipedia, *Natural Hazards and Earth System Sciences*, 26, 2609–2636, <https://doi.org/10.5194/nhess-26-2609-2026>, 2026.
- Maina, J., Ouma, P. O., Macharia, P. M., Alegana, V. A., Mitto, B., Fall, I. S., Noor, A. M., Snow, R. W., and Okiro, E. A.: A spatial database of health facilities managed by the public health sector in sub Saharan Africa, *Scientific Data*, 6, 134, <https://doi.org/10.1038/s41597-019-0142-2>, 2019.
- Mühlhofer, E., Koks, E. E., Kropf, C. M., Sansavini, G., and Bresch, D. N.: A generalized natural hazard risk modelling framework for infrastructure failure cascades, *Reliability Engineering & System Safety*, 234, 109 194, <https://doi.org/10.1016/j.ress.2023.109194>, 2023.
- Riedel, L., Rösli, T., Vogt, T., and Bresch, D. N.: Fluvial flood inundation and socio-economic impact model based on open data, *Geoscientific Model Development*, 17, 5291–5308, <https://doi.org/10.5194/gmd-17-5291-2024>, 2024.
- Scussolini, P., Aerts, J. C. J. H., Jongman, B., Bouwer, L. M., Winsemius, H. C., de Moel, H., and Ward, P. J.: FLOPROS: an evolving global database of flood protection standards, *Natural Hazards and Earth System Sciences*, 16, 1049–1061, <https://doi.org/10.5194/nhess-16-1049-2016>, 2016.
- UNDRR: United Nations DesInventar Open Source Initiative - Official Website, <https://www.desinventar.net/>.
- UNOSAT: Floods in Somalia, November 2023, <https://unosat.org/products/3726>, 2023.