Reviewer 1

The paper is interesting and presents a useful methodology for coastal flood modeling. The validation against a hydrodynamic model is OK but a bit questionable; this should be done against historical flood maps. A few assumptions should be also clarified. Limitations, uncertainties and implications need to be further discussed. I have recommended several edits and some comments in the PDF. Here are some additional comments:

Response: Thank you for the detailed review and the constructive comments.

Please provide more information on the study catchment, particularly those that affect your model results. This includes computational area, soil type, channels' size, ground slope, land use etc.

Response: We have included additional information in the revised manuscript regarding the study catchment (Section 2; Lines 148-159) and model domain (Section 3.1.1; Line 237). In short, the average slope, length, and annual discharge of the Savannah River are 0.00011 m/m, 505 km, and 320 m^3/s, respectively. Also, the river bathymetry was deepened up to 12 m for increasing the capacity of cargo transportation according to the U.S. Army Corps of Engineers. The model domain comprises an area of 1178 km^2 approximately.

Section 2: Present the source of drainage network data. Also, how detailed does that represent the drainage network?

Response: Drainage network data including river streams, tidal channels, and creeks within wetland areas can be obtained from the U.S. National Wetlands Inventory (https://www.fws.gov/wetlands/data/Mapper.html). These publicly available data are continoulsy updated by the U.S. Fish and Wildlife Service (FWS) and are derived from multiple data sources including satellite imagery and aerial photos of 1 m (or less) digital color infrared imagery. We included this information in Section 2 (Lines 157-159).

The verification of the approach against flood maps generated by a hydrodynamic model is questionable. How well is the model calibrated? For what historical events (how large/intense), it has been calibrated? Also, why not using satellite imagery like Dartmouth Flood Observatory?

Response: The primary goal of this manuscript is to propose a low complexity flood mapping (LCFM) method whose accuracy is comparable with a computationally expensive hydrodynamic model. Therefore, we compare our results with a hydrodynamic model to assess if our method can be a proper replacement of these models. This is the idea of proposing surrogate models that mimic the performance of complex physically-based models. A similar approach has been presented recently where surrogate machine learning methods are trained and validated against a well-calibrated hydrodynamic model. The hydrodynamic model has been calibrated for both non-extreme events and two major Hurricanes in the region, namely Hurricanes Matthew and Irma. Please see details of the calibration in Figure 4 and lines 382-398 in the revised manuscript.

Using satellite imagery has major limitations. 1) These maps are rarely available for the peak date of a flood event while we are looking for the maximum flood hazard maps. 2) These maps only provide the extent of flooding (HDC=0) while we need floodwater depths to generate flood hazard maps for different levels of HDCs. For example, here we validate for both HDC=0 and HDC=0.6 m. 3. Daily satellite data, such as Dartmouth Flood Observatory uses coarse-scale satellite imagery, such as MODIS with 250-500 m spatial resolution that is not appropriate for validation. We need a much finer scale (<30 m) for validating our maps.

Please discuss the properties of the high-performance computing system that was used for simulations (Section 3.1.2).

Response: We used available computational resources of the University of Alabama (UA) for running model simulations in parallel. The UAHPC is a 87 node (1660 core) cluster featuring Dell PowerEdge M610s, M620s, and M630s. The nodes contain two Intel 8-Core E5-2650, E5-2640v2, or 10-core E5-2640v3 processors and at least 64GB of RAM per node. More information of UAHPC can be accessed in the following link: https://oit.ua.edu/services/research/. Nevertheles, we consider this information only relevant for the reviewer.

More details on the LHS application are needed. How was it informed by Hurricane Matthew peak WLs? What parameters were considered as uncertain? What probability distributions were used and how were they characterized?

Response: The Latin Hypercube Sampling (LHS) technique was used to sample 200 sets of roughness (n) values for model calibration. We considered a 7-day window around peak water levels (e.g., peak surge of Hurricane Matthew) to evaluate the model's performance. In that way, we identify the optimal combination of n-values (among the 200 model simulations) that accurately represent both non-extreme (low water) and extreme WLs. For simplicity, we only considered n-values as uncertain parameters and assumed that any errors follow a Gaussian distribution as discussed in Helton and Davis (2003). The advantage of LHS over traditional Monte Carlo approaches is that the former results in a denser stratification over the range of each sampled parameter as compared to random sampling. Hence, LHS leads to more stable results that are closer to the true probability density function (PDF) of the parameter. We included this information in Section 3.1.2 (Lines 251-254).

Please discuss how the performance of model was graded based on the fit metrics (RMSE, AUC and R^2). You may refer to Moriasi et al. (2015) and Ahmadisharaf et al. (2019) for streamflow predictions via R^2 or others for flood simulations. Neither RMSE nor R^2 measure bias. Metrics like PBIAS need to be used along to measure the model performance.

Response: We have included additional metrics to evaluate model's performance more rigorously: Kling-Gupta Efficiency (KGE), and Nash-Sutcliffe Efficiency (NSE). In addition, we have replaced R^2 by mean absolute bias (MAB). NSE measures the relative magnitude of the error variance of model simulations compared to the variance of observational data (Nash and Sutcliffe, 1970). NSE ranges between $-\infty$ to 1, where an efficiency of 1 indicates a perfect match between simulated and observed WLs. Kling-Gupta efficiency (KGE) is a robust evaluation metric that accounts for correlation, bias, and ratio of variances (Gupta et al., 2009). KGE can take values between $-\infty$ and 1, where an efficiency of 1 indicates a perfect match. Mean absolute bias (MAB) quantifies the bias of model simulations with respect to observational data. MAB of 0 suggests an absence of bias in the simulations. This information and further discussion of model results are included in Section 3.1.2 (Lines 260-268) and Section 4 (Lines 358-360), respectively.

Please define what 'error' exactly is in the model evaluations under the Results section.

Response: Error is the summation of rate of false positives and rate of false negatives in binary classification problems. We have defined this metric in Equation 5 and lines 349-353 in the revised manuscript as follows:

"In binary classification, positive and negative refer to a value of one and zero, respectively. True positive instances are those positive cells that are correctly predicted by the classifier and false positive instances represent those negative cells that are wrongly classified as positive. The *error,* reflecting all cells that are wrongly predicted by the classifier, is a commonly-used measure for validating the performance of binary classifiers for flood hazard mapping. "

Please discuss what probability distributions exist in the MATLAB allfitdist tool.

Response: 'allfitdist' tool includes the following parametric probability distributions:

<u>Continuous:</u> Beta, Birnbaum-Saunders, Exponential, Extreme value, Gamma, Generalized extreme value, Generalized Pareto, Inverse Gaussian, Logistic, Log-logistic, Lognormal, Nakagami, Normal, Rayleigh, Rician, t location-scale, and Weibull.

<u>Discrete:</u> Binomial, Negative binomial, and Poisson. We believe this information is not relevant for the main goal of this study.

The underlying assumption of a univariate flood frequency analysis is that a peak WL with a given return period leads to a flood event with the same return period. However, studies (e.g., Brunner et al. 2016) have shown that a combination of peak flow and other attributes like volume may lead to a different return period. This limitation should be at least acknowledged in the paper.

Response: Thank you for the great suggestion. We discussed this limitation and added an appropriate reference. Please refer to lines 560-564 in the revised manuscript.

Further details are needed on how TH and HDC are derived. As of now, it appears that they are subjectively derived.

Response: TH is the threshold of the hydrogeomorphic classifier that should be calibrated by optimizing the error measure calculated from the comparison of reference and simulated maps. Therefore, this variable is derived from optimization results and is not derived from subjective decisions. HDC, however, is the hazard depth cutoff that converts the continuous flood depth map to a binary flood hazard map. This is a control variable that the decision maker (emergency responder) should pick from. We use 21 HDC resulting from 0.1 increments in the range of 0-2 m and show the results (TH) for all these HDCs. This provides 21 points for generating a smooth curve (Figure 7) so that the decision-maker can simply use this curve and pick the required TH according to different values of HDCs. Please refer to lines 356-378 in the revised manuscript.

MAB has been reported in the Results section but not in the Methods section. Please either remove it from the Results section or discuss it in the Methods section.

Response: We have included a description of MAB in the revised manuscript (Section 3.1.2, Lines 260-268).

There should be a plot on calibrating w1 and w2 coefficients (for the H and D variables).

Response: Yes, we already included this plot in the manuscript. Please see Figure 6b where we show how calibrated w1 and w2 values change for different HDCs. The higher weight of w1 compared to w2 shows that feature *H* is more important than feature *D*.

L429-432: Reasons for this poor performance need to be discussed in the Discussion Section.

Response: We added more text that explains the potential reasons for the discrepancies between the hydrogeomorphic method and hydrodynamic model results. Please refer to lines 481-488 in the revised manuscript as follows:

"The main discrepancies are some noisy scattered low-hazard areas located in the east and southeast of the study area. These areas can reflect the flooded surface depressions (sinks) resulting from the pluvial impacts of extreme precipitation. Hydrodynamic models simulate the fluvial and coastal processes that occur adjacent to rivers and oceans while disregarding the pluvial impacts. The red circle in the left part of the figures shows a region that the hydrogeomorphic method cannot properly simulate, especially for higher HDCs. This can be due to the inability of the hydrogeomorphic method to properly simulate physical processes."

The comparison of computational time against the hydrodynamic model is unclear to me. Did you compare your static model against an unsteady Delft3D-FM or the steady-state? The runtime of an unsteady hydrodynamic model should not be very long; therefore, this advantage of your presented model is not as strong as it is presented.

Response: The Delft3d-FM simulates the flood in an unsteady condition. Due to the high nonlinearity and complexity of extreme floods, flood modeling in a steady state is highly erroneous. The runtime of a hydrodynamic model depends on the scale of the study area, and the number of grid cells. For a fine scale simulation (<10 m) performed for medium-large scale problems (> 1000 km^2), the computational time of hydrodynamic models can take a few days. The main goal of using LCFM methods is to reduce the computational time while providing acceptable accuracy (improve the efficiency of modeling). For emergency responders, timing is the most important factor, thus having access to more efficient models that estimate the hazardous areas in order of minutes is significantly beneficial.

Broader impacts need to be discussed. The authors should discuss what implications these results have for coastal planners and floodplain managers etc. and what existing programs in the US (e.g., FEMA FIRMs) may benefit from this research.

Response: The Discussion section already touches on this topic a bit. We have expanded this discussion on the implications for coastal planners, floodplain managers, and existing U.S. programs (e.g., the NWS) in the Discussion section. (Lines 580-609)

"Operationally, the Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model (Jelesnianski et al., 1984) is the storm surge model currently used by NWS to perform storm surge forecasting and create probabilistic flood inundation maps for real-time tropical storms (Sea, Lake, and Overland Surges from Hurricanes (SLOSH), 2022). The feature of SLOSH that makes it the preferred model of the NWS for storm surge forecasting and mapping is the model's computational efficiency that allows the model to be run as an ensemble (Forbes et al., 2014). However, SLOSH is just one of several modeling options for storm surge modeling and mapping, each possessing strengths and weaknesses associated with their simulations. The inclusion of additional models that can create flood maps of storm surge for a given event should provide an enhanced understanding of the uncertainty of inundation at a given location (Teng et al., 2015). However, the higher computational burden of alternative models, such as Delft3D-FM, tend to preclude their use in real-time operations and certainly, their use in generating an ensemble necessary for probabilistic flood maps. The methodology we propose in this manuscript may offer the NWS and other agencies a means to utilize alternatives to SLOSH for flood inundation mapping and probabilistic flood inundation mapping on U.S. coastlines. Models such as Delft3D-FM can generate reference maps to train the binary classifer and build the probabilistic operating curves. The probabilistic operative curves would account for the major source of uncertainties and provide a computationally efficient and reliable decision-making tool for coastal planners and floodplain managers. The operative hydrogeomorphic threshold classifiers proposed for real-time coastal flood hazard mapping can be used as an alternative tool for the rapid estimation of hazardous areas during real-time flood events. In an operational mode, water level or meteorological forecasts can be used to estimate the return period of an upcoming coastal flood event and the methodology here can utilize this as an input to perform LCFM flood inundation mapping both deterministically and probabilistically."

Study limitations and potential areas for future research need to be expanded.

Response: We have already included three areas of research for future studies. To expand this, we added more text explaining the study limitations and potential areas for future research. Please refer to lines 520-528 in the revised manuscript.

"The proposed hydrogeomorphic index (I_{HD}) is the primary data for flood hazard mapping in this study. Thus, the quality of two main inputs of this index, namely the DEM and stream network used to calculate features H and D play a vital role in the overall accuracy of the proposed approach. To obtain maximum accuracy, here we used the best available DEM with the finest spatial resolution of 3 m that includes the bathymetry data. However, considering the limited access to such high-quality DEMs in many areas of the world, it is recommended to evaluate the sensitivity of the proposed approach to lower quality DEMs (e.g. 30 m and 90 m DEMs without bathymetry information) in future studies. Another piece of research can investigate the sensitivity of the proposed approach to the density of the drainage network used for calculating the I_{HD} index."

In general here are the areas of research we recommended for future studies:

1. Sensitivity of the hydrogeomorphic index to DEM quality and stream network density (Lines 520-528)
2. Applying the proposed hydrogeomorphic operative curves to inland floods and to other deltas across the US. (Lines 544-550)
3. Improve the flood frequency analysis, considering its uncertainties, incorporating other sources of uncertainties in the modeling to generate probabilistic operative curves (Lines 556-577)
4. A benchmark study that compares the performance of three LCFM methods (Lines 617-620)

Sources of uncertainty and how they may affect your findings need to be discussed.

Response: This has been thoroughly addressed in the discussion section. Please refer to lines 556-579 in the revised manuscript.

"The reference maps used for training the binary classifier are key components for generating reliable results. Since these reference maps are the outcomes of hydrodynamic modeling, they are prone to uncertainties stemming from unrealistic parametrization, imperfect model structure, and erroneous forcing. The design floods used as boundary conditions of the hydrodynamic model are estimated from flood frequency analysis that is prone to uncertainty as well. Here we used a bivariate approach that estimates the design flood based on the water level data. A more comprehensive flood frequency analysis that accounts for other flood attributes, such as volume can improve the reliability of flood frequency analysis in future studies (Brunner et al., 2016). With access to less than 100 years of data for flood frequency analysis, the extreme return levels (i.e. 500- and 1000-year floods) pose high uncertainties due to the extrapolation of annual maxima data. This should warn decision-makers to be more cautious about using operative curves for extreme flood events. For future studies, the uncertainty bounds of flood frequency analysis (especially extrapolations for extreme cases) can be considered in the modeling. In a real-time scenario, the forecasted WL used for flood frequency analysis is also prone to uncertainties originating from imperfect forecasting methods and nonstationary climate data. In addition, the uncertainty of model parametrization can be accounted for by running the hydrodynamic model for different combinations of optimum parameters. Model structure uncertainty can be also considered by using different hydrodynamic models and combining the results. Finally, probabilistic reference maps together with uncertainties involved in WL forecasting and flood frequency analysis can be integrated to develop probabilistic hydrogeomorphic threshold operative curves in future studies. This is in line with the report provided for the NOAA National Weather Service (NWS), showing the NWS stakeholder's preference for utilizing probabilistic storm surge inundation maps (Eastern Research Group, Inc, 2013)."

Please discuss how your presented modeling framework can be used in other study areas. What considerations should be taken to do so? Guidelines should be provided in the Discussion section.

Response: We added the following texts to address this concern of the reviewer. Please refer to lines 550-555 in the revised manuscript.

“To implement this approach, first, a hydrodynamic model should be set up for the new study area and generate reference inundation maps for different return periods. Access to observed water level data (gauges or HWMs) and flood extent maps from past floods is required to properly calibrate the hydrodynamic model. Then the I_{HD} index calculated from a DEM is utilized together with the reference maps to provide the hydrogeomorphic threshold operative curves for future floods.”

Please spell out all the abbreviations in the headings, figures and tables. These need to stand alone.

Response: Done.

Please italicize all variables/parameters in the text.

Response: Done.

