## Summary

In this study, the authors compare three modeling frameworks for mapping inundation extent and flood depth in Clear Creek watershed, a tributary of Galveston Bay in Texas, U.S. They evaluate the performance of AutoRoute, HEC-RAS, and Fathom-US frameworks against in-situ USGS high-water marks of Hurricane Harvey, a well-studied compound flood event in 2017. Also, the authors estimate flood exposure, consequences, and damages to buildings using available data from FEMA. It is shown that both HEC-RAS and Fathom-US outperform AutoRoute due to inherent limitations of the latter framework to simulate flooding in low-lying areas. Although Fathom-US and HEC-RAS achieve high location accuracies and low error and bias, they present some discrepancies regarding the evaluation metrics. The authors suggest an ensemble, multi-model probabilistic methodology to leverage these frameworks and provide more accurate flood maps as discussed in similar studies.

## Major comments

This study presents an inter-model comparison with a practical application in terms of flood exposure and damage assessments, but it does not provide essential information for doing so. In contrast to Fathom-US (Wing et al., 2017, 2019), there is no evidence of model calibration and validation of both AutoRoute and HEC-RAS models for the study area. If the goal is to evaluate model's performance, then input data, forcing, mesh extent, and grid resolution should be identical among the frameworks. This compromises not only the validity of the results, but the analyses presented throughout the manuscript. I suggest the authors to consult or follow other studies that provide guidelines for model comparison (Shustikova et al., 2019; Muñoz et al., 2021; Afshari et al., 2018; Liu et al., 2018).

The authors investigate the performance of the frameworks knowing beforehand that AutoRoute is not suitable in coastal areas (Line 63 in the Introduction). This rises concern about the usefulness of a low-skill model in this study. If the authors want to consider *steady-state* models like AutoRoute (or HAND) in the model comparison, I suggest to follow the approach of Jafarzadegan et al., (2022) to enhance model simulations via hydrogeomorphic classifiers.

## Minor comments

L16: 'Event maps' is too generic for referring to flood inundation maps. 'Event maps' are also used to describe the modeling framework making the manuscript difficult to follow in some sections.

L20: Are you talking about modeling frameworks or flood inundation maps? How can event maps be physically different?

L26: Do you mean flood emergency response instead of flood fights?

L28: We find that the **modeling frameworks** are much different physically…

L43: Typo. Event Maps help emergency managers…

L63: HAND can be adapted to simulate coastal flooding in low-lying areas. See Jafarzadegan et al., (2022).

L88: I can anticipate that you will find substantial differences based on the DEM resolution and forcing data you have chosen for each framework.

L98: Details of the hydrologic and hydraulic modeling are missing. For example, what is the grid size for the 2D component?

L120: Diffusive wave is a simplified version of the shallow water equations. Given the nonlinearities and complexities arising in compound coastal flooding, the complete set of equations (SWL) available in HEC-RAS should be used. This might lead to a better accuracy of the HEC-RAS in terms of inundation extent and flood depth.

L123: What are those mysterious downstream boundary conditions? Figure 1 should include the location of those boundaries for the three modeling frameworks.

L129: Are roughness values calibrated afterwards? These initial 1D and 2D roughness values are event-specific and have to be tuned for future flood events.

L134: Would it not be better to consider the 1-m DEM and so avoid inaccuracies due to DEM resolution? Previously, you suggest using observed meteorological data to avoid limitations in forecast skill...

L150: Evaluation of simulated time series is very informative but missing in this study (e.g., timing and magnitude of peak water level). I strongly suggest assessing model's performance based on time series of available USGS (#08077637) and NOAA stations.

L183: What are the upper and lower bounds?

L185. Typo in the diagram. "Create Kernel density maps".

L197: Diffusion wave does not solve the full mass balance and momentum equations and therefore might have influenced flood inundation extent and depth. In addition, the 1D portion of the model cannot provide 2D flood maps and consequently miss nearby high water marks.

L200. I cannot find the calibration and validation process in this manuscript. The same holds for AutoRoute model.

L218: USGS high-water marks are referenced with respect to NAVD88. There may be uncertainties added in the NAVD88 to MSL conversion process. How did the authors conduct the datum conversion? What is the vertical datum of the DEMs?

L226: … and the steady state assumption. Also, AutoRoute is only forced by streamflow ignoring the contribution of coastal water level (e.g., storm-tide) to compound flooding (Figure 2).

L228: Figure 4. Text size is too small.

L412: I agree that ensemble modeling is the way to go for better compound flood assessments. Nevertheless, I consider Figure 8 unnecessary in this study as you are not actually following this approach for simulating compound flooding due to Hurricane Harvey. A descriptive text is enough for future work in this regard.

**References**

Afshari, S., Tavakoly, A. A., Rajib, M. A., Zheng, X., Follum, M. L., Omranian, E., and Fekete, B. M.: Comparison of new generation low-complexity flood inundation mapping tools with a hydrodynamic model, Journal of Hydrology, 556, 539–556, https://doi.org/10.1016/j.jhydrol.2017.11.036, 2018.

Jafarzadegan, K., Muñoz, D. F., Moftakhari, H., Gutenson, J. L., Savant, G., and Moradkhani, H.: Real-time coastal flood hazard assessment using DEM-based hydrogeomorphic classifiers, 22, 1419–1435, https://doi.org/10.5194/nhess-22-1419-2022, 2022.

Liu, Z., Merwade, V., and Jafarzadegan, K.: Investigating the role of model structure and surface roughness in generating flood inundation extents using one- and two-dimensional hydraulic models, 0, e12347, https://doi.org/10.1111/jfr3.12347, 2018.

Muñoz, D. F., Yin, D., Bakhtyar, R., Moftakhari, H., Xue, Z., Mandli, K., and Ferreira, C.: Inter-Model Comparison of Delft3D-FM and 2D HEC-RAS for Total Water Level Prediction in Coastal to Inland Transition Zones, n/a, https://doi.org/10.1111/1752-1688.12952, 2021.

Shustikova, I., Domeneghetti, A., Neal, J. C., Bates, P., and Castellarin, A.: Comparing 2D capabilities of HEC-RAS and LISFLOOD-FP on complex topography, 64, 1769–1782, https://doi.org/10.1080/02626667.2019.1671982, 2019.

Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., and Erickson, T. A.: Validation of a 30 m resolution flood hazard model of the conterminous United States, 53, 7968–7986, https://doi.org/10.1002/2017WR020917, 2017.

Wing, O. E. J., Sampson, C. C., Bates, P. D., Quinn, N., Smith, A. M., and Neal, J. C.: A flood inundation forecast of Hurricane Harvey using a continental-scale 2D hydrodynamic model, Journal of Hydrology X, 4, 100039, https://doi.org/10.1016/j.hydroa.2019.100039, 2019.