At first we would like to warmly thank both referees for their careful reading of the manuscript, their positive appraisal and their very useful comments and suggestions to improve the overall quality of the manuscript. We provide hereafter a detailed answer explaining how we managed to address each of the raised issues.

Francesco Dottori (Referee)
francesco.dottori@ec.europa.eu
Received and published: 22 December 2020
The manuscript describes the application of three different modelling approaches to map flash flood hazard in three case studies in South France. The topic is undoubtedly worth of interest, considering the potential for near-real-time applications and the possibility to include flash flood impacts in future applications. The manuscript is well structured and reasonably well written. The authors perform a detailed analysis of the model results, including the main types of error found, and this gives the reader a comprehensive overview of the strengths and limitations of each method. In my opinion, the paper will be ready for publications after a moderate revision to correct a few issues.

Main points
L203: "The simulations are all run in steady state regime based on estimated flood peak discharges for each river reach. This leads to neglect the possible dynamic effects related to the inundation of floodplains occurring in unsteady flow regimes. This assumption is considered here as reasonable since the width of the floodplains do not exceed several hundred meters, and the volumes of the floodplains remain limited if compared to the volumes of the floods." I think that the limitations due to this modelling choice need be better explained. Based on the results, this seems indeed a reasonable assumption given that there is no general overestimation for the 1D and 2D models. Still, simulating a steady flow regime using peak flow implies an overestimation of total flood volumes, compared to a real flood wave with increasing and receding limbs. As such, this point should be mentioned in the discussion because it might originate errors in case of flood events where flood volumes are small compared to the floodplain extent. Moreover, steady flow simulations have limitations when modelling the interaction of flood waves at confluences. The underlying assumption is that flood peaks are occurring at the same time (a sort of worst-case scenario), while in reality peaks might occur at different times. This point should also be mentioned.

This initial formulation indeed does not provide much details on the limitations related to the steady state assumption. We propose to replace this with a more developed description: “The simulations are all run in steady state regime based on estimated flood peak discharges for each river reach. The steady state assumption may lead to an overestimation of the inundation extent and depths if the volume of the flood wave is limited in comparison with the storage capacity of the floodplain. This assumption is considered here as reasonable since the widths of the floodplains do not exceed several hundred meters, and therefore the corresponding floodplain storage capacities should remain limited. The computation based on flood peak discharges may also lead to an overestimation of backwater effects at confluences, because of the underlying assumption that maximum peak discharges occur simultaneously for all river branches at a confluence. Lastly, the variations of peak discharges along each river reach are not represented, but these variations are limited since the delineated river reaches have a limited length”

Section 5.3. The presentation of run times would be even more informative if the authors could make a more quantitative comparison with run-time required to actually set up a real-time flood simulation. Often, reliable weather forecasts of flash flood events are available only few hours in adavance, meaning that a real-time simulation should be available to emergency responders in ,say, 2-3 hours to be effective and helpful. Considering the usual speed-up attainable for 2D hydraulic models (see for instance Neal et al., 2018, https://doi.org/10.1016/j.envsoft.2018.05.011) this seems to be feasible objective, provided that the Floodos model can be parallelized in a similar way. Could you please elaborate a bit on this?

We propose here to add a sentence to illustrate to which extent the computation times may be reduced by using parallel computing, and provide a comparison with the current update frequency of short range rainfall nowcasts : “As expected, the SWE 2D approach is computationally the most expensive. But the computation times remain reasonable for the 5 m resolution used here, and first parallel computations achieved using a 32 cores and 128 GB RAM cluster suggest that they may still easily be reduced by a factor 4 with the Floodos model. However, the resulting computation times remain large for real time applications, considering the current refreshment frequency of 1h for short-range rainfall nowcasting products. “

Conclusions: I suggest to elaborate a bit more the discussion on real-time applications, given its importance. In my opinion, real-time applications are meaningful only with the 1D or 2D hydraulic models, which are both able to simulate flood waves in unsteady flow conditions, including the interaction of flood waves with different timings at confluences. On the contrary, if the steady-state approach is deemed appropriate, then an off-line catalogue (similarly to what done by Dottori et al., 2017, https://doi.org/10.5194/nhess-17-1111-2017) would probably be enough. I would be interested in reading the opinion of the authors on this point. In addition, I suggest to mention the possibility of using the described methods to evaluate flash flood impacts (see the recent works by Merz et al., 2020, https://doi.org/10.1029/2020RG000704; and Ritter et al., 2020, https://doi.org/10.1016/j.envint.2019.105375)