Reviewer 2

The authors present an interesting work on flood hazard assessment and mapping. The paper is well-written and easy to follow. However, some issues need to be addressed before the paper can be accepted for publication as follows:

Response: Thank you for your positive feedback and the constructive comments on our manuscript. Please see our detailed response to the comments below:

The abstract should briefly state the purpose of the research, the principal results, and major conclusions. The abstract should be more descriptive rather than informative. More than half of this abstract is allocated to the research gaps which in my opinion is not appropriate (L24-36). Please revise the abstract section with more focus on your methods, and significant results/conclusions.

Response: Thanks for the suggestion on editing the abstract. We removed several lines from the first part of the abstract and added more texts to better describe the method and results of the proposed approach. Please see the revised abstract below:

"In the last decade, DEM-based classifiers based on Height Above Nearest Drainage (HAND) have been widely used for rapid flood hazard assessment demonstrating satisfactory performance for inland floods. The main limitation is the high sensitivity of HAND to the topography which degrades the accuracy of these methods in flat coastal regions. In addition, these methods are mostly used for a given return period and generate static hazard maps for past flood events. To cope with these two limitations, here we modify HAND, propose a composite hydrogeomorphic index and develop hydrogeomorphic threshold operative curves for rapid real-time flood hazard assessment in coastal areas. We select the Savannah river delta as a testbed, calibrate the proposed hydrogeomorphic index on Hurricane Matthew and validate the performance of the developed operative curves for Hurricane Irma. The hydrogeomorphic index is proposed as the multiplication of two normalized geomorphic features, HAND and distance to the nearest drainage. The calibration procedure test different combinations of the weights of these two features and determine the most appropriate index for flood hazard mapping. Reference maps generated by a well-calibrated hydrodynamic model, Delft3D-FM model, are developed for different water level return periods. For each specific return period, a threshold of the proposed hydrogeomoprhic index that provide the maximum fit with the relevant reference map is determined. The collection of hydrogemorphic thresholds developed for different return periods are used to generate the operative curves. Validation results demonstrate that the total cells misclassified by the proposed hydrogeomophic threshold operative curves (summation of overprediction and underprediction) are less than 20% of the total area. The satisfactory accuracy of the validation results indicates the high efficiency of our proposed methodology for fast and reliable estimation of hazard areas for an upcoming coastal flood event which can be beneficial for emergency responders and flood risk managers."

L167. Add one or two sentences to explain about Savanah model in Delft3D-FM.

Response: We have included more details of the Delft3D-FM suite package (Line 186-189). For additional details of the Savannah model, the reviewer is referred to section 3.1.

“Specifically, we used the 2021 Delft3D-FM suite package to model the complex interactions between riverine, estuarine, and intertidal flat hydrodynamics. The suite package can provide detailed information of water level, flow rates, and velocity (Delft3D Flexible Mesh Suite - Deltares, 2021)”

Using a univariate flood frequency analysis in an estuary region should be justified with a detailed analysis that shows there is no correlation between high river flow and sea water level. Otherwise, a bivariate flood frequency analysis should be considered.

Response: In the first steps of this study, we had set up the calibrated Delt3D-FM model for different combinations of upstream flow and downstream water levels. However, we did not find a significant correlation (p-value < 0.05) between river discharge at Clyo station (USGS - 02198500) and coastal water levels at Fort Pulaski station (NOAA - 8670870). The latter was also reported in Ghanbari et al., 2021 and Muñoz et al., 2020. Furthermore, our results demonstrated that high river flow does not affect the inundation area in wetland areas. This indicates that flood inundation is highly dominated by coastal forcing as tides propagate into the Savannah river and lead to flow reversal at upstream gauge stations (see Figure 1 below). The high proximity of wetlands to the Atlantic Ocean shows that the transitional zone, i.e., the area affected by both coastal and inland drivers, is located upstream Port Wentworth station (USGS - 02198920) where the Savannah river trifurcates into the Back River, Middle River, and Front River. Considering the dominant role of sea water level in coastal flooding as well as the negligible effect of river discharge on wetland inundation from the previous analyses, we can justify the proposed univariate flood frequency analysis. For the reviewer’s convenience we also generate a figure of maximum floodwater depth in Savannah under high river flow regimes (10 and 1000-year return period) and mean sea level (Figure 2). The flood maps indicate similar inundation patterns over coastal wetlands and clear differences in upstream zones. . Please refer to Section 3.2 (Lines 277-288) in the revised manuscript for additional justification of the univariate approach.

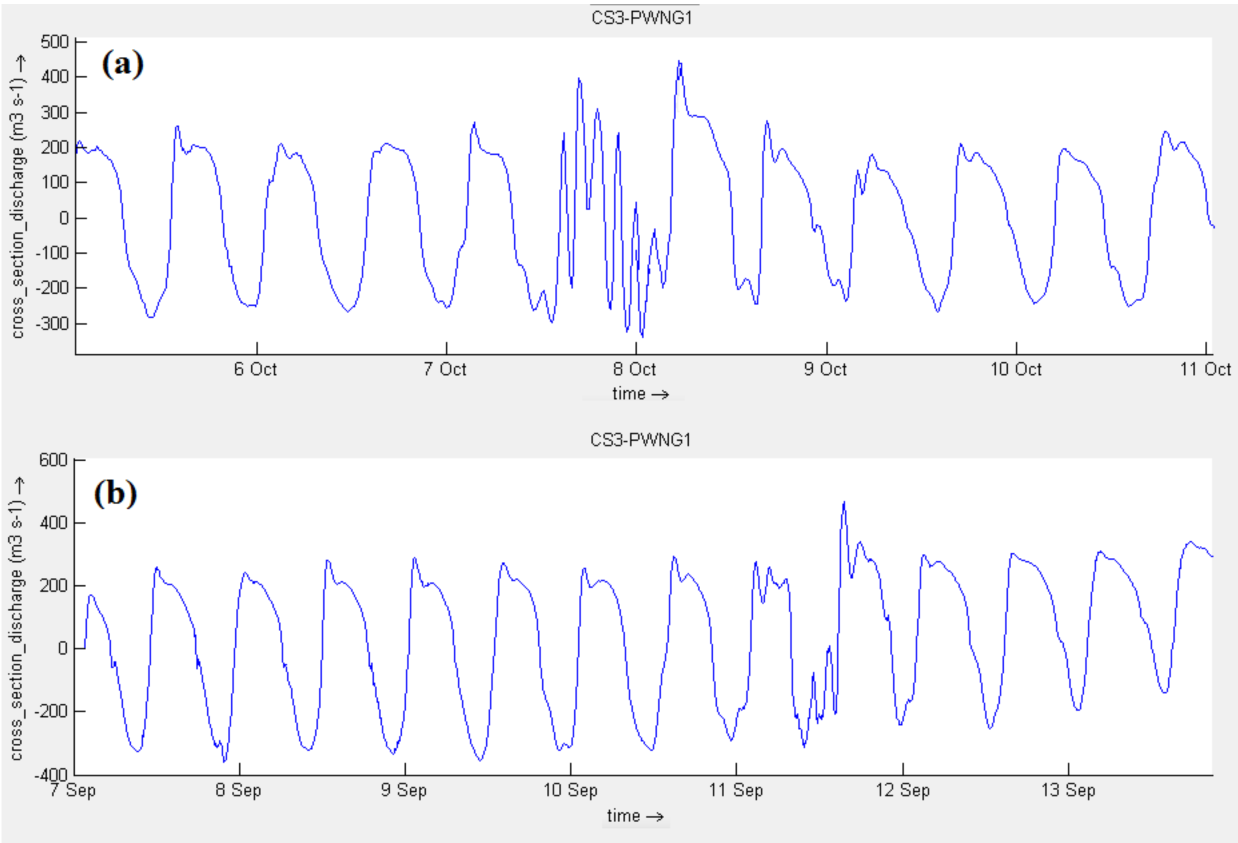

Figure 1. Flow reversal (negative river flow) due to tidal propagation at Port Wentworth station (USGS – 02198920). Simulations of averaged cross section discharge correspond to (a) Hurricane Matthew (Oct/2016) and (b) Hurricane Irma (Sep/2017).

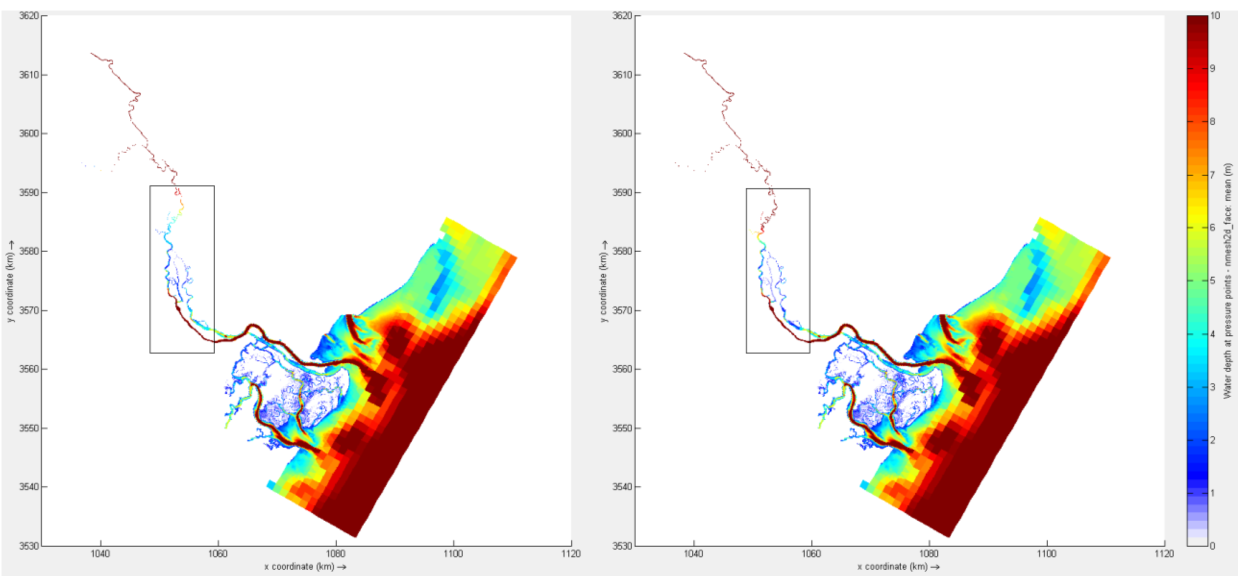

Figure 2. Maximum floodwater depth in Savannah River delta. Simulations of mean sea level and river flow for a return period of (a) 10-year (1413 m^3/s) and (b) 1000-year (2273 m^3/s). Black boxes outline differences of floodwater depth in the transitional zone. The

water depth maps created for the lower parts of the transitional zone (wetland) suggest the negligible effect of river discharge on coastal wetland inundation.

How did you test different combinations of W1 and W2 (Weight parameters)? Please clarify.

Response: Knowing the condition of W1+W2=1, we uniformly pick 100 random w1 from the range of 0-1 which results in 100 set of w1 and w2 (1-w2) for our calibration. Please refer to lines 321-322 in the revised manuscript.

It is not clear how the parameter of TH is derived. Please clarify.

Response: The TH parameter is the result of solving a simple optimization problem by minimizing the total error. We added more information to better explain how to optimize the parameter TH in the revised manuscript. Please refer to lines 357-359.

"To calibrate the binary classifier we minimize the error while searching for the optimum TH value. This means, we use a hundred TH values uniformly picked from the range of $I_{HD}^{min}$and I_{HD}^{max}. For each TH, we use Eq. 2 to generate a binary hazard map and then compare this map with the reference map by calculating the error from Eqs. 3-5."

The manuscript would be significantly improved by providing more discussion about the broader contribution of the study. (e.g., How coastal planners and managers could benefit from the proposed methodology? How the proposed methodology can be utilized in other coastal regions?)

Response: We provided more discussion on the broader impacts of this study and implementation of if in other coastal regions. Please refer to lines 550-555 and 580-609 in the revised manuscript.

"To implement this approach, first, a hydrodynamic model should be set up for the new study area and generate reference inundation maps for different return periods. Access to observed water level data (gauges or HWMs) and flood extent maps from past floods is required to properly calibrate the hydrodynamic model. Then the I_{HD} index calculated from a DEM is utilized together with the reference maps to provide the hydrogeomorphic threshold operative curves for future floods."

"Operationally, the Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model (Jelesnianski et al., 1984) is the storm surge model currently used by NWS to perform storm surge forecasting and create probabilistic flood inundation maps for real-time tropical storms (Sea, Lake, and Overland Surges from Hurricanes (SLOSH, 2022). The feature of SLOSH that makes it the preferred model of the NWS for storm surge forecasting and mapping is the model's computational efficiency that allows the model to be run as an ensemble (Forbes et al., 2014). However, SLOSH is just one of several modeling options for storm surge modeling and mapping, each possessing strengths and weaknesses associated with their simulations. The inclusion of additional models that can create flood maps of storm surge for a given event should provide an enhanced understanding of the uncertainty of inundation at a given location (Teng et al., 2015). However, the higher computational burden of alternative models, such as Delft3D-FM, tend to preclude their use in real-time operations and certainly, their use in generating an ensemble necessary for probabilistic flood maps. The methodology we propose in this manuscript may offer the NWS and other agencies a means to utilize alternatives to SLOSH for flood inundation mapping and probabilistic flood inundation mapping on U.S. coastlines. Models such as Delft3D-FM can generate reference maps to train the binary classifer and build the probabilistic operating curves. The probabilistic operative curves would account for the major source of uncertainties and provide a computationally efficient and reliable decision-making tool for coastal planners and floodplain managers. The operative hydrogeomorphic threshold classifiers proposed for real-time coastal flood hazard mapping can be used as an alternative tool for the rapid estimation of hazardous areas during real-time flood events. In an operational mode, water level or meteorological forecasts can be used to estimate the return period of an upcoming coastal flood event and the methodology here can utilize this as an input to perform LCFM flood inundation mapping both deterministically and probabilistically."

The limitations of the study and the possible enhancements of the proposed methodology should be discussed clearly

Response: We have already included three areas of research for future studies. To expand this, we added more text explaining the study limitations and potential areas for future research. Please refer to lines 520-528 in the revised manuscript.

"The proposed hydrogeomorphic index (I_{HD}) is the primary data for flood hazard mapping in this study. Thus, the quality of two main inputs of this index, namely the DEM and stream network used to calculate features H and D play a vital role in the overall accuracy of the proposed approach. To obtain maximum accuracy, here we used the best available DEM with the finest spatial resolution of 3 m that includes the bathymetry data. However, considering the limited access to such high-quality DEMs in many areas of the world, it is recommended to evaluate the sensitivity of the proposed approach to lower quality DEMs (e.g. 30 m and 90 m DEMs without bathymetry information) in future studies. Another piece of research can investigate the sensitivity of the proposed approach to the density of the drainage network used for calculating the I_{HD} index."

In general here are the areas of research we recommended for future studies:

1. Sensitivity of the hydrogeomorphic index to DEM quality and stream network density (Lines 520-528)
2. Applying the proposed hydrogeomorphic operative curves to inland floods and to other deltas across the US. (Lines 544-550)
3. Improve the flood frequency analysis, considering its uncertainties, incorporating other sources of uncertainties in the modeling to generate probabilistic operative curves (Lines 556-577)
4. A benchmark study that compares the performance of three LCFM methods (Lines 617-620)

Real-time coastal flood hazard assessment using DEM-based hydrogeomorphic classifiers

Keighobad Jafarzadegan[1], David F. Muñoz[1], Hamed Moftakhari[1], Joseph L Gutenson[2], Gaurav Savant[2], Hamid Moradkhani[1]

[1] Center for Complex Hydrosystems Research, Department of Civil, Construction, and Environmental Engineering, University of Alabama, Tuscaloosa, AL
[2] US Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Vicksburg, MS, USA

Correspondence to: Keighobad Jafarzadegan kjafarzadegan@ua.edu

Key Points

- A DEM-based approach is developed for rapid flood hazard assessment in coastal regions.
- The Height Above Nearest Drainage (HAND) is modified for flood mapping in flat areas.
- Operative hydrogeomorphic curves are proposed for real-time flood hazard mapping.

