Journal of Hydrology 625 (2023) 130157

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# Unraveling the 2021 Central Tennessee flood event using a hierarchical multi-model inundation modeling framework☆

Sudershan Gangrade [a,*], Ganesh R. Ghimire [a], Shih-Chieh Kao [a], Mario Morales-Hernández [b], Ahmad A. Tavakoly [c,d], Joseph L. Gutenson [c], Kent H. Sparrow [c], George K. Darkwah [e], Alfred J. Kalyanapu [e], Michael L. Follum [f]

[a] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA
[b] I3A-University of Zaragoza, Spain
[c] U.S. Army Engineer Research and Development Center, Vicksburg, MS 39180, USA
[d] Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD, USA
[e] Tennessee Technological University, Cookeville, TN 38505, USA
[f] Follum Hydrologic Solutions, LLC, Casper, WY 82601, USA

ARTICLE INFO

*Keywords:*
Flood inundation
TRITON
AutoRoute
Central Tennessee floods
Error propagation
High-performance computing

ABSTRACT

Flood prediction systems need hierarchical atmospheric, hydrologic, and hydraulic models to predict rainfall, runoff, streamflow, and floodplain inundation. The accuracy of such systems depends on the error propagation through the modeling chain, sensitivity to input data, and choice of models. In this study, we used multiple precipitation forcings (hindcast and forecast) to drive hydrologic and hydrodynamic models to analyze the impacts of various drivers on the estimates of flood inundation depth and extent. We implement this framework to unravel the August 2021 extreme flooding event that occurred in Central Tennessee, USA. We used two radar-based quantitative precipitation estimates (STAGE4 and MRMS) as well as quantitative precipitation forecasts (QPF) from the National Weather Service Weather Prediction Center (WPC) to drive a series of models in the hierarchical framework, including the Variable Infiltration Capacity (VIC) land surface model, the Routing Application for Parallel Computation of Discharge (RAPID) river routing model, and the AutoRoute and TRITON inundation models. An evaluation with observed high-water marks demonstrates that the framework can reasonably simulate flood inundation. Despite the complex error propagation mechanism of the modeling chain, we show that inundation estimates are most sensitive to rainfall estimates. Most notably, QPF significantly underestimates flood magnitudes and inundations leading to unanticipated severe flooding for all stakeholders involved in the event. Finally, we discuss the implications of the hydrodynamic modeling framework for real-time flood forecasting.

## 1. Introduction

Flooding is one of the most devastating natural disasters in the United States (US) posing a significant threat to society, ecosystem, and infrastructure (Bates et al., 2021; Swain et al., 2020). Flood fatalities, in particular, are a key reason for weather-related deaths in the US accounting for nearly 104 deaths per year since 2010 (NOAA/NWS, 2022). The magnitude and frequency of extreme weather events are projected to intensify due to climate change leading to greater flood risks (e.g., Dankers and Feyen, 2008; Davenport et al., 2021; Swain et al., 2020; Villarini and Zhang, 2020; Wing et al., 2022; Wobus et al., 2021). Extreme weather events are also exacerbated by other factors such as urbanization and land use/land cover change (Hemmati et al., 2021; Konrad, 2003; Rogger et al., 2016; Zhang et al., 2018). Therefore, flood

☆ Notice: This manuscript has been authored by UT-Battelle, LLC, under contract DE-AC05-00OR22725 with the US Department of Energy (DOE). The US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for US government purposes. DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (https://energy.gov/downloads/doe-public-access-plan).

* Corresponding author.
*E-mail address:* gangrades@ornl.gov (S. Gangrade).

https://doi.org/10.1016/j.jhydrol.2023.130157
Received 20 December 2022; Received in revised form 30 June 2023; Accepted 25 August 2023
Available online 14 September 2023

forecasting and prediction are of the utmost importance for timely flood warnings, hazards and risk characterization, emergency preparedness, flood mitigation, and resiliency.

A recent example of one such catastrophic disaster is the 2021 Central Tennessee flood, a widespread flash flood event, triggered as a consequence of extreme precipitation on August 21st, 2021, claiming 21 lives with hundreds of homes and properties damaged (Hineman, 2021; Rick Rojas, 2021). The City of Waverly in Humphrey County and the surrounding areas were severely impacted (England et al., 2022; Hineman, 2021; Rick Rojas, 2021). The rain gauge in the town of McEwen, nearest to Waverly, measured a state-record 24-hour precipitation total of about 17 in. on August 21. Considering a basin size of ~ 800 $km^2$ (300 $mi^2$), the rainfall amount corresponds to approximately 65% of the 24-hour probable maximum precipitation for the Waverly region (Spring et al., 1986). The rainfall amount is roughly 1.5 times larger than the 100-year return period 24-hour rainfall at the Waverly-4 W rain gauge (HDSC, 2021). Despite the severity, this event occurred unexpectedly, decreasing flood warning lead times, and leading to questions about the causes and predictability of such an event. Given the disastrous outcomes and the lack of hydrologic observations/monitoring in the area, the purpose of this study is to provide a retrospective diagnosis of the event and to simulate the flood inundation using an end-to-end hierarchical modeling framework.

Physics-based, hierarchical atmospheric, hydrologic, and hydraulic models (hereinafter referred to as "modeling framework") are desired for reliable flood simulation (e.g., Rodríguez-Rincón et al., 2015) and can help reveal the propagation of uncertainty. The rainfall estimates obtained from quantitative precipitation estimates (QPE), or quantitative precipitation forecasts (QPF) may serve as the primary meteorological input to a hydrologic model. The hydrologic model outputs, such as runoff, can then be used to drive a river routing model to simulate streamflow discharge. Finally, the simulated runoff and/or streamflow can be used to drive hydraulic models to simulate the resulting flood inundation. At each stage of modeling, evaluation can be made to understand the reasonableness of the input/output data. In areas without sufficient observations, this physics-based approach may provide the best available reconstruction of a major hydrologic extreme event.

Despite significant advances in numerical weather prediction and real-time forecasting, flood inundation forecasts inherit significant uncertainties from each component of the modeling framework (Merwade et al., 2008). Past studies have explored some aspects of uncertainties in modeling framework and resulting uncertainties, e.g., the role of inputs such as precipitation on streamflow forecasts (Dymond and Adams, 2019; Ghimire et al., 2022; Ghimire et al., 2021b; Li et al., 2021; e.g., Vivoni et al., 2007), the role of reanalysis runoff and streamflow prediction (Mohanty and Simonovic, 2021; Rajib et al., 2020), different hydraulic models (Hocini et al., 2021; Zarzar et al., 2018), and model inputs, parameters and initial conditions (e.g., Afshari et al., 2018; Fernández-Pato et al., 2016; Pakoksung and Takagi, 2020; Rodríguez-Rincón et al., 2015; Sayama et al., 2015; Zarzar et al., 2018) on inundation predictions/forecasts. Due to the computationally intense nature of the modeling framework, specifically, the two-dimensional (2D) hydrodynamic modeling, the uncertainty evaluation, and characterization are usually conducted at hydrometeorological levels. Only a few studies (e.g., Pakoksung and Takagi, 2020; Rodríguez-Rincón et al., 2015; Sayama et al., 2015; Zarzar et al., 2018) have extended them to inundation predictions. Since these uncertainties can propagate through the modeling framework and lead to biases in the streamflow and inundation prediction/forecasts (Zappa et al., 2011; Li et al., 2021), careful investigation of error propagation at each element of the modeling framework is desired (Rodríguez-Rincón et al., 2015). Particularly for small-scale watersheds and localized areas affected by pluvial floods, such a comprehensive evaluation of error propagation become more important as they are more sensitive to the quality of QPFs in capturing intense rainfall events.

Furthermore, the choice of hydraulic models can also introduce significant differences in flood predictions. Several methods and modeling tools with diverse complexity are available for inundation mapping. Terrain-based, non-physical approaches such as Height Above Nearest Drainage (HAND; Liu et al., 2018; Michael Johnson et al., 2019; Rennó et al., 2008; Viterbo et al., 2020) and AutoRoute (Follum, 2013; Follum et al., 2017; Follum et al., 2020) can be suitable to produce first-order approximations of riverine flooding. These approaches are computationally efficient and can be implemented in data-scarce regions (McGrath et al., 2018) for real-time operations. For more process-based modeling of flood regimes, hydrodynamic inundation models that solve the full shallow water equations are typically utilized. However, 2D inundation modeling can be computationally expensive by several orders of magnitudes compared to low-complexity approaches. Graphics Processing Units (GPU) - accelerated inundation models such as Two-dimensional Runoff Inundation Toolkit for Operational Needs (TRITON; Morales-Hernández et al., 2020; Morales-Hernández et al., 2021) that are tailored to utilize the modern heterogeneous high-performance computing (HPC) system have shown promising computing speed in hierarchical modeling frameworks (Dullo et al., 2021a; Dullo et al., 2021b; Gangrade et al., 2019; Li et al., 2021). For such implementations, a better understanding of the tradeoff between accuracy and computational costs between different processes should also be carefully evaluated.

To fill the above-mentioned knowledge gaps, we 1) develop a multi-model hierarchical modeling framework to reconstruct the August 2021 Central Tennessee flood inundation event, 2) explore the sensitivity of input precipitation (QPEs and QPFs) on flood inundation and evaluate how these differences (errors) propagate through the modeling framework, and 3) assess the accuracy and value of different inundation models of varying complexity on flood predictions/forecasting. In particular, we utilize four forecast and hindcast precipitation datasets, a hydrologic model - Variable Infiltration Capacity (VIC; Liang et al., 1994), a streamflow routing model - Routing Application for Parallel Computation of Discharge (RAPID; David et al., 2011), and two inundation models of varying complexity (AutoRoute and TRITON) to simulate, compare, and validate the simulated Central Tennessee flood event. In addition to the evaluation of the errors (uncertainties), the outcome of this study provides insights into the flood event and a modeling capability for future flood mitigation planning.

## 2. Study area

We consider US Hydrologic Unit Code Subregion (HUC) 0604 which encompasses the Tennessee River Basin below Pickwick Dam (~21,500 $km^2$) and the surrounding areas as the modeling domain (Fig. 1). The elevation of HUC 0604 ranges from approximately 67 m to 418 m, stretching from the hilly Highland Rim in the west and the Cumberland Plateau in the east with the enclosed Central Basin in the middle (TCO, 2022). The basin is dominated by a humid subtropic climate, with average annual precipitation varying between 1,143 mm in the Central Basin to 1,270 – 1,400 mm in the hilly Highland Rim from south to north. The largest precipitation occurs during the winter and early spring due to the more frequent movement of large-scale storms, while the next largest precipitation occurs in midsummer due to active thunderstorms (TCO, 2022). The main stem of the Tennessee River, which drains the entire watershed into the northwest, is mostly regulated by the Tennessee Valley Authority (TVA) providing flood control, and hydropower, among other benefits. The basin is predominantly forest covered (56%) followed by pasture (19%), crop cover (8%), and developed areas (7%), respectively (Figure S1; USGS, 2021a). Waverly, TN, which is the main focus of our study, is located in the center of the modeling domain. The tributary passing Waverly is not under any dam regulation and can be considered as unregulated stream reaches during simulation.

