AGU ADVANCING EARTH AND SPACE SCIENCES

Water Resources Research

RESEARCH ARTICLE
10.1029/2024WR038032

Uncertainties in Simulating Flooding During Hurricane Harvey Using 2D Shallow Water Equations

Donghui Xu[1], Gautam Bisht[1], Darren Engwirda[2], Dongyu Feng[1], Zeli Tan[1], and Valeriy Y. Ivanov[3]

[1]Atmospheric, Climate, & Earth Sciences Division, Pacific Northwest National Laboratory, Richland, WA, USA, [2]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM, USA, [3]Department of Civil and Environmental Engineering, University of Michigan, Ann Arbor, MI, USA

Key Points:

- Using variable resolution mesh for flood modeling can balance computational efficiency and simulation accuracy
- Mesh resolution and precipitation cause significant uncertainties in inundation dynamics; streamflow uncertainty is mainly due to precipitation
- The compound effect results in higher flooding depth and broader extent but only impacts the area within 15–20 km upstream from the outlet

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

D. Xu,
donghui.xu@pnnl.gov

Citation:

Xu, D., Bisht, G., Engwirda, D., Feng, D., Tan, Z., & Ivanov, V. Y. (2025). Uncertainties in simulating flooding during Hurricane Harvey using 2D shallow water equations. *Water Resources Research*, *61*, e2024WR038032. https://doi.org/10.1029/2024WR038032

Received 23 MAY 2024
Accepted 20 DEC 2024

Author Contributions:

Conceptualization: Donghui Xu, Gautam Bisht
Data curation: Donghui Xu, Dongyu Feng
Formal analysis: Donghui Xu
Funding acquisition: Gautam Bisht, Zeli Tan, Valeriy Y. Ivanov
Investigation: Donghui Xu, Gautam Bisht, Darren Engwirda, Dongyu Feng, Zeli Tan
Methodology: Donghui Xu, Valeriy Y. Ivanov
Project administration: Gautam Bisht
Resources: Donghui Xu
Software: Donghui Xu, Darren Engwirda

Abstract Flooding is one of the most impactful weather-related natural hazards. Numerical models that solve the two dimensional (2D) shallow water equations (SWE) represent the first-principles approach to simulate all types of spatial flooding, such as pluvial, fluvial, and coastal flooding, and their compound dynamics. High spatial resolution (e.g., $\mathcal{O}(10^0 - 10^1)$ m) is needed in 2D SWE simulations to capture flood dynamics accurately, resulting in formidable computational challenges. Thus, relatively coarser spatial resolutions are used for large-scale simulations of flooding, which introduce uncertainties in the results. It is unclear how the uncertainty associated with the model resolution compares to the uncertainties in precipitation data sets and assumptions regarding boundary conditions when channelized flows interact with other water bodies. In this study, we compare these three sources of uncertainties in 2D SWE simulations for the 2017 Houston flooding event. Our results show that precipitation uncertainty and mesh resolution have more significant impacts on the simulated streamflow and inundation dynamics than the choice of the downstream boundary condition at the watershed outlet. We point out the viability to confine the uncertainty of coarsening mesh resolution by using the variable resolution mesh (VRM) which refines critical topographic features with far fewer grid cells. Specifically, in simulations with VRM, the simulated inundation depths over the refined region are comparable to that use the finest uniform mesh. This study contributes to understanding the challenges and pathways for applying 2D SWE models to improve the realism of flood simulations over large scales.

Plain Language Summary Flooding, a significant natural hazard driven by weather events, is often simulated using computer models that simulate how water moves. High model resolution (e.g., 1–10 m) is vital for accurately simulating flood dynamics, but it poses considerable computational challenges, particularly for large-scale applications. Running flood models at lower resolutions (e.g., 100–1,000 m) is computationally feasible for large areas but introduces uncertainties. This study compares the model errors arising from model resolution, precipitation data, and boundary condition using the 2017 Hurricane Harvey flood as a case study. Results indicate that precipitation bias and model resolution are the most important sources of uncertainty during Hurricane Harvey. Using a mesh with variable resolutions in the domain that preserves critical landscape features helps mitigate resolution-related errors. This mesh balances simulation accuracy with computational efficiency. Understanding these challenges and solutions can enhance the realism of large-scale flood simulations, leading to better flood prediction and management.

1. Introduction

Flooding stands as one of the most perilous and devastating natural disasters (Doocy et al., 2013). Unfortunately, the associated risks have escalated in recent decades (Ivanov et al., 2021) and are projected to continue rising in the future because of global warming (Gudmundsson et al., 2021; Hirabayashi et al., 2013; Tabari, 2020) and growing urbanization (Ajjur & Al-Ghamdi, 2021; Huong & Pathirana, 2013). While the increasing trend of extreme flooding appears inevitable (Huang & Swain, 2022), proactive emergency preparedness and appropriate adaptation strategies offer avenues for mitigating its consequences. The effectiveness of such strategies to respond to flooding events relies heavily on the accuracy and detail of flood simulations to inform decision-making processes (Sanders et al., 2020).

However, most current large-scale (e.g., national, continental, or global scales) flood forecasts (Hirpa et al., 2018; Kauffeldt et al., 2016; Viterbo et al., 2020) and flood risk projections (Alfieri et al., 2017; Hirabayashi et al., 2013;

Supervision: Gautam Bisht
Validation: Donghui Xu, Dongyu Feng
Visualization: Donghui Xu
Writing – original draft: Donghui Xu
Writing – review & editing: Donghui Xu, Gautam Bisht, Darren Engwirda, Dongyu Feng, Zeli Tan, Valeriy Y. Ivanov

Xu et al., 2024) rely on simplified one-dimensional (1D) runoff routing models (Luo et al., 2017; Yamazaki et al., 2011). For the sake of computational feasibility in large-scale applications, simplified 1D routing models typically use kinematic or diffusion wave approximations of the shallow water equations (SWE; Katopodes, 2018) at coarse spatial resolutions in the order of $10^4 - 10^5$ m. These models route the runoff to the outlet through a river network by only accounting for the main channel, which has a conceptual channel geometry. Consequently, most large-scale runoff routing models ignore the overland flow process (i.e., flow in the non-channel areas within the domain), and thus cannot simulate pluvial flooding that can influence larger areas than fluvial flooding (Tanaka et al., 2020), especially in urban environments. Recent development of large-scale forecasting systems operated by the National Oceanic and Atmospheric Administration (NOAA) include overland flow routing process (Cosgrove et al., 2024), however, a reduced physics formulation (i.e., 1D Muskingum-Cunge method) was implemented for channel routing (Read et al., 2023). Parameterizations of the channel geometry in the 1D river routing models introduce uncertainties in streamflow simulations (Xu, Bisht, Zhou, et al., 2022) due to the upscaling of information on real-world rivers to "representative" rivers at the model resolution (Wu et al., 2011). It remains challenging to reduce the uncertainty associated with river geometry because of the resolution mismatch between real-world rivers and the river network representation in models (Liao et al., 2022). Overall, while the 1D simplified river routing models are essential tools for simulating streamflow, their representativeness and ability to predict flood dynamics are limited (Gharbi et al., 2016).

Full sets of two dimensional (2D) SWE are promising to simulate flood wave propagation on arbitrary topography (Bradford & Sanders, 2002; Guo et al., 2021), though they may not capture the velocity accurately around buildings in urban environments, where 3D SWE models are required (Dewals et al., 2023). While numerous studies have developed flood models based on 2D SWE (Begnudelli & Sanders, 2006; Cea & Bladé, 2015; Dullo et al., 2021; Ming et al., 2020; Morales-Hernández et al., 2021; Wright et al., 2022), these models have been typically applied in small watersheds at ultra-high spatial resolution at the order of $\mathcal{O}(10^0 - 10^1)$ m. Many efforts have been dedicated to enhancing simulation efficiency by leveraging recent computational advancements (Sharifian et al., 2023; Vacondio et al., 2014; Wittmann et al., 2017) or implementing efficient numerical schemes (Bates et al., 2010; Xia & Liang, 2018). Nevertheless, using 2D SWE models to simulate flood dynamics at large scales (e.g., state, continental, and global) presents formidable computational and data storage challenges, especially when uncertainty quantification is needed (Ivanov et al., 2021). Computationally cheap surrogate modeling of the flooding solver has been used to facilitate real-time forecasting, but thousands of ultra-high resolution flood simulations are commonly required to train the surrogate model (Fraehr et al., 2023; Ivanov et al., 2021; Y. Zhou et al., 2021). Training such a surrogate model framework can still be computationally infeasible for large-scale applications, particularly without access to ample high-performance computing resources. In addition, surrogate models may not be able to capture future extreme flooding events accurately because climate non-stationarity (Milly et al., 2008) suggests that historical records may be too limited for surrogate model training.

Resolving the intricacies of flood dynamics at large scales necessitates striking a balance between computational feasibility and modeling accuracy. As it is computationally infeasible to solve 2D SWE for large-scale domains at the resolution of $\mathcal{O}(10^0 - 10^1)$ m and the simplified 1D routing models cannot provide sufficient details and accuracy for flood prediction, running the 2D SWE simulation at coarser resolutions of $\mathcal{O}(10^2 - 10^3)$ m holds promise in bridging this gap. It remains a question whether such relatively coarse 2D SWE simulations are possible at large scales and how much accuracy can be achieved. Maxwell and Condon (2016); Maxwell et al. (2015); O'Neill et al. (2021) simulated runoff routing using the kinematic wave approximation for the continental United States at 1 km resolution with reasonable performance for the streamflow simulations. A simple and efficient approximation of SWE with inertial terms considered (Bates et al., 2010) was successfully applied to analyze pluvial and fluvial flood hazards at the continental (Bates et al., 2021; Wing et al., 2017) and global scales (Sampson et al., 2015). However, uncertainty quantification was not the focus of the above-mentioned studies.

In addition to computational burdens, precipitation uncertainty poses another challenge for large-scale flooding simulation (Cloke & Pappenberger, 2009; Hjelmstad et al., 2021). Specifically, existing precipitation forecasts can be highly biased due to input data and model parameter uncertainties (Xu, Bisht, Sargsyan, et al., 2022). The commonly used global precipitation reanalysis forcings (Sun et al., 2018) exhibit relatively coarse spatial resolutions (e.g., $\mathcal{O}(10^4 - 10^5)$ m), which cannot resolve spatial heterogeneity critical for flooding simulations (Z.

Zhou et al., 2021). While a few precipitation data sets have a higher spatial resolution (e.g., 1,000 m), they are available at coarse temporal daily scales (Karger et al., 2021; Thornton et al., 2021). Considering that a flooding event can last from a few hours to several days, the sub-daily temporal variation of precipitation is critical for accurately simulating flood dynamics. In addition, high-resolution flooding can be coupled with state-of-the-science climate models for projecting flooding conditions in a warmer climate. Due to the coarse spatial resolutions and parametric uncertainty, climate models may significantly underestimate extreme precipitation events (Jong et al., 2023).

Over the coastal regions, hurricanes can cause devastating flooding hazards due to the compound effects when heavy precipitation and storm surges occur simultaneously or in close succession (Feng et al., 2023; Santiago-Collazo et al., 2019, 2024; Valle-Levinson et al., 2020). Increasing efforts have been made to apply flooding models to estimate the magnitude of compound flood events over coastal urban areas (Karamouz et al., 2017; Muñoz et al., 2024; Sarhadi et al., 2024). Accurate water levels are needed in inland flooding models as boundary conditions to capture the nonlinear interactions between fluvial and coastal processes at the river-ocean interface (Olbert et al., 2023). Specifically, flooding models tend to underestimate the inundated extent without including the interactions between high streamflow accumulated from upstream and high tides and storm surge (Huang et al., 2021; Kumbier et al., 2018). However, the uncertainty arising from the ocean boundary has not yet been well addressed for inland flooding models. It remains challenging for the current ocean modeling framework to forecast storm surges accurately before hurricane landfall with enough leading time because of the lack of an accurate storm forecast (Dullaart et al., 2020) and insufficient spatial resolution in global ocean model (Saint-Amand et al., 2023). Therefore, questions persist regarding the extent to which coastal processes may influence inland flood simulations and the degree of impact relative to other sources of uncertainty.

Given the inherent uncertainties in the current atmospheric and oceanic models' simulations and data sets, it is crucial to assess whether running the flooding model at a high spatial resolution is necessary to improve the prediction accuracy. Muñoz et al. (2024) quantified the cascading uncertainty in compound flood modeling without accounting for the impact of mesh resolution and found model structure and forcing conditions are the dominant sources of uncertainties. Although several other studies evaluated the impacts of mesh resolution on simulation accuracy (Caviedes-Voullième et al., 2012; Feng et al., 2019; Neal et al., 2011), such uncertainty has not been explicitly compared to the uncertainties arising from precipitation and boundary conditions. For instance, if the uncertainty stemming from the mesh resolution is significantly smaller than precipitation uncertainty, we cannot simulate flood dynamics more accurately by increasing mesh resolution. In addition, while the use of variable resolution mesh has been proposed as a solution to reduce the uncertainty of mesh resolution (Vivoni et al., 2004), the sensitivity of flood simulations to variable resolution mesh has not been systematically examined. In this study, we aim to compare the aforementioned sources of uncertainty in flood modeling that solves 2D SWE for a highly urbanized watershed in the Houston area during the 2017 Hurricane Harvey event. We generated variable resolution meshes with regions refined at different levels to understand the sensitivity of mesh resolution. We designed several simulation experiments to quantify and compare the impacts of (a) mesh resolution, (b) precipitation uncertainty, and (c) boundary conditions on the simulated streamflow and inundation dynamics.

The manuscript is structured as follows: Section 2 outlines the flood model formulation, the mesh generation process, and the forcing data used in this study. Section 3 presents the results. Section 4 discusses the findings, while Section 5 highlights the study's limitations. Finally, Section 6 provides the conclusions.