Abstract

~~Deltas, estuaries, and wetlands are prone to frequent coastal flooding throughout the world. In addition, a large number of people in the United States have settled in these low-lying regions. Therefore, the ecological merit of wetlands for maintaining sustainable ecosystems highlights the importance of flood risk and hazard management in these regions. Typically, hydrodynamic models are used for coastal flood hazard mapping. The huge computational resources required for hydrodynamic modeling and the long-running time of these models (order of hours or days) are two major drawbacks that limit the application of these models for prompt decision-making by emergency responders.~~ In the last decade, DEM-based classifiers based on Height Above Nearest Drainage (HAND) have been widely used for rapid flood hazard assessment demonstrating satisfactory performance for inland floods. The main limitation is the high sensitivity of HAND to the topography which degrades the accuracy of these methods in flat coastal regions. In addition, these methods are mostly used for a given return period and generate static hazard maps for past flood events. To cope with these two limitations, here we modify HAND, ~~and~~ propose a composite hydrogeomorphic index and develop hydrogeomorphic threshold operative curves for rapid real-time flood hazard assessment in coastal areas. ~~We also propose the development of hydrogeomorphic threshold operative curves for real-time flood hazard mapping.~~ We select the Savannah river delta as a testbed, calibrate the proposed hydrogeomorphic index on Hurricane Matthew and validate the performance of the developed operative curves for Hurricane Irma. The hydrogeomorphic index is proposed as the multiplication of two normalized geomorphic features, HAND and distance to the nearest drainage. The calibration procedure tests different combinations of the weights of these two features and determines the most appropriate index for flood hazard mapping. Reference maps generated by a well-calibrated hydrodynamic model, Delft3D-FM model, are developed for different water level return periods. For each specific return period, a

threshold of the proposed hydrogeomo~~pr~~rphic index that provides the maximum fit with the relevant reference map is determined. The collection of hydrogeomorphic thresholds developed for different return periods ~~are~~is used to generate the operative curves. Validation results demonstrate that the total cells misclassified by the proposed hydrogeomorphic threshold operative curves (summation of overprediction and underprediction) are less than 20% of the total area. ~~can rapidly generate flood hazard maps with satisfactory accuracy.~~ The satisfactory accuracy of the validation results ~~This~~ indicates the high efficiency of our proposed methodology for fast and ~~accurate~~ reliable estimation of hazard areas for an upcoming coastal flood event which can be beneficial for emergency responders and flood risk managers.

1. Introduction

Densely populated coastal areas are some of the most productive ecosystems on Earth. Coastal wetlands provide important services to society, including flood attenuation, water storage, carbon sequestration, nutrient cycling, pollutant removal, and wildlife habitat (Barbier, 2019; Land et al., 2019; Wamsley et al., 2010). Characterizing the hydrological processes unique to coastal areas is tremendously important for ensuring the sustainability of these ecosystem services. Endangered coastal ecosystems are threatened by anthropogenic effects, including direct impacts of human activities (i.e. urbanization and navigational development) or indirect impacts (e.g. sea level rise (SLR), and hydroclimate extremes (e.g. floods) exacerbated by climate change (Alizad et al., 2018; Kirwan and Megonigal, 2013; Wu et al., 2017). Nearly 70% of global wetlands have been lost since the 1900s and rates of wetland loss have increased by a factor of ~~4~~ four in the late 20^{th} and early 21^{st} century (Davidson, 2014). Urbanization hinders wetland migration toward upland areas in an effort to cope with rising water levels (WLs) (Schieder et al., 2018). Likewise, moderate to high Relative Sea Level Rise (RSLR) rates can influence the fate of sediments and nutrient availability to coastal wetlands (Schile et al., 2014); and eventually transform low marsh regions into open water or mudflat areas (Alizad et al., 2018). SLR and navigational development can alter the tidal regime and long-wave propagation characteristics inside estuaries/bays and ~~so~~ subsequently change the flooding inundation patterns (Familkhalili et al., 2020; Khojasteh et al., 2021a, b). Similarly, hurricane impacts can create interior ponds, trigger shoreline erosion, and denude marshes (Morton and Barras, 2011). People and assets located in low-lying coastal regions and river deltas are frequently exposed to compound flooding. Challenges for flood hazard assessment unique to these systems include compounding effects of multiple flooding mechanisms, complex drainage systems with relatively low slopes, and periodically saturated soils. ~~it~~ It is expected that between 0.2-4.6% of the global population may be exposed to coastal flooding if no strategic adaptation takes place (Kulp and Strauss, 2019).

Efficient ~~flood~~ risk reduction strategies require accurate real-time assessment of flood~~ing~~ hazards (Gutenson, 2020; USGS Surface Water Information, 2021). ~~In order t~~To simulate the coastal flood hazard in wetlands, two-dimensional (2D) hydrodynamic models are commonly used for flood inundation mapping, as they allow for simulating complex oceanic, hydrological, meteorological~~,~~ and anthropogenic processes based on process-based numerical schemes. The advanced circulation model (ADCIRC) (Luettich et al., 1992), DELFT3D (Roelvink and Banning, 1995), and LISFLOOD-FP (Bates et al., 2010) are among the most commonly used 2D hydrodynamic models for coastal flood hazard assessment in low-lying areas at local and regional scales (Bates et al., 2021; Muis et al., 2019; Thomas et al., 2019). Nonetheless, hydrodynamic modeling approaches require huge computational resources to conduct flood hazard assessments at a large scale. This is even more challenging when emergency responders need timely flood risk information at a desirable accuracy and resolution on a real-time basis. Therefore, while 2D hydrodynamic models are still a key component in many frameworks for detailed analyses of the flood hazard, the use of low-complexity flood mapping (LCFM) methods is essential for the preliminary estimation of areas exposed to flooding in a short time. Applying LCFM methods together with detailed hydrodynamic models provide a more comprehensive set of information for emergency responders and improve the efficiency of flood risk management in practice.

The advent of Digital Elevation Models (DEMs) has led to the development of a series of GIS-based LCFM methods for rapid estimation of flood hazard in the last couple of decades (Afshari et al., 2018; Dodov and Foufoula-Georgiou, 2006; Manfreda et al., 2011; McGlynn and McDonnell, 2003; McGlynn and Seibert, 2003; Nardi et al., 2006; Samela et al., 2016; Teng et al., 2015; Williams et al., 2000). Among these methods, binary classification of a hydrogeomorphic raster has been shown to be an efficient approach for reliable delineation of floodplains (Degiorgis et al., 2012; Manfreda et al., 2014). In a binary hydrogeomorphic classification approach, the study area is examined as a grid of cells, then a threshold of a hydrogeomorphic feature, typically calculated from a DEM, is chosen. Comparing the hydrogeomorphic feature value of cells with the threshold, the entire study area is classified into flooded and non-flooded cells.

The Federal Emergency Management Agency (FEMA) provides flood hazard maps across the United States. These maps, also referred to as Flood Insurance Rate Maps (FIRMs) identify flood-prone areas corresponding to specific return periods. While these hazard maps provide useful information for a few recurrence intervals, they are no longer reliable for extreme flood events characterized by lower frequencies or longer return periods. In 2015, the National Water Center Innovators Program initiated the national flood interoperability experiment (NFIE) for real-time flood inundation mapping across the United States (Maidment, 2017; Maidment et al., 2014). The plan highlighted the tendency for event-based flood mapping which is more valuable and practical for emergency response and warning systems. Unlike past DEM-based methods that mostly focused on flood hazard mapping, Zheng et al., (2018b) proposed the development of DEM-based synthetic rating curves for real-time flood inundation mapping. In most current, real-time flood mapping methods, the forecasted river flows and/or water surface elevation are typically fed into flood inundation libraries to simulate the upcoming flood inundation areas (IWRSS, 2015, 2013; Maidment, 2017; Wing et al., 2019; Zheng et al., 2018a). The computationally intensive and time-consuming nature of detailed hydrodynamic models to numerically route flood waves typically restricts their usage in supporting emergency response activities (Gutenson et al., 2021; Longenecker et al., 2020).

An LCFM method based on Height Above Nearest Drainage (HAND) has been widely used and recognized as one of the best classifiers for identifying flood hazard areas (Degiorgis et al., 2012; Jafarzadegan et al., 2018; Jafarzadegan and Merwade, 2019; McGrath et al., 2018; Samela et al., 2017; Zheng et al., 2018a). The performance assessment of HAND classifiers in different topographic settings suggests, despite an acceptable performance in most locations, the accuracy of hazard maps is significantly lower in low-lying coastal regions (Jafarzadegan and Merwade, (2017) and Samela et al., (2017)). While the majority of DEM-based flood hazard mapping methods have been developed and tested for inland floods, access to an appropriate DEM-based method for coastal flooding is lacking in the literature. Since coastal flooding occurs rapidly and the time for hydrodynamic modeling and designing flood mitigation strategies is limited especially in data-scarce regions, efficient DEM-based approaches can be significantly beneficial for emergency and response-related decision-makers.

The overarching goal of this study is to propose a DEM-based LCFM method for coastal wetlands, estuaries, and deltas. To our knowledge, this is the first study that investigates the application of hydrogeomorphic binary classifiers for flooding in semi-flat coastal zones. We modify the HAND commonly used for riverine inland flooding (Degiorgis et al., 2013; Jafarzadegan et al., 2020; Samela et al., 2017) and propose a composite hydrogeomorphic index for tidally-influenced coastal regions. We enhance the applicability of the proposed method by developing hydrogeomorphic threshold operative curves for coastal flood hazard mapping. Unlike previous studies that rely on binary classifiers for specific return periods, the operative curves here offer a unique opportunity for rapid assessment of hazardous areas in real-time. These curves have substantial benefits for emergency responders when wetlands are prone to coastal flooding.

2. Study area and data

We study the Savannah River delta located ~~in the Southeast United States,~~ at the border of Georgia and South Carolina in the southeast United States (Figure 1a). The Savannah River originates at the confluence of Tulagoo and Seneca rivers and drains the Lower Savannah watershed (HU08_03060109) comprising an area of 2603.96 km^2. The morphology of this region is relatively complex due to the existence of a braided river followed by a dense drainage network of interior rivers and tidal creeks. The average slope, length, and annual discharge of the Savannah River are 0.00011 m/m, 505 km, and 320 m^3/s, respectively (Carlston, 1969). Moreover, the river bathymetry was deepened up to 12 m for increasing the capacity of cargo transportation (U.S. Army Corps of Engineers, 2017). This region is mostly characterized by its unique ecology, including vast wetlands and saltmarsh ecosystems. We obtain detailed drainage network data including river streams, tidal channels, and creeks within wetland areas from the U.S. National Wetlands Inventory (https://www.fws.gov/wetlands/data/Mapper.html).

To simulate the flood hazard in this region, a mesh boundary encompassing the Savannah River delta, surrounding areas, and a portion of the Atlantic Ocean is generated (Figure 1b). Two U.S.

Geological Survey (USGS) gauges, located at the Savannah River (#02198500, #02198690) and the Fort Pulaski station of the National Oceanic and Atmospheric Administration (NOAA) are used as upstream and downstream boundary conditions of the hydrodynamic model, respectively. Fort Pulaski station (NOAA – 8670870) counts with an 85-year length of records (since 1935) that enables a proper characterization of coastal flooding for design levels at lower frequencies or relatively large return periods. We select this region as a testbed because of 1) frequent coastal flooding induced by large semidiurnal tidal amplitudes at the estuary mouth (Cowardin et al., 2013) and 2) exposure of more than twenty thousand people settled in four developed areas, the Whitemarsh, Talahi, Wilmington, and Tybee Islands located in this region (Figure 1c).

The high-resolution DEM used as the base of our proposed hydrogeomorphic index is a 3 m light detection and ranging (LiDAR) that includes topographic and bathymetric (topobathy) data. This dataset has been developed by the NOAA's National Centers for Environmental Information (NCEI) and is available at the NOAA's Data Access Viewer repository (https://coast.noaa.gov/dataviewer/). The topobathy data was further corrected for wetland elevation error using the DEM-correction tool developed by Muñoz et al., (2019) ~~in order~~ to minimize vertical bias errors commonly found in LiDAR-derived coastal DEMs (Alizad et al., 2018; Medeiros et al., 2015; Rogers et al., 2018). The vertical and horizontal accuracy of the DEM are 50 and 100 cm, respectively and its vertical datum is the North American Vertical Datum 1988 (NAVD88). Land cover maps are obtained from the 2016 National Land Cover Database (NLCD) available at (https://www.mrlc.gov/). River discharge and WL records are obtained from the USGS (https://maps.waterdata.usgs.gov/mapper/index.html) and NOAA (https://tidesandcurrents.noaa.gov/), respectively. In addition, post-flood high water marks (HWMs) of Hurricane Irma and Matthew are obtained from the USGS Flood Event Viewer platform (https://stn.wim.usgs.gov/FEV/). These high-water marks are used for calibration and validation of the Savannah model in Delft3D-FM. Specifically, we used the 2021 Delft3D-FM suite package to model the complex interactions between riverine, estuarine, and intertidal flat hydrodynamics. The suite package can provide detailed information o~~f~~n water level, flow rates, and velocity (Delft3D Flexible Mesh Suite - Deltares, 2021).

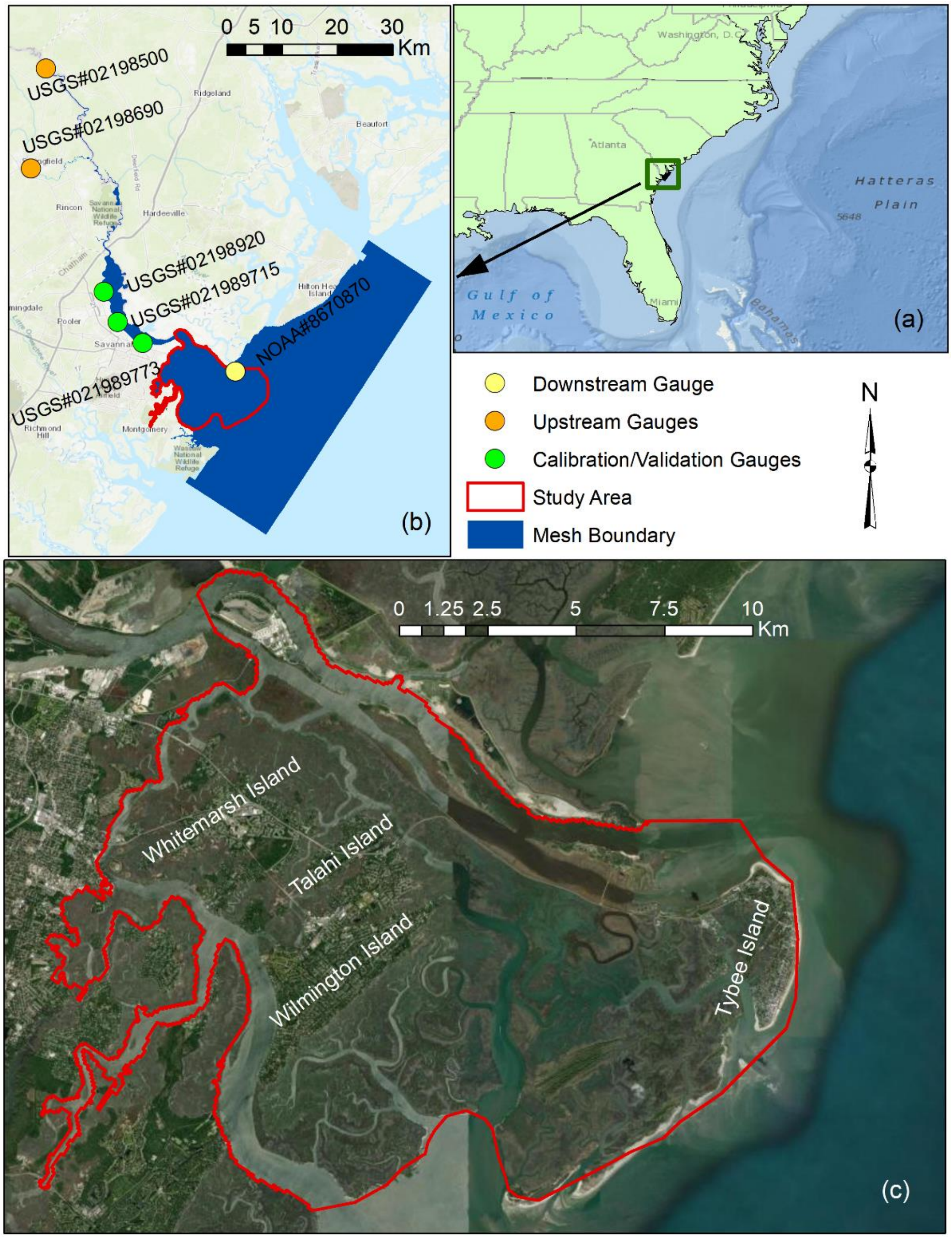

Figure 1. Map of the study area and mesh boundary of the hydrodynamic model**.** (a) the geographic location of the study area in the southeast U.S., (b) The mesh boundary used by the hydrodynamic model (blue) for flood inundation mapping as well as the location of upstream (orange),

downstream (yellow), and calibration/validation (green) gauges, and (c) the boundary of Savannah wetlands used as the case study along with urbanized areas. (ESRI 2018)