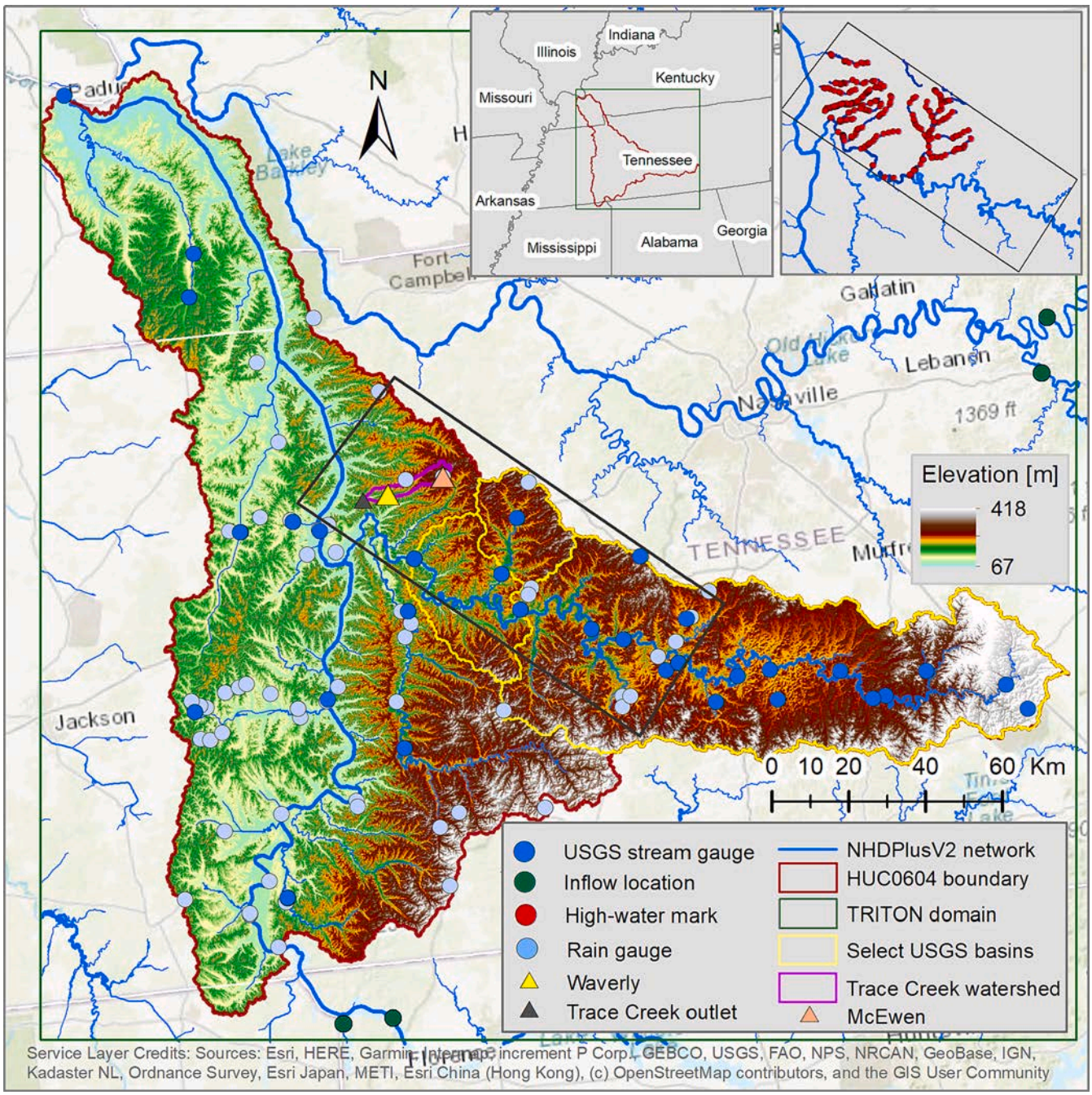


**Fig. 1.** Study area map showing the computational model domain. A solid dark rectangle in the middle depicts the region affected the most by the August 2021 extreme flooding event including Waverly, TN.

## 3. Data and methods

### 3.1. Hierarchical modeling framework

Model predictions are uncertain and can be influenced by a variety of factors, such as natural, unavoidable errors (aleatory) and imperfect knowledge of the physical processes involved in a phenomenon, input data, and limitations in the computational methods (epistemic). For flood modeling, risk assessments, and informing decision-making, it is important to understand how these errors (uncertainties) propagate through the modeling framework. We implement the multi-model hierarchical approach using two forecast and two hindcast precipitation datasets to drive a hydrologic model (VIC) and a river routing model (RAPID) to simulate runoff and streamflow. The simulated runoff and streamflow are then fed into two inundation models of varying complexity (AutoRoute and TRITON). To understand how error (uncertainty) propagates through the modeling framework, we evaluate the modeling skills of precipitation, streamflow, and inundation relative to their respective observations. Fig. 2 presents a graphical summary of the hierarchical modeling framework implemented to investigate the August 2021 Central Tennessee flood. Further details are discussed in the following subsections.

### 3.2. Precipitation inputs

Precipitation is expected to be the most significant input variable for the hierarchical modeling framework. We used two QPEs: STAGE4 (MESONET, 2021; NCEP-EMC, 2021) and MRMS (Muliti-Radar Multi-

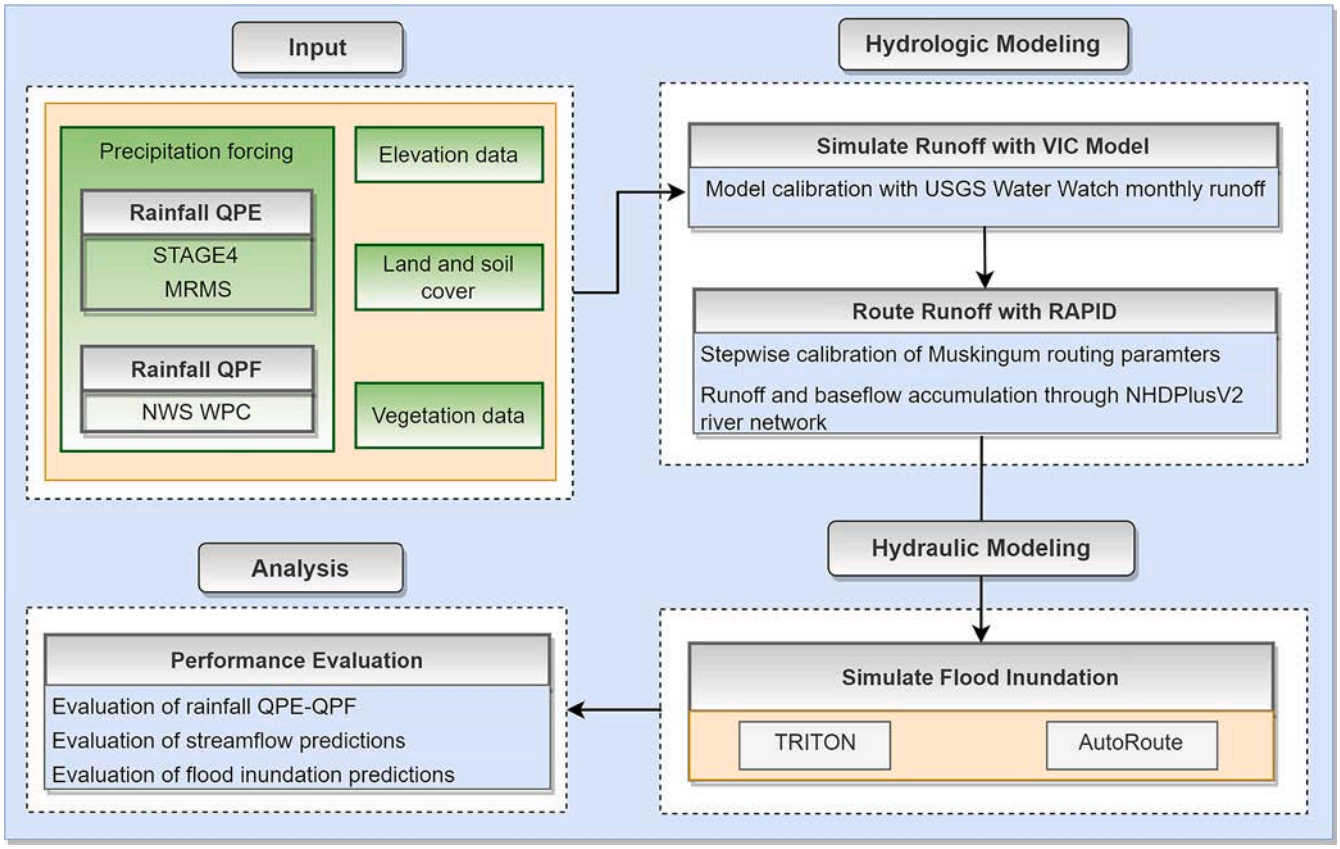


**Fig. 2.** Graphical description of the hierarchical modeling framework used in this study.

Sensor; Zhang et al., 2016), and two QPFs from the National Weather Service (NWS) Weather Prediction Center (WPC; NCEP-NOAA, 2021) to drive the hydrologic and hydraulic models. STAGE4 is a 4-km resolution gridded rainfall product from National Center for Environmental Prediction (NCEP) produced from regional hourly multi-sensor (radar and gauges) precipitation analyses by the 12 River Forecast Centers over the US (NCEP, 2021). MRMS rainfall on the other hand is a 1-km hourly resolution gauge-corrected rainfall product derived from Level II data collected by the WSR-88DP radars. We used the two most recent WPC QPFs prior to the August 21 event, issued 6 h apart at 00:00 and 06:00 UTC, for the simulation referred to as QPF082100 and QPF082106, respectively. These QPFs are for lead times up to 3 days with 6-hour rainfall accumulation intervals.

## 4. Hydrologic and routing model

The VIC model solves water and energy balances to simulate surface runoff, baseflow, evapotranspiration, and other hydrologic processes at each grid cell. We use a set of calibrated parameters derived from Oubeidillah et al. (2014) and Naz et al. (2016) to simulate runoff at each 1/24° (~4 km grid). We then use a calibrated RAPID model to route the simulated total runoff (i.e., VIC surface runoff plus baseflow) through the NHDPlusV2 river network to produce streamflow (McKay et al., 2012). The applicability of VIC to drive RAPID (David et al., 2015; Tavakoly et al., 2021; Tavakoly et al., 2017) has been demonstrated in several previous studies (David et al., 2011; Tavakoly et al., 2017, 2021). In this study, the VIC-RAPID framework is driven by the above-mentioned precipitation inputs spatially interpolated to 1/24° (~4 km) spatial resolution. Temperature, wind speed, elevation, soil, land cover, and vegetation serve as other key inputs (Table 1). For more details on VIC-RAPID, refer to Oubeidillah et al. (2014), Naz et al. (2016), and Ghimire et al. (2023). While the VIC-RAPID modeling framework has been implemented in several studies for streamflow simulations (David et al., 2015; Tavakoly et al., 2023; Tavakoly et al., 2021; Tavakoly et al., 2017), their integration with computationally intensive hydrodynamic model is less explored (Dullo et al., 2021b; Nandi and Reddy, 2022). We use VIC-RAPID to drive two inundation models to demonstrate their applicability for large-scale and high-resolution inundation modeling.

**Table 1**
Summary of input datasets used for hierarchical modeling framework.

| Data | Space-time resolution | Source |
|---|---|---|
| STAGE4 | 4 km – 1 h | National Center for Environmental Prediction (MESONET, 2021; NCEP-EMC, 2021) |
| MRMS | 1 km – 1 h | Zhang et al. (2016) |
| WPC QPF | 5 km – 6 h | National Weather Service Weather Prediction Center (NCEP-NOAA, 2021) |
| Rain gauge data | 1 h | Iowa Environment Mesonet ( MESONET, 2021) |
| CONUS-SOIL | 1 km | Pennsylvania State University (Miller and White, 1998) |
| Land Cover Classification | 1 km | University of Maryland (Hansen et al., 2000) |
| Land Cover | 1 km | USGS (USGS, 2021a) |
| MODIS Leaf Area Index (LAI) | 1 km | NASA (Myneni et al., 2015) |
| USGS National Elevation Dataset (NED) | 10 *m* | USGS (USGS, 2018) |
| NARR Wind Speed | 32 km | NCEP (Mesinger et al., 2006) |
| National Water Information System | 15 min | USGS (USGS, 2021b) |
| NHDPlusV2 | – | USGS (USGS, 2020) |

## 5. Inundation models

### 5.1. AutoRoute

AutoRoute is a raster-based, steady state, uniform flow hydraulic model that can quickly simulate flood hydraulic conditions for a given discharge which can then be fed to a flood mapping software to generate flood depth grids. AutoRoute can be coupled with RAPID (AutoRAPID) to generate flood hydraulic conditions over large-scale vector river networks, such as the NHDPlus (Afshari et al., 2018; Follum, 2013; Follum et al., 2017; Follum et al., 2020). AutoRoute solves Manning's equation at automatically generated cross sections along the river network and requires topography and land cover datasets as inputs. The RAPID streamflow data for each NHDPlus river reach discussed in the previous section provides AutoRoute with the appropriate streamflow data. Because AutoRoute is a steady-state model, the maximum discharge from the RAPID time series is the flow utilized to estimate the flood inundation. The US Geological Survey (USGS) one-third arc-second National Elevation Dataset (NED) digital elevation model (DEM) and the 2016 National Land Cover Database (NLCD; Dewitz, 2019) represent the topography and surface roughness of the study domain, respectively. All roughness coefficients utilized by AutoRoute are consistent with those deployed in Follum et al. (2020). In a manner consistent with Follum et al. (2020), the minimum flow for each river reach in RAPID becomes a baseflow estimate with which AutoRoute estimates a parabolic bathymetry for each cross-section sampled from the DEM. AutoRoute estimates of riverine hydraulic conditions (depth, velocity, and top width) then pass to the flood mapping software, FloodSpreader (or AutoRoute postprocessing script, ARPP) in Follum et al. (2020) to estimate flood inundation depths and extents. The simulation time for AutoRoute/FloodSpreader for the entire domain (21,500 km$^2$) is ~ 4 min using an Intel Core i7-865U processor for one event.