2. Methods

2.1. Flood Models

In this study, we used the Overland Flow Model (OFM; Kim et al., 2012) that solves 2D SWE to simulate flood dynamics:

$$\frac{\partial h}{\partial t}+\frac{\partial uh}{\partial x}+\frac{\partial vh}{\partial y}=i, \tag{1}$$

$$\frac{\partial uh}{\partial t}+\frac{\partial\left(u^2h+\frac{1}{2}gh^2\right)}{\partial x}+\frac{\partial uvh}{\partial y}=-gh\frac{\partial z_b}{\partial x}-C_Du\sqrt{u^2+v^2}, \tag{2}$$

$$\frac{\partial vh}{\partial t}+\frac{\partial(uvh)}{\partial x}+\frac{\partial\left(v^2h+\frac{1}{2}gh^2\right)}{\partial y}=-gh\frac{\partial z_b}{\partial y}-C_Dv\sqrt{u^2+v^2} \quad (3)$$

where h denotes the water depth, u and v represent the velocity in the x and y directions in the Cartesian coordinate system, i is the source (positive) or sink (negative) of water on the land surface, g is the gravitational acceleration constant, z_b is the bed elevation, and C_D is the bed drag coefficient estimated using the Manning equation: $C_D = gn^2h^{-1/3}$, and n is the Manning roughness coefficient. In brief, OFM is a finite volume model that implements a 1st order Godunov-type upwind scheme on triangular mesh and uses Roe's approximate Riemann solver for computing the fluxes. The readers can find detailed numerical methods in Kim et al. (2012); Schubert et al. (2008). Later, Ivanov et al. (2021) parallelized OFM using Portable, Extensible Toolkit for Scientific Computation (PETSc; Balay et al., 2019) for the domain decomposition and time stepping.

2.2. Study Domain

Hurricane Harvey made landfall on the Texas coast on 25 August 2017, as a category 4, resulting in devastating flooding on August 27–28. Hurricane Harvey brought extensive precipitation to southeastern Texas in 2017, resulting in widespread and devastating flooding and enormous socioeconomic consequences (Amadeo, 2018). The Buffalo Bayou at Turning Basin with the basin outlet at the USGS gage of 08074710 was selected as the study domain as it includes 19 gages with streamflow records and 48 high water marks (Arctur, 2018) that can be used for validating the simulations (Figure 1). The outlet gage is a tidal stream gage, thus the observed gage height can be used as a boundary condition to investigate the compounding effects of precipitation and storm surge on flooding. We used the 30 m Digital Elevation Model (DEM) from the National Elevation Data set (NED) for the surface elevation. Although there exists a higher resolution DEM data set at 1 m resolution, the 30 m DEM represents a better option for large-scale simulations, as it leads to computational units with larger sizes and manageable computational burdens. Instead of establishing a high accuracy flood simulation, this study aims to identify a balance between model accuracy and computational efficiency for large-scale applications.

2.3. Mesh Generation

Two types of meshes were used in this study: uniform resolution mesh (URM) and variable resolution mesh (VRM). Both URM and VRM were generated at eight spatial resolutions to investigate the impacts of mesh resolution on the simulated flood dynamics. The finest resolution base URM was generated by splitting each 30 m DEM quadrilateral grid cell into two triangles, which resulted in about 3 million triangular grid cells (Table 1). The coarser resolution URMs were generated by first upscaling the 30 m DEM values to 60, 90, 120, 180, 240, 480, and 990 m by averaging the elevations of all the 30 m cells within each coarse cell. We then split each coarse resolution quadrilateral cells into two triangles (Table 1).

Topographic features such as river geometry (Kim et al., 2022; Xu, Bisht, Zhou, et al., 2022), floodplain (Schrapffer et al., 2020; Yamazaki et al., 2011), and water management structure (Schmutz & Moog, 2018) are important in shaping the runoff routing. Therefore, simulation using a VRM with higher resolution in rivers, floodplains, and dams, while a lower resolution in other regions, is more computationally efficient and performs similarly to a simulation with URM (Vivoni et al., 2004). The equivalent mesh resolution of a VRM to compare with the mesh resolution in URM is given by:

$$\text{Equivalent resolution} = \sqrt{\frac{A_{\text{watershed}}}{N_{\text{cell}}}\times 2}, \quad (4)$$

where $A_{\text{watershed}}$ is the total watershed area, N_{cell} is the number of grid cells in a VRM, and the multiplier 2 is from the URM that each square is split into two triangles. We note that the proposed equivalent resolution formula is not applicable for the calculation of the resolution of quadrilateral cells, but Equation 4 provides a consistent metric to compare the mesh resolution between URMs and VRMs that were used in this study. In this study, we used a Delaunay-based unstructured mesh generator, JIGSAW (Engwirda, 2017), to generate VRM with varying spatial resolution in the domain. We extracted the streamlines from the original 30 m DEM using the MATLAB package of Schwanghart and Scherler (2014) at different levels of detail by changing the threshold of flow

10.1029/2024WR038032

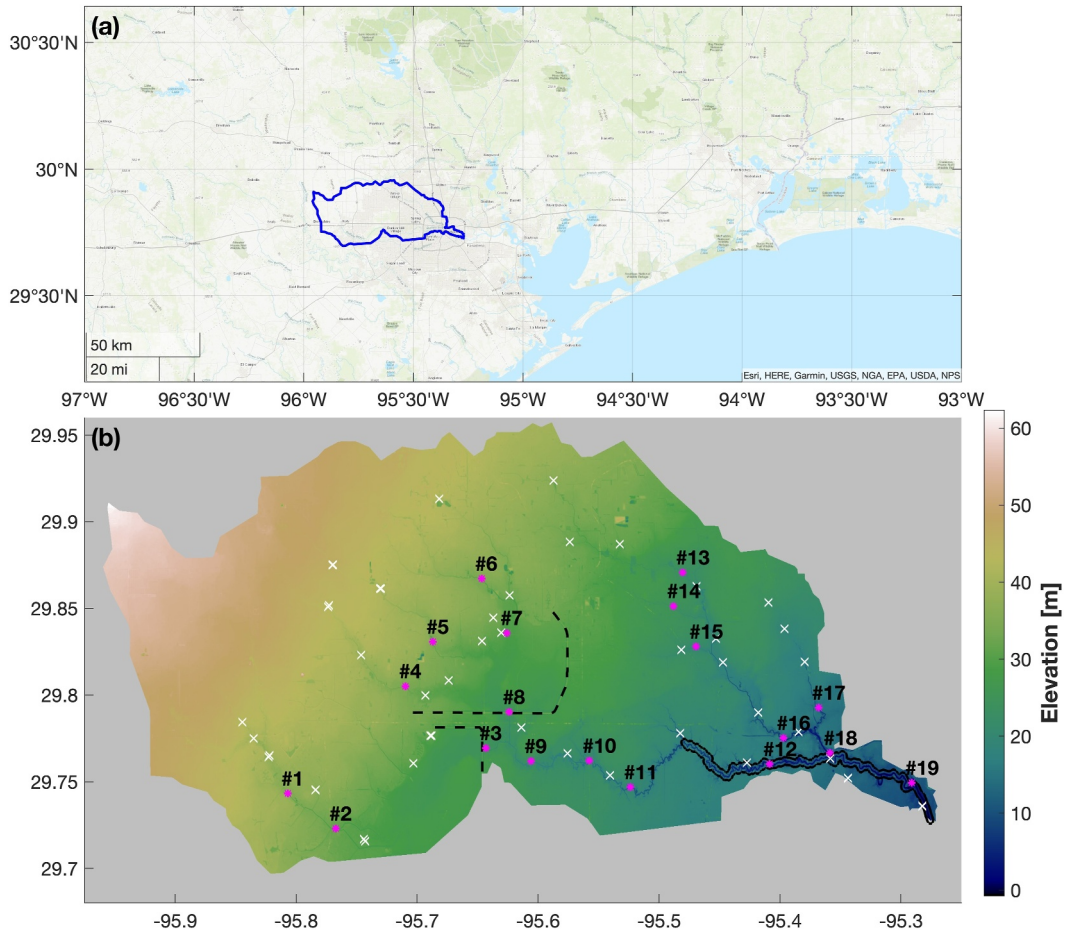

Figure 1. (a). Study domain. (b). A 30 m DEM data set illustrating the topography of the study domain. The white crosses denote the locations of high water mark measurements from USGS, and the red asterisks are the USGS gages with the gage number. The black dashed lines represent two dams used as flood control: Addicks (the upper one) and Barker (the lower one). The black solid line shows the extent of major channels. The USGS gage information can be found in Table S1 in Supporting Information S1.

Table 1
A Summary of Mesh Configurations

Mesh#		URM1	URM2	URM3	URM4	URM5	URM6	URM7	URM8
Resolution m		30	60	90	120	180	240	480	990
Number of cells		2,926,532	792,265	351,419	197,309	87,351	48,946	12,036	2,746
Mesh#		VRM1	VRM2	VRM3	VRM4	VRM5	VRM6	VRM7	VRM8
Main Channel	Floodplain width m	500	250	250	250	250	250	250	250
	Resolution m	30	30	30	30	30	30	1,000	1,000
Tributary	Floodplain width m	300	150	150	150	150	150	1,000	1,000
	Resolution m	60	90	90	90	90	90	1,000	1,000
Dams	Resolution m	30	30	30	30	30	30	30	1,000
Other	Resolution m	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Number of cells		664,724	427,478	254,953	148,514	109,358	58,749	14,536	2,643
Equivalent Resolution m		~60	~80	~100	~130	~150	~200	~420	~1,000

Note. Note each column represents setup for a mesh.

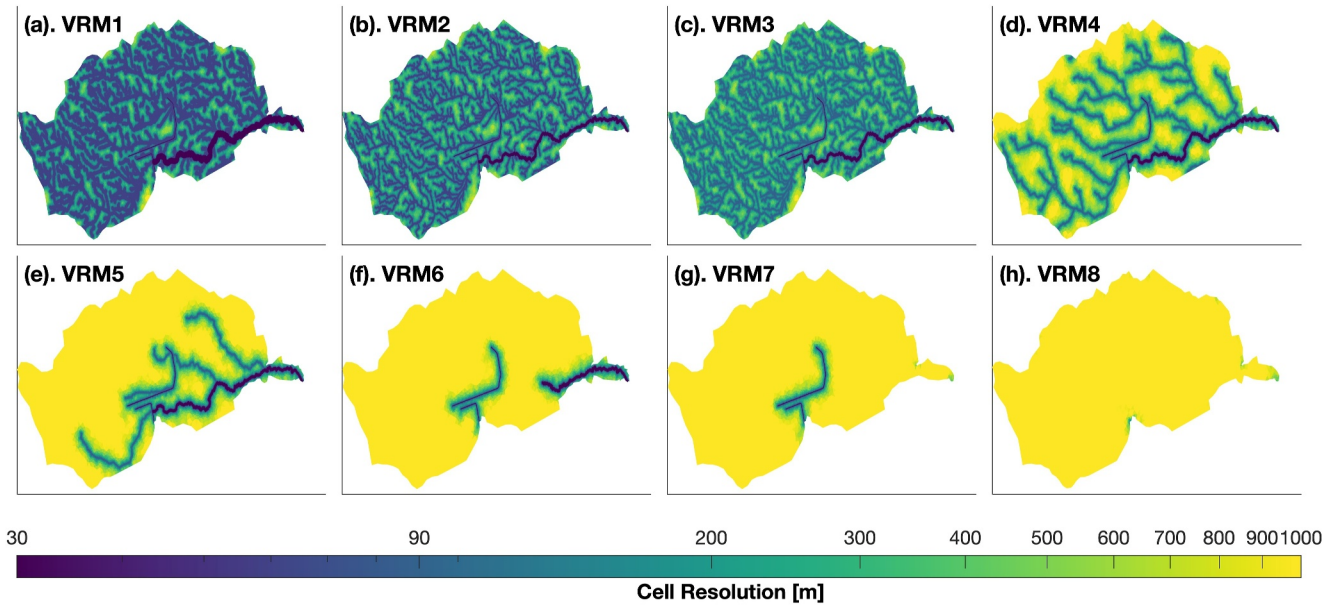

Figure 2. Spatial distribution of the cell resolution for variable resolution meshes. Note that the cell resolution is approximated by the averaged triangle edge length. The logarithmic scale is used in the color bar.

accumulation. Although the DEM may have bias in the river bed elevation, it captures the stream network well. The floodplain was delineated along the streamlines with different buffer widths for the main channel and tributaries. Specifically, the main channel is assumed to have a wider floodplain than the tributaries, as the main channel has a higher streamflow than the tributaries (Table 1). In addition, a coarser spatial resolution was used for the tributary floodplain (i.e., 60–90 m) than the main channel floodplain. Figure 2 illustrates the spatial density of the mesh cells at different resolution levels used for generating VRMs. Furthermore, there are two flooding control reservoirs in the study domain, Addicks and Barker reservoirs, which are critical for flood simulations (Du et al., 2019). Although the dam operation has not been implemented in OFM, the corresponding flooding control effects can be simulated if the mesh correctly captures the dam elevation, as the high elevation blocks the flooding water. Thus, a higher resolution was used around the corresponding dams (black lines in Figure 1b) in VRM1—7 to capture the higher elevation of the dams. Although the dam regions were not zoomed in for VRM8, the mesh nodes are forced to fall exactly on the dam locations (black lines in Figure 1b). Overall, VRM configurations are described in Table 1.