3. Methods

We propose a DEM-based LCFM approach for the rapid assessment of flood hazard areas in real-time. The proposed approach consists of two phases (Figure 2). In phase 1, a 2D hydrodynamic model is calibrated based on observed WLs at USGS gauges and HWMs that are available during Hurricane Matthew in 2017. We then use the calibrated hydrodynamic model to generate a flood inundation map that serves as a reference map in the next phase. In addition, for flood frequency ~~analysis~~analyses, we perform block maxima sampling approach to select the annual WL maxima at Fort Pulaski station. The selected samples are then used to estimate ~~return levels~~WLs for six return periods of 10, 50, 100, 200, 500, and ~~1000 year~~1000-year floods. Using these estimated WLs as the main boundary conditions of the hydrodynamic model, we also generate six flood inundation maps corresponding to these return periods. In phase 2, we use a high-resolution DEM together with the drainage network data to calculate the hydrogeomorphic index. Subsequently, the flood inundation map generated for Hurricane Matthew in phase 1, is used as a reference map to calibrate the hydrogeomorphic index. Then, the calibrated index uses the flood inundation maps provided for different return periods in phase 1 to develop the operative curves. These curves form the basis for the rapid assessment of flood hazard areas for any upcoming coastal flood event in the future. To validate the effectiveness and reliability of the developed operative curves, we use them to identify hazard areas corresponding to Hurricane Irma, and then we compare their accuracy with the reference map provided by the hydrodynamic model for this flood event. In the following sections, we further explain the hydrodynamic model, flood frequency analysis, and hydrogeomorphic method, respectively.

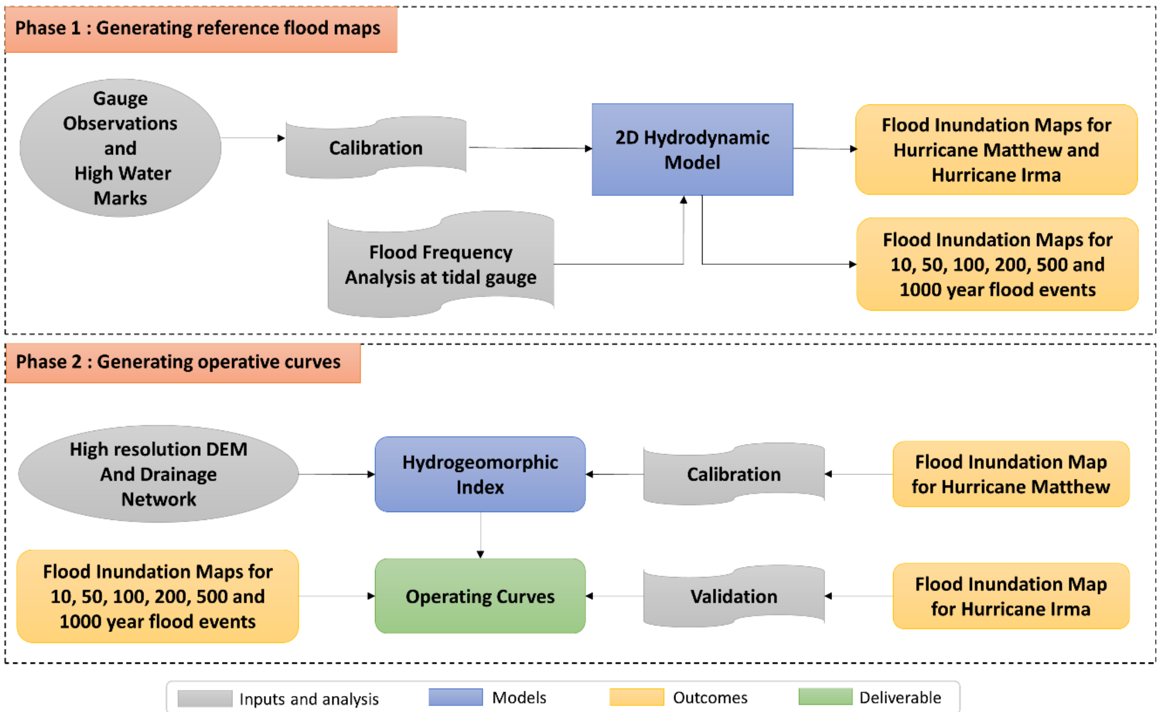

Figure 2. Flowchart of the proposed approach for generating hydrogeomorphic threshold operative curves. In Phase 1, the 2D hydrodynamic model is calibrated and generates the required reference maps for the next phase. In Phase 2, the reference maps are used in conjunction with the hydrogeomorphic index to generate the operative curves for fast and real-time coastal flood hazard assessment.

3.1 Hydrodynamic Model

3.1.1 Model setup

We use the 2019 Delft3D-FM suite package (Deltares, 2019) to model the complex riverine, estuarine, and intertidal flat hydrodynamics in the Savannah River delta and wetland regions. The suite package has been used in similar coastal studies characterized by vast wetland regions with satisfactory results (Fagherazzi et al., 2014; Kumbier et al., 2018; Sullivan et al., 2019). Moreover, the model developed for Savannah has been used in other studies to simulate extreme and non-extreme events including Hurricane Matthew that hit the southeast Atlantic Coast in October 2016

(Muñoz et al., 2021, 2020). The 2D hydrodynamic model comprises nearly 85 km of the Savannah River extending from Fort Pulaski station (NOAA – 8670870) at the coast up to Clyo station (USGS – 02198500) and cover an area of 1178 km^2 approximately. The model consists of an unstructured triangular mesh to ensure a correct representation of geomorphological settings including sinuous and braided river waterways and relatively narrow tidal inlets. Furthermore, the mesh has a spatially varying cell size ranging from 1.5 m in the upstream riverine area, 10 m over wetland regions, 120 m along the coast, and up to 1.4 km over the Atlantic Ocean (Figure 1b).

3.1.2 Model calibration

For calibration purposes, the model was forced with time series of river flow obtained from Clyo station as an upstream boundary condition (BC), coastal WL from Fort Pulaski station as a downstream BC, and with spatially varying Manning's roughness values (*n*) classified into open water, wetland, urban, and riverine areas. The optimal (or calibrated) set of *n*-values ~~weres~~ was inferred from 200 model simulations of Hurricane Matthew, as this event reported the highest peak WL at Fort Pulaski station since the year 1935 (2.59 m w.r.t. NAVD88). Each simulation was conducted in a high-performance computing system and included a one-month warm-up period and a unique set of *n*-values for each land cover generated with the Latin Hypercube Sampling (LHS) technique (Helton and Davis, 2003). The advantage of LHS over traditional Monte Carlo approaches is that the former results in a denser stratification over the range of each sampled parameter and is therefore superior to random sampling. LHS leads to more stable results that are closer to the true probability density function of the parameter and has been used in similar studies (Jafarzadegan et al., 2021; Muñoz et al., 2022). The range of *n*-values was derived from pertinent literature and included hydrodynamic modeling and open channel flow studies (Arcement and Schneider, 1989; Chow Ven, 1959; Liu et al., 2019). The set of values achieving both the lowest

Root Mean Square Error (RMSE) ~~and highest correlation coefficient (R^2)~~ and highest Kling-Gupta Efficiency (KGE) around the peak WL (e.g., 7-day window) was selected as the optimal one and further used for coastal flood simulations. KGE is a robust evaluation metric that accounts for correlation, bias, and the ratio of variances and can take values between -∞ and 1 (Gupta et al., 2009). An efficiency of 1 indicates a perfect match between model simulations and observations. In addition to those metrics, we evaluate the model's performance using the Nash-Sutcliffe Efficiency (NSE) and Mean Absolute Bias (MAB). NSE measures the relative magnitude of the error variance of model simulations compared to the variance of observational data (Nash and Sutcliffe, 1970). NSE ranges between -∞ to 1, where an efficiency of 1 indicates a perfect match. MAB quantifies the bias of model simulations with respect to observational data. MAB of 0 suggests an absence of bias in the simulation. The calibrated *n*-values used in ~~this~~ the Savannah model are~~:~~ open water (n = 0.027), wetland (n = 0.221), urban (n = 0.03), and downstream/upstream riverine areas (n = 0.037 and n = 0.086, respectively).

3.2 Flood Frequency Analysis

Preliminary model simulations indicate a negligible influence of river flow on coastal wetland inundation as compared to storm surge ~~at on~~in the Wassaw Sound, Wilmington, and Tybee islands (Figure 1c). This can be explained by the proximity of the islands to the Atlantic Ocean as well as freshwater runoff regulation and flood controls by three large dams located upstream of the Clyo station (USGS – 02198500), namely J. Strom Thurmond, Richard B. Russell, and Hartwell (Zurqani et al., 2018). In addition, bivariate statistical analysis via copulas suggests no significant correlation between river flow at Clyo station (USGS - 02198500) and coastal water levels at Fort Pulaski station (NOAA - 8670870). The latter was also reported in Ghanbari et al., (2021) and Muñoz et al., (2020). Furthermore, our analysis demonstrated that high river flow does not affect the inundation area in wetlands~~ areas~~. This indicates that flood inundation is highly dominated by coastal forcing as tides propagate into the Savannah River and lead to flow reversal at upstream gauge stations (see Figure 1 below). The high proximity of wetlands to the Atlantic Ocean shows that the transitional zone, i.e., the area affected by both coastal and inland drivers, is located upstream of Port Wentworth station (USGS - 02198920) where the Savannah river trifurcates into the Back River, Middle River, and Front River. Considering the dominant role of sea water level in coastal flooding as well as the negligible effect of river discharge on wetland inundation from the previous analyses, we can justify the proposed univariate flood frequency analysis. We~~,~~ therefore~~,~~ conduct a univariate flood frequency analysis based on annual block maxima sampling of WLs observed at the Fort Pulaski station~~ (NOAA – 8670870)~~. We use the *'allfitdist'* tool in MATLAB to find the best parametric probability distribution fit to the data, based on Maximum Likelihood, Bayesian information criterion (BIC), or Akaike information criterion (AIK).

3.3 Hydrogeomorphic index

Among different hydrogeomorphic features used for flood hazard mapping, HAND (sometimes also referred to as feature *H*) has been widely used as one of the best indicators of floodplains. However, due to the weakness of this feature for proper characterization of floodplains in flat regions and coastal areas, here we develop a composite hydrogeomorphic index that considers *H* as well as the distance to the nearest drainage (*D*). Although the overall performance of feature *D* is less than *H* in most case studies (Degiorgis et al., 2012; Manfreda et al., 2015a; Samela et al., 2016), feature *D* can be a better descriptor of floodplains in highly flat regions according to the study conducted by Samela et al., (2017). In another study, Gharari et al., (2011) proposed a composite index by multiplying both features *H* and *D* and demonstrated that *H* is a better feature compared to the case that both features are used for landscape classification. The main drawback of their proposed index was that they used the same weights for both features which result in degrading the classification performance. To overcome the limitation of the proposed index and to consider the key role of feature D in flat areas, we maintain feature D in our composite index and add different weights to H and D using Eq. 1 as follows:

$$I_{HD} = \left(\frac{H}{H_{max}}\right)^{w1} \times \left(\frac{D}{D_{max}}\right)^{w2} \qquad where \qquad w1 + w2 = 1 \qquad (1)$$

In Eq.1, H_{max} and D_{max} denote the maximum value of raster H and D used for normalizing the hydrogeomorphic index whereas $w1$ and $w2$ refer to the weights of feature H and D, respectively. The conditional equation of $w1 + w2 = 1$ helps lower the computational burden of the calibration procedure by reducing the number of unknown parameters from two to one. Figure 3 illustrates an example of calculating the I_{HD} index with a given set of weights ($w1$=0.6 and $w2$=0.4) for the study area. Using a high-resolution coastal DEM (Figure 3a), raster H and D are calculated (Figures 3b and 3c). Considering a DEM with N cells, the main step is to find a coordinate matrix that indicates the location of the nearest stream cell to each grid cell. Knowing this matrix and the number of cells required to cross the nearest stream cell, the feature D is calculated. The coordinate matrix can also be used in conjunction with the DEM to calculate the feature H. ~~In order t~~To calculate the ~~hydrogeomorphic index~~ I_{HD} index, the weights in Eq. 1 are calibrated using a reference flood hazard map obtained from hydrodynamic simulation (e.g., Hurricane Matthew). We tested ~~different~~ a hundred combinations of weight parameters ($w1$ and $w2 = 1 - w1$), derived from random generation of 100 $w1$ in the range of (0-1), to find the importance of features H and D, and then finalized the I_{HD} ~~hydrogeomorphic~~ index with known parameters for future flood hazard mapping. We further validated the weight parameters ~~with~~ through the simulations of Hurricane Irma.

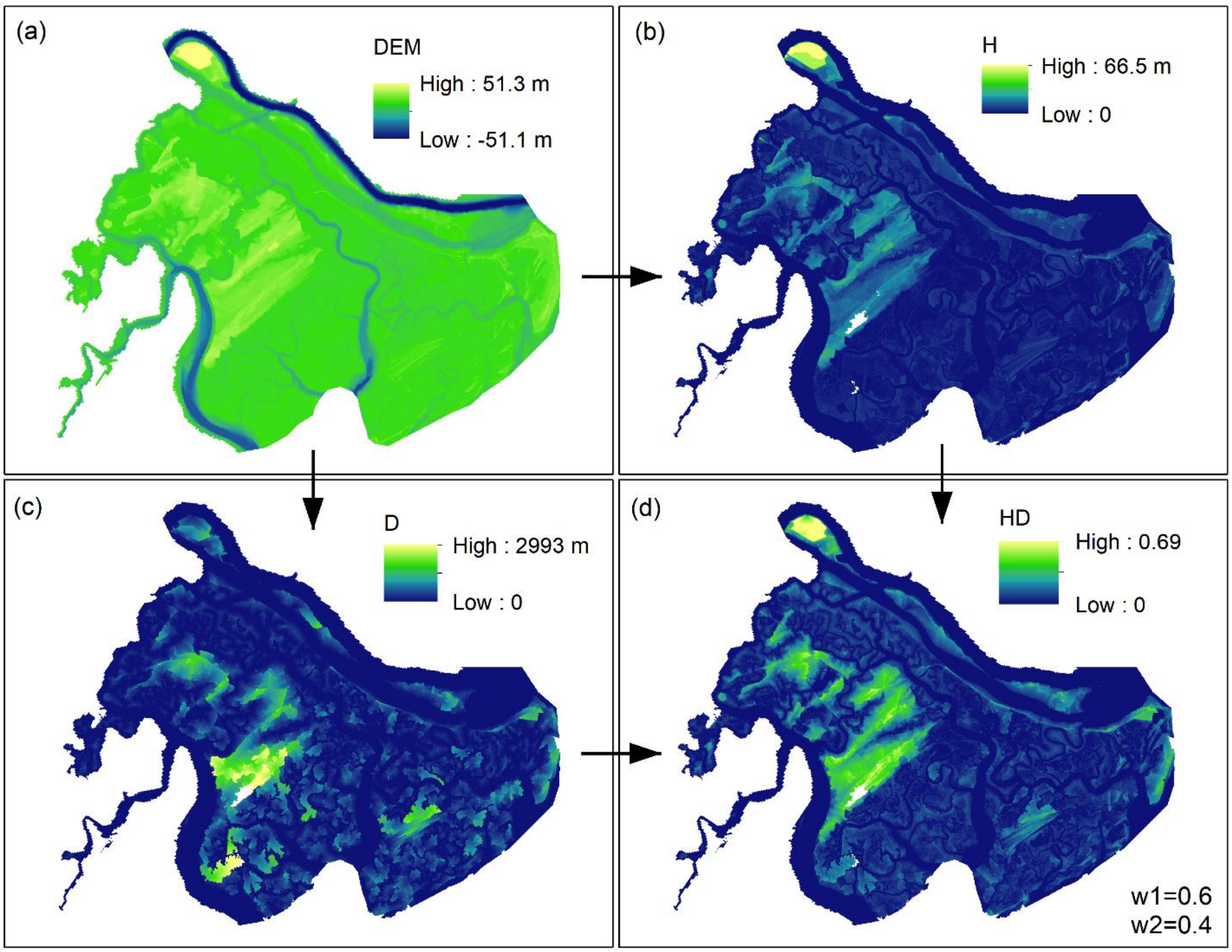

Figure 3. The required steps for calculating the proposed hydrogeomorphic index. A high-resolution coastal DEM (3 m) is used as the source data to (a) generate the Height Above ~~n~~Nearest Drainage (*H*) and the Distance to the nearest Drainage (*D*), respectively (b, c). Using Eq. 1, the normalized features *H* and *D* are multiplied with different weights to generate the I_{HD} ~~hydrogeomorphic~~ index (d).