We agree that questions related to real time applications and representation of FF impacts are of great importance. We therefore developed the last paragraph of the conclusion in the following way: " Finally, the methods presented here should be of great help to provide realistic inundation scenarios and develop information about possible flash-flood impacts as a support of flood risk management policies (Merz et al., 2020; Ritter et al., 2020). However, further work is still needed to integrate these methods into real-time forecasting chains and assess their performance in this context. The errors on discharge forecasts may indeed be dominating the other sources of uncertainties, and the computation times may also be another important limiting factor. Depending on the considered inundation mapping methods, real time computations may be feasible and may improve the representation of flood-wave volumes and flood dynamics at confluences, whereas off-line libraries of inundation scenarios can be generated and sampled in real time (Dottori et al., 2017), which may help representing discharge uncertainties by selecting multiple scenarios (Leedal et al., 2010). The definition of the best real-time computation strategy is even more complex in the case of flash-floods, because of their very fast evolution dynamics. The delay necessary to run and provide forecasts may indeed highly limit the capacity of emergency services to analyse forecasts and adapt their response strategies by reference to inundation scenarios they are prepared for. Finally, an optimal compromise has probably to be found in the case of flash floods between the accuracy of inundation forecasts and the rapidity of forecast delivery."

Minor issues

The Title is maybe a bit redundant, consider shortening , e.g.: "Performance of automated methods for mapping flash flood hazard: a comparison of hydrodynamic and geomorphologic methods" or something similar

The title has been modified as follows: "Performance of automated methods for flash flood inundation mapping: a comparison of a DTM filling and two hydrodynamic methods"

Abstract L13-14: "With these methods, the inundated areas are overall well retrieved..." Here I would suggest replacing the qualitative evaluation with some quantitative metrics, as done for the water levels

We replaced the sentence with "With these methods, a good retrieval of the inundated areas is illustrated by Critical Success Index median values close to 80%, and …"

L 19 "Flash floods represent a significant part of flood related damages worldwide". Do you have a quantitative assessment of the share of flash flood damages, for instance in France? You might for instance look at the HANZE dataset by Paprotny et al (2018,https://doi.org/10.5194/essd-10-565-2018)

A reference to a report edited by CCR (Caisse Centrale de Réassurance) on natural disasters in France since 1982 has been added: "For instance, in France eight floods caused insurance losses exceeding 500 million euros over the period 1989-2018, among which 4 were flash floods (CCR, 2019)".

L 37: "For instance, in France it is estimated that a river network of about 100.000 km should be documented for a comprehensive coverage of the small streams". Is there a reference for this statement?

We added an explanation for this statement: "For instance, in France the entire stream network includes 120.000 km of rivers of more than 1m width, whereas flood hazard information is concentrated on the 23.000 km of main rivers, corresponding to the network covered by the Vigicrues national flood forecasting service. It can thus be estimated that about 100.000 km of small rivers should be documented with hazard information to ensure a comprehensive coverage."

L62-69: This paragraph doesn't read well due to many references and lists of models. Please try to rearrange the information (e.g. I would put first the sentence "All these methods determine a local discharge/height relationship from..." and then "These methods are applied either directly from the DTM for the AutoRoute method...")

We reformulated this paragraph in the suggested manner: "Direct DTM filling approaches have been developed more recently. All these methods are based on a local discharge/water height relationship determined from i) the cross-section and longitudinal profile geometries, and ii) a local hydraulic formula: Manning-Strickler (ZhengXing et al., 2018; Zheng at al., 2018; Johnson et al. 2019, GarousiNejad et al., 2019} or Debord (Rebolho et al., 2018). The cross-section geometry is either extracted locally from the DTM for the AutoRoute method (Follum et al., 2017, 2020), or averaged at the river reach scale based on a Height Above Nearest Drainage raster (Nobre et al., 2011) for the following methods : f2HAND (Speckhann et al., 2017); Geoflood (Zheng et al., 2018); MHYST (Rebolho et al., 2018); Hydrogeomorphic FHM (TavaresdaCosta et al., 2019). "

L83: "A significant evaluation and validation effort is proposed..." Maybe better rewrite as "A comprehensive evaluation and validation exercise is proposed..."
The sentence has been reformulated as suggested.

