The manuscript describes an approach to combine data analysis and simulation as a service. To offer simulations as services is in principle a new way forward and may pave the way for many applications in natural systems analysis.

The manuscript shows a case study of a re-simulated flood event in 2016 in the Agueda river basin. In my opinion, the innovative aspect of the presented case study lays in the combination of data and simulation services, rather than in the dataset itself.

The dataset itself offers an opportunity for studying the flood impacts, e.g. by overlaying flood depths with the damaged buildings to elaborate vulnerability functions or to study the disruption of human mobility during the flood event. Thus, the dataset is worth to describe and being published.

However, the manuscript and also the data need to be improved before being useable by other scientists or practitioners. In the following, I am listing the main critical points that should be considered on the revision of the manuscript:

Authors: We thank the reviewer for the positive assessment of our article. We will place our best efforts to further improve the work accommodating the valuable suggestions.

- A more stringent division/distinction between data and modeling is needed. At the moment, the method (the model) is in the foreground. A restructuring of the manuscript in that the data itself becomes a more prominent position might be worth to consider.

Authors: We will change the manuscript' structure to separate the descriptions of the dataset and the software tools employed. Moreover, we will further elaborate on the significance of the proposed dataset. We note that the dataset we are proposing is a synthesis of udometric, hydrometric, and numerically-produced data. The latter is a closure – we know the maximum flow depth and the boundary conditions but we do not know the discharge at Águeda bridge because the hydrometric station was malfunctioning during the event. The discharge at Águeda bridge had to be estimated with the hydrodynamic model. This task that requires a time-consuming assemblage of topo-bathymetric data and calibration procedure. The relevance of this numerically-produced data may have been overemphasized due to the relatively higher difficulty in producing it. We will correct this aspect in the final version of the manuscript.

-The validation of the output data must be much more improved. At least, the simulated flow depths must be compared with observed ones at sample sites. If there is a dataset of affected buildings or blocked roads available, the output dataset can be compared with these proxy data that can be used for validation.

Authors: This is a crucial issue indeed. We have testimonials of local civil protection agents and photos that show, with no uncertainty, that the peak water level, registered on the 12th February between 16:00 and 19:00, never overtopped Águeda's protection walls. The level stayed within 10 to 20 cm below the crest of the walls at Águeda bridge. This is the only water level estimate we have since the hydrometric station was malfunctioning. Since we have the bathymetry and since we have reliable hydrometric data at upstream boundaries, we used this information to calibrate the model and compute velocities and discharges at Águeda bridge. The dataset comprises boundary information, topo-bathymetric information, calibrated roughness coefficients and computed discharges and velocities. So, to be clear, and answering the reviewer, the simulated flow levels have been compared with the observed ones at the only sample site available – downtown Águeda. But not to validate the model – such comparison is what allows

for model calibration and is thus the condition of possibility of having the dataset complemented with numerically-produced discharges at Águeda bridge. In our opinion this is a strong point of the manuscript – we provide a synthesis of hydrometric, udometric and numerically-produced data from a calibrated model that constitute a relevant and complete validation test for other flood models and a tool to better mitigate floods in this river and in similar rivers. Per se, none of the data sets (hydrometric, udometric or numerically-produced) would be sufficient. It is the synthesis that matters and explaining it is the driver of this paper.

Following the reviewer' suggestion, we will improve the text so that our arguments come across more clearly.

- It would be of interest if the data portal can also store sensor data that document the flood event (eg., event documentation data, drone images, flow depth measurements, sediment transport and deposition, changes in cros-section profiles, etc.). This would make the dataset more interesting and valuable for developing and evaluating (benchmarking) newly developed simulation models.

Authors: The RCP is currently not storing that type of data. However, we are working along these lines to include data from "social sensors" (dedicated channels to collect images from social networks and its metadata). The aim is to transform such data into information that can be assimilated into hydrodynamic tools. This, however, is not the main focus of this paper – this is meant to describe the Feb 2016 Águeda flood, synthetizing all available data, including informal data. The planned advances should be published in a different type of publication.

- The main criticism of the simulation model is the poor representation of the river channel geometry. It is not clear which cross-section data was used. The simplification of the overall assumed channel depth of 5 m is not accurate. The presentation of local case study data is valuable only if it is really accurate and at a high spatial resolution. This approach of river geometry simplification is used for global models but not for local-scale models - or for river reaches that are dominated by downstream sea or lake levels. River conveyance capacity is one of the most sensitive factors in flood risk analysis. Moreover, please add more information on the mesh size. This all would make the data more comprehensible and more useful for others.