3.4 Binary classifiers for flood hazard mapping

Considering the study area as a grid of cells, a binary classifier assigns a value of zero or one to each cell and generates a map of two different classes. In flood hazard mapping, the common approach is to define a threshold on a hydrogeomorphic index (e.g. I_{HD}) and use the following if-and-else rule for the classification:

$$f(i) = \begin{cases} 1 & I_{HD}^{i} \leq TH \\ 0 & I_{HD}^{i} > TH \end{cases} \tag{2}$$

where $f(i)$ and I_{HD}^{i} denote the label of flood hazard map and the proposed hydrogeomorphic index value at cell *i*, respectively, and *TH* denotes the threshold of the hydrogeomorphic classifier that should be calibrated. The flood hazard map generated with the binary classifier is compared with a binary reference hazard map, and the rate of true positive (*rtp*), rate of false positive (*rfp*), and *error* are calculated as follows (Jafarzadegan and Merwade, 2017):

$$rtp = \frac{True\ positive\ instances}{Total\ positives} \tag{3}$$

$$rfp = \frac{False\ positive\ instances}{Total\ negatives} \tag{4}$$

$$error = rfp + (1 - rtp) \tag{5}$$

In binary classification, positive and negative refer to a value of one and zero, respectively. True positive instances are those positive cells that are correctly predicted by the classifier and false positive instances represent those negative cells that are wrongly classified as positive. The *error*, reflecting all cells that are wrongly predicted by the classifier, is a commonly-used measure for validating the performance of binary classifiers for flood hazard mapping. Another useful performance measure to validate the binary classifier is the area under the curve (AUC) of the Receiver Operating Characteristic (ROC) graph proposed by Fawcett, (2006).

To calibrate the binary classifier we minimize the *error* while searching for the optimum *TH* value. This means~~,~~ we use a hundred *TH* values uniformly picked from the range of I_{HD}^{min} and I_{HD}^{max}. For each *TH*, we use Eq. 2 to generate a binary hazard map and then compare this map with the reference map by calculating the error from Eqs. 3-5. In this optimization problem, the reference flood hazard map used for calculating the *error* is the key input that should be further described. The flood inundation maps generated by the hydrodynamic model indicate WLs at different cells in different time steps and should be converted to a single binary map. A common approach used for inland floods is to find the maximum inundation area over the entire flooding period and then assign all cells with zero WL to "dry" or "non-flooded" and other cells with positive values as "wet" or "flooded". In delta estuaries and coastal regions nearby the ocean, however, almost all cells can be flooded with small WL values. Therefore, after finding the maximum inundation over the flooding period, we use another set of binary labels as "low hazard" vs "high hazard" and define the hazard depth cutoff (HDC) as a threshold used to convert a continuous map of WL to a binary map with only two labels. Depending on the HDC used for distinguishing low from high hazard regions, the reference flood map is changed which results in a different calibrated *TH*. In addition to HDC, the intensity of the flood event shown with the return period ~~(T)~~ also changes the reference flood hazard map. Therefore, the calibrated parameter *TH* is a function of both HDC and T and the main goal of this study is to provide operative curves showing the variation of *TH* with these two factors. We run the hydrodynamic model for 6 different return periods of 10, 50, 100, 200, 500, and 1000 year events and then convert the WL maps to binary maps using 21 HDC resulting from 0.1 increments in the range of 0-2 m. The binary classification and calibration of

TH are performed for different reference maps generated from various combinations of T and HDC.

4. Results

A comprehensive calibration and validation of the Savannah River model is shown in Figure 4. This step is crucial to ensure that the flood hazard maps provided by the model are reliable enough to be used as the reference of the hydrogeomorphic method. We assess the performance of the model by first comparing simulated and observed WLs at four USGS stations along the Savannah River (Figure 1b, green circles). For convenience, we only present simulated and observed WLs of Hurricane Matthew and Irma at Garden City (Figure 4c and 4d, respectively) located at ~29.5 km from the river mouth (Figure 4a, yellow square). The results of the remaining stations are included in the supplementary material (Figure S1). ~~The~~ RMSE and MA~~E~~B remain below 30 cm and 25 cm whereas KGE and NSE ~~and R^2 of the gauges stations remain below 30 cm and above~~achieve values above ~~0.90~~0.75 and 0.85, respectively, for the two hurricane events, which is reflective of satisfactory model performance. Overall, the magnitude and timing of the highest peak WL observed during the hurricanes are well captured by the Savannah River model. To further evaluate the model performance in coastal flood propagation analysis, we compare maximum WLs resulting from model simulations with the USGS HWMs collected in urban and surrounding wetland areas (Figure 4b). The 1:1 line represents a perfect fit between simulated and observed maximum WLs and helps visualize overestimation (above the 1:1 line) and underestimation of the model. Similarly, the evaluation metrics indicate a satisfactory performance of the model with a slightly over- and underestimation during Matthew and Irma. Moreover, the model achieves a relatively small RMSE (< 35 cm)~~and MAE (< 30 cm)~~.

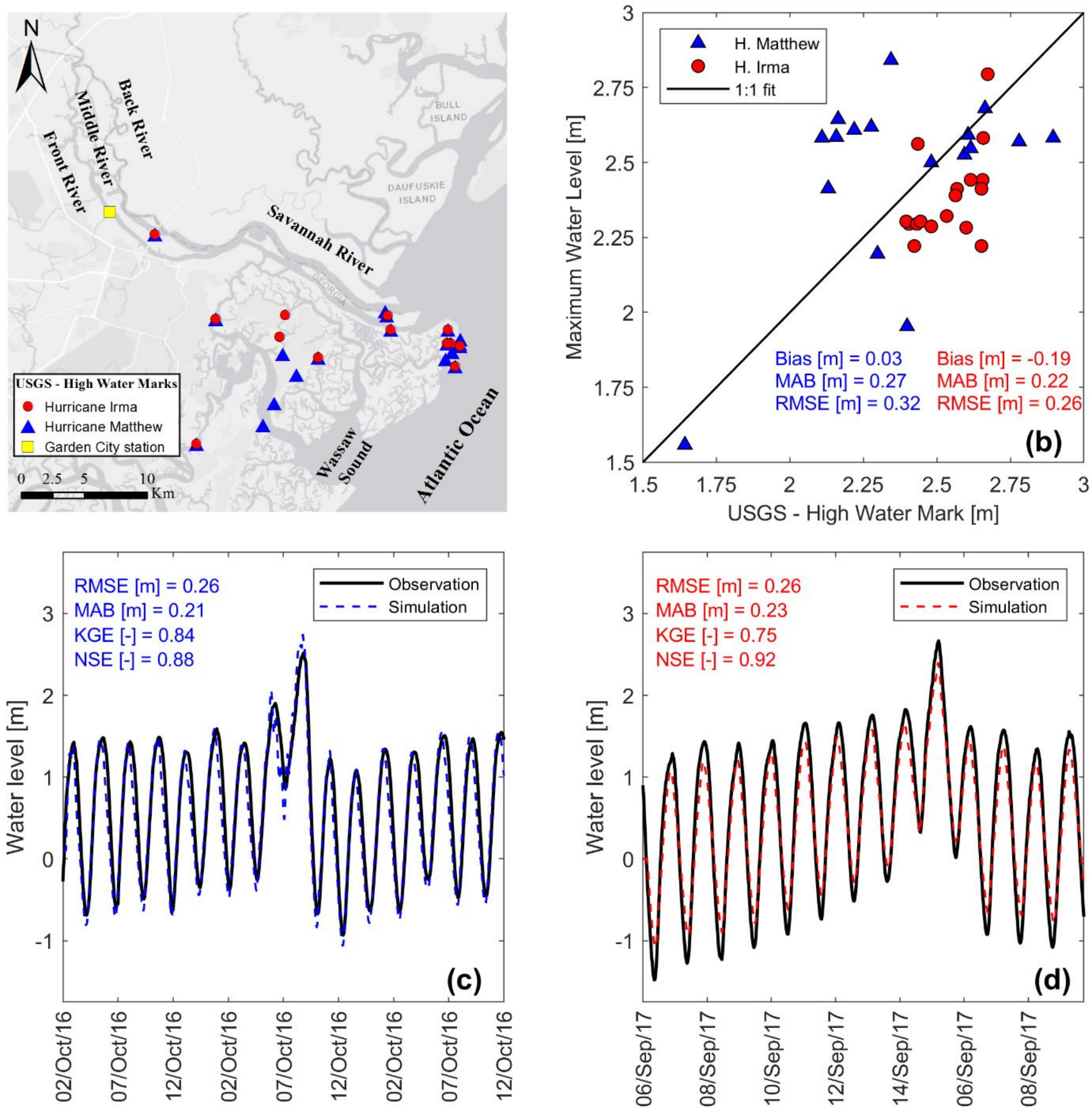

Figure 4. Calibration and validation of the Savannah Delft3D-FM model. (a) Location of high-water marks (HWMs) in the Savannah River delta for Hurricane Mathew (blue triangles) and Hurricane Irma (red circles). (b) Comparison between simulated maximum water levels (WLs) and HWMs in Savannah. (c and d) Time series of simulated and observed WLs at Garden station for Hurricane Matthew and Hurricane Irma, respectively.

To generate boundary conditions for coastal flood modeling simulations associated with the proposed return periods, we perform flood frequency analysis of coastal WL at the Fort Pulaski station (in Figure 5) located at the mouth of the Savannah River (Figure 1b, yellow circle). In this study, we select Generalized Extreme Value (GEV) because of its smallest estimated BIC compared to other parametric distributions available at the Matlab '*allfitdist*' tool. In addition, we show the 95% confidence bounds of the GEV distribution and fit a non-parametric Weibull distribution to the data for comparison purposes. Hereinafter, we will use the GEV distribution to estimate WLs for 10, 50, 100, 200, 500, and 1000-year return periods.

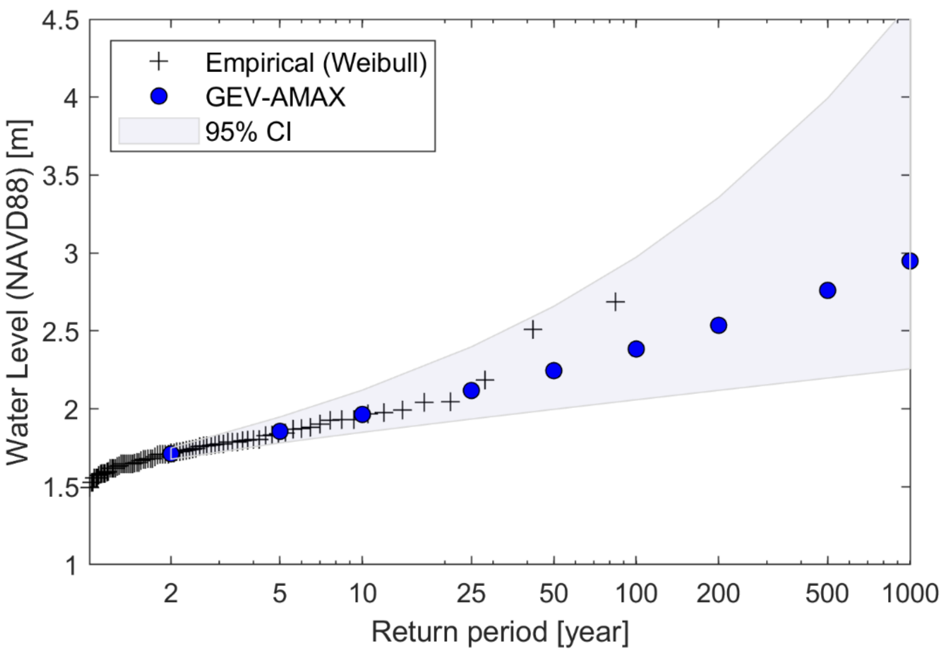

Figure 5. Return water levels (WLs) for Fort Pulaski station in Savannah GA (NOAA - 8670870). Plotting positions (black crosses) are derived from the Weibull formula based on annual block maxima time series (AMAX) and comparable to the Generalized Extreme Value (GEV) distribution (blue circles). 95% confidence intervals (CI) for the distribution parameters of the GEV distribution are shown with a shaded blue band.

After calibrating the Delft3D-FM model, we generate daily flood inundation maps for Hurricane Matthew, determine the maximum flood extent among all days, and then use an HDC to convert the maximum inundation map to a binary map of low and high hazard classes. Using 21 different HDCs ranging from 0 to 2 m, we perform 21 calibrations corresponding to a given reference flood hazard map generated from a specific HDC value. Figure 6a shows the *error* and AUC of calibration corresponding to different HDC values. As can be seen, increasing the HDC decreases the accuracy of the hydrogeomorphic method for flood hazard mapping. Looking into the errors and AUC values reported in the literature of binary flood hazard mapping studies, we consider an error of 0.2 and an AUC of 0.9 (dash lines) as the limits for distinguishing acceptable models from unacceptable ones. The grey region indicates the rejected HDC values above 1.1 m that result in unacceptable accuracy (e.g., Error> 0.2 or AUC<0.9). Figure 6b indicates the optimum weights calculated from the calibration of the hydrogeomorphic method corresponding to different HDC values. The higher value of $w1$ compared to $w2$ demonstrates that feature H is a more important factor than feature D in representing the flood hazard areas, and a combination of both features is the best indicator of floodplains compared to using each feature individually ($w1$= 0 or $w2$=0). Figure 6b also shows that for the HDC=0 (wet vs dry classification), feature D shows the highest contribution (30%) while using the high HDC value of 2 m decreases the contribution of this feature to almost zero.