### 5.2. Triton

TRITON (https://triton.ornl.gov) is a 2D open-source flood simulation tool designed for modern GPU-centric HPC. The core of TRITON is a computationally efficient, physics-based hydraulic model that operates on a regular/structured grid and solves the full 2D shallow water equations. TRITON can operate on multiple computer platforms and utilize modern HPC environments, including (1) multi-core shared memory platform using OpenMP, (2) multi-node cluster using MPI or MPI + OpenMP, (3) single node GPU using CUDA, and (4) multi-node GPU cluster using MPI + CUDA, in which the highest TRITON computational efficiency can be achieved by using multi-GPU implementation. TRITON utilizes topographical data (e.g., DEM, LiDAR) as its base input in a uniform (Cartesian) grid structure. The model can be driven by streamflow hydrographs at specified locations or gridded runoff hydrographs, or both which serve as the model's hydrological forcing.

For the TRITON setup, we use a one-third arc-second NED DEM from USGS resampled to 10 m spatial resolution over the TRITON domain (Fig. 1). The domain extends well beyond the area of interest (i.e., the HUC 0604 boundary) to minimize any potential artifacts arising due to boundaries. The setup includes 702 million grid cells (26,000 rows × 27,000 columns). The flood simulation is conducted for 10 days from August 14th, 2021, through August 23rd, 2021, driven by RAPID streamflow outputs at the southern boundary of the TRITON domain, and via a ~ 4 km VIC runoff output distributed over the entire domain (see Fig. 1). We used a spatially varying Manning's n value map which was prepared based on the 2016 NLCD (Dewitz, 2019) as a proxy of ground roughness features following the approach used by Kalyanapu et al. (2009). The downstream boundary conditions are defined for the normal flow conditions using the bed slope value for the outlet of the watershed obtained from the NHDPlus network (USGS, 2020). The simulation outputs are generated every 2 h, with a run time of ~ 2.5 h for one event using 384 GPUs (NVIDIA V100) on a cluster computer supported by the Oak Ridge Leadership Computing Facility.

### 5.3. Observation data

Using the corresponding observations, we evaluate the outputs of each component within the modeling framework (sources are listed in Table 1). We used rainfall data from three networks (MESONET, 2021): NWS Cooperative Observer Program (COOP), National Oceanic and Atmospheric Administration (NOAA) Automated Surface Observing System (ASOS), and Data Collection Platforms (DCP) to evaluate QPE and QPF at co-located rain gauge stations (Figs. 1 & 3). We used 15-min streamflow data from the USGS stream gauge (see Fig. 1; USGS, 2021b). The Trace Creek watershed covering both towns of Waverly and McEwen, which was engulfed by the August 2021 flood, does not have streamflow monitoring gauges. Therefore, for the VIC-RAPID streamflow evaluation, we focus on three USGS gauges from the surrounding areas that were most affected by the flood event (see Fig. 1). Next, we evaluated both TRITON and AutoRoute inundation outputs (resampled to TRITON domain to enable the grid to grid comparison) driven by QPEs and QPFs, against surveyed high-water marks (USACE-Nashville, 2021). These high-water marks (HWM; see Fig. 1) correspond to the locations of maximum inundation on the flood plains of the most affected streams by the 2021 Central Tennessee floods. These HWM data contain an estimate of water surface elevation (WSE) at each HWM. There are about 320 such locations that we used for flood inundation evaluation.

### 5.4. Evaluation metrics

We used several popular metrics to evaluate the performance at different levels of the hierarchical modeling framework. The use of four rainfall QPEs and QPFs, and two inundation models provide 8 unique combinations to assess the error propagation in the modeling chain. For continuous verification, we report Kling-Gupta Efficiency (KGE; Gupta et al., 2009) and Mean Absolute Error (MAE). KGE encompasses three components representing different characteristics: Pearson's correlation ($r$), the mean ratio $\left(\beta = \frac{\mu_s}{\mu_o}\right)$, and the variance ratio $\left(\alpha = \frac{\sigma_s}{\sigma_o}\right)$, where μ and σ denote mean and standard deviation, respectively. The subscripts '$s$' and '$o$' represent simulation (estimation) and observation, respectively. The ideal value of KGE is equal to 1, which is achieved only if each component is equal to 1.

$$KGE = 1 - \sqrt{(r-1)^2 + (\alpha-1)^2 + (\beta-1)^2} \tag{1}$$

MAE is computed as,

$$MAE = \frac{\sum_{i=1}^{N} |Q_s - Q_o|}{N} \tag{2}$$

where $Q_s$ and $Q_o$ represent simulated (estimated) and observed variables, respectively. $N$ represents the length of the time series.

We also report a multiplicative bias, $B$, which depicts a systematic under or over-prediction of the simulation (estimation) relative to the reference. For the uniform space–time grid, $B$ is similar to $\beta$. The ideal value of $B$ is 1.

$$B = \frac{\sum_{x,t}^{N} Q_s(x,t)}{\sum_{x,t}^{N} Q_o(x,t)} \tag{3}$$

where x and $t$ represent the grid location and time, respectively.

To capture the peak flow performance, we compute the percent peak difference (PPD) between the simulated and observed streamflow time series.

$$PPD = 100 \times \frac{Q_{s,peak} - Q_{o,peak}}{Q_{o,peak}} \tag{4}$$

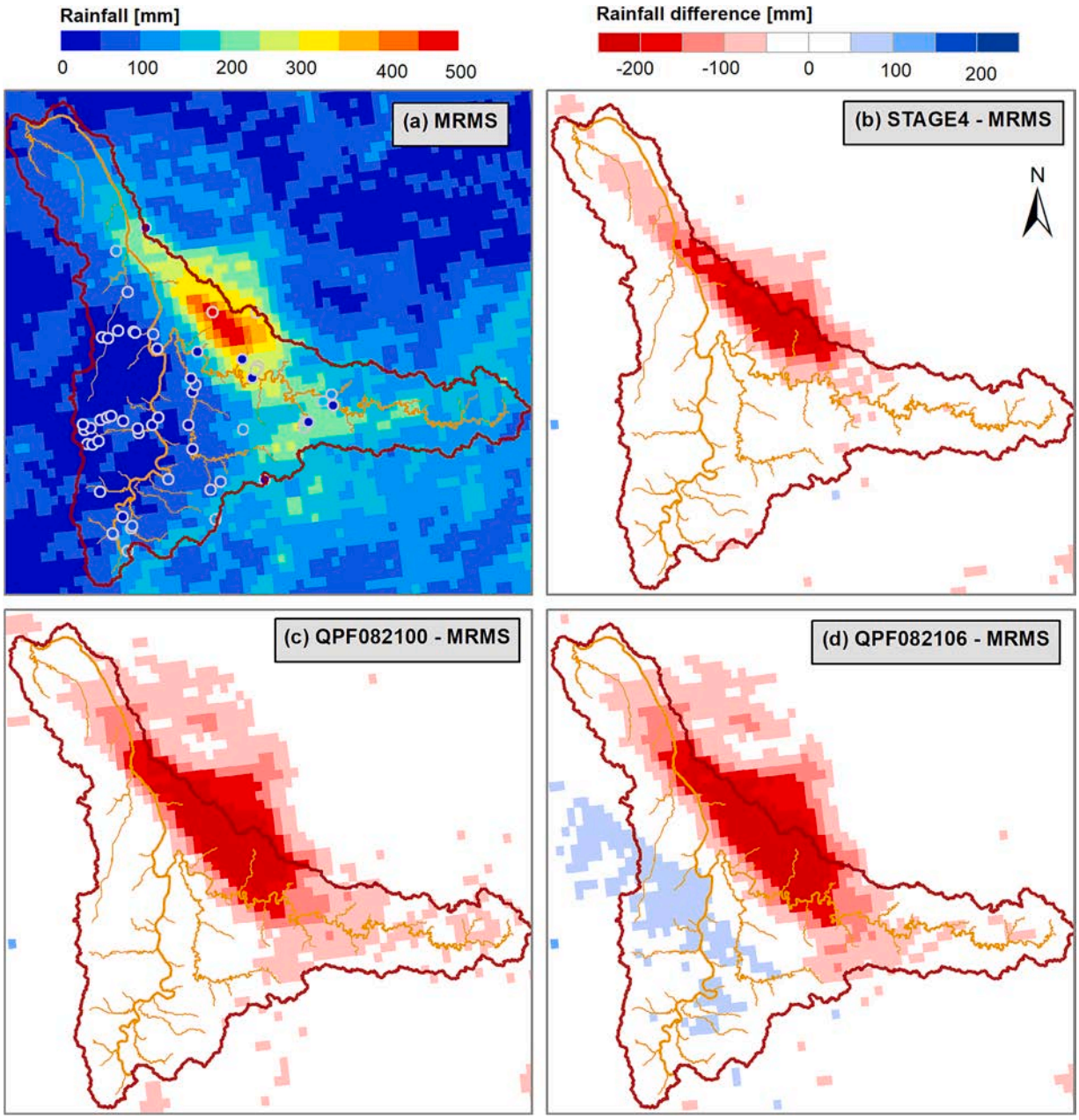


**Fig. 3.** Rainfall differences (error) with respect to MRMS QPE. (a) MRMS rainfall accumulation over the August 2021 extreme event. Color-coded circles show rainfall accumulation at co-located rain gauges. (b), (c), and (d) demonstrate the differences of STAGE4, QPF082100, and QPF082106 rainfall accumulations with respect to MRMS shown in (a).

where $Q_{s,peak,}$ and $Q_{o,peak}$ represent peak flow from simulated and observed streamflow time series, respectively.

We evaluate rainfall and flood inundation using both categorical and continuous verification measures. The categorical verification measures represent the detectability of rainfall/flood inundation and are computed based on rain/no rain (inundation/ no inundation) comparison between the two raster grids using the contingency table presented in Table S1 of Supplementary Information (e.g., CAWCR, 2017; Ghimire et al., 2021b). For categorical verification, we use Hit Rate (HR), False Alarm Ratio (FAR), Frequency Bias (FB), and Gilbert Skill Score (GSS). Table 2 presents a detailed description of these measures. In addition to the error metrics presented above, we use correlation (*R*), and Root Mean Squared Error (*RMSE*) for continuous verification that quantifies the forecasting skill in terms of the magnitude of rainfall/inundation relative to the corresponding reference.

## 6. Results and discussion

### 6.1. Performance evaluation

#### 6.1.1. Rainfall QPE and QPF

Both rainfall QPEs and QPFs are known to have significant errors (uncertainties). Since precipitation is the main driver of the rainfall-runoff-inundation models, understanding how these errors propagate to the model outputs is important. First, we conducted both categorical (detection) and continuous (estimation) verifications for QPEs and QPFs to understand associated errors and uncertainties. A one-to-one comparison of rainfall accumulation at rain gauges and corresponding QPE and QPFs (Fig. S3) reveals that among all four rainfall forcings, MRMS results in the least MAE across all rain gauges followed by STAGE4, QPF082100, and QPF082106 respectively. Fig. 3a shows the spatial

**Table 2**
Binary evaluation metrics for categorical verification of rainfall and inundations.

| Criterion | Formula | Range | Description |
|---|---|---|---|
| **Hit rate (HR)** | $HR = \frac{M1B1}{M1B1 + M0B1}$ | 0 – 1 | Measure of the tendency of the model to accurately predict the benchmark flood extents |
| **False alarm ratio (FAR)** | $FAR = \frac{M1B0}{M1B0 + M1B1}$ | 0 – 1 | Measure of the tendency to overpredict flood extent |
| **Critical success index (CSI)** | $CSI = \frac{M1B1}{M1B1 + M0B1 + M1B0}$ | 0 – 1 | Measure of fit with the penalty for overprediction and underprediction |
| **Error (E)** | $E = \frac{M1B0}{M0B1}$ | 0 – ∞ | Measure of the tendency toward overprediction or underprediction |
| **Frequency bias (FB)** | $FB = \frac{M1B1 + M1B0}{M1B1 + M0B1}$ | 0 – ∞ | Measure of the tendency toward overprediction or underprediction |
| **Gilbert skill score (GSS)** | $GSS = \frac{N}{D}$ | −1/3 – 1 | Measure of the tendency to accurately predict accounting for hits related to random chance |

where.$N = M1B1 \times M0B0 - M0B1 \times M1B0$ $D = (M0B1 + M1B0)(M1B1 + M0B1 + M1B0 + M0B0) + (M1B1 \times M0B0 - M0B1 \times M1B0)$

distribution of MRMS-based rainfall accumulation for the event overlaid with color-coded rainfall gauge accumulated data, suggesting that MRMS captures the total rainfall accumulation pattern well when compared to that of the rain gauges. A comparison of four representative rain gauges in the most affected region shows that MRMS estimates are also closest to the rain gauge observations (Fig. 3a and S3) with bias ($B$) values close to 1. STAGE4 and both QPFs show a significant underestimation of rainfall i.e., $B$ is much smaller than 1 at these locations. Fig. 3 (panels b through d, also see Fig. S2) further shows the spatial distribution of rainfall accumulation differences for STAGE4 and QPFs, relative to MRMS rainfall for the August 2021 event. STAGE4 and two QPFs show underestimation in the most affected region while also showing the spatial shift. Note that MRMS rainfall considered here as a reference is not free of uncertainties either as revealed by the comparison with co-located rain gauge stations. The gridded products will likely include differences due to differences in spatial scale when compared to in-situ observations. Also, note that MRMS implements rain gauge adjustments to its estimation potentially leading to better performance.