2.4. Precipitation Data Sets

Five precipitation data sets with different spatial and temporal resolutions were used to drive OFM to assess the uncertainty of precipitation inputs to flood simulation (Table 2). Multi-Radar Multi-Sensor System (MRMS) and IMERG were upscaled from their native temporal resolution to the hourly resolution, and DAYMET and MSWEP were downscaled to hourly resolution, assuming a constant precipitation rate. During the event, the total precipitation amount was more than 760 mm (Tate, 2017), which is captured more accurately by MRMS than the

Table 2
Precipitation Data Sets

Data set	Spatial resolution	Native temporal resolution	Reference
Multi-radar multi-sensor system (MRMS)	$1{,}000\ m$	2-min	Zhang et al. (2016)
North American land data assimilation system (NLDAS)	$\sim 12{,}500\ m$	Hourly	Xia et al. (2012)
DAYMET	$1{,}000\ m$	Daily	Thornton et al. (2021)
Integrated multi-satellite retrievals for global precipitation measurement (IMERG)	$\sim 1{,}000\ m$	30-min	Huffman et al. (2020)
Multi-source weighted-ensemble precipitation (MSWEP)	$\sim 1{,}000\ m$	3-hourly	Beck, Wood, et al. (2019)

Table 3
Simulation Experiments

Experiment set #	Mesh	Precipitation	Boundary condition
EXP-CTL	URM1	MRMS	On
EXP-URM	URM2–8	MRMS	On
EXP-VRM	VRM1–8	MRMS	On
EXP-PREC	URM1	Spatially averaged MRMS, NLDAS, DAYMET, IMERG, MSWEP	On
EXP-BC	URM1	MRMS	OFF

other data sets (Figure S1 in Supporting Information S1). The precipitation in all data sets is assumed to become runoff because the study domain is highly urbanized (more than 70% of the study domain are developed areas, Figure S2 in Supporting Information S1), and precipitation loss is negligible during such an extreme hurricane event. The nearest neighbor interpolation method was used to map the precipitation data on the various URMs and VRMs. Note that the interpolation may result in small changes in the total precipitation read into different meshes.

2.5. Simulation Experiments

We ran multiple simulation experiment sets (EXP) listed in Table 3, and selected the simulation of EXP-CTL as a benchmark, which uses the most detailed information, such as 30 m resolution mesh, including the boundary condition at the outlet that connects the watershed to the ocean, and the most accurate precipitation forcing (i.e., MRMS; Habibi et al., 2021) among the tested precipitation data sets. Specifically, the impacts of the mesh resolution on flooding simulations are assessed by comparing EXP-URM and EXP-VRM with EXP-CTL. EXP-PREC ran the model forced by multiple precipitation data sets to investigate the impacts of precipitation uncertainty on the flood simulation. The difference between EXP-BC and EXP-CTL is used to examine the compounding effect of heavy precipitation and high storm surge level on the inundation dynamics during the hurricane period. In Table 3, the boundary condition set to "On" indicates the application of the Dirichlet boundary condition at the basin outlet, using the water level measurements from the tidal gage. The boundary condition set to "Off" indicates the use of the critical flow condition to estimate the discharge, relying solely on upstream water levels, which excludes the effect of storm surge on inundation. Using critical flow condition assumes the flow exits the outlet as free outfall, and only the water depth from the outlet cell is needed for estimating the discharge. All the simulations were conducted over 6 days from 25 August 2017, to 30 August 2017, with the dry domain as an initial condition. The first day was treated as a model spin-up period and excluded from analysis. Due to the high urbanization, a constant Manning coefficient of 0.015 was used for all the simulations (Ivanov et al., 2021). We further tested the sensitivity of the roughness parameter by performing six additional EXP-CTL simulations with Manning coefficients equal to 0.03, 0.045, 0.060, 0.075, 0.090, 0.105. The impacts of precipitation on simulation accuracy were evaluated by similarly performing an additional set of six EXP-CTL simulations with uniformly changing MRMS at each grid cell by −30%, −20%, −10%, 10%, 20%, and 30%.

2.6. Evaluation Metrics

Five metrics were used in this study to evaluate EXP-CTL simulation against the USGS water level observations and to compare the remaining simulations listed in Table 3 to EXP-CTL. The Pearson correlation coefficient (ρ) was computed as:

$$\rho = \frac{\sum_{i=1}^{N}\left(\mathrm{ref}_i - \mu_{\mathrm{ref}}\right)\left(\mathrm{sim}_i - \mu_{\mathrm{sim}}\right)}{\sqrt{\sum_{i=1}^{N}\left(\mathrm{ref}_i - \mu_{\mathrm{ref}}\right)^2 \sum_{i=1}^{N}\left(\mathrm{sim}_i - \mu_{\mathrm{sim}}\right)^2}}, \tag{5}$$

where i denotes a sample, ref_i is the reference data, sim_i is the simulation, and N is the total number of samples. Note, N can be the number of data points in a time series or a spatial field. Nash–Sutcliffe model efficiency (NSE; Nash & Sutcliffe, 1970) was computed as:

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{N} (\text{ref}_i - \text{sim}_i)^2}{\sum_{i=1}^{N} (\text{ref}_i - \mu_{\text{sim}})^2}, \tag{6}$$

NSE varies from $-\infty$ to 1, with a NSE value of 1 indicating a perfect simulation. When NSE is less than 0, it means that the model's predictive skill is worse than the mean value. The root mean square error (RMSE) was computed as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} (\text{ref}_i - \text{sim}_i)^2}{N}}, \tag{7}$$

RMSE varies from 0 to $+\infty$, with RMSE $= 0$ representing a perfect simulation. The Kling-Gupta Efficiency (KGE; Gupta et al., 2009) was computed as:

$$\text{KGE} = 1 - \sqrt{(\rho - 1)^2 + \left(\frac{\sigma_{\text{ref}}}{\sigma_{\text{sim}}} - 1\right)^2 + \left(\frac{\mu_{\text{ref}}}{\mu_{\text{sim}}} - 1\right)^2}, \tag{8}$$

where σ_{ref} is the standard deviation of the reference data, σ_{sim} is the standard deviation of the simulation, μ_{ref} is the mean of the reference data, and μ_{sim} is the mean of the simulation. KGE varies from $-\infty$ to 1, with a KGE value of 1 indicating a perfect simulation. When KGE is less than -0.41, it means model performance is worse than the mean value (Knoben et al., 2019). We selected KGE to provide a comprehensive assessment of model performance by combining correlation, bias, and variability measures, while NSE and RMSE were chosen to specifically evaluate prediction bias. Together, these metrics provide a robust evaluation framework.

To evaluate the model performance in simulating flooding extent, we used the hit rate (H), false rate (F), and success index (S) as defined in Wing et al. (2017). The hit rate represents the percentage of wet cells in the benchmark case that are correctly reproduced by the simulation, while the false rate represents the percentage of dry cells in the benchmark case that are incorrectly simulated as wet. In this study, wet cells are defined as the cells with an inundated water depth larger than the ankle height (e.g., 0.12 m; Sanders et al., 2020), and dry cells have a simulated water depth lower than that. The success index is to quantify the overall performance with both overestimation and underestimation in simulating flooding extent accounted for:

$$S = \frac{M_1 B_1}{M_1 B_1 + M_0 B_1 + M_1 B_0} \times 100\%, \tag{9}$$

where M and B are the grid cells from model simulations and benchmark data, respectively, and the subscripts 1 and 0 represent flooded and dry cells, respectively.

3. Results

3.1. Validation of the OFM Simulation for EXP-CTL

The EXP-CTL simulation demonstrates satisfactory performance in capturing the water level dynamics at 19 selected USGS gages within the study domain (Figure 3). Specifically, EXP-CTL simulation captures the temporal variations and magnitudes of peak water levels relatively well for most gages. The median KGE and NSE between the simulated water levels and observations are 0.46 and 0.55, respectively. Our model performance for the gages at the main channel is promising and comparable to previous studies that focused on the Harvey flooding simulation (Chen et al., 2021; Dullo et al., 2021; X. Li et al., 2023). The less promising performance of EXP-CTL can be found from the gages located in upstream and small tributaries (Figure 1b), which were not selected for validation in other studies. Specifically, the simulations at gages #1, #2, #4, #6, #14, and #15 are relatively poor, with a significant underestimation of water level variations. Those gages are on tributaries with a

10.1029/2024WR038032

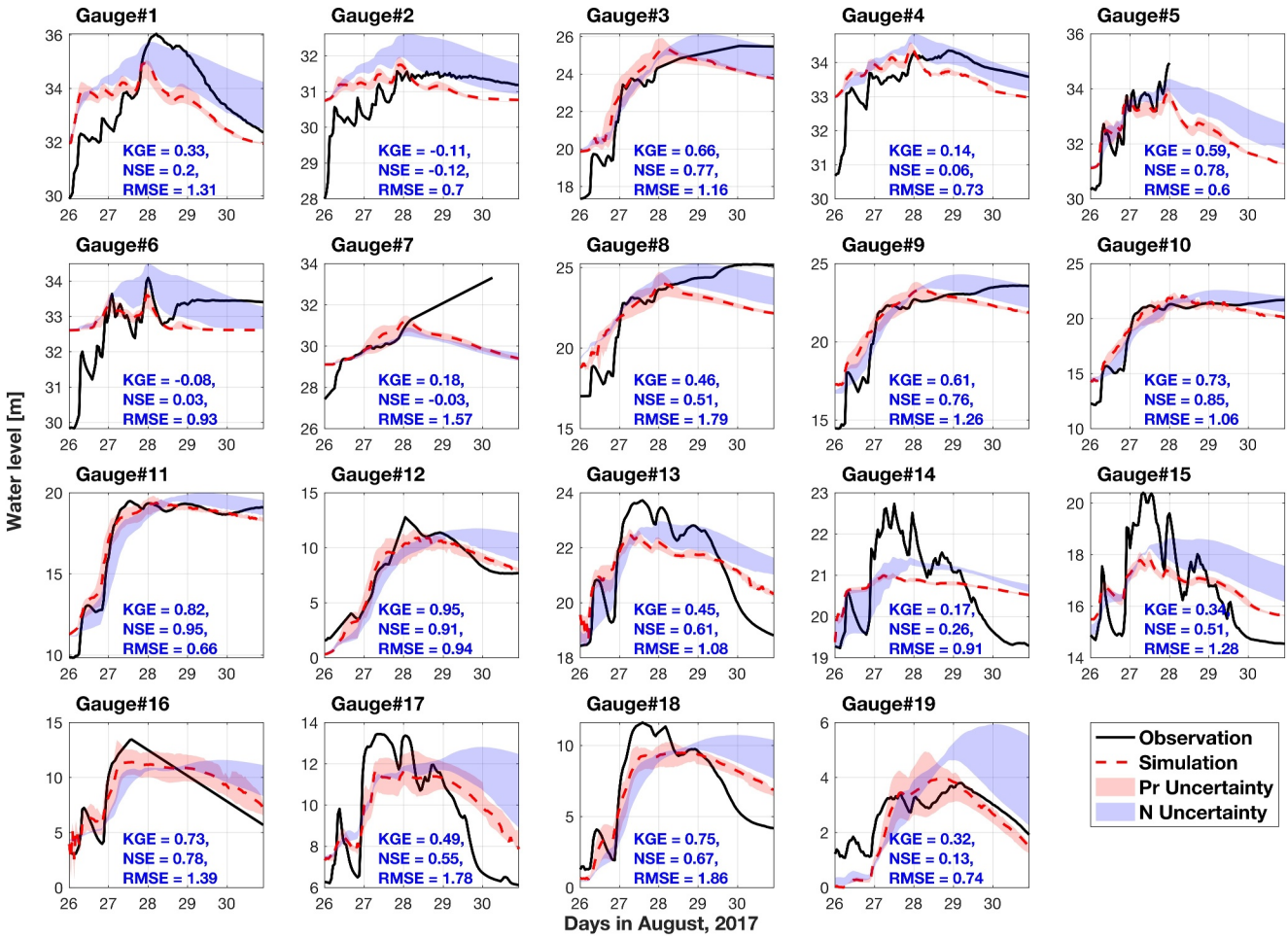

Figure 3. Validation of the EXP-CTL against water level measurements at selected USGS gages within the study domain at hourly scales. The *X*-Axis are the days from August 26 to 30, 2017. The text inserted in each subplot shows evaluation metrics of the Kling–Gupta efficiency (KGE), the Nash-Sutcliffe model efficiency (NSE), and Root mean square deviation (RMSE). The gage locations can be found in Figure 1. The red shaded area and blue shaded area are the sensitivity analysis with EXP-CTL configuration for precipitation (Pr) and Manning coefficient (N).

narrower channel, and the 30 m DEM may have insufficient resolution to accurately capture the geometry of those channels, resulting in a model bias. Performance is expected to improve with higher resolution DEMs (e.g., 1 m), better capturing the topography for small tributaries (Dullo et al., 2021). However, running the model at a higher resolution for the study domain or larger areas is computationally infeasible. Other factors contributing to the large error include the lack of representation of buildings and stormwater structures in the mesh, which significantly affects flow in the highly urbanized watershed (Iliadis et al., 2024).

We further implement sensitivity analysis for the precipitation and Manning coefficient in EXP-CTL. Perturbing the precipitation amount by −30%–30% without changing the temporal variability did not improve the simulated water depth at gages #1, #2, #4, #6, #14, and #15 (red shade areas in Figure 3). Due to the small contributing areas for those gages, it is difficult for the selected precipitation data set to accurately capture the precipitation temporal variability, which could explain the underestimation of the water depth variation. The choice of Manning coefficient has a significant impact on the simulated water depth as well, especially when the floodwater recedes (green shade areas in Figure 3). However, the default Manning coefficient value led to better performance than other test values in the sensitivity analysis (Figure S3 in Supporting Information S1). In addition, OFM is unable to capture the increasing trends of water level after August 28th at gages # 7 and #8, which is likely due to the precipitation uncertainty and the elevation of dams not being accurately represented in the mesh (e.g., the floodwater leaks).