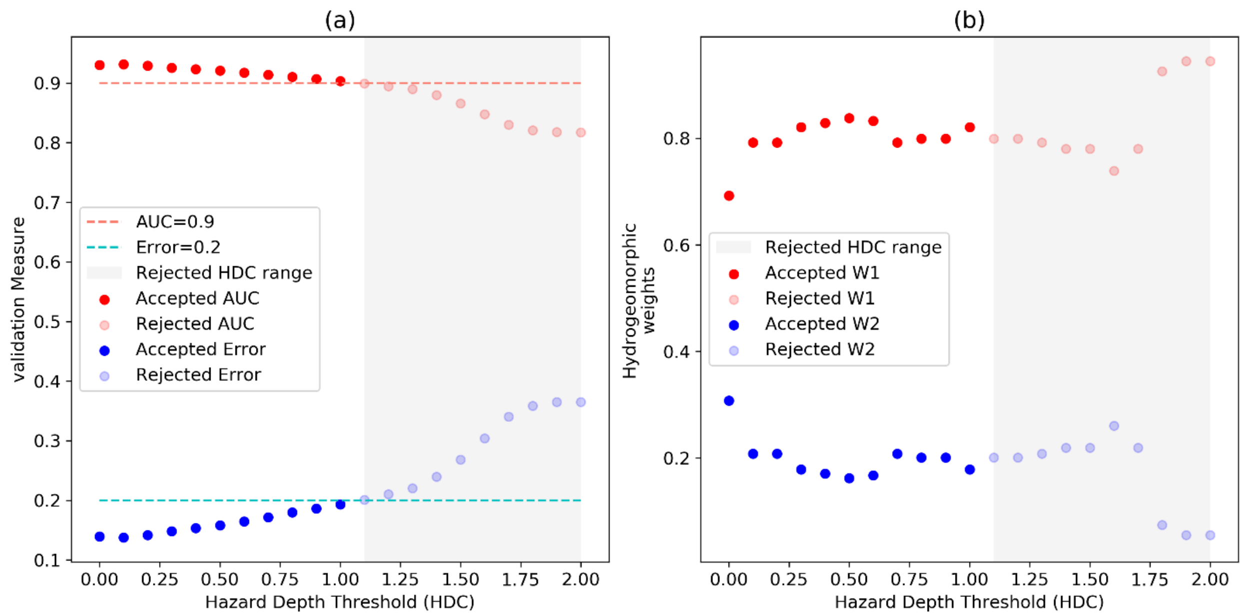

Figure 6. Calibration of I_{HD} ~~Hydrogeomorphic~~ index for Hurricane Matthew. (a) the variation of performance measures AUC (red) and error (blue) for different hazard depth cutoff (HDC) values and (b) the optimum weights of the I_{HD} ~~hydrogeomorphic~~ index for different HDC values. The dash lines show the maximum error (0.2) and minimum AUC (0.9) that are acceptable for flood hazard mapping. Using these criteria, the gray regions show that the hydrogeomorphic model cannot provide acceptable results for HDC values higher than 1.1 m.

To generate the operative curves for future flood events, we design 36 scenarios that include ~~6~~ six HDCs (0, 0.2, 0.4, 0.6, 0.8, 1 m) from the acceptable range of 0-1 m for six different reference hazard maps, provided by the Delft3D-FM model for return periods of 10, 50, 100, 200, 500, and 1000 years. Each scenario provides a reference hazard map, so a binary classification is performed to estimate TH corresponding to each scenario. Figure 7a indicates the error curves for different return period events. For low HDCs, increasing the magnitude of the flood (higher return period) results in more accuracy of the hydrogeomorphic method. This pattern is opposite for high HDCs where flood event with a 10 year return period provides the highest accuracy. In general, the grey region shows that for high HDCs, the performance of the hydrogeomorphic method is poor for almost all return periods while for low HDCs, all flood events can be accurately used for flood hazard mapping. Figure 7b illustrates the hydrogeomorphic threshold operative curves for future flood hazard mapping. The *TH* in the y-axis is the key value that can be estimated for each combination of HDC and return period. Knowing this threshold, Eq. 2 can be used to rapidly estimate the hazard areas for future floods. As expected, a higher magnitude of flood needs a higher hydrogeomorphic threshold while increasing HDC (smaller high-hazard areas) requires a smaller threshold for binary classification. The grey parts of the curves refer to those scenarios that have unacceptable accuracy so it is recommended to not use HDCs corresponding to these parts.

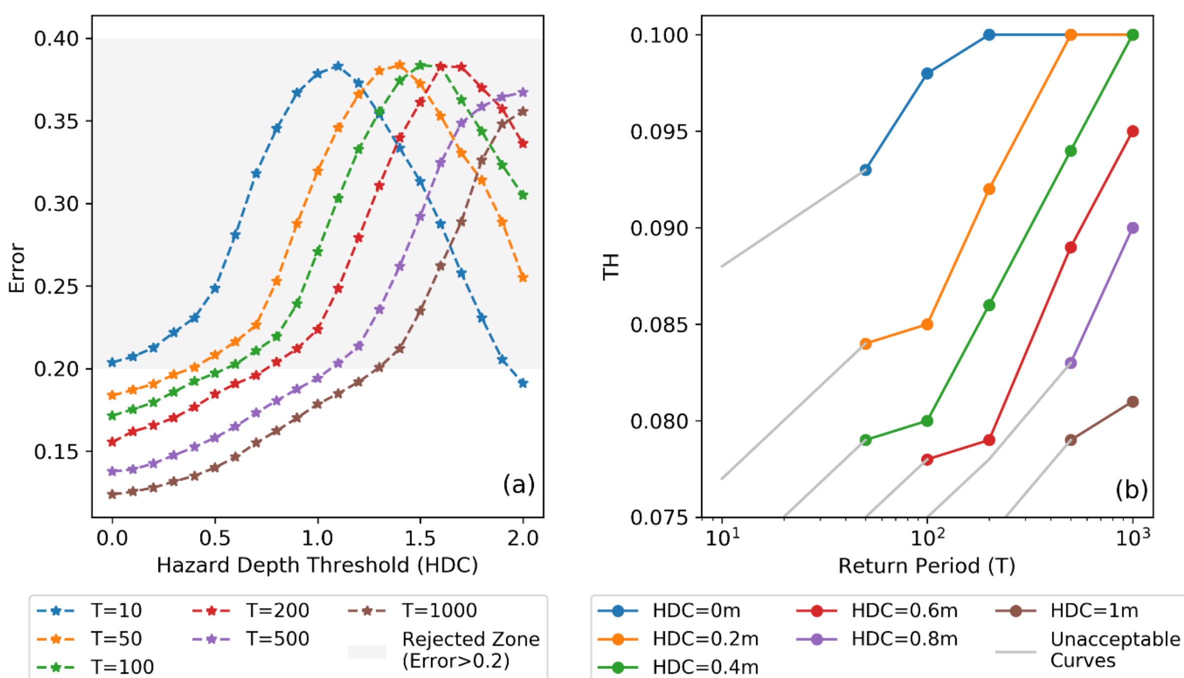

Figure 7. (a) The errors of flood hazard maps generated by the calibrated hydrogeomorphic method for different return period flood events and hazard depth cutoff (HDC) values. (b) The hydrogeomorphic threshold operative curves provided for different HDC values. These operative

curves are the major tool for fast flood hazard mapping as depending on the return period of a future flood event and the HDC value chosen by the decision-maker, the operative curves estimate the hydrogeomorphic threshold. Knowing this threshold, the flood hazard map will be generated in a few minutes.

Finally, we evaluate the accuracy and effectiveness of the proposed operative curves by validating their performance in generating flood hazard areas during Hurricane Irma. The maximum WL during this flood event was 2.49 m ~~that~~ which corresponds to a 223-year flood event according to our flood frequency analysis (e.g., GEV distribution). For two HDCs of 0 and 0.6 m, the operative curves suggest the hydrogeomorphic thresholds of 0.1 and 0.08, respectively. Using these thresholds and Eq.2, the flood hazard maps corresponding to Hurricane Irma can be generated. Figure 8 indicates a side by side comparison of flood hazard maps generated by the Delft3D-FM model (Figures 8a and 8c) and the hydrogeomorphic threshold operative curves (Figures 8b and 8d) for two different HDCs of 0 (Figures 8a and 8b) and 0.6 m (Figures 8c and 8d). For both HDCs, errors (0.152 and 0.186) are less than a 0.2 limit used for reliable flood hazard mapping. The main ~~errors of the hydrogeomorphic method~~discrepancies are some noisy scattered low-hazard areas located in the east and southeast of the study area. These areas can reflect the flooded surface depressions (sinks) resulting from the pluvial impacts of extreme precipitation. Hydrodynamic models simulate the fluvial and coastal processes that occur adjacent to rivers and oceans while disregarding the pluvial impacts. The red circle in the left part of the figures ~~also~~ shows a region that the hydrogeomorphic method cannot properly simulate, especially for higher HDCs. This can be due to the inability of the hydrogeomorphic method to properly simulate physical processes. On the other hand, the red eclipse at the right side of the figures illustrates an urbanized region where the hydrogeomorphic method properly classifies the area compared to the reference map.

Overall, the high overlap of the flood hazard maps provided by the hydrogeomorphic method with the reference maps provided by the hydrodynamic model (error <0.2) illustrates the reliability and effectiveness of the proposed hydrogeomorphic method for flood hazard mapping. Besides, the high efficiency of this approach for rapid estimation of flood hazard maps (order of minutes) compared to the long computational time required for detailed hydrodynamic modeling (order of hours) suggests the proposed hydrogeomorphic method as an alternative for efficient flood hazard mapping during emergencies.

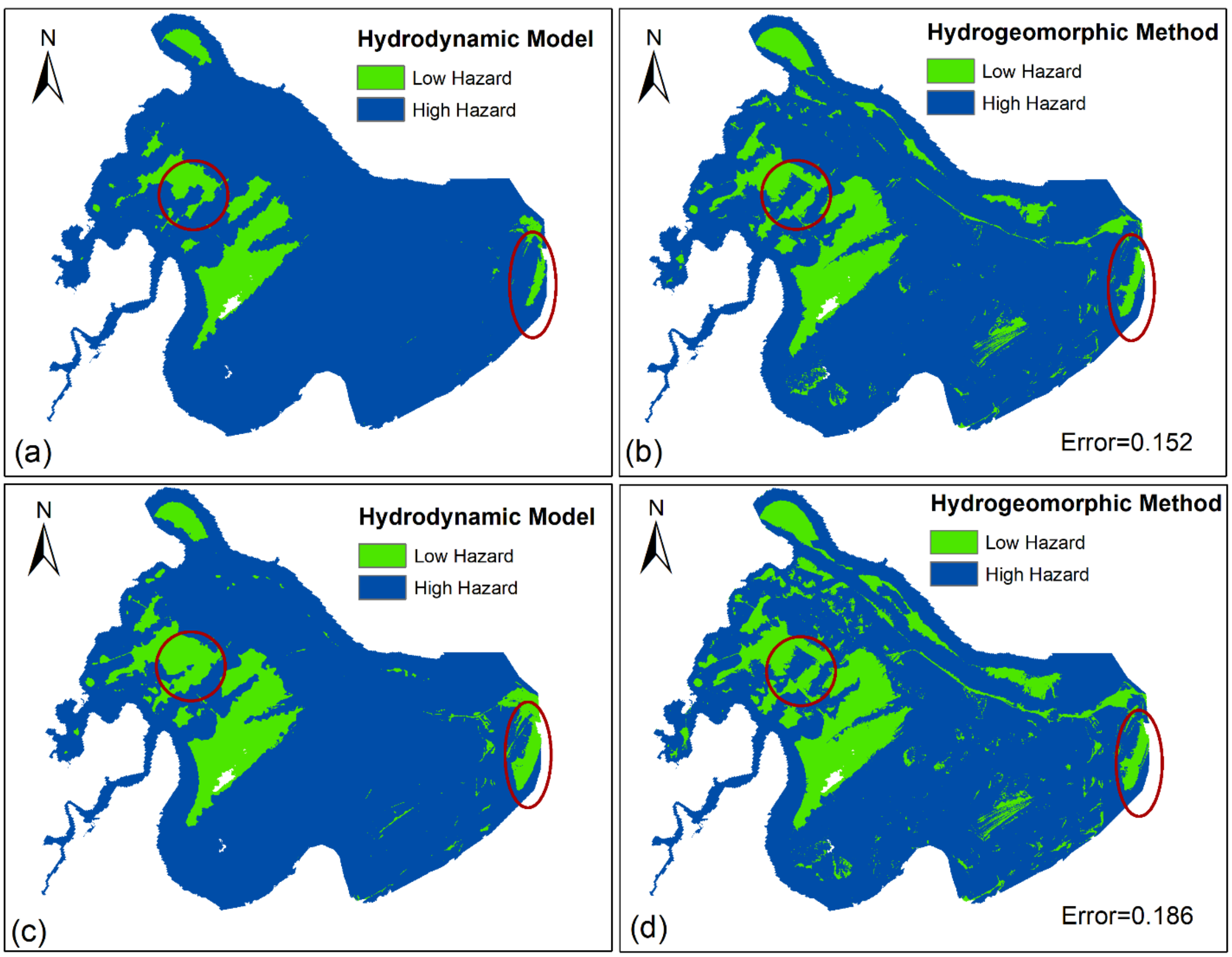

Figure 8. Validation results for Hurricane Irma showing a side-by-side comparison of flood hazard maps generated by the hydrodynamic model and hydrogeomorphic method for two

different hazard depth cutoffs (HDCs), HDC=0 (a, b) and HDC=0.6 m (c, d). To generate the flood hazard maps by the hydrogeomorphic method, the operative curves estimate two hydrogeomorphic thresholds of 0.1 and 0.08 for HDC= 0 m and HDC= 0.6 m, respectively while the return period of Hurricane Irma is estimated as a 223 years flood event.

5. Discussion

This study develops hydrogeomorphic threshold operative curves for rapid estimation of hazardous areas during emergencies of future coastal floods in deltas and estuaries. The low errors (<0.2) of estimated hazard maps for Hurricane Irma generated by the proposed approach compared to the reference hydrodynamic model results demonstrate the high accuracy of the proposed operative curves for future flood events in this region. According to studies conducted on the binary classification of hydrogeomorphic features in the literature, the errors of the best classifiers were mostly in the range of 0.2-0.3 for inland floods (Degiorgis et al., 2012; Manfreda et al., 2014). Therefore, given the more complexity of terrain and drainage network in deltas, predicting the hazard maps with errors less than 0.2 (e.g. error of 0.152 for HDC=0) is a promising achievement. The potential reasons explaining a high accuracy of the proposed binary classifier for wetlands include the high-resolution DEM used for mapping (~3m)~~,~~ and the incorporation of bathymetry into DEM. In addition, the flexible structure of the proposed hydrogeomorphic index, with two varying weights of H and D features, allows for calibrating the index with the optimum contribution of each feature, which in return results in the highest accuracy.

The proposed hydrogeomorphic index (I_{HD}) is the primary data for flood hazard mapping in this study. Thus, the quality of two main inputs of this index, namely the DEM and stream network used to calculate features H and D play a vital role in the overall accuracy of the proposed approach. To obtain maximum accuracy, here we used the best available DEM with the finest spatial resolution of 3 m that includes the bathymetry data. However, considering the limited access to such high-quality DEMs in many areas of the world, it is recommended to evaluate the sensitivity of the proposed approach to lower quality DEMs (e.g. 30 m and 90 m DEMs without bathymetry information) in future studies. Another piece of research can investigate the sensitivity of the proposed approach to the density of the drainage network used for calculating the I_{HD} index.

Unlike past studies that used binary classifiers for detecting hazard areas corresponding to past floods or generated static maps for a specific return period ~~event~~ (Degiorgis et al., 2012; Jafarzadegan et al., 2018; Manfreda et al., 2015b; Samela et al., 2017), here we propose the hydrogeomorphic threshold operative curves for real-time flood hazard mapping. Considering the rapid occurrence of hurricane-induced flooding in deltas and estuaries, these curves can be highly beneficial for emergency responders to provide a preliminary estimation of hazard areas for an upcoming flood in these regions and design the appropriate evacuation strategies. In addition, the proposed operative curves demonstrate the hydrogeomorphic threshold variations with HDCs. This feature of the operative curves gives additional flexibility to decision-makers for estimating the hazard maps based on the HDC that is considered given the momentary safety issues. For example, identifying the hazard map based on HDC<0.3 is useful for checking the operability and accessibility of essential facilities and infrastructure, while a hazard map corresponding to HDC=1 indicates those areas that experience high WLs above 1 m as hazardous areas, with greater potential for casualties and significant structural damage. Overall, the hydrogeomorphic threshold operative curves are a function of both the return period (flood severity) and HDC (a decision-making option that controls the definition of high hazard). Using a similar approach, future studies can provide these curves for inland floods as well. In addition, due to the practical benefits of these curves for efficient coastal flood hazard assessment, the hydrogeomorphic threshold operative curves can be extended to other deltas and estuaries that experience frequent flooding across the US (e.g., Mississippi ~~-~~– Louisiana (LA), Galveston Bay ~~-~~– Texas (TX), Delaware Bay ~~-~~– Delaware (DE), Chesapeake Bay ~~-~~– Virginia (VA), among others) and the world (e.g. Yangtze - China, Brahmaputra - Bangladesh, among others). To implement this approach, first, a hydrodynamic model should be set up for the new study area and generate reference inundation maps for different return periods. Access to observed water level data (gauges or HWMs) and flood extent maps from past floods is required to properly calibrate the hydrodynamic model. Then the I_{HD} index calculated from a DEM is utilized together with the reference maps to provide the hydrogeomorphic threshold operative curves for future floods.