For more robust estimates of rainfall errors, while accounting for its spatial variability, we present the grid-based verification results in Fig. 4a. Given the superior performance of MRMS based on the comparison with rain gauge observations (see Fig. S3, and as discussed above), we evaluate STAGE4 and two QPFs against MRMS. During the categorical evaluation, a threshold of 0.5 mm is used, which corresponds to a drizzle, to distinguish between rainfall and no rainfall. All categorical verification measures consistently demonstrate a higher performance of STAGE4 compared to both most recent QPFs (Fig. 4a). Though HR is not significantly different among the three, FB and FAR are significantly higher which explains much higher overall skill scores in terms of GSS by STAGE4. Much larger values of FB for both QPFs suggest that more pixels detect rainfall but significantly underestimate the values as demonstrated by corresponding continuous verification measures $B$, $R$, $RMSE$, and $MAE$ (Fig. 4b-4d) across three USGS monitored basins (see Fig. 1). Among QPFs, the values of $B$, $R$, and $RMSE$ depicted by QPF issued at UTC-00:00 are slightly better than the one issued at UTC-00:06, August 21, which somewhat agrees with the numerical weather prediction model's difficulty with short lead time forecasts to capture initial rainfall amount and distribution (Lin et al., 2005; Seo et al., 2018; Viterbo et al., 2020). Also, note a general spatial scale dependence of verification measures for QPFs, $B$ and $R$ in particular,

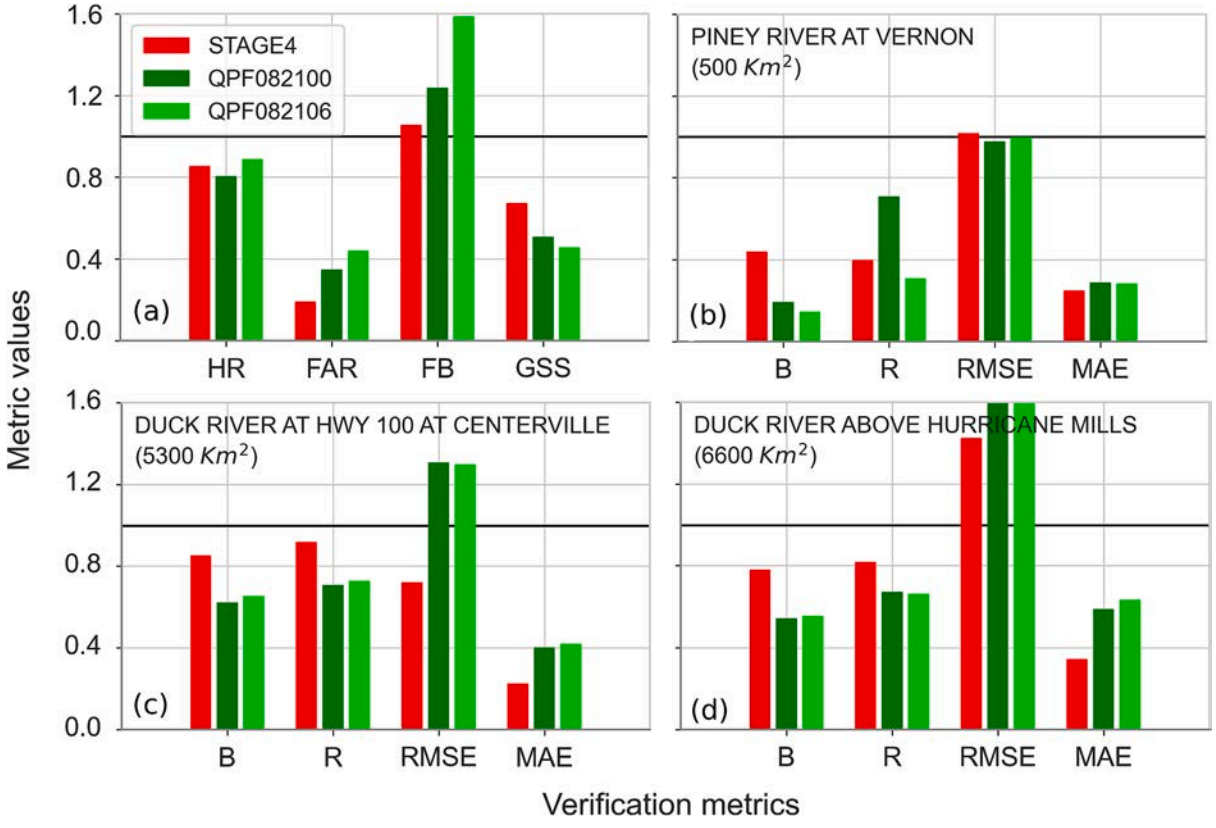


**Fig. 4.** Verification results of rainfall QPEs and QPFs. The verifications of STAGE4, QPF082100, and QPF082106 are with respect to MRMS QPE. (a) shows the computational domain scale categorical verification while (b), (c), and (d) demonstrate the continuous verification measures across three USGS-monitored river basins. RMSE and MAE are in the units of mm.

consistent with results from previous studies (e.g., Ghimire et al., 2021b; Seo et al., 2018).

### 6.2. Streamflow

Similar patterns can be observed in the performance indicators linked to each QPE/QPF forcing as we have observed when evaluating the rainfall forcings across basins (Figs. 4 and 5). The MRMS-driven VIC-RAPID demonstrates a superior performance with a KGE $>$ 0.5 and $B$ close to 1. STAGE4 and QPF-driven VIC-RAPID show significantly lower performance. Most importantly, MRMS captures both the timing and magnitude of the streamflow peaks reasonably, which is crucial from a real-time flood forecasting perspective. The QPF-driven VIC-RAPID, however, does not capture them well. Note that for the smaller basin (e.g., Piney River), the QPFs bias is the largest (Fig. 5), hence propagating to the streamflow forecasts showing the values of both *KGE* and *B* much smaller than 1. Though we do not show it in Fig. 5, the streamflow forecasts for the Trace Creek watershed (~70 km$^2$; Fig. 1) show similar behavior as Piney River with *KGE* = 0.14 and *B* = 0.45 (details are described in Section 4.2). These results further depict the influence of QPE/QPF spatial scale on streamflow forecasting, also demonstrated in Vivoni et al. (2007) and Ghimire et al. (2021b) as well as the sensitivity

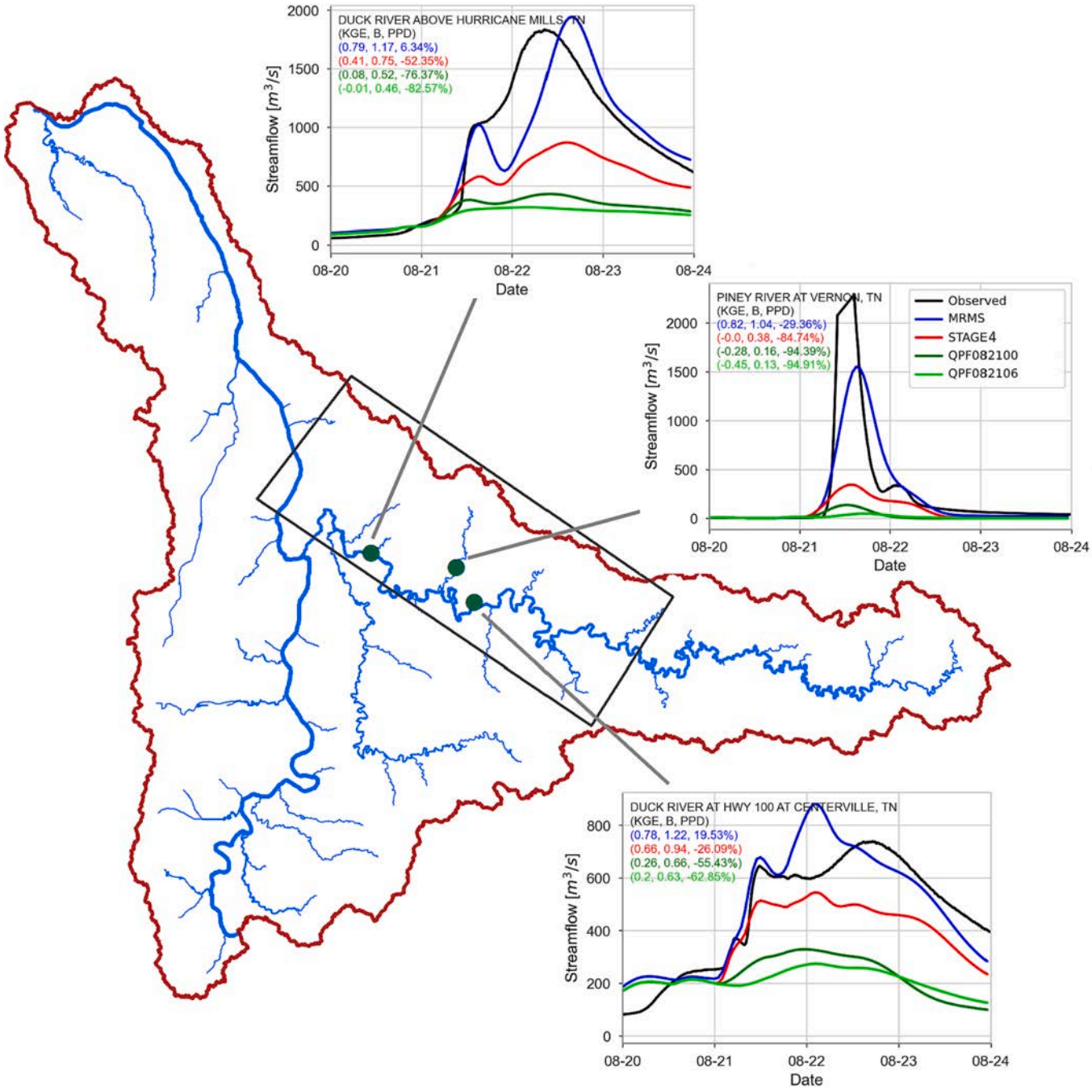


**Fig. 5.** Simulated streamflow hydrographs forced with rainfall QPEs and QPFs in the VIC-RAPID modeling framework. Green dots correspond to three select USGS stream gauges in the region (solid dark rectangle in the middle) most affected by the extreme flooding in August 2021. The values shown on each plot at the top left corner depict corresponding performance metrics in terms of Kling-Gupta Efficiency, multiplicative bias, and percent peak difference (*KGE, B, PPD*). The colors represent respective streamflow hydrographs.

of streamflow to spatial and temporal displacement in meteorological forcings as highlighted in Viterbo et al. (2020).

### 6.3. Flood inundation

In Fig. 6, we overlay the HWM locations on the maximum inundation extents. A binary evaluation is conducted to determine wet pixel fraction (WPF) i.e., a ratio of the number of HWM locations identified as wet/flooded under the simulation (using a threshold of 0.1 m) to the total number of HWM locations. MRMS-driven flood inundation with TRITON achieves the highest performance (WPF $\sim$ 94%) in the region, followed by STAGE4-based inundations (WPF $\sim$ 49%). AutoRoute performs in a similar pattern for MRMS and STAGE4, but with lower performance (WPF $\sim$ 65% and 31%, respectively). Both TRITON and AutoRoute when driven by QPFs result in a lower performance with WPF in the range of 13–25%. The performance of inundation predictions follows a similar pattern as that of rainfall errors and streamflow prediction errors. We discuss more on the error propagation across model cascades in the following section.