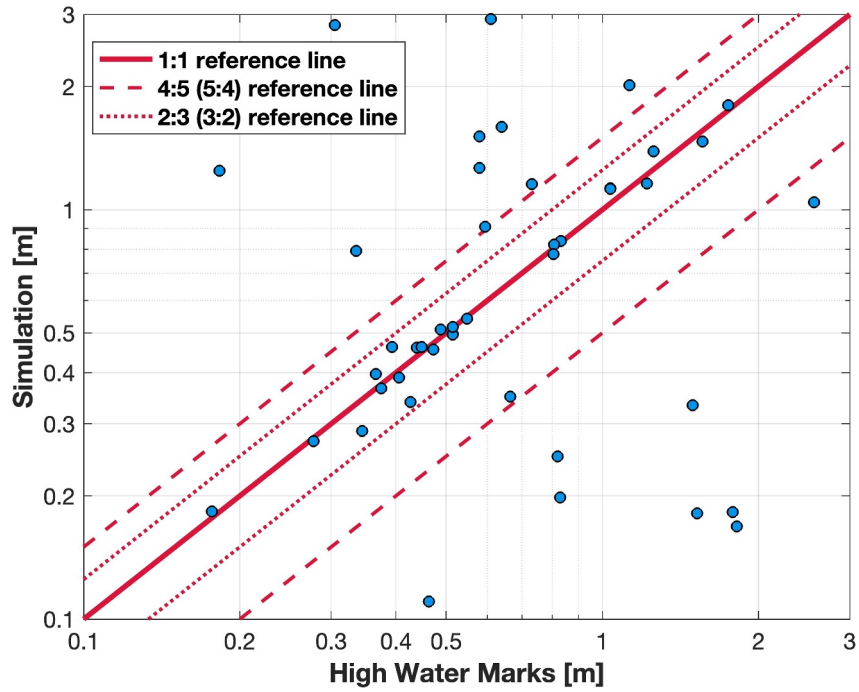

Figure 4. Validation of the EXP-CTL simulation results against the USGS High Water Marks. The red solid line is the 1:1 reference line. The red dotted and dashed lines delineate the regions with the relative error less than 25% and 50%, respectively.

We further validate the EXP-CTL simulation using USGS high water marks. The relative errors between the maximum simulated water levels during the simulation period and the high water mark measurements are smaller than 25% at 25 out of 48 sites (Figure 4). On average, the RMSE of the simulations for all selected high water marks is 0.9 m. These relatively large errors are primarily caused by DEM errors (e.g., insufficient spatial resolution), since high water marks were measured at point level, while the flooding model simulates the averaged water depth for a grid cell at the spatial resolution of 30 m. Figure S4 in Supporting Information S1 shows the maximum elevation difference in 30 m DEM compared to 1 m DEM for a smaller region in the study domain. The average elevation difference for the small region is 0.6 m, with a standard deviation equal to 0.7 m. Additional reasons for the high bias include the uncertainty in the high water mark measurements (Feaster & Koenig, 2017), and a mismatch of the high water mark locations in the model. However, it is unclear how to attribute the discrepancies between the simulation and high water marks observations to different factors. Due to the challenges of accurately simulating water depths at high water marks, the performance of EXP-CTL is similar to that reported in a previous study (RMSE = ~0.9 m; Chen et al., 2021).

Overall, EXP-CTL's performance is satisfactory for large-scale flooding prediction. It serves as a robust benchmark for quantifying the uncertainties arising from precipitation sources, mesh resolutions, and boundary condition types in large-scale flooding simulation.

3.2. Impacts of Resolution in URMs

A comparison of EXP-URM and EXP-CTL simulation results reveals that uniformly upscaling the spatial resolution has significant impacts on the simulated streamflow and spatial patterns of maximum inundation (Figure 5). In EXP-URM, all seven simulations with URM2-8 exhibited higher peak streamflow, occurring earlier than in the EXP-CTL simulation result (Figure 5a). Even a single level upscaling of the spatial resolution from 30 m (URM1) to 60 m (URM2) resulted in significant differences in the simulated streamflow. For example, the URM2 overestimates the magnitude of the peak streamflow by about 50% than the URM1 simulation, and the timing of the peak streamflow is simulated 1 day earlier. This discrepancy is caused by the smoothing of elevation around the dam locations (e.g., black dashed lines in Figure 1) in the upscaled resolution mesh. Specifically, as the width of the dams can be less than 30 m, upscaling the mesh resolution results in “lower” dams being represented

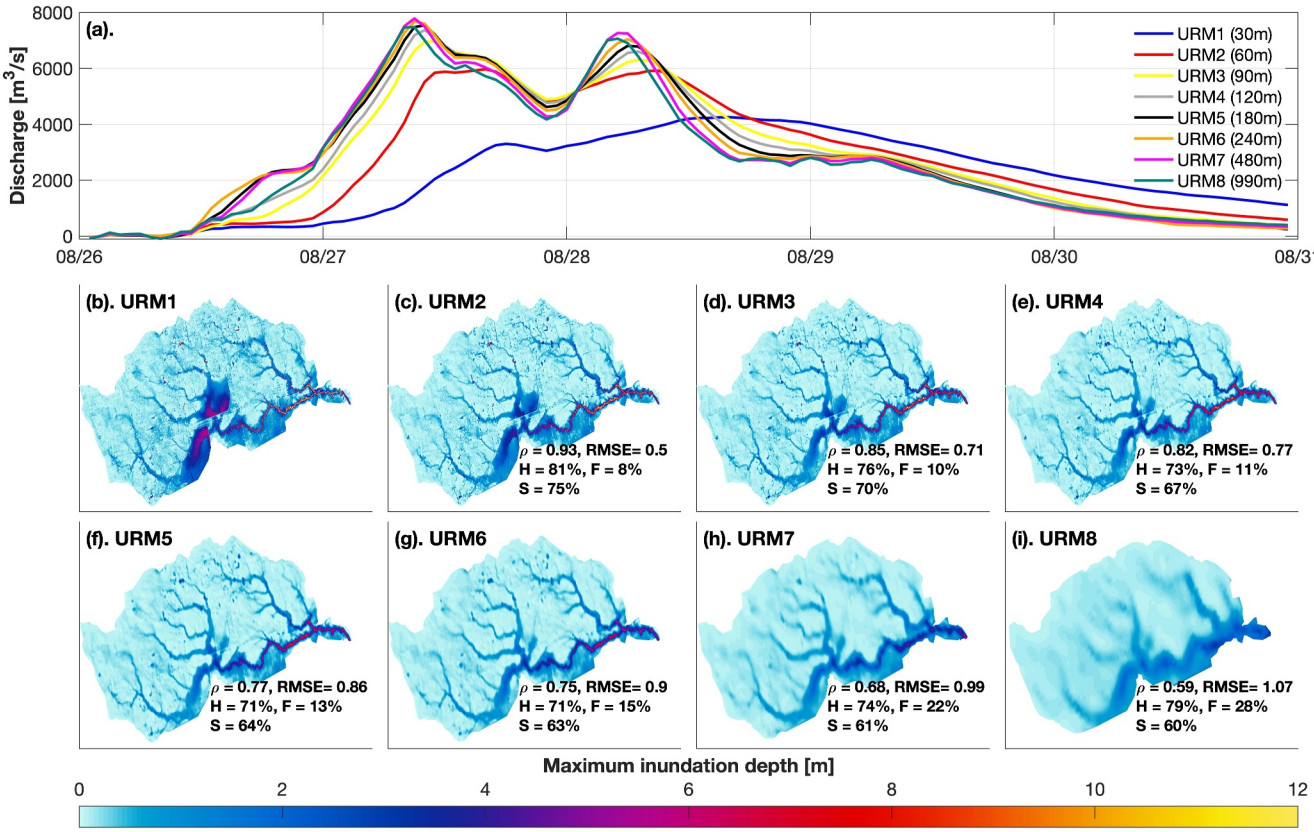

Figure 5. A comparison of simulations with the uniform resolution meshes (URM) at different spatial resolutions. Subplot (a) shows a comparison of the simulated streamflow at the outlet. Subplot (b)–(i) illustrate the maximum inundation depth for different mesh resolutions. The inserted texts are the performance metrics when compared against the URM1 simulation. ρ is the correlation coefficient, RMSE is the root mean square deviation, H represents the hit rate, F denotes the false rate, and S is the success index.

in the mesh, thus underestimating their flooding control effects. This is further evidenced by the underestimated water depths behind the dams in EXP-URM simulations compared to that in EXP-CTL (Figures 5b and 5c). In addition, the simulations with a spatial resolution coarser than 30 m exhibit similar temporal variability in simulated streamflow, suggesting that the streamflow at the outlet is not highly sensitive to mesh resolution except in the vicinity of dams.

The maximum simulated inundation depths are very sensitive to mesh resolution, as model performance degrades at coarser resolution compared EXP-CTL with a decreasing ρ and increasing *RMSE* (Figures 5b–5i). The simulated water depths behind the dams and in the river channels are underestimated in all EXP-URM simulations, as expected due to the smoothing of the elevation at the coarser resolutions. However, the spatial pattern of the inundation is relatively well captured, with the correlation coefficient between simulations in EXP-URM and EXP-CTL simulation larger than ∼0.6 (Figures 5c–5i). This suggests that OFM with a coarser mesh can still reasonably simulate the general inundation spatial pattern. We further evaluate the simulated flooding extent in EXP-URM using hit rate and false rate for the inundation depth larger than the ankle height (i.e., 0.12 m). Specifically, all the simulations in EXP-URM have a hit rate larger than 70% with similar performance among different simulations. However, the false rate increases monotonically as the mesh resolution decreases, from 8% in the 60m-resolution simulation (Figures 5b) to 28% in the 990 m simulation (Figure 5i). The success rate decreases from 75% to 60%.

3.3. Impacts of Resolution in VRMs

Simulations with VRMs show overall improved performance in both streamflow at the outlet and inundation dynamics compared to simulations with URMs (Figure 6). Specifically, the simulated streamflow closely matches the benchmark simulation of EXP-CTL (Figure 6a) since the topography around the dam regions is well captured

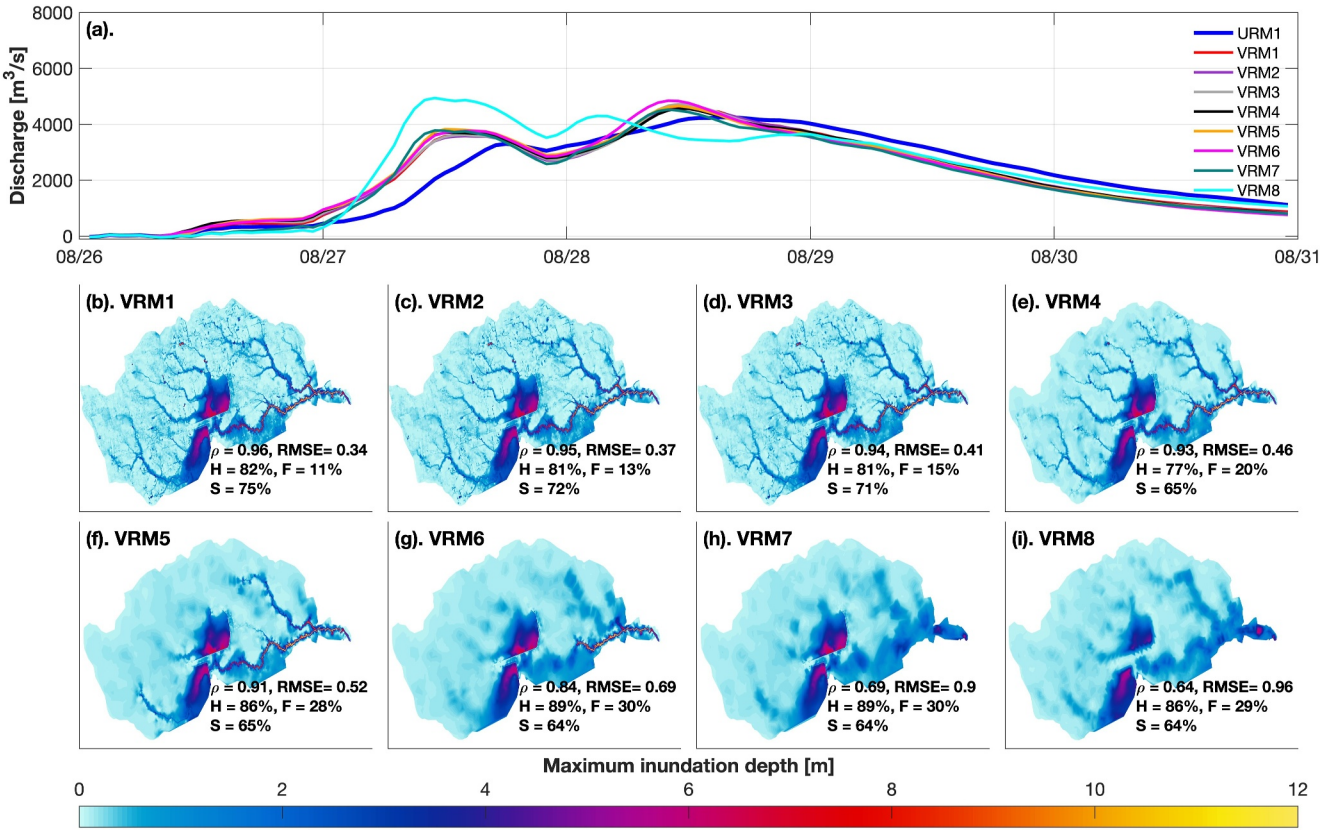

Figure 6. A comparison of simulations with variable resolution meshes (VRM) against the benchmark simulation that uses URM1 (30 m). Subplot (a) shows a comparison of the simulated streamflow series at the outlet. Subplots (b)–(i) illustrate the maximum inundation depth for different simulation meshes. The inserted text shows the performance metrics when comparing simulation results with that of URM1. ρ is the correlation coefficient, RMSE is the root mean square deviation, H represents the hit rate, F denotes the false rate, and S is the success index.

in the regionally refined VRM1–VRM7 meshes (Figure 2). The simulated streamflow with VRM8 however is not as good as that with VRM1—VRM7 since VRM8 is not refined around the dam regions. Instead, the dams are "burnt" into VRM8 to capture dam heights, forcing the nodes of VRM8 to align with the dam locations (i.e., the black dashed lines). While the averaged spatial resolution of VRM8 is approximately 1,000 m (e.g., cell width), the corresponding simulated streamflow is still better than EXP-URM simulations that used URM2—URM8. Simplifying the representation of rivers and floodplains from VRM1 to VRM7 does not significantly degrade the performance of the simulated streamflow (Figure 6a). The model performance for VRM7, which does not refine rivers and floodplains, remains comparable to the simulations with the other VRMs. This suggests that the flood control structures exert more control on the hydrograph than the representation of river and floodplain in this study domain. This is consistent with the EXP-URM simulations that the simulated streamflow is similar for all the URMs with spatial resolution coarser than 30 m (Figure 5a). For the natural watersheds (i.e., no flooding control structures), the impacts of the representation of rivers in mesh remain to be investigated.