L88-90 Please replace "first section", "second section" etc with "Section 2", Section 3"etc
The replacements have been made.

Title of Section 2: I'd rather use "description" than "presentation".
The replacement has been made.

L103-104: "A conventionnal Dinf approach is used here instead of the Geonet approach used in GeoFlood." Could you please either specify the difference or provide references for the two approaches? Also, correct the typo (conventional)
A reference has been added for the Dinf approach: Tarboton, D. G., (1997), A New Method for the Determination of Flow Directions and Contributing Areas in Grid Digital Elevation Models, Water Resources Research, 33(2): 309-319. The Geotnet approach is described in Zheng et al (2018).

Section 2.2: I suggest renaming the approach as CaRtino-1D HECRAS ,given that HECRAS is the actual hydraulic model applied.
The approach has been renamed as suggested

L116: please provide a reference for the Mascaret model here (or remove the mention if not relevant for the study)
A reference has been added: MASCARET : a 1-D Open-Source Software for Flow Hydrodynamic and Water Quality in Open Channel Networks, N. Goutal, J.-M. Lacombe, F. Zaoui and K. El-Kadi-Abderrezzak, River Flow 2012 – Murillo (Ed.), pp. 1169-1174, doi:10.1201/b13250

L125-126: "Its main limits, already identified in previous works, lie in the 1D scheme which may not be adapted in areas with complex hydraulic features". Please name some of these works here.
A reference to Le Bihan et al. (2017) has been added here: https://doi.org/10.5194/hess-21-5911-2017

L 140: "The model has been compared with the widely used 2D LISFLOOD-FP model (Bates et al., 2010), showing equivalent results and faster computation times." Were these tests performed by Davy et al. as well? Please specify also the reduction in computational time as compared with LISFLOOD-FP.
Yes the comparison is included in Davy et al. (2017), but according to the authors this comparison should not be considered as a benchmark since their use of the LISFLOOD-FP may still be optimized. Thus it is thus difficult to conclude on the reduction level of computation times. This has been more explicitly mentioned: "Davy et al. (2017) indicate the CPU time changes approximately linearly with the number of pixels of the computation domain. They compared Floodos with the widely used 2D LISFLOOD-FP model (Bates et al., 2010). They obtained similar results and faster computation times with Floodos, although they mention this comparison should not be considered as a benchmark."

Section 4 L243-248: This paragraph and Figure 4 might be better placed in a separate subsection after subsection 4.1
Figure 4 provides a generic illustration of both evaluations of simulated flood areas (subsection 4.1) and simulated water levels (subsection 4.2). Therefore we prefer to present it at the beginning of section 4.

Figure 4: it is not clear where these two areas are located within the study area, Please add a smaller map of the study area showing the location of the two boxes
We updated the figure to explicitly indicate the location of the two presented areas.

Figure 8: Is it simulated water level in panels b-d?
Yes indeed, this has been mentioned in the figure captions.

Figures 8 and 9: I assume that you are using Floodos simulations here right? Please specify this in the text and captions
Yes indeed, this has been specified in the captions and the text.

Section 5.2.1: Accounting for protection structures is indeed a major challenge in any large-scale flood risk assessment. Could you tell how much of the study areas is protected by dykes or other defence structures?
A sentence has been added to provide this information: "This is a specificity of the Aude case study, where numerous flood defense structures have been built, especially along the Frequel river and in the downstream floodplains of the Aude river".

L358 typo: feasibility of reasonably accurate

This has been corrected.

L375 "The sensitivity to roughness values has also to be further investigated for an appropriate representation of uncertainties". Using variable roughness values according to land cover could be an option for future studies. This is actually a standard practice for large-scale flood models (see Sampson et al., 2015, https://doi.org/10.1002/2015WR016954; and Dottori et al., 2016, https://doi.org/10.1016/j.advwatres.2016.05.002)
This possibility has been explicitly mentioned "The sensitivity to roughness values has also to be further investigated for an appropriate representation of uncertainties, and variable roughness values may also be defined depending on land cover (Sampson et al., 2015 ; Dottori et al., 2016)."