WPF-based evaluations alone can overstate the performance and therefore, flood extent comparison alone is not the best benchmark for evaluation (Hunter et al., 2005; Rodríguez-Rincón et al., 2015). For a comprehensive evaluation, we supplement our analysis by comparing simulated inundation depths by TRITON and AutoRoute with water depths at HWM locations. Fig. 7 shows an evaluation of simulated water depths at 320 co-located HWM locations for each rainfall forcing. Both hydrodynamic models can capture the inundation depths reasonably well when driven by MRMS, with MAE in the range of 0.8–1.0 m. With the use of other forcings (i.e., STAGE4 and QPFs), both models result in lower performance (MAE in the range of 1.25–1.50 m). TRITON demonstrates an overall better performance (i.e., smaller *MAE*) relative to AutoRoute.

Further, for a more robust evaluation, we also explore the distribution of the WSE errors across 320 HWM locations. WSE can be more reliable compared to water depth comparison at gauging stations mainly due to the uncertainty in the exact location and elevation of the gauge stations and the uncertainties in the DEM topography. The violin plots in Fig. 8 help capture the variability of the inundation prediction errors. The variability of WSE errors across models and rainfall forcings is almost similar except for the systematic biases. As discussed above, MRMS specifically shows the smallest median WSE error for both models. Systematic underestimation of WSE by the QPFs (absolute median error $>$ 1 m) shows that the rainfall volume bias (underestimation) propagated along model cascades (i.e., to streamflow and inundation forecasts), which is consistent with previous results (Ghimire et al., 2022). This underestimating of inundation estimates largely explains why the risk of the August 2021 flood was misjudged, which led to a surprise for both flood emergency managers and the local population. Also, note that there is a large variability in inundation depth errors which highlights the sensitivity of depth estimations in valley-filling events. Insufficient representation of local geomorphologic characteristics such as slope between floodplain and channel, low-lying areas, and roughness features contribute to such variability in the peak flood magnitudes (Rodríguez-Rincón et al., 2015).

Further, we conduct a grid-by-grid evaluation of flood inundation between TRITON and AutoRoute-generated depth estimates. Assuming AutoRoute as a baseline, we compared TRITON inundation estimates for each of the forcings (e.g., Fig. 9 for MRMS, Figures S5 through S7 for STAGE4 and QPFs). Based on Fig. 9a, the color-coded grid cells in blue indicate all the pixels that can be classified as wet (threshold of 0.1 m) by both AutoRoute and TRITON. The pixels in the red and green represent wet cells for TRITON only and AutoRoute only, respectively. We find that both models can predict wet cells along the main stem of the river in the northwest quadrant of the watershed. The TRITON-only values are generally more prominent in the eastern headwaters and the southern portion along the main channel, with overall Hit Rate = 0.82 and FAR = 0.44. Here, a higher value of FAR can be associated with the fact that AutoRoute simulates fluvial floods only while TRITON simulates both fluvial and pluvial floods. Further, water depth differences in panel (b) and Figure S8 indicate that a major difference is observed along the main channel where TRITON simulated depths are much higher compared to AutoRoute. The difference can be attributed to the fact that TRITON can utilize a downstream boundary condition during the simulation and therefore can simulate the backwater effect.

Overall, TRITON performs better, perhaps due to its high computational demand while being still efficient (run time of 2.5 h for the entire event) to solve the full shallow water equations. AutoRoute also performs reasonably well, especially considering its lower computational demands and high speed (using a single Intel processor and a runtime of $\sim$ 4 min). Similar findings are also reported in other studies such as Hocini et al. (2021) showing an increased grading in model prediction capability between low-complexity models and the ones which solve 2D shallow water equations. Given their unique features, both types of models may provide their unique contributions during operational applications, i.e., a mix of both types of models may provide an effective strategy to balance the needs for speed and accuracy under constrained resources.

### 6.4. Error propagation in the framework

We present the error propagation along the three levels of model cascades in terms of model performance owing to the complexity of disaggregating many sources of uncertainty. In Fig. 10, we show the evolution of the bias, *B,* through model cascades. Errors in rainfall QPFs are generally amplified when propagated to the streamflow predictions as indicated by both median *B* and its variability. Increased variability of *B* in streamflow demonstrates higher sensitivity of the rainfall-runoff process to the precipitation forcings, and basin scales, among other characteristics. Despite increased variability of *B* for streamflow, the QPFs show improved performance in terms of the median. This could be due to the potential QPF random error aggregation by the river network, in addition to the sensitivity of rainfall-runoff processes, particularly for larger basin scales.

The propagation of errors to the inundation simulations by two hydraulic models reveals an interesting pattern. The median bias for MRMS-forced inundation prediction by TRITON is close to 1 demonstrating its potential to reproduce inundation observed during the August 21 event (Fig. 10). STAGE4 and QPFs-forced TRITON simulations generally show underpredictions in a similar pattern as that of precipitation forcings. Similar behavior in bias propagation was also observed by Rodríguez-Rincón et al. (2015). This suggests that bias propagation to streamflow is more sensitive to that of inundation, and further, the bias propagation to streamflow can be reverted in the next modeling step i.e., inundation, and therefore highlights the need to systematically study the modeling frameworks and associated uncertainties. The AutoRoute inundation simulations show a very similar pattern of bias propagation except that they generally show inundation depth underestimation, notably for MRMS. One key explanation for such behavior is due to the ability of TRITON to simulate both fluvial and pluvial floods as opposed to AutoRoute. Note, however, that AutoRoute's ability to reasonably simulate large-scale flood inundations using fewer computational resources makes it a good candidate for real-time flood inundation forecasting. It is particularly true when adequate computational resources are not readily available for TRITON.

We use the Analysis of Variance (ANOVA) approach to identify the dominant sources contributing to the variance of the water surface depths within an eight-member ensemble. ANOVA approach, utilized by several hydrologic modeling studies (Bosshard et al., 2013; Gangrade et al., 2020; Meresa and Romanowicz, 2017), allows for explaining the total variance as a sum of variances from the individual factors and residual error. Here, the two main factors included in the ensembles are the four rainfall forcings and two flood inundation models. The grid-by-

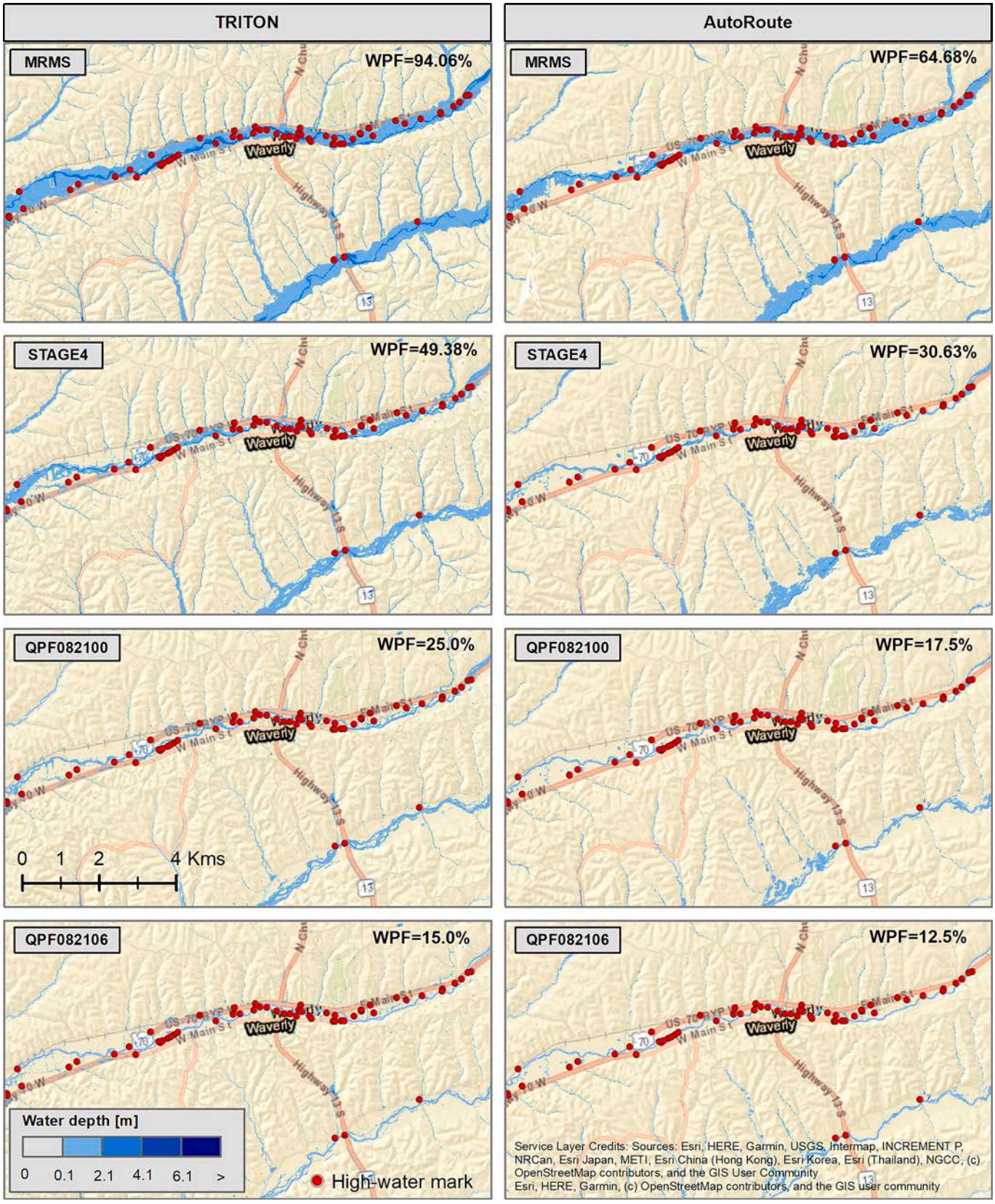


**Fig. 6.** Simulated flood inundation depth and extent comparison between TRITON (left column) and AutoRoute (right column) forced with rainfall QPEs and QPFs. Flood extents are shown for the Waverly region. The values on the top right corner show the wet pixel fraction (WPF) metric. The red dots represent the high-water marks.

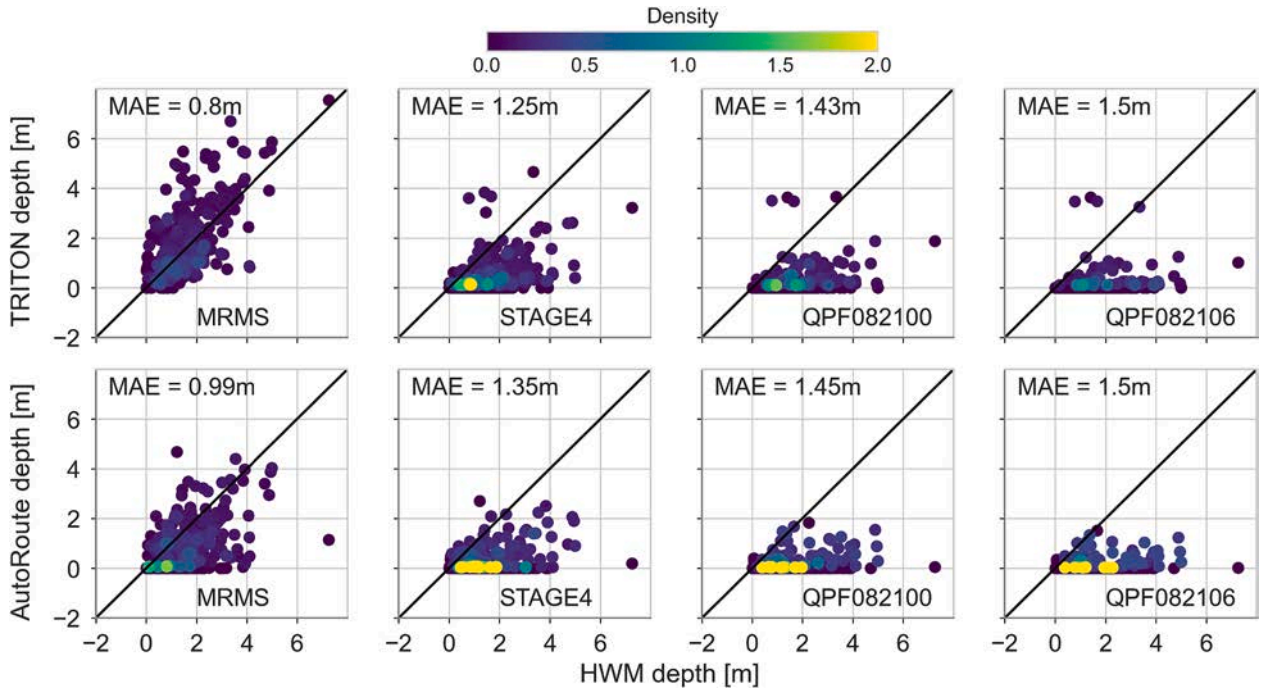


**Fig. 7.** Evaluation of simulated water depths at collocated high-water marks (HWM) locations. Each dot corresponds to the surveyed HWM locations (n = 320). The first row corresponds to the TRITON simulations while the second row represents the AutoRoute simulations. The color depicts the density of points with yellow color pointing to the largest number of data points.