The reference maps used for training the binary classifier are key components for generating reliable results. Since these reference maps are the outcomes of hydrodynamic modeling, they are prone to uncertainties stemming from unrealistic parametrization, imperfect model structure, and erroneous forcing. The design floods used as boundary conditions of the hydrodynamic model are estimated from flood frequency analysis that is prone to uncertainty as well. Here, we used a bivariate approach that estimates the design flood based on the water level data. A more comprehensive flood frequency analysis that accounts for other flood attributes, such as volume, spatial dependencies, or nonstationarity~~–~~ can improve the reliability of flood frequency analysis in future studies (Brunner et al., 2016; Yan and Moradkhani, 2015; Bracken et al., 2018). With access to less than 100 years of data for flood frequency analysis, the extreme return levels (i.e. ~~500 and 1000 year~~500- and 1000-year floods) pose high uncertainties due to the extrapolation of annual maxima data. This should warn decision-makers to be more cautious about using operative curves for extreme flood events. For future studies, the uncertainty bounds of flood frequency analysis (especially extrapolations for extreme cases) can be considered in the modeling. In a real-time scenario, the forecasted WL used for flood frequency analysis is also prone to uncertainties originating from imperfect forecasting methods and nonstationary climate data. In addition, the uncertainty of model parametrization can be accounted for by running the hydrodynamic model for different combinations of optimum parameters. Model structure uncertainty can be also considered by using different hydrodynamic models and combining the results. Finally, probabilistic reference maps together with uncertainties involved in WL forecasting and flood frequency analysis can be integrated to develop probabilistic hydrogeomorphic threshold operative curves in future studies. This is in line with the report provided for the NOAA National Weather Service (NWS), showing the NWS stakeholder's preference for utilizing probabilistic storm surge inundation maps ~~in the future~~ (Eastern Research Group, Inc, 2013).

Operationally, ~~The probabilistic operative curves account for the major source of uncertainties and provide a more reliable decision-making tool for coastal flood hazard mapping.~~

~~The operative hydrogeomorphic threshold classifiers proposed for real-time coastal flood hazard mapping can be used as an alternative tool for the rapid estimation of hazardous areas. In the operational mode, the water level forecasts provided by the NWS can be used to estimate the return period of an upcoming coastal flood event.~~ the Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model (Jelesnianski et al., 1984) is the storm surge model currently used by NWS to perform storm surge forecasting and create probabilistic flood inundation maps for real-time tropical storms (Sea, Lake, and Overland Surges from Hurricanes (SLOSH), 2022). The feature of SLOSH that makes it the preferred model of the NWS for storm surge forecasting and mapping is the model's computational efficiency that allows the model to be ru~~a~~n as an ensemble (Forbes et al., 2014). However, SLOSH is just one of several modeling options for storm surge modeling and mapping, each possessing strengths and weaknesses associated with their simulations. The inclusion of additional models that can create flood maps of storm surges for a given event should provide an enhanced understanding of the uncertainty of inundation at a given location (Teng et al., 2015). However, the higher computational burden of alternative models, such as Delft3D-FM, tend to preclude their use in real-time operations and certainly, their use in generating an ensemble necessary for probabilistic flood maps. The methodology we propose in this manuscript may offer the NWS and other agencies a means to utilize alternatives to SLOSH for flood inundation mapping and probabilistic flood inundation mapping on U.S. coastlines. Models such as Delft3D-FM can generate reference maps to train the binary classifier and build the probabilistic operating curves. ~~Using the proposed operative curves, the hydrogeomorphic threshold is determined and t~~ The probabilistic operative curves would account for the major source of uncertainties and provide a computationally efficient and reliable decision-making tool for coastal planners and floodplain managers. The operative hydrogeomorphic threshold classifiers proposed for real-time coastal flood hazard mapping can be used as an alternative tool for the rapid estimation of hazardous areas during real-time flood events. In an operational mode, water level or meteorological forecasts can be used to estimate the return period of an upcoming coastal flood event and the methodology here can utilize this as an input to perform LCFM flood inundation mapping both deterministically and probabilistically.

~~he flood hazard map is generated. The Sea, Lake, and Overland Surges from Hurricanes (SLOSH) model is an LCFM tool currently used by NWS to estimate probabilistic storm surge forecasts. The flood inundation maps generated by this model are the results of overlaying storm surge forecast with DEM. The model doesn't consider the streamflow network and riverine flood mechanisms. On the other hand, our proposed hydrogeomorphic index uses both streamflow network and DEM to provide a more detailed representation of the flooding in coastal areas.~~

Another LCFM approach is to train machine learning algorithms on reference inundation maps provided by well-calibrated hydrodynamic models (Bass and Bedient, 2018). A benchmark study that compares the performance (accuracy and efficiency) of ~~three~~ two LCFM methods, including our proposed DEM-based hydrogeomorphic classifier~~,~~ and the surrogate machine learning-based algorithm, and the SLOSH model is highly recommended for future studies.

~~The probabilistic operative curves account for the major source of uncertainties and provide a more reliable decision-making tool for coastal flood hazard mapping. The operative hydrogeomorphic threshold classifiers proposed for real-time coastal flood hazard mapping can be used as an alternative tool for the rapid estimation of hazardous areas. In the operational mode, the water level forecasts provided by the NWS can be used to estimate the return period of an upcoming coastal flood event.~~

6. Summary and Conclusions

In this study, we proposed binary classifiers for efficient flood hazard mapping in deltas and estuaries. The HAND, typically used for modeling inland ~~flooding~~floods, is modified for flat regions along the coastline, and a new hydrogeomorphic index (I_{HD}) that comprises both HAND and distance to nearest drainage was developed. The DEM used as the base of these binary classifiers is a 3 m Lidar that includes bathymetric information. This is another improvement compared to previous DEM-based classifiers that commonly used 10-30 m DEMs without bathymetric data. The I_{HD} index has two unknown weights that show the contribution of both HAND and feature *D*. We simulated Hurricane Matthew with the Delft3D-FM model and used the results as a reference flood hazard map to calibrate the I_{HD} index. Using Delft3D-FM again, we generated six flood hazard maps corresponding to different return periods and employed these maps as a reference to generate the hydrogeomorphic threshold operative curves. Finally, we validated the proposed operative curves for reliable and efficient flood hazard mapping by comparing the flood hazard maps generated for Hurricane Irma with the proposed curves and the Delft3D-FM model. The high accuracy of validation results (<0.2 error) together with the ~~rapid fashion~~computational efficiency of this approach for real-time flood hazard mapping suggests the proposed operative curves as a practical decision-making tool for on-time and reliable estimation of hazard areas in estuaries.

Data availability

All the data used in this study, including the gauge streamflow and water stage data are publicly available from the USGS and NOAA websites. The High Water Marks provided for Hurricanes Irma and Matthew are available from the USGS Flood Event Viewer platform.

Author contribution

KJafarzadegan and HMoradkhani conceptualized the study. KJafarzadegan designed the whole framework and implemented the hydrogeomorphic methodology. DMuñoz implemented the hydrodynamic model~~ing~~. KJafarzadegan and DMuñoz wrote the first draft of the manuscript. HMoradkhani, HMoftakhari,~~ a~~ JGutenson, and GSavant provided comments and edited the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgment

This work was financially supported by the USACE award # A20-0545-001. We would like to thank the anonymous reviewers for their constructive comments on the original version of the manuscript.

Supplementary Material

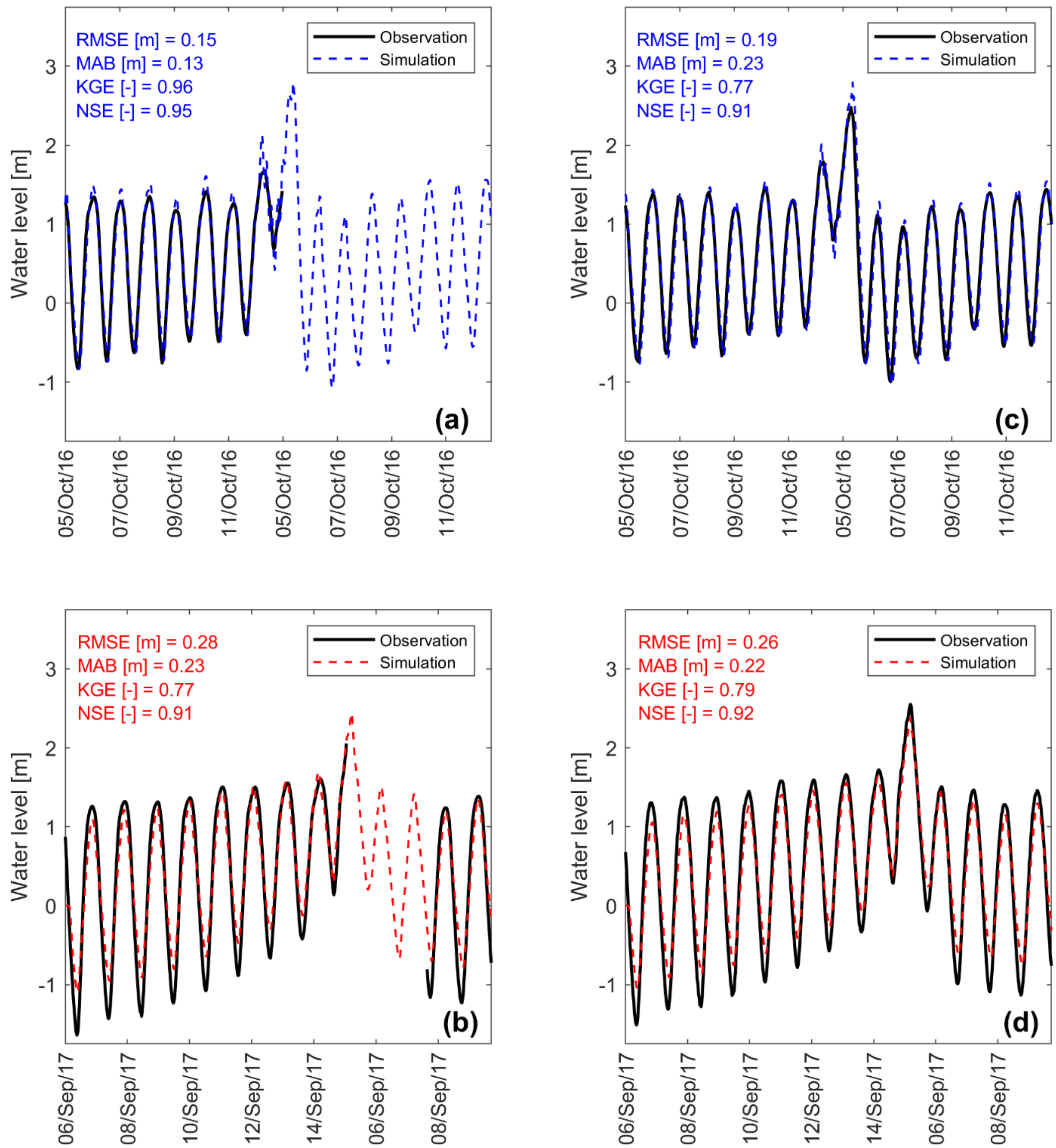

Figure S1. Calibration and validation of the Savannah Delft3D-FM model. Time series of simulated and observed WLs at (a, b) Port Wentworth and (c, d) Savannah River at USACE dock. Top and bottom panel show times series of Hurricane Matthew and Irma, respectively. Note that the model can simulate water level variability, and so fill the data gaps observed during the Hurricanes.

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, J. Hydrol., 556, 539–556, https://doi.org/10.1016/j.jhydrol.2017.11.036, 2018.

Alizad, K., Hagen, S. C., Medeiros, S. C., Bilskie, M. V., Morris, J. T., Balthis, L., and Buckel, C. A.: Dynamic responses and implications to coastal wetlands and the surrounding regions under sea level rise, PLOS ONE, 13, e0205176, https://doi.org/10.1371/journal.pone.0205176, 2018.

Sea, Lake, and Overland Surges from Hurricanes (SLOSH): https://www.nhc.noaa.gov/surge/slosh.php, last access: 18 January 2022.

USGS Surface Water Information: https://water.usgs.gov/osw/iwrss/, last access: 16 November 2021.

Arcement, G. J. and Schneider, V. R.: Guide for selecting Manning's roughness coefficients for natural channels and flood plains, 1989.

Barbier, E. B.: Chapter 27 - The Value of Coastal Wetland Ecosystem Services, in: Coastal Wetlands, edited by: Perillo, G. M. E., Wolanski, E., Cahoon, D. R., and Hopkinson, C. S., Elsevier, 947–964, https://doi.org/10.1016/B978-0-444-63893-9.00027-7, 2019.

Bass, B. and Bedient, P.: Surrogate modeling of joint flood risk across coastal watersheds, J. Hydrol., 558, 159–173, https://doi.org/10.1016/j.jhydrol.2018.01.014, 2018.

Bates, P. D., Horritt, M. S., and Fewtrell, T. J.: A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling, J. Hydrol., 387, 33–45, https://doi.org/10.1016/j.jhydrol.2010.03.027, 2010.

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., Savage, J., Olcese, G., Neal, J., Schumann, G., Giustarini, L., Coxon, G., Porter, J. R., Amodeo, M. F., Chu, Z., Lewis-Gruss, S., Freeman, N. B., Houser, T., Delgado, M., Hamidi, A., Bolliger, I., E. McCusker, K., Emanuel, K., Ferreira, C. M., Khalid, A., Haigh, I. D., Couasnon, A., E. Kopp, R., Hsiang, S., and Krajewski, W. F.: Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates, Water Resour. Res., 57, e2020WR028673, https://doi.org/10.1029/2020WR028673, 2021.

Bracken, C., Holman, K. D., Rajagopalan, B., and Moradkhani, H.: A Bayesian hierarchical approach to multivariate nonstationary hydrologic frequency analysis, Water Resour. Res., 54, 243–255, 2018.

Brunner, M. I., Seibert, J., and Favre, A.-C.: Bivariate return periods and their importance for flood peak and volume estimation, WIREs Water, 3, 819–833, https://doi.org/10.1002/wat2.1173, 2016.

Carlston, C. W.: Longitudinal Slope Characteristics of Rivers of the Midcontinent and the Atlantic East Gulf Slopes, Int. Assoc. Sci. Hydrol. Bull., 14, 21–31, https://doi.org/10.1080/02626666909493751, 1969.

Chow Ven, T.: Open channel hydraulics, 1959.

Davidson, N. C.: How much wetland has the world lost? Long-term and recent trends in global wetland area, Mar. Freshw. Res., 65, 934–941, https://doi.org/10.1071/MF14173, 2014.

Degiorgis, M., Gnecco, G., Gorni, S., Roth, G., Sanguineti, M., and Taramasso, A. C.: Classifiers for the detection of flood-prone areas using remote sensed elevation data, J. Hydrol., 470–471, 302–315, https://doi.org/10.1016/j.jhydrol.2012.09.006, 2012.

Degiorgis, M., Gnecco, G., Gorni, S., Roth, G., Sanguineti, M., and Taramasso, A. C.: Flood Hazard Assessment Via Threshold Binary Classifiers: Case Study of the Tanaro River Basin, Irrig. Drain., 62, 1–10, https://doi.org/10.1002/ird.1806, 2013.