However, representing rivers and floodplains in more detail in VRMs improves the simulation of the inundation dynamics (Figures 6b–6i). For example, compared to the maximum inundation depth in EXP-CTL, the ρ decreases from 0.96 with VRM1 to 0.64 with VRM8, and the RMSE increases from 0.34 with VRM1 to 0.96 with VRM8. The simulated flooding extent improves as well, with a hit rate greater than 80% using VRMs, which is better than that using URMs (i.e., about 70%). However, the false rates are slightly higher in simulations with VRMs than with URMs at a similar number of grid cells, leading to a similar success index. This is because the false rate is strongly related to the spatial resolution, and the higher false rates are from the coarsely resolved regions in VRMs. We note that the simulation with a higher resolution mesh may not lead to an improved hit rate, for example, the hit rate of VRM4 is much lower than that of VRM5. This is because the flooding model tends to overestimate inundated areas with a coarser resolution mesh than the simulation with a finer resolution mesh, thus

10.1029/2024WR038032

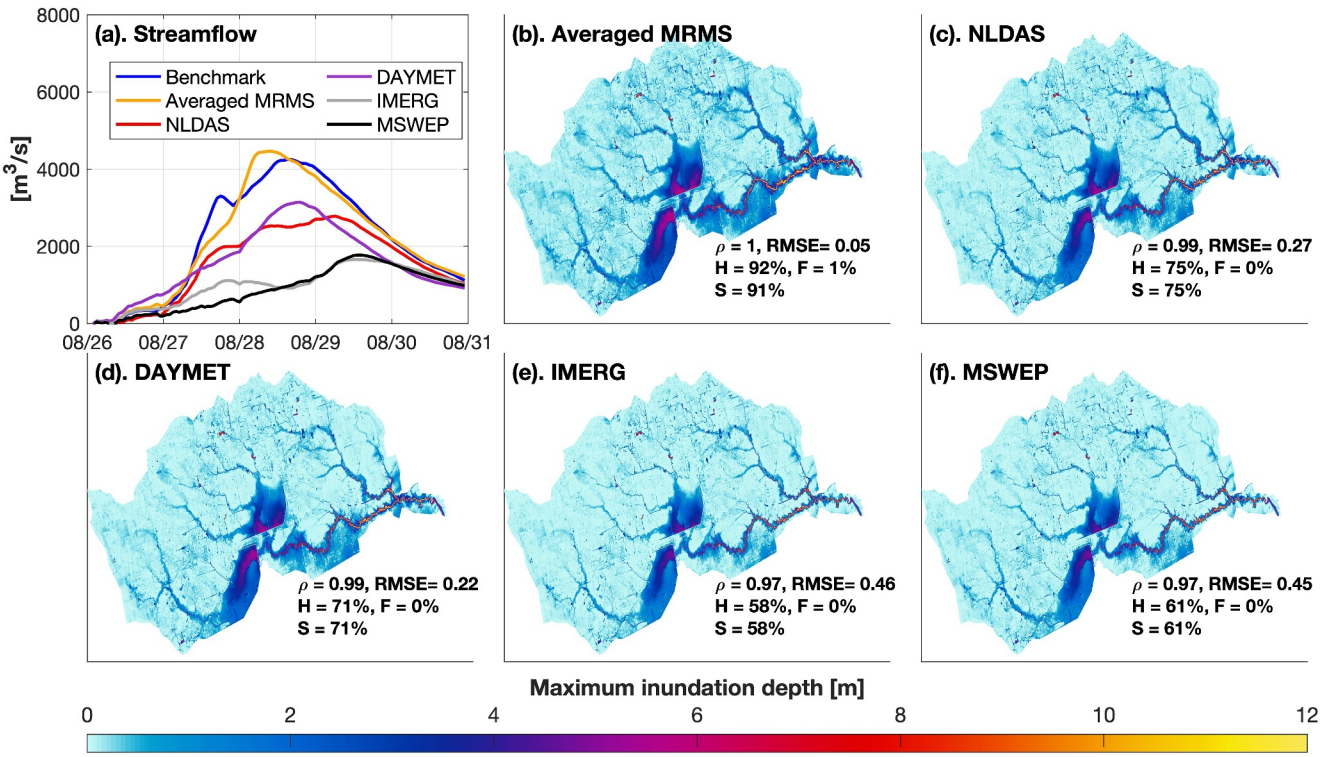

Figure 7. A comparison of simulations with different precipitation forcings in EXP-BC against the benchmark simulation that uses the 1,000 m resolution multi-radar multi-sensor system precipitation product. Subplot (a) shows a comparison of the simulated streamflow series at the outlet. Subplots (b)–(f) illustrate the maximum inundation depth for different precipitation forcings. Note that all of the simulations in panels (b)–(f) use the URM1 (30 m) mesh. The inserted text shows the performance metrics when compared against the simulation of EXP-1. ρ is the correlation coefficient, RMSE is the root mean square deviation, H represents the hit rate, F denotes the false rate, and S is the success index.

it may include more inundated areas that are presented in the benchmark simulation. However, there could be other factors contributing to the higher hit rate in VRM5, which are not fully understood.

3.4. Impacts of Precipitation Uncertainty

Simulations in EXP-PREC underestimate the magnitude of peak streamflow and cannot capture the corresponding timing as compared to the benchmark simulation of EXP-CTL (Figure 7a). This is because of substantial uncertainty in the temporal (Figure S1a in Supporting Information S1) and spatial scales (Figures S1b–S1f in Supporting Information S1) of the selected precipitation products. For example, compared to the precipitation data set used in EXP-CTL (i.e., MRMS), other precipitation data sets used in EXP-PREC notably underestimate precipitation amounts during the Harvey event. This aligns with the well-known challenge of the current precipitation data sets in accurately capturing extreme events (Alexopoulos et al., 2023; Beck, Pan, et al., 2019; Hu & Franzke, 2020). The spatial distribution of the precipitation impacts the simulation as well. When the spatial structure of precipitation is disregarded, as in the simulation with spatially averaged MRMS, the first streamflow peak before Aug 28th is missing, and the second peak shifts to an earlier time (Figure 7a). Although the simulation forced by DAYMET underestimated the peak streamflow magnitude of EXP-CTL, it correctly captures the timing of the streamflow peak. This underestimation of peak streamflow magnitude is attributed to the overall lower total precipitation by the DAYMET data set. Given that DAYMET is a daily data set, this finding implies that the accumulated precipitation amount but not the temporal variation of precipitation can significantly impact peak streamflow magnitude and timing in this case study. In contrast, simulations with other hourly precipitation forcings not only underestimate the peak streamflow magnitude but also delay the timing (Figure 7a).

Despite significant differences among simulations using different precipitation data sets, all EXP-PREC simulations capture the spatial pattern of the inundation depth quite well, with $\rho > 0.97$, as compared to the EXP-CTL simulation (Figures 7b–7f). This suggests that mesh resolution is the most critical factor for the spatial pattern of

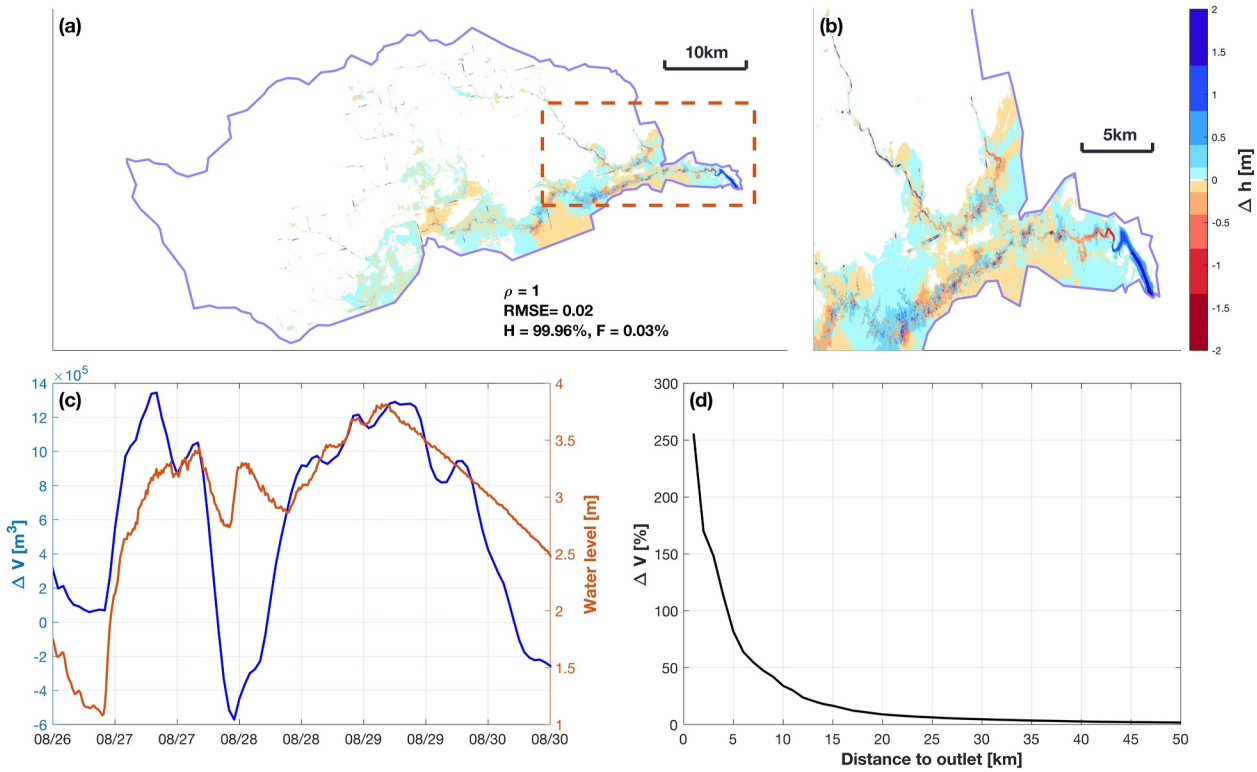

Figure 8. Impact of boundary condition on the simulated inundation depth. Subplot (a) shows the difference of maximum inundation depth between EXP-BC and EXP-CTL simulations (Δh, e.g., EXP-BC—EXP-CTL). Subplot (b) illustrates the zoom-in region of the red dashed box in subplot (a). Subplot (c) shows the difference of water volume in the domain (ΔV, left *Y*-Axis) and water level at the outlet (right *Y*-Axis) between EXP-BC and EXP-CTL simulations. Subplot (d) shows the relative change of regional water volume between EXP-BC and EXP-CTL simulations from the outlet to upstream.

the inundation. However, the precipitation uncertainty results in a large bias in the simulated flooding extent, resulting RMSE up to 0.46 m. Except for using the averaged MRMS precipitation, other simulations in EXP-PREC have lower hit rates than the EXP-URM and EXP-VRM simulations. The false rate is however much lower than in EXP-URM and EXP-VRM simulations with coarser meshes, indicating that the false rate is more sensitive to the spatial resolution than the hit rate.

The spatial variation of precipitation has a small impact on the simulation, as using basin-averaged MRMS precipitation in EXP-PREC leads to very similar simulated streamflow and inundation dynamics as in the EXP-CTL simulation. MRMS has the largest spatial variation among the selected precipitation data sets (Figure S1), thus the spatial structure of precipitation in other data sets will not significantly influence the simulation. It should be noted that the small impact associated with the precipitation heterogeneity is likely due to the relatively small size of the study domain (i.e., ~1,500 km^2). Precipitation heterogeneity will play a more important role in flood simulation at larger domains.

3.5. Impacts of Boundary Condition

The differences between EXP-BC and EXP-CTL show that including ocean boundary conditions for pluvial/fluvial flood simulations results in higher inundation depths and a larger inundation extent (Figure 8a). These effects are mainly observed in the downstream region and align with the main channel (Figure 8b). Without including the boundary condition from the storm surge, the flood simulation underestimates the water depth by up to 2 m in the main channel downstream. This is because streamflow cannot flow out of the domain in the simulation when the storm surge water level is higher than that in the river channel, causing water to accumulate in the domain and leading to a backwater effect. However, the backwater effect cannot propagate to upstream, where the riverbed elevation is higher than the storm surge level. Specifically, the change of water volume in the entire domain strongly correlates with the water level from the boundary condition (Figure 8c). The relative change of the water volume is over 250% around the outlet, and it decreases very quickly from downstream to

upstream, with a negligible impact further than ∼15–20 km from the outlet (Figure 8d). Consequently, for the whole domain, the EXP-BC simulated inundation depth is very close to that of EXP-CTL, with $\rho \approx 1$, RMSE $= 0.02$, $H > 99.9\%$, and $F < 0.1\%$. This suggests that the uncertainty from the boundary conditions in this case study is restricted to the basin outlet, but its impact on the entire basin is minimal compared to the uncertainties from mesh resolution and precipitation sources. Our simulation probably underestimates the backwater effects as only water level was used for the boundary condition, without including other factors, such as momentum exchange, pressure gradient, and wind.