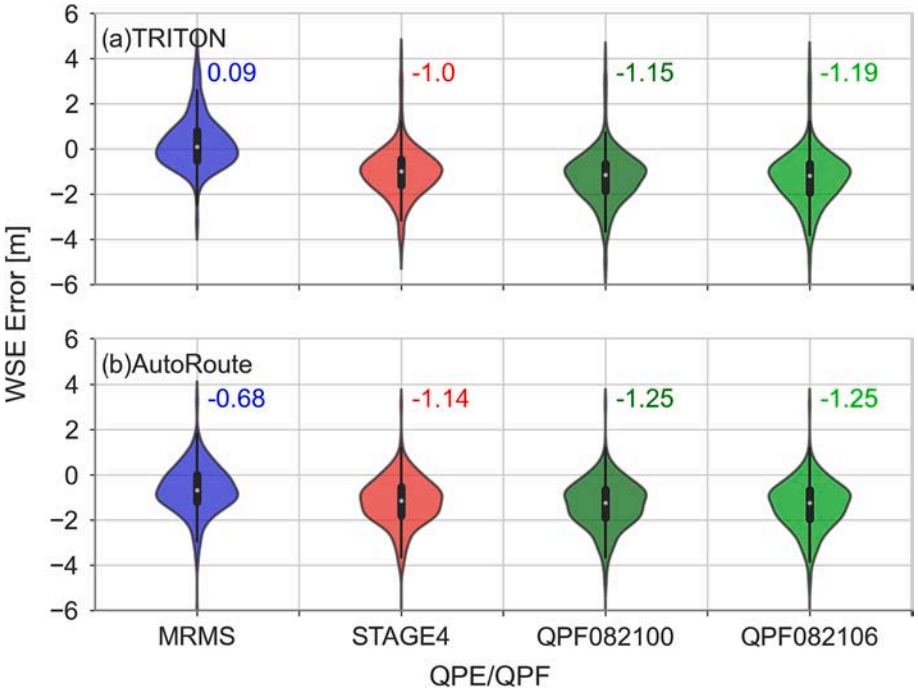


**Fig. 8.** Violin plots demonstrating the variability of water surface elevation (WSE) errors across (a) TRITON and (b) AutoRoute model simulations forced with different precipitation forcings. The color-coded labels represent the median WSE error in [m] i.e., the white dot in each violin where the dark solid line in the middle of each violin represents the corresponding interquartile range.

grid ANOVA results are presented in Fig. 11. The results highlight that for the region along the main river channel (south to north in the TRITON domain), as well as upstream catchment areas (eastern side of the TRITON domain) greater variability is caused by the choice of flood model. However, in the areas like Waverly, McEwen, and Centerville, the choice of meteorological forcings results in greater variability.

### 6.5. *Limitations*

Despite promising results, our experimental setup has some limitations. We use a default model setup for AutoRoute and TRITON. Model calibration and validation can be an important step in improving the accuracy of inundation predictions. More sophisticated approaches to

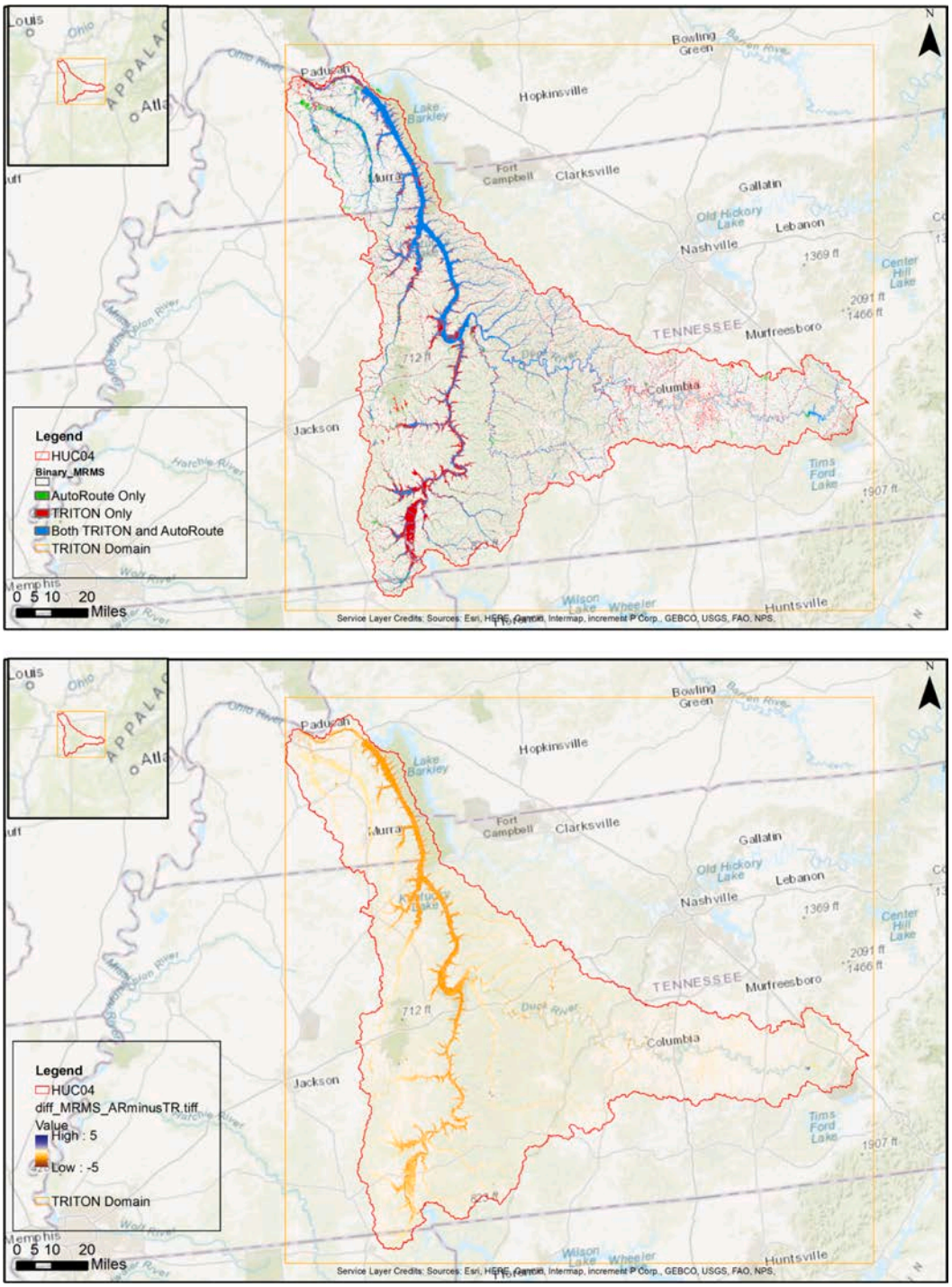


**Fig. 9.** Binary evaluation of flood inundation simulations across TRITON and AutoRoute forced with MRMS QPE. (a) Pixel-to-pixel comparison and (b) water depth difference.

establishing the model's initial conditions, model parameterization, and external river boundary conditions can improve prediction accuracy. As our results in Fig. 10 indicated, inundation depth and extent predictions depend more on the characterization of the river and floodplain than on the characterization of the rainfall-runoff relationship. Approaches to improving the DEM such as burning the large rivers' bathymetry (e.g., Bates et al., 2021), subgrid-scale representation of smaller streams (Neal et al., 2012), local adaptation of the inundation model (e.g., Bates et al., 2021) accounting for many human interventions such as dams, levees, and adaptation infrastructures could lead to improvement in future model applications. Note that this framework does not isolate the uncertainties arising from the hydrologic model, streamflow routing, and observational data as our objective is to explore the error propagation to inundation predictions. Also, the effect of storm direction relative to the basin orientation, shape, and size on the hydrologic response (e.g., Ghimire et al., 2021a; Perez et al., 2021; Seo et al., 2012; Volpi et al., 2013) is not disaggregated. Future evaluations can include additional choices at all levels of the cascading modeling framework.