Delft3D Flexible Mesh Suite - Deltares: https://www.deltares.nl/en/software/delft3d-flexible-mesh-suite/, last access: 15 November 2021.

Dodov, B. A. and Foufoula-Georgiou, E.: Floodplain morphometry extraction from a high-resolution digital elevation model: a simple algorithm for regional analysis studies, IEEE Geosci. Remote Sens. Lett., 3, 410–413, https://doi.org/10.1109/LGRS.2006.874161, 2006.

Eastern Research Group, Inc: Hurricane Forecast Improvment Program Socio-Economic Research and Recommendations:, NOAA National Weather Service, 2013.

Fagherazzi, S., Mariotti, G., Banks, A. T., Morgan, E. J., and Fulweiler, R. W.: The relationships among hydrodynamics, sediment distribution, and chlorophyll in a mesotidal estuary, Estuar. Coast. Shelf Sci., 144, 54–64, https://doi.org/10.1016/j.ecss.2014.04.003, 2014.

Familkhalili, R., Talke, S. A., and Jay, D. A.: Tide-Storm Surge Interactions in Highly Altered Estuaries: How Channel Deepening Increases Surge Vulnerability, J. Geophys. Res. Oceans, 125, e2019JC015286, https://doi.org/10.1029/2019JC015286, 2020.

Fawcett, T.: An introduction to ROC analysis, Pattern Recognit. Lett., 27, 861–874, https://doi.org/10.1016/j.patrec.2005.10.010, 2006.

Forbes, C., Rhome, J., Mattocks, C., and Taylor, A.: Predicting the Storm Surge Threat of Hurricane Sandy with the National Weather Service SLOSH Model, J. Mar. Sci. Eng., 2, 437–476, https://doi.org/10.3390/jmse2020437, 2014.

Ghanbari, M., Arabi, M., Kao, S.-C., Obeysekera, J., and Sweet, W.: Climate Change and Changes in Compound Coastal-Riverine Flooding Hazard Along the U.S. Coasts, Earths Future, 9, e2021EF002055, https://doi.org/10.1029/2021EF002055, 2021.

Gharari, S., Hrachowitz, M., Fenicia, F., and Savenije, H. H. G.: Hydrological landscape classification: investigating the performance of HAND based landscape classifications in a central European meso-scale catchment, Hydrol. Earth Syst. Sci., 15, 3275–3291, https://doi.org/10.5194/hess-15-3275-2011, 2011.

Gutenson, J.: A Review of Current and Future NWS Services, 2020.

Gutenson, J. L., Tavakoly, A. A., Massey, T. C., Savant, G., Tritinger, A. S., Owensby, M. B., Wahl, M. D., and Islam, M. S.: Investigating Modeling Strategies to Couple Inland Hydrology and Coastal Hydraulics to Better Understand Compound Flood Risk, 64–75, https://doi.org/10.1061/9780784483466.006, 2021.

Helton, J. C. and Davis, F. J.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliab. Eng. Syst. Saf., 81, 23–69, 2003.

IWRSS: Requirements for the National Flood Inundation Mapping Services, National Oceanic and Atmospheric Administration United States Army Corps of Engineers United States Geological Survey, 2013.

IWRSS: Design for the National Flood Inundation Mapping Services, National Oceanic and Atmospheric Administration United States Army Corps of Engineers United States Geological Survey, 2015.

Jafarzadegan, K. and Merwade, V.: A DEM-based approach for large-scale floodplain mapping in ungauged watersheds, J. Hydrol., 550, 650–662, https://doi.org/10.1016/j.jhydrol.2017.04.053, 2017.

Jafarzadegan, K. and Merwade, V.: Probabilistic floodplain mapping using HAND-based statistical approach, Geomorphology, 324, 48–61, https://doi.org/10.1016/j.geomorph.2018.09.024, 2019.

Jafarzadegan, K., Merwade, V., and Saksena, S.: A geomorphic approach to 100-year floodplain mapping for the Conterminous United States, J. Hydrol., 561, 43–58, https://doi.org/10.1016/j.jhydrol.2018.03.061, 2018.

Jafarzadegan, K., Merwade, V., and Moradkhani, H.: Combining clustering and classification for the regionalization of environmental model parameters: Application to floodplain mapping in data-scarce regions, Environ. Model. Softw., 125, 104613, https://doi.org/10.1016/j.envsoft.2019.104613, 2020.

Jafarzadegan, K., Abbaszadeh, P., and Moradkhani, H.: Sequential Data Assimilation for Real-Time Probabilistic Flood Inundation Mapping, Hydrol. Earth Syst. Sci. Discuss., 1–39, https://doi.org/10.5194/hess-2021-181, 2021.

Jelesnianski, C., Chen, J., Shaffer, W., and Gilad, A.: SLOSH - A Hurricane Storm Surge Forecast Model, in: OCEANS 1984, OCEANS 1984, 314–317, https://doi.org/10.1109/OCEANS.1984.1152341, 1984.

Khojasteh, D., Chen, S., Felder, S., Heimhuber, V., and Glamore, W.: Estuarine tidal range dynamics under rising sea levels, PLOS ONE, 16, e0257538, https://doi.org/10.1371/journal.pone.0257538, 2021a.

Khojasteh, D., Glamore, W., Heimhuber, V., and Felder, S.: Sea level rise impacts on estuarine dynamics: A review, Sci. Total Environ., 780, 146470, https://doi.org/10.1016/j.scitotenv.2021.146470, 2021b.

Kirwan, M. L. and Megonigal, J. P.: Tidal wetland stability in the face of human impacts and sea-level rise, Nature, 504, 53–60, https://doi.org/10.1038/nature12856, 2013.

Kulp, S. A. and Strauss, B. H.: New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding, Nat. Commun., 10, 4844, https://doi.org/10.1038/s41467-019-12808-z, 2019.

Kumbier, K., Carvalho, R. C., Vafeidis, A. T., and Woodroffe, C. D.: Investigating compound flooding in an estuary using hydrodynamic modelling: a case study from the Shoalhaven River, Australia, Nat. Hazards Earth Syst. Sci., 18, 463–477, https://doi.org/10.5194/nhess-18-463-2018, 2018.

Land, M., Tonderski, K., and Verhoeven, J. T. A.: Wetlands as Biogeochemical Hotspots Affecting Water Quality in Catchments, in: Wetlands: Ecosystem Services, Restoration and Wise Use, edited by: An, S. and Verhoeven, J. T. A., Springer International Publishing, Cham, 13–37, https://doi.org/10.1007/978-3-030-14861-4_2, 2019.

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, J. Flood Risk Manag., 12, e12347, 2019.

Longenecker, H. E., Graeden, E., Kluskiewicz, D., Zuzak, C., Rozelle, J., and Aziz, A. L.: A rapid flood risk assessment method for response operations and nonsubject-matter-expert community planning, J. Flood Risk Manag., 13, e12579, https://doi.org/10.1111/jfr3.12579, 2020.

Luettich, R. A. (Richard A., Westerink, J. J., and Scheffner, N. W.: ADCIRC : an advanced three-dimensional circulation model for shelves, coasts, and estuaries. Report 1, Theory and methodology of ADCIRC-2DD1 and ADCIRC-3DL, This Digit. Resour. Was Creat. Scans Print Resour., 1992.

Maidment, D. R.: Conceptual Framework for the National Flood Interoperability Experiment, JAWRA J. Am. Water Resour. Assoc., 53, 245–257, https://doi.org/10.1111/1752-1688.12474, 2017.

Maidment, D. R., Clark, E., Hooper, R., and Ernest, A.: National Flood Interoperability Experiment, in: AGU Fall Meeting Abstracts, 2014.

Manfreda, S., Di Leo, M., and Sole, A.: Detection of Flood-Prone Areas Using Digital Elevation Models, J. Hydrol. Eng., 16, 781–790, https://doi.org/10.1061/(ASCE)HE.1943-5584.0000367, 2011.

Manfreda, S., Nardi, F., Samela, C., Grimaldi, S., Taramasso, A. C., Roth, G., and Sole, A.: Investigation on the use of geomorphic approaches for the delineation of flood prone areas, J. Hydrol., 517, 863–876, https://doi.org/10.1016/j.jhydrol.2014.06.009, 2014.

Manfreda, S., Samela, C., Gioia, A., Consoli, G. G., Iacobellis, V., Giuzio, L., Cantisani, A., and Sole, A.: Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1D and 2D hydraulic models, Nat. Hazards, 79, 735–754, https://doi.org/10.1007/s11069-015-1869-5, 2015a.

Manfreda, S., Samela, C., Gioia, A., Consoli, G. G., Iacobellis, V., Giuzio, L., Cantisani, A., and Sole, A.: Flood-prone areas assessment using linear binary classifiers based on flood maps obtained from 1D and 2D hydraulic models, Nat. Hazards, 79, 735–754, https://doi.org/10.1007/s11069-015-1869-5, 2015b.

McGlynn, B. L. and McDonnell, J. J.: Quantifying the relative contributions of riparian and hillslope zones to catchment runoff, Water Resour. Res., 39, https://doi.org/10.1029/2003WR002091, 2003.

McGlynn, B. L. and Seibert, J.: Distributed assessment of contributing area and riparian buffering along stream networks, Water Resour. Res., 39, https://doi.org/10.1029/2002WR001521, 2003.

McGrath, H., Bourgon, J.-F., Proulx-Bourque, J.-S., Nastev, M., and Abo El Ezz, A.: A comparison of simplified conceptual models for rapid web-based flood inundation mapping, Nat. Hazards, 93, 905–920, https://doi.org/10.1007/s11069-018-3331-y, 2018.

Medeiros, S., Hagen, S., Weishampel, J., and Angelo, J.: Adjusting Lidar-Derived Digital Terrain Models in Coastal Marshes Based on Estimated Aboveground Biomass Density, Remote Sens., 7, 3507–3525, https://doi.org/10.3390/rs70403507, 2015.

Morton, R. A. and Barras, J. A.: Hurricane Impacts on Coastal Wetlands: A Half-Century Record of Storm-Generated Features from Southern Louisiana, J. Coast. Res., 27, 27–43, https://doi.org/10.2112/JCOASTRES-D-10-00185.1, 2011.

Muis, S., Lin, N., Verlaan, M., Winsemius, H. C., Ward, P. J., and Aerts, J. C. J. H.: Spatiotemporal patterns of extreme sea levels along the western North-Atlantic coasts, Sci. Rep., 9, 3391, https://doi.org/10.1038/s41598-019-40157-w, 2019.

Muñoz, D. F., Cissell, J. R., and Moftakhari, H.: Adjusting Emergent Herbaceous Wetland Elevation with Object-Based Image Analysis, Random Forest and the 2016 NLCD, Remote Sens., 11, 2346, https://doi.org/10.3390/rs11202346, 2019.

Muñoz, D. F., Moftakhari, H., and Moradkhani, H.: Compound effects of flood drivers and wetland elevation correction on coastal flood hazard assessment, Water Resour. Res., 56, e2020WR027544, 2020.

Muñoz, D. F., Muñoz, P., Moftakhari, H., and Moradkhani, H.: From local to regional compound flood mapping with deep learning and data fusion techniques, Sci. Total Environ., 782, 146927, 2021.

Muñoz, D. F., Abbaszadeh, P., Moftakhari, H., and Moradkhani, H.: Accounting for uncertainties in compound flood hazard assessment: The value of data assimilation, Coast. Eng., 171, 104057, https://doi.org/10.1016/j.coastaleng.2021.104057, 2022.

Nardi, F., Vivoni, E. R., and Grimaldi, S.: Investigating a floodplain scaling relation using a hydrogeomorphic delineation method, Water Resour. Res., 42, https://doi.org/10.1029/2005WR004155, 2006.

Roelvink, J. A. and Banning, G. K. F. M. V.: Design and development of DELFT3D and application to coastal morphodynamics, Oceanogr. Lit. Rev., 11, 925, 1995.

Rogers, J. N., Parrish, C. E., Ward, L. G., and Burdick, D. M.: Improving salt marsh digital elevation model accuracy with full-waveform lidar and nonparametric predictive modeling, Estuar. Coast. Shelf Sci., 202, 193–211, https://doi.org/10.1016/j.ecss.2017.11.034, 2018.

Samela, C., Manfreda, S., Paola, F. D., Giugni, M., Sole, A., and Fiorentino, M.: DEM-Based Approaches for the Delineation of Flood-Prone Areas in an Ungauged Basin in Africa, J. Hydrol. Eng., 21, 06015010, https://doi.org/10.1061/(ASCE)HE.1943-5584.0001272, 2016.

Samela, C., Troy, T. J., and Manfreda, S.: Geomorphic classifiers for flood-prone areas delineation for data-scarce environments, Adv. Water Resour., 102, 13–28, https://doi.org/10.1016/j.advwatres.2017.01.007, 2017.

Schieder, N. W., Walters, D. C., and Kirwan, M. L.: Massive Upland to Wetland Conversion Compensated for Historical Marsh Loss in Chesapeake Bay, USA, Estuaries Coasts, 41, 940–951, https://doi.org/10.1007/s12237-017-0336-9, 2018.

Sullivan, J. C., Torres, R., and Garrett, A.: Intertidal Creeks and Overmarsh Circulation in a Small Salt Marsh Basin, J. Geophys. Res. Earth Surf., 124, 447–463, https://doi.org/10.1029/2018JF004861, 2019.

Teng, J., Vaze, J., Dutta, D., and Marvanek, S.: Rapid Inundation Modelling in Large Floodplains Using LiDAR DEM, Water Resour. Manag., 29, 2619–2636, https://doi.org/10.1007/s11269-015-0960-8, 2015.

Thomas, A., Dietrich, J., Asher, T., Bell, M., Blanton, B., Copeland, J., Cox, A., Dawson, C., Fleming, J., and Luettich, R.: Influence of storm timing and forward speed on tides and storm surge during Hurricane Matthew, Ocean Model., 137, 1–19, https://doi.org/10.1016/j.ocemod.2019.03.004, 2019.

U.S. Army Corps of Engineers: Current Channel Condition Survey Reports and Charts. Savannah Harbor., 2017.

Wamsley, T. V., Cialone, M. A., Smith, J. M., Atkinson, J. H., and Rosati, J. D.: The potential of wetlands in reducing storm surge, Ocean Eng., 37, 59–68, https://doi.org/10.1016/j.oceaneng.2009.07.018, 2010.

Williams, W. A., Jensen, M. E., Winne, J. C., and Redmond, R. L.: An Automated Technique for Delineating and Characterizing Valley-Bottom Settings, in: Monitoring Ecological Condition in the Western United States: Proceedings of the Fourth Symposium on the Environmental Monitoring and Assessment Program (EMAP), San Francisco, CA, April 6–8, 1999, edited by: Sandhu, S. S., Melzian, B. D., Long, E. R., Whitford, W. G., and Walton, B. T., Springer Netherlands, Dordrecht, 105–114, https://doi.org/10.1007/978-94-011-4343-1_10, 2000.

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, J. Hydrol. X, 4, 100039, https://doi.org/10.1016/j.hydroa.2019.100039, 2019.

Wu, W., Zhou, Y., and Tian, B.: Coastal wetlands facing climate change and anthropogenic activities: A remote sensing analysis and modelling application, Ocean Coast. Manag., 138, 1–10, https://doi.org/10.1016/j.ocecoaman.2017.01.005, 2017.

Yan, H. and Moradkhani, H.: A regional Bayesian hierarchical model for flood frequency analysis, Stoch. Environ. Res. Risk Assess., 29, 1019–1036, 2015.

Zheng, X., Maidment, D. R., Tarboton, D. G., Liu, Y. Y., and Passalacqua, P.: GeoFlood: Large-Scale Flood Inundation Mapping Based on High-Resolution Terrain Analysis, Water Resour. Res., 54, 10,013-10,033, https://doi.org/10.1029/2018WR023457, 2018a.

Zheng, X., Tarboton, D. G., Maidment, D. R., Liu, Y. Y., and Passalacqua, P.: River Channel Geometry and Rating Curve Estimation Using Height above the Nearest Drainage, JAWRA J. Am. Water Resour. Assoc., 54, 785–806, https://doi.org/10.1111/1752-1688.12661, 2018b.