4. Discussion

Although the simulation with 30 m DEM shows satisfactory performance, conducting a long-term simulation with the study domain or a simulation with a larger domain is not computationally feasible. Specifically, the 5-day Harvey flooding simulation with the time step of 0.25 s took about 8.5 hr with 20 CPU nodes (Intel Xeon Skylake CPUs with 192 GB of DDR4 DRAM), with 40 cores per node. The size of the hourly outputs for the 5-day simulation requires a storage of ∼10 Gigabytes. Considering the study domain area is only ∼1,500 km^2, representing 0.22% of Texas state and 0.04% of the continental United States, any large-scale (state or continent level) simulations at this spatial resolution are computationally challenging. Therefore, an overall coarser spatial resolution should be preferred for large-scale flood simulations, and this study provided an evaluation of the loss of accuracy in the upscaling.

The uncertainty of precipitation forcing is another challenge for a large-scale application of a 2D SWE flood model, as current global precipitation data sets exhibit substantial bias in both spatial and temporal scales (Sun et al., 2018). Such precipitation uncertainty can result in high bias in the simulated water depth (Figure 4), as our sensitivity analysis suggests that the simulated water depths at the high water marks are highly sensitive to the changes in precipitation (Figure S5 in Supporting Information S1). Thus, we compared the uncertainties arising due to the mesh resolution (EXP-URM and EXP-VRM) to the uncertainties associated with precipitation (EXP-PREC) to examine the necessity of running the flood model at a high spatial resolution (e.g., <30 m) when the predicted precipitation is highly uncertain. The simulations with URMs at spatial resolutions coarser than 30 m show higher bias in the streamflow than the bias caused by the precipitation uncertainty (Figure 9a, red line vs. symbols). However, using VRMs in 2D flood models can potentially improve the simulated streamflow (e.g., EXP-VRM). Even the coarsest VRM has a satisfactory performance (e.g., NSE >0.5) and is better than any simulations with URMs in Table 1. This suggests that the resolution uncertainty has a larger impact on streamflow than the precipitation uncertainty for URM, but the impact of mesh resolution can be significantly reduced by using a well-designed VRM. The advantage of VRM is to capture critical geometric features of the flooded domain without introducing a large number of grid cells, with the accuracy and efficiency being well balanced. Specifically, the computation time and storage with VRM8 is ∼1,000 times less than using the mesh at 30 m resolution. The computation time with VRM8 can be further reduced by using a larger time step, which is not explored here. In the VRM, accurately representing flooding control structures in the mesh is the first priority to improve the streamflow simulation. The representations of flooding control structures are also important for large-scale simulations because about half of the global rivers are regulated (Nilsson et al., 2005; Schmutz & Moog, 2018).

Relative to streamflow, the inundation dynamics are more sensitive to mesh resolution in our experiments. For example, although all the simulated streamflows in EXP-VRM are well validated, the performance of simulated inundation depth decreases quickly as the mesh resolution decreases (Figure 9b). Using URMs results in higher bias in the simulated maximum inundation depth than precipitation uncertainty. Such bias due to mesh resolution is reduced by using VRMs, though it is still much higher than that caused by precipitation uncertainty at coarser resolution (e.g., VRM5 - VRM8). However, simulations with VRMs show remarkable performance in capturing inundation depth over the refined regions (e.g., main channel), which is not significantly affected by the non-refined regions (Figure 9c and Figure S6 in Supporting Information S1).

Precipitation is found to be the dominant source of uncertainty for the flooding extent indicated by the hit rate (Figure 9d). The hit rates using VRMs are better than those using URMs, which is likely attributed to the improved streamflow dynamics. While the hit rate is not affected by mesh resolution (Figure 9d), the false rate is strongly controlled by the mesh resolution (Figure 9d). Specifically, the flooding model with coarse mesh overestimates the inundation extent as the topography is smoothed, reducing the capacity of the local storage.

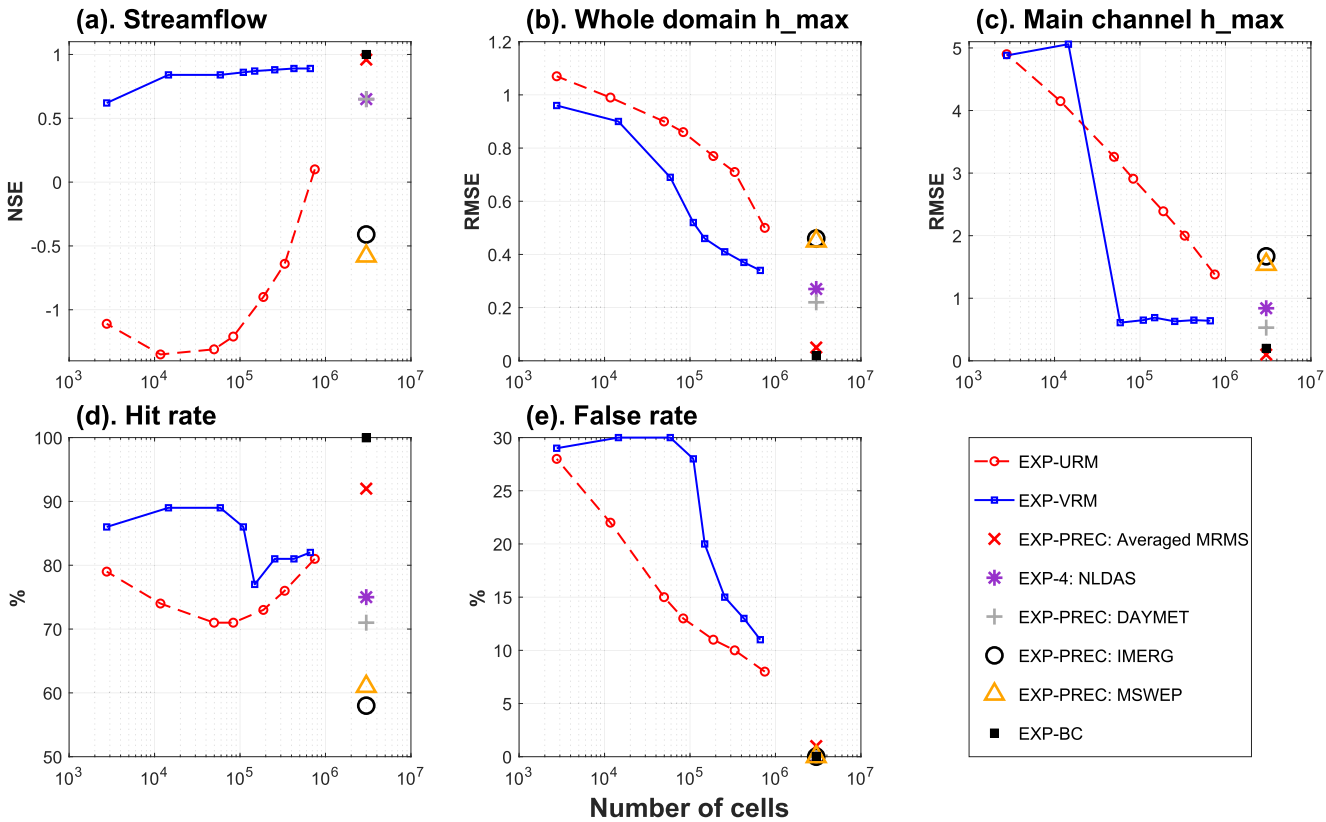

Figure 9. A comparison of simulation experiment performance in terms of (a). Streamflow, (b). Whole domain maximum inundation depth (h_max), (c). Main channel maximum inundation depth, (d) hit rate, and (e) false rate. *X*-Axis denotes the number of grid cells in the mesh that are used in each experiment, and *Y*-Axis denotes the evaluation metric for each experimental simulation against EXP-1. NSE is Nash-Sutcliffe model efficiency, and RMSE is root mean square error.

Although the impact of boundary condition is not as significant as that of precipitation uncertainty and mesh resolution (Figure 9), it results in higher inundation depth within ~15–20 km upstream from the outlet. Such compounding effects can impact broader areas in flatter regions and regions experience stronger storm surges and tidal processes (Feng et al., 2023). In addition, the compounding effect was underestimated in our model as the velocity of the seawater is not available in the measurements and the inertia is ignored.

5. Limitations

We briefly summarize a few limitations of our study. The benchmark model performance at the gauges on tributaries is not as good as that on main channel. Simulation with a finer mesh resolution may further improve the model performance but require considerably longer computational time. Although the selected Manning coefficient in the benchmark simulation is reasonable according to the sensitivity analysis and the high urbanization ratio, a spatially varying Manning coefficient should be used in future application.

All precipitation is assumed to be converted directly to runoff, ignoring hydrological processes (e.g., infiltration, evapotranspiration, etc.). While this simplification might introduce some uncertainty, it is likely minimal for our simulation due to the highly urbanized nature of the study area (Figure S2 in Supporting Information S1) and the large total precipitation amounts. In such scenarios, infiltration and evaporation losses are negligible compared to the total precipitation. This assumption is supported by the close agreement between our simulation results and observations, as the model does not overestimate water levels (Figure 3). However, for less extreme precipitation events than those like Hurricane Harvey, hydrological processes could play a more significant role in flood generation (Berghuijs et al., 2016; Jafarzadegan et al., 2023; Xu et al., 2021). To address this limitation, future work will focus on coupling the 2D SWE model with a land surface model.

Another limitation of this study is excluding stormwater infrastructure and dam operations from the model. Hossain Anni et al. (2020) found stormwater infrastructures can reduce the flooding volume by up to 20%, although their mitigation effects decrease for extreme events. Dam operations can significantly reduce flooding damage as well (Jafarzadegan et al., 2023; Shrestha & Kawasaki, 2020). While our model does account for the reduction in flood levels due to dams, as the dam elevations are well represented in the mesh, the absence of dynamic dam operation could result in significant biases during the flood recession period when floodwaters are released from reservoirs.

Additionally, our simulation may underestimate compounding effects at the river-ocean interface, as only water level data was used as the boundary condition. Other critical drivers for the compounding effects, such as momentum exchange (Santiago-Collazo et al., 2024), pressure gradient, and wind effects, were not considered. In future work, this limitation should be addressed by fully coupling OFM with a storm surge model (Jafarzadegan et al., 2023; Santiago-Collazo et al., 2019), which will exchange mass and momentum simultaneously. The wind effects on the flood wave propagation can be incorporated into the model by adding a momentum source in the governing equations (e.g., Equation 2 and Equation 3) as a function of wind velocity.

6. Conclusions

In this study, we examined the impacts of different uncertainty sources on flooding simulation that solves 2D SWE to identify configurations that balance the simulation accuracy and computational efficiency. The uncertainty sources include the mesh spatial resolution, the precipitation uncertainty, and the watershed outlet boundary condition. According to the uncertainty types classified in Abbaszadeh et al. (2022), mesh spatial resolution is one component of the model structural uncertainty, and precipitation and outlet boundary conditions correspond to observational and forcing data uncertainty, respectively. Using measured water levels at the outlet as the boundary condition, the simulation captures the nonlinear interactions between the high streamflow and storm surge, resulting in intense downstream flooding. For example, the simulation with boundary condition is smaller than sum of the precipitation induced and surge induced inundation (Figure S7 in Supporting Information S1), which has also been shown in Bilskie and Hagen (2018). However, such compounding effects are minimal when evaluating the streamflow and whole domain inundation dynamics (Figure 9), which is likely because the flooding during Hurricane Harvey was dominated by pluvial and fluvial processes. We acknowledge the compounding effect may have a larger impact in other coastal areas and events. Both mesh resolution and precipitation uncertainty can significantly impact streamflow and inundation dynamics in the flooding simulations. However, the uncertainty associated with the mesh resolution can be mitigated by using a variable resolution within the mesh that is carefully designed to resolve critical topographical features and drainage connectivity. The simulation with VRM with an equivalent spatial resolution of 1,000 m cannot accurately simulate the maximum water depth but can simulate the streamflow and inundation extent relatively well. Considering that the precipitation uncertainty results in larger biases, increasing the mesh resolution from ∼1,000–30 m in a 2D hydrodynamic model may not significantly improve the model performance. Furthermore, a spatial resolution of 1,000 m is computationally feasible from the continental to global scales as demonstrated in previous studies using 1,000 m land surface model simulations (L. Li et al., 2023; O'Neill et al., 2021; Yang et al., 2023). This encourages the application of 2D SWE in simulating flood events, and surface water dynamics at large scales (e.g., continental, and global) instead of using the 1D simplified physics (Xu et al., 2024; Yamazaki et al., 2011).

Acknowledgments
This work was supported by Scientific Discovery through Advanced Computing 5 (Capturing the Dynamics of Compound Flooding in E3SM), funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research. The Pacific Northwest National Laboratory is operated by Battelle for the U.S. Department of Energy under Contract DE-AC05-76RLO1830. V. Y. Ivanov acknowledges the support of the U.S. National Science Foundation CMMI program award 2053429 and the Department of Defense, Department of the Navy, the Office of Naval Research award N00014-23-1-2735.

Data Availability Statement

The watershed boundary and Digital Elevation Model data were downloaded from https://www.usgs.gov/the-national-map-data-delivery. The parallel OFM code is available at GitHub (Xu, 2024b). The USGS high water marks were downloaded from https://www.hydroshare.org/resource/615d426f70cc4346875c725b4b8fdc59/. The USGS water levels were downloaded from https://waterdata.usgs.gov/nwis/rt. The scripts, USGS observations, and simulation results that were used to plot the figures in the manuscript are available at Xu (2024a).

References

Abbaszadeh, P., Munoz, D. F., Moftakhari, H., Jafarzadegan, K., & Moradkhani, H. (2022). Perspective on uncertainty quantification and reduction in compound flood modeling and forecasting. *IScience*, *25*(10), 105201. https://doi.org/10.1016/j.isci.2022.105201

Ajjur, S. B., & Al-Ghamdi, S. G. (2021). Evapotranspiration and water availability response to climate change in the Middle East and North Africa. *Climatic Change*, *166*(3), 28. https://doi.org/10.1007/s10584-021-03122-z

Alexopoulos, M. J., Müller-Thomy, H., Nistahl, P., Šraj, M., & Bezak, N. (2023). Validation of precipitation reanalysis products for rainfall-runoff modelling in Slovenia. *Hydrology and Earth System Sciences*, *27*(13), 2559–2578. https://doi.org/10.5194/hess-27-2559-2023

Alfieri, L., Bisselink, B., Dottori, F., Naumann, G., de Roo, A., Salamon, P., et al. (2017). Global projections of river flood risk in a warmer world. *Earth's Future*, *5*(2), 171–182. https://doi.org/10.1002/2016ef000485

Amadeo, K. (2018). Hurricane Harvey facts, damage and costs. *Balance*.