Authors: We agree with the reviewer on the importance of simulating adequately river conveyance. But we disagree that we have a poor representation of the river channel geometry. River Águeda in front of Águeda town has suffered many interventions that resulted in a rather simple geometry – River Águeda is indeed a channel with a rectangular cross-section with protection walls. In the revised manuscript, we will provide a more detailed description of the process for defining the river bathymetry obtained by elements provided by Águeda municipality. These elements confirm the accuracy and the adequacy of the spatial resolution of the DEM in what concerns the river bottom.

Minor points:

-The abbreviation HiSTAV must be explained before being used the first time and the model must be cited.

Authors: HiSTAV stands for the High Performance Computing version of Strong Transients in Alluvial Channels. The model was originally designed to tackle river morphology problems,

especially when involving subcritical-supercritical shocks or transitions, as is the case in anthropic floods such as those resulting from a dam-break. The morphological module is not explored in this work but we see no reason to change the name of the model.

-Figure 1: Add a few explanations in the figure captions, explain abbreviations.

Authors: The caption of Figure 1 will be improved with further details.

-line 86: what is an "in-house" model

Authors: It is a model developed by the authors of the study, as opposed to a model (freeware or proprietary) developed elsewhere. In this case, the efforts of developing HiSTAV have been coordinated by co-author Rui M.L. Ferreira. Details of its HPC implementation can be consulted in Conde et al. (2020).

-line 138: Please explain why the correction factor +0.5 has been chosen

Authors: The physical parameter used to define human stability and manoeuvrability thresholds is the mass flux per unit width and per unit density in the direction of the flow. We named as hazard index the variable defined as $h\times(v+0.5)$ included in HiSTAV, where h is the flow depth, and v is the flow velocity, following research by colleagues in the UK and TU Delft (see for example Penning-Rowsell et al, 2015 or Jonkman and Penning-Rowsell, 2008).

-line 179: "bam break"?

Authors: It was a typo. We meant 'dam-break'.

-line 182: Who used the data from the media to map the limits of the flooded areas? Has this been done for validation?

Authors: The extensive media coverage of the 2016 flood event in Águeda allowed a significant collection of images with vertical references. Data from the media was used primarily to ascertain the water level at Águeda bridge, once the hydrometric station was malfunctioning. It was also used to verify the extent of inundation along Rua Luís de Camões and Rua Vasco da Gama, as there are photos taken around 16:00 on the 12th February. The simulated flow levels have been compared with the observed ones at the sample site available but have not been used to validate the model. That comparison allows for model calibration and is thus the condition of the possibility of having the dataset complemented with numerically-produced discharges at Águeda bridge.

-figure 3: add scale bar and a better legend. Please add also the site of the gauging station.

Authors: The legend of Figure 3 will be improved. The location of the gauging station will be added as a georeferenced layer into the input dataset.

-figure 4: add scale bar and a proper legend

Authors: Figure 4 will be improved.

-How did you asess the role of insflow 1 and 2 from the measured flow in the gauging station?

Authors: The discharge time-series at the two gauges were input boundary conditions of the model.

-Figure 5: Please add the toponyms "Ponte Redonda" and "Ribeiro" in figure 3

Authors: Figure 3 will be changed to incorporate this suggestion.

- Please describe the REST API in more detail. This would be an innovative feature for data sharing and especially for interoperability.

Authors: The article's main goal is to present a well-documented, unique and relevant dataset, that is publicly available through HydroShare, a worldwide repository. The REST API is a feature of the repository, therefore, its detailed explanation will be out of the scope of the article. Nevertheless, we may include some extra information.

References:

Conde, D., Canelas, R.B., and Ferreira, R.M.L. (2020). A unified object-oriented framework for CPU+GPU explicit hyperbolic solvers. *Advances in Engineering Software* 148, 102802. DOI: j.advengsoft.2020.102802

Penning-Rowsell, E.C., P. Floyd, D. Ramsbottom, and S. Surendran (2005). Estimating Injury and Loss of Life in Floods: A Deterministic Framework. *Natural Hazards* 36:43-64. DOI: 10.1007/s11069-004-4538-7

Jonkman, S. N. and Penning-Rowsell, E. (2008). Human Instability in Flood Flows. *JAWRA Journal of the American Water Resources Association* 44(5), 1208–1218. DOI:10.1111/j.1752-1688.2008.00217.x