## 7. Summary and conclusions

In this study, we demonstrate the need to use a physics-based

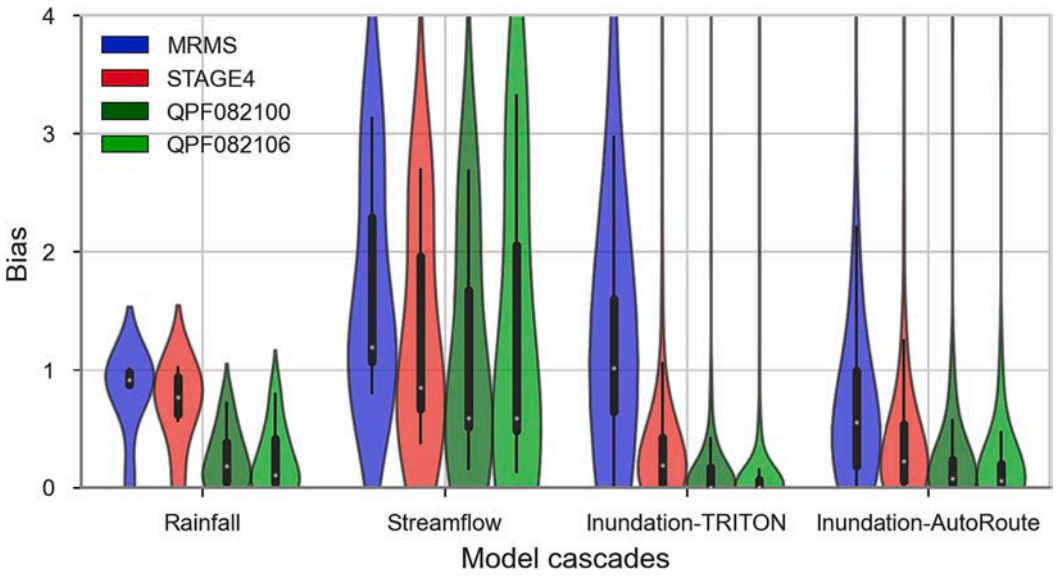


**Fig. 10.** Error propagation in terms of bias, *B* through model cascades of hierarchical modeling framework.

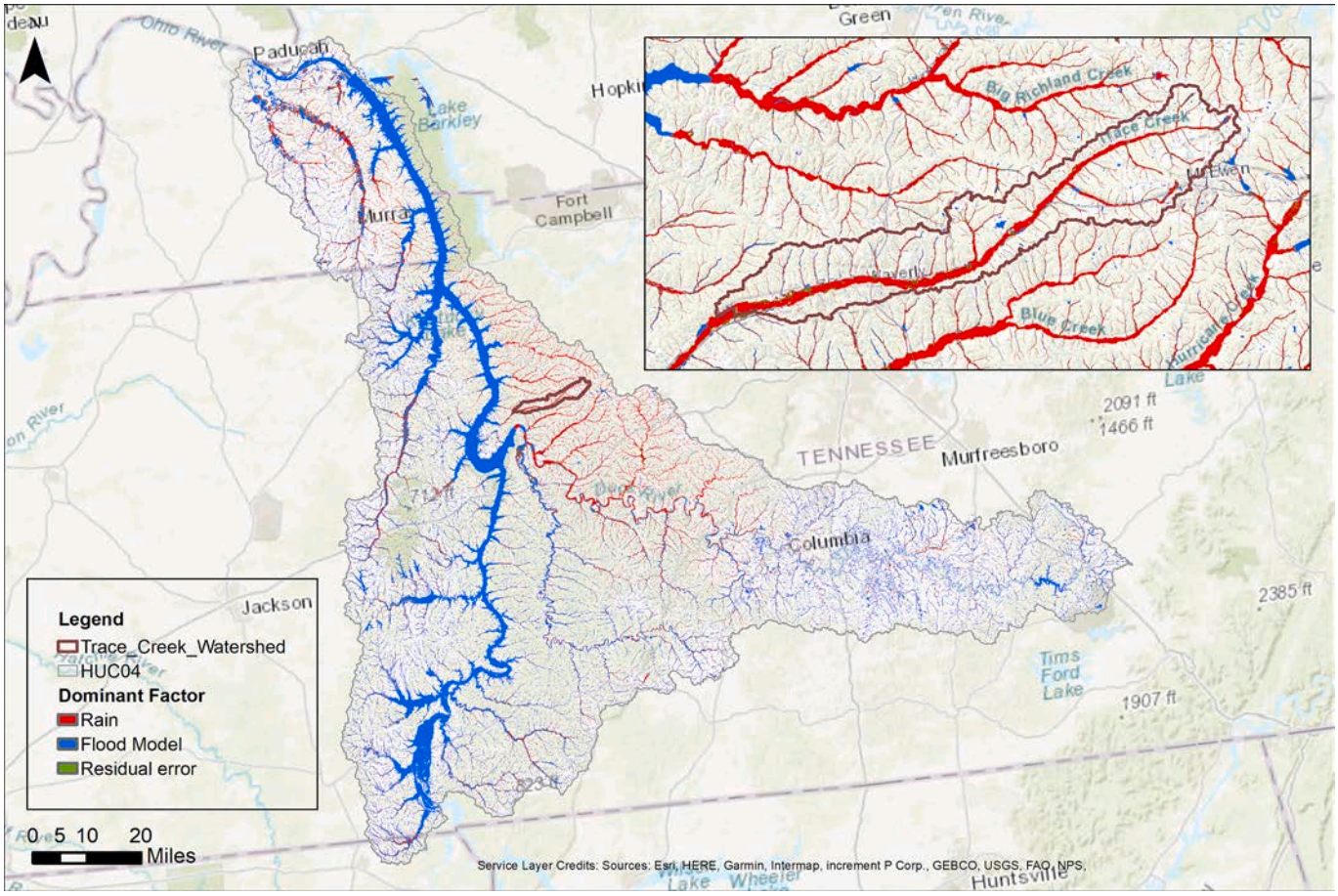


**Fig. 11.** Dominant contributing factor in explaining the total variance of simulated water depths as obtained from ANOVA analysis. The inset shows Trace Creek Watershed.

hierarchical modeling framework that integrates atmospheric, hydrologic, and hydraulics models for accurate flood inundation prediction and forecasting. We implement the framework to unravel an extreme flooding event that occurred in Central Tennessee in August 2021. We use two radar-based QPEs (i.e., STAGE4 and MRMS) and two QPFs from WPC to drive a series of models in the framework, that includes the VIC hydrologic model, RAPID routing model, and AutoRoute and TRITON inundation models. We evaluate skill at each level of the modeling chain and demonstrate the propagation of rainfall errors. We list below some key conclusions derived from the study:

1) Significant rainfall QPF bias was observed for the August 21st event which propagated to the inundation forecasts, resulting in unanticipated severe flooding to both emergency flood managers and the people living in the surrounding area.
2) The flood inundation depth and extent prediction for the event vary across inundation models. However, over most flood-affected areas such as Waverly, rainfall forcings emerge as the most dominant source contributing to the variance of inundation estimates.
3) The hierarchical modeling framework can reproduce flood inundation as well as reveal the propagation of rainfall differences in the inundation modeling chain.
4) There are relative benefits of using varying complexity inundation models in operational settings of flood forecasting, particularly under computational resource constraints.

In addition to providing insights from the event-scale retrospective assessment, the study also provides the region with a new modeling capability for developing future flood mitigation and resilience efforts. The modeling framework presented in this study is scalable and can be applied anywhere, particularly with recent computational advancements.

## CRediT authorship contribution statement

**Sudershan Gangrade:** Conceptualization, Methodology, Writing – review & editing, Visualization, Writing – original draft. **Ganesh R. Ghimire:** Conceptualization, Methodology, Writing – review & editing, Visualization, Writing – original draft. **Shih-Chieh Kao:** Conceptualization, Methodology, Writing – review & editing, Funding acquisition, Writing – original draft. **Mario Morales-Hernández:** Conceptualization, Methodology, Writing – review & editing. **Ahmad A. Tavakoly:** Conceptualization, Methodology, Writing – review & editing. **Joseph L. Gutenson:** Conceptualization, Methodology, Writing – review & editing. **Kent H. Sparrow:** Conceptualization, Methodology, Writing – review & editing. **George K. Darkwah:** Conceptualization, Methodology, Writing – review & editing. **Alfred J. Kalyanapu:** Conceptualization, Methodology, Writing – review & editing. **Michael L. Follum:** Conceptualization, Methodology, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Data will be made available on request.

## Acknowledgment

This study was supported by the US Air Force Numerical Weather Modeling Program. The research used resources of the Oak Ridge Leadership Computing Facility at Oak Ridge National Laboratory (ORNL), which is a Department of Energy (DOE) Office of Science User Facility. The High-Water Marks data was collected and compiled by US Army Corps of Engineers Nashville District. We thank Araz Baranji (USACE) for sharing the high water marks data with us. ORNL is managed by UT-Battelle, LLC for DOE under Contract DE-AC05-00OR22725. Accordingly, the US government retains and the publisher, by accepting the article for publication, acknowledges that the US government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript or allow others to do so, for US Government purposes.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2023.130157.

## References

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

Bates, P.D., et al., 2021. Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates. Water Resources Research 57 (2). https://doi.org/10.1029/2020WR028673.

Bosshard, T., Carambia, M., Goergen, K., Kotlarski, S., Krahe, P., Zappa, M., Schär, C., 2013. Quantifying uncertainty sources in an ensemble of hydrological climate-impact projections. Water Resources Research 49 (3), 1523–1536.https://doi.org/10.1029/2011wr011533.

CAWCR, 2017. WWRP/WGNE Joint Working Group on forecast verification research.

Dankers, R., Feyen, L., 2008. Climate change impact on flood hazard in Europe: An assessment based on high-resolution climate simulations. Journal of Geophysical Research-Atmospheres 113 (D19), 1–17. https://doi.org/10.1029/2007jd009719.

Davenport, F.V., Burke, M., Diffenbaugh, N.S., 2021. Contribution of historical precipitation change to US flood damages. Proc Natl Acad Sci U S A 118 (4), 1–7. https://doi.org/10.1073/pnas.2017524118.

David, C.H., Famiglietti, J.S., Yang, Z.L., Eijkhout, V., 2015. Enhanced fixed-size parallel speedup with the Muskingum method using a trans-boundary approach and a large subbasins approximation. Water Resources Research 51 (9), 7547–7571. https://doi.org/10.1002/2014wr016650.

Dewitz, J., 2019. National Land Cover Database (NLCD) 2016 Products: US Geological Survey data release.

David, C.H., Maidment, D.R., Niu, G.Y., Yang, Z.L., Habets, F., Eijkhout, V., 2011. River network routing on the NHDPlus dataset. Journal of Hydrometeorology 12 (5), 913–934. https://doi.org/10.1175/2011JHM1345.1.

Dullo, T.T., Darkwah, G.K., Gangrade, S., Morales-Hernández, M., Sharif, M.B., Kalyanapu, A.J., Kao, S.-C., Ghafoor, S., Ashfaq, M., 2021a. Assessing climate-change-induced flood risk in the Conasauga River watershed: an application of ensemble hydrodynamic inundation modeling. Natural Hazards and Earth System Sciences 21 (6), 1739–1757.

Dullo, T.T., Gangrade, S., Morales-Hernández, M., Sharif, M.B., Kao, S.-C., Kalyanapu, A. J., Ghafoor, S., Evans, K.J., 2021b. 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.

Dymond, R., Adams, T.E., 2019. The Effect of QPF on Real-Time Deterministic Hydrologic Forecast Uncertainty. Journal of Hydrometeorology 20 (8), 1687–1705. https://doi.org/10.1175/jhm-d-18-0202.1.

England, B., Arms, M., Kalyanapu, A.J., 2022. Understanding Flood Risk: The Trace Creek Watershed.

Fernández-Pato, J., Caviedes-Voullième, D., García-Navarro, P., 2016. Rainfall/runoff simulation with 2D full shallow water equations: Sensitivity analysis and calibration of infiltration parameters. J Hydrol 536, 496–513. https://doi.org/10.1016/j.jhydrol.2016.03.021.

Follum, M.L., Tavakoly, A.A., Niemann, J.D., Snow, A.D., 2017. Autorapid: A Model for Prompt Streamflow Estimation and Flood Inundation Mapping over Regional to Continental Extents. J Am Water Resour As 53 (2), 280–299. https://doi.org/10.1111/1752-1688.12476.

Follum, M.L., Vera, R., Tavakoly, A.A., Gutenson, J.L., 2020. Improved accuracy and efficiency of flood inundation mapping of low-, medium-, and high-flow events using the AutoRoute model. Natural Hazards and Earth System Sciences 20 (2), 625–641. https://doi.org/10.5194/nhess-20-625-2020.

Follum, M.L., 2013. Autoroute rapid flood inundation model, Engineer Research and Development Center Vicksburg MS Coastal and Hydraulics Lab.

Gangrade, S., Kao, S.C., Dullo, T.T., Kalyanapu, A.J., Preston, B.L., 2019. Ensemble-based flood vulnerability assessment for probable maximum flood in a changing environment. J Hydrol 576, 342–355. https://doi.org/10.1016/j.jhydrol.2019.06.027.

Gangrade, S., Kao, S.C., McManamay, R.A., 2020. Multi-model Hydroclimate Projections for the Alabama-Coosa-Tallapoosa River Basin in the Southeastern United States. Sci Rep 10 (1), 2870. https://doi.org/10.1038/s41598-020-59806-6.

Ghimire, G.R., Jadidoleslam, N., Goska, R., Krajewski, W.F., 2021a. Insights into storm direction effect on flood response. J Hydrol 600, 126683.

Ghimire, G.R., Krajewski, W.F., Quintero, F., 2021b. Scale-Dependent Value of QPF for Real-Time Streamflow Forecasting. Journal of Hydrometeorology 22 (7), 1931–1947. https://doi.org/10.1175/Jhm-D-20-0297.1.

Ghimire, G.R., Krajewski, W.F., Ayalew, T.B., Goska, R., 2022. Hydrologic Investigations of Radar-Rainfall Error Propagation to Rainfall-Runoff Model Hydrographs, Advances in Water Resources. Advances in Water Resources 161, 104145. https://doi.org/10.1016/j.advwatres.2022.104145.

Ghimire, G.R., Hansen, C., Gangrade, S., Kao, S.-C., Thornton, P.E., Singh, D., 2023. Insights From Dayflow: A Historical Streamflow Reanalysis Dataset for the Conterminous United States. Water Resources Research 59 (2). https://doi.org/10.1029/2022wr032312.