Arctur, K. D. (2018). *USGS - Harvey high water marks*. HydroShare.

Balay, S., Abhyankar, S., Adams, M., Brown, J., Brune, P., Buschelman, K., et al. (2019). PETSc users manual.

Bates, P. D., Horritt, M. S., & Fewtrell, T. J. (2010). A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. *Journal of Hydrology*, *387*(1), 33–45. https://doi.org/10.1016/j.jhydrol.2010.03.027

Bates, P. D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., et al. (2021). Combined modeling of US fluvial, pluvial, and coastal flood hazard under current and future climates. *Water Resources Research*, *57*(2), e2020WR028673. https://doi.org/10.1029/2020wr028673

Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., et al. (2019a). Daily evaluation of 26 precipitation datasets using Stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, *23*(1), 207–224. https://doi.org/10.5194/hess-23-207-2019

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., van Dijk, A. I. J. M., et al. (2019b). MSWEP V2 global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bulletin America Meteorology Social*, *100*(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Begnudelli, L., & Sanders, B. F. (2006). Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying. *Journal of Hydraulic Engineering*, *132*(4), 371–384. https://doi.org/10.1061/(asce)0733-9429(2006)132:4(371)

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters*, *43*(9), 4382–4390. https://doi.org/10.1002/2016gl068070

Bilskie, M. V., & Hagen, S. C. (2018). Defining flood zone transitions in low-gradient coastal regions. *Geophysical Research Letters*, *45*(6), 2761–2770. https://doi.org/10.1002/2018gl077524

Bradford, S. F., & Sanders, B. F. (2002). Finite-volume model for shallow-water flooding of arbitrary topography. *Journal of Hydraulic Engineering*, *128*(3), 289–298. https://doi.org/10.1061/(asce)0733-9429(2002)128:3(289)

Caviedes-Voullième, D., García-Navarro, P., & Murillo, J. (2012). Influence of mesh structure on 2D full shallow water equations and SCS Curve Number simulation of rainfall/runoff events. *Journal of Hydrology*, *448–449*, 39–59. https://doi.org/10.1016/j.jhydrol.2012.04.006

Cea, L., & Bladé, E. (2015). A simple and efficient unstructured finite volume scheme for solving the shallow water equations in overland flow applications. *Water Resources Research*, *51*(7), 5464–5486. https://doi.org/10.1002/2014wr016547

Chen, M., Li, Z., Gao, S., Luo, X., Wing, O. E. J., Shen, X., et al. (2021). A comprehensive flood inundation mapping for hurricane Harvey using an integrated hydrological and hydraulic model. *Journal of Hydrometeorology*, *22*(7), 1713–1726.

Cloke, H. L., & Pappenberger, F. (2009). Ensemble flood forecasting: A review. *Journal of Hydrology*, *375*(3), 613–626. https://doi.org/10.1016/j.jhydrol.2009.06.005

Cosgrove, B., Gochis, D., Flowers, T., Dugger, A., Ogden, F., Graziano, T., et al. (2024). NOAA's National Water Model: Advancing operational hydrology through continental-scale modeling. *JAWRA Journal of the American Water Resources Association*, *60*(2), 247–272. https://doi.org/10.1111/1752-1688.13184

Dewals, B., Kitsikoudis, V., Angel Mejía-Morales, M., Archambeau, P., Mignot, E., Proust, S., et al. (2023). Can the 2D shallow water equations model flow intrusion into buildings during urban floods? *Journal of Hydrology*, *619*, 129231.

Doocy, S., Daniels, A., Murray, S., & Kirsch, T. D. (2013). The human impact of floods: A historical review of events 1980-2009 and systematic literature review. *PLoS Curr*, *5*. https://doi.org/10.1371/currents.dis.f4deb457904936b07c09daa98ee8171a

Du, J., Park, K., Dellapenna, T. M., & Clay, J. M. (2019). Dramatic hydrodynamic and sedimentary responses in Galveston Bay and adjacent inner shelf to Hurricane Harvey. *Science of The Total Environment*, *653*, 554–564. https://doi.org/10.1016/j.scitotenv.2018.10.403

Dullaart, J. C. M., Muis, S., Bloemendaal, N., & Aerts, J. C. J. H. (2020). Advancing global storm surge modelling using the new ERA5 climate reanalysis. *Climate Dynamics*, *54*(1), 1007–1021. https://doi.org/10.1007/s00382-019-05044-0

Dullo, T. T., Gangrade, S., Morales-Hernández, M., Sharif, M. B., Kao, S.-C., Kalyanapu, A. J., et al. (2021). Simulation of Hurricane Harvey flood event through coupled hydrologic-hydraulic models: Challenges and next steps. *Journal of Flood Risk Management*, *14*(3), e12716. https://doi.org/10.1111/jfr3.12716

Engwirda, D. (2017). JIGSAW-GEO (1.0): Locally orthogonal staggered unstructured grid generation for general circulation modelling on the sphere. *Geoscientific Model Development*, *10*(6), 2117–2140. https://doi.org/10.5194/gmd-10-2117-2017

Feaster, T. D., & Koenig, T. A. (2017). Field manual for identifying and preserving high-water mark data (No. 2017-1105). *US Geological Survey*.

Feng, D., Passalacqua, P., & Hodges, B. R. (2019). Innovative approaches for geometric uncertainty quantification in an operational oil spill modeling system. *Journal of Marine Science and Engineering*, *7*(8), 259. https://doi.org/10.3390/jmse7080259

Feng, D., Tan, Z., Xu, D., & Leung, L. R. (2023). Understanding the compound flood risk along the coast of the contiguous United States. *Hydrology and Earth System Sciences*, *27*(21), 3911–3934. https://doi.org/10.5194/hess-27-3911-2023

Fraehr, N., Wang, Q. J., Wu, W., & Nathan, R. (2023). Supercharging hydrodynamic inundation models for instant flood insight. *Nature Water*, *1*(10), 835–843. https://doi.org/10.1038/s44221-023-00132-2

Gharbi, M., Soualmia, A., Dartus, D., & Masbernat, L. (2016). Comparison of 1D and 2D hydraulic models for floods simulation on the Medjerda Riverin Tunisia. *Journal of Materials and Environmental Science*, *7*(8), 3017–3026.

Gudmundsson, L., Boulange, J., Do, H. X., Gosling, S. N., Grillakis, M. G., Koutroulis, A. G., et al. (2021). Globally observed trends in mean and extreme river flow attributed to climate change. *Science*, *371*(6534), 1159–1162. https://doi.org/10.1126/science.aba3996

Guo, K., Guan, M., & Yu, D. (2021). Urban surface water flood modelling – A comprehensive review of current models and future challenges. *Hydrology and Earth System Sciences*, *25*(5), 2843–2860. https://doi.org/10.5194/hess-25-2843-2021

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Habibi, H., Awal, R., Fares, A., & Temimi, M. (2021). Performance of Multi-Radar Multi-Sensor (MRMS) product in monitoring precipitation under extreme events in Harris County, Texas. *Journal of Hydrology*, *598*, 126385. https://doi.org/10.1016/j.jhydrol.2021.126385

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., et al. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. https://doi.org/10.1038/nclimate1911

Hirpa, F. A., Salamon, P., Beck, H. E., Lorini, V., Alfieri, L., Zsoter, E., & Dadson, S. J. (2018). Calibration of the global flood awareness system (GloFAS) using daily streamflow data. *Journal of Hydrology*, *566*, 595–606. https://doi.org/10.1016/j.jhydrol.2018.09.052

Hjelmstad, A., Shrestha, A., Garcia, M., & Mascaro, G. (2021). Propagation of radar rainfall uncertainties into urban pluvial flood modeling during the North American monsoon. *Hydrological Sciences Journal*, *66*(15), 2232–2248. https://doi.org/10.1080/02626667.2021.1980216

Hossain Anni, A., Cohen, S., & Praskievicz, S. (2020). Sensitivity of urban flood simulations to stormwater infrastructure and soil infiltration. *Journal of Hydrology*, *588*, 125028. https://doi.org/10.1016/j.jhydrol.2020.125028

Huffman, G. J. Bolvin, D. T., Braithwaite, D., Hsu, K. L., Joyce, R. J., Kidd, C., et al. (2020). Integrated multi-satellite retrievals for the global precipitation measurement (GPM) mission (IMERG). In V. Levizzani, C. Kidd, D. B. Kirschbaum, C. D. Kummerow, K. Nakamura, & F. J. Turk, (Eds.), *Satellite Precipitation Measurement*, *Advances in Global Change Research* (Vol. 67). Springer. https://doi.org/10.1007/978-3-030-24568-9_19

Hu, G., & Franzke, C. L. E. (2020). Evaluation of daily precipitation extremes in reanalysis and gridded observation-based data sets over Germany. *Geophysical Research Letters*, *47*(18), e2020GL089624. https://doi.org/10.1029/2020gl089624

Huang, W., Ye, F., Zhang, Y. J., Park, K., Du, J., Moghimi, S., et al. (2021). Compounding factors for extreme flooding around Galveston Bay during Hurricane Harvey. *Ocean Modelling*, *158*, 101735. https://doi.org/10.1016/j.ocemod.2020.101735

Huang, X., & Swain, D. L. (2022). Climate change is increasing the risk of a California megaflood. *Science Advances*, *8*(32), eabq0995. https://doi.org/10.1126/sciadv.abq0995

Huong, H. T. L., & Pathirana, A. (2013). Urbanization and climate change impacts on future urban flooding in Can Tho city, Vietnam. *Hydrology and Earth System Sciences*, *17*(1), 379–394. https://doi.org/10.5194/hess-17-379-2013

Iliadis, C., Glenis, V., & Kilsby, C. (2024). Representing buildings and urban features in hydrodynamic flood models. *Journal of Flood Risk Management*, *17*(1), e12950. https://doi.org/10.1111/jfr3.12950

Ivanov, V. Y., Xu, D., Dwelle, M. C., Sargsyan, K., Wright, D. B., Katopodes, N., et al. (2021). Breaking down the computational barriers to real-time urban flood forecasting. *Geophysical Research Letters*, *48*(20), e2021GL093585. https://doi.org/10.1029/2021gl093585

Jafarzadegan, K., Moradkhani, H., Pappenberger, F., Moftakhari, H., Bates, P., Abbaszadeh, P., et al. (2023). Recent advances and new frontiers in riverine and coastal flood modeling. *Reviews of Geophysics*, *61*(2), e2022RG000788. https://doi.org/10.1029/2022rg000788

Jong, B.-T., Delworth, T. L., Cooke, W. F., Tseng, K.-C., & Murakami, H. (2023). Increases in extreme precipitation over the Northeast United States using high-resolution climate model simulations. *Npj Climate and Atmospheric Science*, *6*(1), 18. https://doi.org/10.1038/s41612-023-00347-w

Karamouz, M., Razmi, A., Nazif, S., & Zahmatkesh, Z. (2017). Integration of inland and coastal storms for flood hazard assessment using a distributed hydrologic model. *Environmental Earth Sciences*, *76*(11), 395. https://doi.org/10.1007/s12665-017-6722-6

Karger, D. N., Wilson, A. M., Mahony, C., Zimmermann, N. E., & Jetz, W. (2021). Global daily 1 km land surface precipitation based on cloud cover-informed downscaling. *Scientific Data*, *8*(1), 307. https://doi.org/10.1038/s41597-021-01084-6

Katopodes, N. D. (2018). *Free-surface flow:: Shallow water dynamics*. Butterworth-Heinemann.

Kauffeldt, A., Wetterhall, F., Pappenberger, F., Salamon, P., & Thielen, J. (2016). Technical review of large-scale hydrological models for implementation in operational flood forecasting schemes on continental level. *Environmental Modelling & Software*, *75*, 68–76. https://doi.org/10.1016/j.envsoft.2015.09.009

Kim, D.-W., Chung, E. G., Kim, K., & Kim, Y. (2022). Impact of riverbed topography on hydrology in small watersheds using Soil and Water Assessment Tool. *Environmental Modelling & Software*, *152*, 105383. https://doi.org/10.1016/j.envsoft.2022.105383

Kim, J., Warnock, A., Ivanov, V. Y., & Katopodes, N. D. (2012). Coupled modeling of hydrologic and hydrodynamic processes including overland and channel flow. *Advances in Water Resources*, *37*, 104–126. https://doi.org/10.1016/j.advwatres.2011.11.009

Knoben, W. J. M., Freer, J. E., & Woods, R. A. (2019). Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores. *Hydrology and Earth System Sciences*, *23*(10), 4323–4331. https://doi.org/10.5194/hess-23-4323-2019

Kumbier, K., Carvalho, R. C., Vafeidis, A. T., & Woodroffe, C. D. (2018). Investigating compound flooding in an estuary using hydrodynamic modelling: A case study from the Shoalhaven River, Australia. *Natural Hazards and Earth System Sciences*, *18*(2), 463–477. https://doi.org/10.5194/nhess-18-463-2018

Li, L., Bisht, G., Hao, D., & Leung, L. Y. R. (2023a). Global 1km land surface parameters for kilometer-scale earth system modeling. *Earth System Science Data Discussions*, *2023*, 1–47.

Li, X., Fu, D., Nielsen-Gammon, J., Gangrade, S., Kao, S.-C., Chang, P., et al. (2023b). Impacts of climate change on future hurricane induced rainfall and flooding in a coastal watershed: A case study on Hurricane Harvey. *Journal of Hydrology*, *616*, 128774. https://doi.org/10.1016/j.jhydrol.2022.128774

Liao, C., Zhou, T., Xu, D., Barnes, R., Bisht, G., Li, H. Y., et al. (2022). Advances in hexagon mesh-based flow direction modeling. *Advances in Water Resources*, *160*, 104099. https://doi.org/10.1016/j.advwatres.2021.104099

Luo, X., Li, H. Y., Leung, L. R., Tesfa, T. K., Getirana, A., Papa, F., & Hess, L. L. (2017). Modeling surface water dynamics in the Amazon Basin using MOSART-inundation v1.0: Impacts of geomorphological parameters and river flow representation. *Geoscientific Model Development*, *10*(3), 1233–1259. https://doi.org/10.5194/gmd-10-1233-2017

Maxwell, R. M., & Condon, L. E. (2016). Connections between groundwater flow and transpiration partitioning. *Science*, *353*(6297), 377–380. https://doi.org/10.1126/science.aaf7891

Maxwell, R. M., Condon, L. E., & Kollet, S. J. (2015). A high-resolution simulation of groundwater and surface water over most of the continental US with the integrated hydrologic model ParFlow v3. *Geoscientific Model Development*, *8*(3), 923–937. https://doi.org/10.5194/gmd-8-923-2015

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Stationarity is dead: Whither water management? *Science*, *319*(5863), 573–574. https://doi.org/10.1126/science.1151915

Ming, X., Liang, Q., Xia, X., Li, D., & Fowler, H. J. (2020). Real-time flood forecasting based on a high-performance 2-D hydrodynamic model and numerical weather predictions. *Water Resources Research*, *56*(7), e2019WR025583. https://doi.org/10.1029/2019wr025583

Morales-Hernández, M., Sharif, M. B., Kalyanapu, A., Ghafoor, S. K., Dullo, T. T., Gangrade, S., et al. (2021). TRITON: A multi-GPU open source 2D hydrodynamic flood model. *Environmental Modelling & Software*, *141*, 105034. https://doi.org/10.1016/j.envsoft.2021.105034

Muñoz, D. F., Moftakhari, H., & Moradkhani, H. (2024). Quantifying cascading uncertainty in compound flood modeling with linked process-based and machine learning models. *Hydrology and Earth System Sciences Discussions*, *2024*(11), 1–45. https://doi.org/10.5194/hess-28-2531-2024

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I — A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Neal, J., Schumann, G., Fewtrell, T., Budimir, M., Bates, P., & Mason, D. (2011). Evaluating a new LISFLOOD-FP formulation with data from the summer 2007 floods in Tewkesbury, UK. *Journal of Flood Risk Management*, *4*(2), 88–95. https://doi.org/10.1111/j.1753-318x.2011.01093.x

Nilsson, C., Reidy, C. A., Dynesius, M., & Revenga, C. (2005). Fragmentation and flow regulation of the world's large river systems. *Science*, *308*(5720), 405–408. https://doi.org/10.1126/science.1107887

Olbert, A. I., Moradian, S., Nash, S., Comer, J., Kazmierczak, B., Falconer, R. A., & Hartnett, M. (2023). Combined statistical and hydrodynamic modelling of compound flooding in coastal areas - Methodology and application. *Journal of Hydrology*, *620*, 129383. https://doi.org/10.1016/j.jhydrol.2023.129383

O'Neill, M. M. F., Tijerina, D. T., Condon, L. E., & Maxwell, R. M. (2021). Assessment of the ParFlow–CLM CONUS 1.0 integrated hydrologic model: Evaluation of hyper-resolution water balance components across the contiguous United States. *Geoscientific Model Development*, *14*(12), 7223–7254. https://doi.org/10.5194/gmd-14-7223-2021

Read, L. K., Yates, D. N., McCreight, J. M., Rafieeinasab, A., Sampson, K., & Gochis, D. J. (2023). Development and evaluation of the channel routing model and parameters within the National Water Model. *JAWRA Journal of the American Water Resources Association*, *59*(5), 1051–1066. https://doi.org/10.1111/1752-1688.13134

Saint-Amand, A., Lambrechts, J., Thomas, C. J., & Hanert, E. (2023). How fine is fine enough? Effect of mesh resolution on hydrodynamic simulations in coral reef environments. *Ocean Modelling*, *186*, 102254. https://doi.org/10.1016/j.ocemod.2023.102254

Sampson, C. C., Smith, A. M., Bates, P. D., Neal, J. C., Alfieri, L., & Freer, J. E. (2015). A high-resolution global flood hazard model. *Water Resources Research*, *51*(9), 7358–7381. https://doi.org/10.1002/2015wr016954

Sanders, B. F., Schubert, J. E., Goodrich, K. A., Houston, D., Feldman, D. L., Basolo, V., et al. (2020). Collaborative modeling with fine-resolution data enhances flood awareness, minimizes differences in flood perception, and produces actionable flood maps. *Earth's Future*, *8*(1), e2019EF001391. https://doi.org/10.1029/2019ef001391

Santiago-Collazo, F. L., Bilskie, M. V., Bacopoulos, P., & Hagen, S. C. (2024). Compound inundation modeling of a 1-D idealized coastal watershed using a reduced-physics approach. *Water Resources Research*, *60*(5), e2023WR035718. https://doi.org/10.1029/2023wr035718

Santiago-Collazo, F. L., Bilskie, M. V., & Hagen, S. C. (2019). A comprehensive review of compound inundation models in low-gradient coastal watersheds. *Environmental Modelling & Software*, *119*, 166–181. https://doi.org/10.1016/j.envsoft.2019.06.002

Sarhadi, A., Rousseau-Rizzi, R., Mandli, K., Neal, J., Wiper, M. P., Feldmann, M., & Emanuel, K. (2024). Climate change contributions to increasing compound flooding risk in New York city. *Bulletin America Meteorology Social*, *105*(2), E337–E356. https://doi.org/10.1175/bams-d-23-0177.1

Schmutz, S., & Moog, O. (2018). Dams: Ecological impacts and management, Riverine ecosystem management: Science for governing towards a sustainable future (pp. 111–127).

Schrapffer, A., Sörensson, A., Polcher, J., & Fita, L. (2020). Benefits of representing floodplains in a land surface model: Pantanal simulated with ORCHIDEE CMIP6 version. *Climate Dynamics*, *55*(5), 1303–1323. https://doi.org/10.1007/s00382-020-05324-0

Schubert, J. E., Sanders, B. F., Smith, M. J., & Wright, N. G. (2008). Unstructured mesh generation and landcover-based resistance for hydrodynamic modeling of urban flooding. *Advances in Water Resources*, *31*(12), 1603–1621. https://doi.org/10.1016/j.advwatres.2008.07.012

Schwanghart, W., & Scherler, D. (2014). Short communication: TopoToolbox 2 – MATLAB-based software for topographic analysis and modeling in Earth surface sciences. *Earth Surface Dynamics*, *2*(1), 1–7. https://doi.org/10.5194/esurf-2-1-2014

Sharifian, M. K., Kesserwani, G., Chowdhury, A. A., Neal, J., & Bates, P. (2023). LISFLOOD-FP 8.1: New GPU-accelerated solvers for faster fluvial/pluvial flood simulations. *Geoscientific Model Development*, *16*(9), 2391–2413. https://doi.org/10.5194/gmd-16-2391-2023

Shrestha, B. B., & Kawasaki, A. (2020). Quantitative assessment of flood risk with evaluation of the effectiveness of dam operation for flood control: A case of the Bago River Basin of Myanmar. *International Journal of Disaster Risk Reduction*, *50*, 101707. https://doi.org/10.1016/j.ijdrr.2020.101707

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., & Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, *56*(1), 79–107. https://doi.org/10.1002/2017rg000574

Tabari, H. (2020). Climate change impact on flood and extreme precipitation increases with water availability. *Scientific Reports*, *10*(1), 13768. https://doi.org/10.1038/s41598-020-70816-2

Tanaka, T., Kiyohara, K., & Tachikawa, Y. (2020). Comparison of fluvial and pluvial flood risk curves in urban cities derived from a large ensemble climate simulation dataset: A case study in Nagoya, Japan. *Journal of Hydrology*, *584*, 124706. https://doi.org/10.1016/j.jhydrol.2020.124706

Tate, J. (2017). *Storm summary number 14 for tropical storm Harvey rainfall and wind*. Weather Prediction Center, United States Government.

Thornton, P. E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., & Wilson, B. E. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data*, *8*(1), 190. https://doi.org/10.1038/s41597-021-00973-0

Vacondio, R., Dal Palù, A., & Mignosa, P. (2014). GPU-enhanced Finite Volume Shallow Water solver for fast flood simulations. *Environmental Modelling & Software*, *57*, 60–75. https://doi.org/10.1016/j.envsoft.2014.02.003

Valle-Levinson, A., Olabarrieta, M., & Heilman, L. (2020). Compound flooding in Houston-Galveston Bay during Hurricane Harvey. *Science of The Total Environment*, *747*, 141272. https://doi.org/10.1016/j.scitotenv.2020.141272

Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., et al. (2020). A multiscale, hydrometeorological forecast evaluation of national water model forecasts of the May 2018 Ellicott City, Maryland, flood. *Journal of Hydrometeorology*, *21*(3), 475–499. https://doi.org/10.1175/jhm-d-19-0125.1

Vivoni, E. R., Ivanov, V. Y., Bras, R. L., & Entekhabi, D. (2004). Generation of triangulated irregular networks based on hydrological similarity. *Journal of Hydrologic Engineering*, *9*(4), 288–302. https://doi.org/10.1061/(asce)1084-0699(2004)9:4(288)

Wing, O. E. J., Bates, P. D., Sampson, C. C., Smith, A. M., Johnson, K. A., & Erickson, T. A. (2017). Validation of a 30 m resolution flood hazard model of the conterminous United States. *Water Resources Research*, *53*(9), 7968–7986. https://doi.org/10.1002/2017wr020917

Wittmann, R., Bungartz, H.-J., & Neumann, P. (2017). High performance shallow water kernels for parallel overland flow simulations based on FullSWOF2D. *Computers & Mathematics with Applications*, *74*(1), 110–125. https://doi.org/10.1016/j.camwa.2017.01.005

Wright, K., Passalacqua, P., Simard, M., & Jones, C. E. (2022). Integrating connectivity into hydrodynamic models: An automated open-source method to refine an unstructured mesh using remote sensing. *Journal of Advances in Modeling Earth Systems*, *14*(8), e2022MS003025. https://doi.org/10.1029/2022ms003025

Wu, H., Kimball, J. S., Mantua, N., & Stanford, J. (2011). Automated upscaling of river networks for macroscale hydrological modeling. *Water Resources Research*, *47*(3), W03517. https://doi.org/10.1029/2009wr008871

Xia, X., & Liang, Q. (2018). A new efficient implicit scheme for discretising the stiff friction terms in the shallow water equations. *Advances in Water Resources*, *117*, 87–97. https://doi.org/10.1016/j.advwatres.2018.05.004

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., et al. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research*, *117*, D03109. https://doi.org/10.1029/2011JD016048

Xu, D. (2024a). Scripts and data for uncertainties in coastal flood simulations using 2D shallow water equations: A case study of the 2017 Harvey event [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.11238099

Xu, D. (2024b). xdongh/ofm_petsc: V1.0 [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.13891851

Xu, D., Bisht, G., Sargsyan, K., Liao, C., & Leung, L. R. (2022). Using a surrogate-assisted Bayesian framework to calibrate the runoff-generation scheme in the energy exascale earth system model (E3SM) v1. *Geoscientific Model Development*, *15*(12), 5021–5043. https://doi.org/10.5194/gmd-15-5021-2022

Xu, D., Bisht, G., Tan, Z., Sinha, E., Di Vittorio, A. V., Zhou, T., et al. (2024). Climate change will reduce North American inland wetland areas and disrupt their seasonal regimes. *Nature Communications*, *15*(1), 2438. https://doi.org/10.1038/s41467-024-45286-z

Xu, D., Bisht, G., Zhou, T., Leung, L. R., & Pan, M. (2022). Development of land-river two-way hydrologic coupling for floodplain inundation in the energy exascale earth system model. *Journal of Advances in Modeling Earth Systems*, *14*(8), e2021MS002772. https://doi.org/10.1029/2021ms002772

Xu, D., Ivanov, V. Y., Li, X., & Troy, T. J. (2021). Peak runoff timing is linked to global warming trajectories. *Earth's Future*, *9*(8), e2021EF002083. https://doi.org/10.1029/2021ef002083

Yamazaki, D., Kanae, S., Kim, H., & Oki, T. (2011). A physically based description of floodplain inundation dynamics in a global river routing model. *Water Resources Research*, *47*(4), W04501. https://doi.org/10.1029/2010wr009726

Yang, C., Tijerina-Kreuzer, D. T., Tran, H. V., Condon, L. E., & Maxwell, R. M. (2023). A high-resolution, 3D groundwater-surface water simulation of the contiguous US: Advances in the integrated ParFlow CONUS 2.0 modeling platform. *Journal of Hydrology*, *626*, 130294. https://doi.org/10.1016/j.jhydrol.2023.130294

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., et al. (2016). Multi-Radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin America Meteorology Social*, *97*(4), 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhou, Y., Wu, W., Nathan, R., & Wang, Q. J. (2021a). A rapid flood inundation modelling framework using deep learning with spatial reduction and reconstruction. *Environmental Modelling & Software*, *143*, 105112. https://doi.org/10.1016/j.envsoft.2021.105112

Zhou, Z., Smith, J. A., Baeck, M. L., Wright, D. B., Smith, B. K., & Liu, S. (2021b). The impact of the spatiotemporal structure of rainfall on flood frequency over a small urban watershed: An approach coupling stochastic storm transposition and hydrologic modeling. *Hydrology and Earth System Sciences*, *25*(9), 4701–4717. https://doi.org/10.5194/hess-25-4701-2021