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. J Hydrol 377 (1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003.
Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Global land cover classification at 1 km spatial resolution using a classification tree. approach 21 (6–7), 1331–1364. https://doi.org/10.1080/014311600210209.
HDSC, 2021. NOAA Atlas 14 Point Precipitation Frequency Estimates. In: National Weather Service, H.D.S.C. (Ed.).
Hemmati, M., Mahmoud, H.N., Ellingwood, B.R., Crooks, A.T., 2021. Unraveling the complexity of human behavior and urbanization on community vulnerability to floods. Sci Rep 11 (1), 20085. https://doi.org/10.1038/s41598-021-99587-0.
Hineman, B., 2021. A look at the deadly Waverly, Tennessee, flood by the numbers, Tennessean.
Hocini, N., Payrastre, O., Bourgin, F., Gaume, E., Davy, P., Lague, D., Poinsignon, L., Pons, F., 2021. Performance of automated methods for flash flood inundation mapping: a comparison of a digital terrain model (DTM) filling and two hydrodynamic methods. Hydrology and Earth System Sciences 25 (6), 2979–2995. https://doi.org/10.5194/hess-25-2979-2021.
Hunter, N.M., Bates, P.D., Horritt, M.S., De Roo, P.J., Werner, M.G.F., 2005. Utility of different data types for calibrating flood inundation models within a GLUE framework. Hydrology and Earth System Sciences 9 (4), 412–430. https://doi.org/10.5194/hess-9-412-2005.
Kalyanapu, A.J., Burian, S.J., McPherson, T.N., 2009. Effect of land use-based surface roughness on hydrologic model output. Journal of Spatial Hydrology 9 (2).
Konrad, C.P., 2003. Effects of Urban Development on Floods. U.S. Geological Survey d, 1–4.
Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415-14428. https://doi.org/10.1029/94JD00483.
Li, X., Rankin, C., Gangrade, S., Zhao, G., Lander, K., Voisin, N., Shao, M., Morales-Hernández, M., Kao, S.-C., Gao, H., 2021. Evaluating precipitation, streamflow, and inundation forecasting skills during extreme weather events: A case study for an urban watershed. J Hydrol 603. https://doi.org/10.1016/j.jhydrol.2021.127126.
Lin, C., Vasić, S., Kilambi, A., Turner, B., Zawadzki, I., 2005. Precipitation forecast skill of numerical weather prediction models and radar nowcasts, Geophysical Research Letters, pp. 1-4. 10.1029/2005GL023451.
Liu, Y.Y., Maidment, D.R., Tarboton, D.G., Zheng, X., Wang, S., 2018. A CyberGIS Integration and Computation Framework for High-Resolution Continental-Scale Flood Inundation Mapping. J Am Water Resour As 54 (4), 770–784. https://doi.org/10.1111/1752-1688.12660.
McGrath, H., Bourgon, J.F., Proulx-Bourque, J.S., Nastev, M., Abo El Ezz, A., 2018. A comparison of simplified conceptual models for rapid web-based flood inundation mapping. Natural Hazards 93 (2), 905–920. https://doi.org/10.1007/s11069-018-3331-y.
McKay, L. et al., 2012. NHD Plus Version 2 : User Guide.
Meresa, H.K., Romanowicz, R.J., 2017. The critical role of uncertainty in projections of hydrological extremes. Hydrology and Earth System Sciences 21 (8), 4245–4258. https://doi.org/10.5194/hess-21-4245-2017.
Merwade, V., Olivera, F., Arabi, M., Edleman, S., 2008. Uncertainty in flood inundation mapping: Current issues and future directions. Journal of Hydrologic Engineering 13 (7), 608–620. https://doi.org/10.1061/(Asce)1084-0699(2008)13:7(608).
Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P.C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E.H., Ek, M.B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., Shi, W., 2006. North American regional reanalysis. Bulletin of the American Meteorological Society 87 (3), 343–360. https://doi.org/10.1175/Bams-87-3-343.
MESONET, 2021. Iowa Environmental Mesonet. In: University, I.S. (Ed.).
Michael Johnson, J., Munasinghe, D., Eyelade, D., Cohen, S., 2019. An integrated evaluation of the National Water Model (NWM)-Height above nearest drainage (HAND) flood mapping methodology. Natural Hazards and Earth System Sciences 2405–2420. https://doi.org/10.5194/nhess-19-2405-2019.
Miller, D.A., White, R.A., 1998. A conterminous United States multilayer soil characteristics dataset for regional climate and hydrology modeling. Earth interactions 2 (2), 1–26. https://doi.org/10.1175/1087-3562(1998)002<0001:ACUSMS>2.3.CO;2.
Mohanty, M.P., Simonovic, S.P., 2021. Fidelity of reanalysis datasets in floodplain mapping: Investigating performance at inundation level over large regions. J Hydrol 597, 125757. https://doi.org/10.1016/j.jhydrol.2020.125757.
Morales-Hernández, M., Sharif, M.B., Gangrade, S., Dullo, T.T., Kao, S.-C., Kalyanapu, A., Ghafoor, S.K., Evans, K.J., Madadi-Kandjani, E., Hodges, B.R., 2020. High-performance computing in water resources hydrodynamics. Journal of Hydroinformatics 22 (5), 1217–1235. https://doi.org/10.2166/hydro.2020.163.
Morales-Hernández, M., Sharif, M.B., Kalyanapu, A., Ghafoor, S.K., Dullo, T.T., Gangrade, S., Kao, S.-C., Norman, M.R., Evans, K.J., 2021. TRITON: A Multi-GPU open source 2D hydrodynamic flood model. Environmental Modelling & Software 141. https://doi.org/10.1016/j.envsoft.2021.105034.
Myneni, R., Knyazikhin, Y., Park, T., 2015. MCD15A2H MODIS/Terra+Aqua Leaf Area Index/FPAR 8-day L4 Global 500m SIN Grid V006. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MCD15A2H.006.
Nandi, S., Reddy, M.J., 2022. An integrated approach to streamflow estimation and flood inundation mapping using VIC. RAPID and LISFLOOD-FP. J Hydrol 610, 127842. https://doi.org/10.1016/j.jhydrol.2022.127842.
NCEP, 2021. National Stage IV QPE Product.
NCEP-EMC, 2021. NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data.
NCEP-NOAA, 2021. Weather Prediction Center (WPC) Quantitative Precipitation Forecasts.
Neal, J., Schumann, G., Bates, P., 2012. A subgrid channel model for simulating river hydraulics and floodplain inundation over large and data sparse areas, Water Resources Research, pp. 1-16. 10.1029/2012WR012514.
NOAA/NWS, 2022. NWS Preliminary US Flood Fatality Statistics.
Pakoksung, K., Takagi, M., 2020. Effect of DEM sources on distributed hydrological model to results of runoff and inundation area. Modeling Earth Systems and Environment 7 (3), 1891–1905. https://doi.org/10.1007/s40808-020-00914-7.
Perez, G., Gomez-Velez, J.D., Mantilla, R., Wright, D.B., Li, Z., 2021. The Effect of Storm Direction on Flood Frequency Analysis, Geophysical Research Letters. Geophysical Research Letters 48 (9). https://doi.org/10.1029/2020GL091918.
Rajib, A., Liu, Z., Merwade, V., Tavakoly, A.A., Follum, M.L., 2020. Towards a large-scale locally relevant flood inundation modeling framework using SWAT and LISFLOOD-FP. J Hydrol 581, 124406. https://doi.org/10.1016/j.jhydrol.2019.124406.
Rennó, C.D., Nobre, A.D., Cuartas, L.A., Soares, J.V., Hodnett, M.G., Tomasella, J., Waterloo, M.J., 2008. HAND, a new terrain descriptor using SRTM-DEM: Mapping terra-firme rainforest environments in Amazonia, Remote Sensing of Environment. Remote Sensing of Environment 112 (9), 3469–3481. https://doi.org/10.1016/j.rse.2008.03.018.
Rick Rojas, T.M.a.J.M., 2021. Tennessee City Searches for Flood Victims While Grappling With All It Has Lost, The New York Times.
Rodríguez-Rincón, J.P., Pedrozo-Acuña, A., Breña-Naranjo, J.A., 2015. Propagation of hydro-meteorological uncertainty in a model cascade framework to inundation prediction. Hydrology and Earth System Sciences 19 (7), 2981–2998. https://doi.org/10.5194/hess-19-2981-2015.
Rogger, M. et al., 2016. Water Resources Research. 5209-5219. 10.1002/2017WR020723.Received.
Sayama, T., Tatebe, Y., Iwami, Y., Tanaka, S., 2015. Hydrologic sensitivity of flood runoff and inundation: 2011 Thailand floods in the Chao Phraya River basin. Natural Hazards and Earth System Sciences 15 (7), 1617–1630. https://doi.org/10.5194/nhess-15-1617-2015.
Seo, B.C., Quintero, F., Krajewski, W.F., 2018. High-resolution QPF uncertainty and its implications for flood prediction: A case study for the eastern Iowa flood of 2016. Journal of Hydrometeorology 1289–1304. https://doi.org/10.1175/JHM-D-18-0046.1.
Seo, Y., Schmidt, A.R., Sivapalan, M., 2012. Effect of storm movement on flood peaks: Analysis framework based on characteristic timescales. Water Resources Research 1–12. https://doi.org/10.1029/2011WR011761.
Spring, S. et al., 1986. Probable Maximum and TVA Precipitation Estimates With Areal Distribution for Tennessee River Drainages .~ess Than 3,000 Mi 2 in Area TENNESSEE VALLEY AUTHORITY TENNESSEE VALLEY AUTHORITY Probable Maximum and TVA Precipitation Estimates With Areal Distribution for Tennessee River Drainages Less Than 3,000 Mi 2 in Area.
Swain, D.L., Wing, O.E.J., Bates, P.D., Done, J.M., Johnson, K.A., Cameron, D.R., 2020. Increased Flood Exposure Due to Climate Change and Population Growth in the United States. Earths. Future 8 (11). https://doi.org/10.1029/2020EF001778.
Tavakoly, A.A., Snow, A.D., David, C.H., Follum, M.L., Maidment, D.R., Yang, Z.-L., 2017. Continental-Scale River Flow Modeling of the Mississippi River Basin Using High-Resolution Nhdplus Dataset. J Am Water Resour As 53 (2), 258–279. https://doi.org/10.1111/1752-1688.12456.
Tavakoly, A.A., Gutenson, J.L., Lewis, J.W., Follum, M.L., Rajib, A., LaHatte, W.C., Hamilton, C.O., 2021. Direct Integration of Numerous Dams and Reservoirs Outflow in Continental Scale Hydrologic Modeling. Water Resources Research 57 (9). https://doi.org/10.1029/2020WR029544.
Tavakoly, A.A., David, C.H., Gutenson, J.L., Wahl, M.W., Follum, M., 2023. Development of non-data driven reservoir routing in the routing application for parallel computatIon of discharge (RAPID) model. Environmental Modelling & Software 161, 105631. https://doi.org/10.1016/j.envsoft.2023.105631.
TCO, 2022. Tennesse Climate Office: Tennessee Climate.
USACE-Nashville, 2021. High-Water Marks Survey by the US Army Corps of Engineers, Nashville District.
USGS, 2018. USGS 3D Elevation Program Digital Elevation Model.
USGS, 2020. National Hydrography Dataset V2.
USGS, 2021a. National Land Cover Database 2019.
USGS, 2021b. USGS Water data for the nation.
Villarini, G., Zhang, W., 2020. Projected changes in flooding: a continental U.S. perspective. Ann N Y Acad Sci 1472 (1), 95–103. https://doi.org/10.1111/nyas.14359.
Viterbo, F., Mahoney, K., Read, L., Salas, F., Bates, B., Elliott, J., Cosgrove, B., Dugger, A., Gochis, D., Cifelli, R., 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., Entekhabi, D., Hoffman, R.N., 2007. Error propagation of radar rainfall nowcasting fields through a fully distributed flood forecasting model. J Appl Meteorol Clim 46 (6), 932–940. https://doi.org/10.1175/Jam2506.1.
Volpi, E., Di Lazzaro, M., Fiori, A., 2013. Analytical modeling of the hydrologic response under moving rainstorms: Storm-catchment interaction and resonance. J Hydrol. Elsevier B.V. 132–139. https://doi.org/10.1016/j.jhydrol.2013.04.025.
Wing, O.E.J., Lehman, W., Bates, P.D., Sampson, C.C., Quinn, N., Smith, A.M., Neal, J.C., Porter, J.R., Kousky, C., 2022. Inequitable patterns of US flood risk in the Anthropocene. Nature. Climate Change 12 (2), 156–162. https://doi.org/10.1038/s41558-021-01265-6.

Wobus, C., Porter, J., Lorie, M., Martinich, J., Bash, R., 2021. Climate change, riverine flood risk and adaptation for the conterminous United States. Environ Res Lett 16 (9), 094034. https://doi.org/10.1088/1748-9326/ac1bd7.

Zappa, M., Jaun, S., Germann, U., Walser, A., Fundel, F., 2011. Superposition of three sources of uncertainties in operational flood forecasting chains. Atmospheric Research 100 (2–3), 246–262. https://doi.org/10.1016/j.atmosres.2010.12.005.

Zarzar, C.M., Hosseiny, H., Siddique, R., Gomez, M., Smith, V., Mejia, A., Dyer, J., 2018. A Hydraulic MultiModel Ensemble Framework for Visualizing Flood Inundation Uncertainty. J Am Water Resour As 54 (4), 807–819. https://doi.org/10.1111/1752-1688.12656.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller, D., 2016. MULTI-RADAR MULTI-SENSOR (MRMS) QUANTITATIVE PRECIPITATION ESTIMATION Initial Operating Capabilities. Bulletin of the American Meteorological Society 97 (4), 621–638. https://doi.org/10.1175/Bams-D-14-00174.1.

Zhang, W., Villarini, G., Vecchi, G.A., Smith, J.A., 2018. Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. Nature 563 (7731), 384–388. https://doi.org/10.1038/s41586-018-0676-z.