Journal of Hydrology: Regional Studies 64 (2026) 103234

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

A hybrid CNN-LSTM surrogate model for hyper-resolution spatiotemporal flood forecasting in Norfolk, Virginia

Yidi Wang [a,b,*], Jonathan L. Goodall [a,b], Chetan Kumar [c], Diana McSpadden [d], Sergio A. Barbosa [e], Binata Roy [a,b], Ali Shahabi [f], Navid Tahvildari [f]

[a] *Department of Civil and Environmental Engineering, University of Virginia, Charlottesville, VA 22904, United States*
[b] *Link Lab, School of Engineering and Applied Science, University of Virginia, Charlottesville, VA 22904, United States*
[c] *Old Dominion University - Thomas Jefferson National Accelerator Facility Joint Institute on Advanced Computing for Environmental Studies, Portsmouth, VA 23703, United States*
[d] *Chief Data Office, Thomas Jefferson National Accelerator Facility, Newport News, VA 23606, United States*
[e] *Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN 37831, United States*
[f] *Department of Civil and Environmental Engineering, Florida International University, Miami, FL 33174, United States*

ARTICLE INFO

Keywords:
Coastal urban flooding
Spatiotemporal forecasting
Deep learning
Hybrid model
Deeplabv3+

ABSTRACT

Study region: Norfolk, Virginia, United States
Study focus: Accurate and timely flood forecasting is essential for enhancing resilience in coastal urban areas in the context of increasing frequency and intensity of rainfall, sea level rise and urbanization. This study presents a hybrid deep learning-based surrogate model that integrates Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to enable real-time spatiotemporal flood forecasting. The model leverages CNN to capture spatial features from inputs such as elevation and Topographic Wetness Index (TWI), while LSTM processes time-series inputs of rainfall and tide data to capture temporal features.
New hydrologic insights for the region: The hybrid CNN-LSTM model was trained using the physics-based hydrodynamic model simulations obtained from the Two-dimensional Unsteady FLOW (TUFLOW) model for Norfolk, Virginia, and achieved high predictive accuracy across diverse flood-prone areas. The reduced computational time from four to six hours using TUFLOW to 3.2 min per event using CNN-LSTM enables rapid flood inundation mapping and early warning applications. The model effectively captured both spatial flood extents and their temporal evolution across different flooding scenarios, providing forecasts at a 2.5-m spatial resolution and 15-min temporal resolution and a one-hour-ahead prediction horizon. While challenges remain in terms of transferability to new regions and real-time data assimilation, this approach demonstrates strong potential for supporting operational flood risk management in coastal urban environments.

1. Introduction

Coastal cities across the globe are becoming increasingly vulnerable to flooding primarily due to increasing frequency and intensity of rainfall, sea level rise and rapid urbanization, which motivates the need to enhance flood resilience for those areas (Hallegatte et al.,

* Corresponding author at: Department of Civil and Environmental Engineering, University of Virginia, Charlottesville, VA 22904, United States.
E-mail address: yw5vq@virginia.edu (Y. Wang).

https://doi.org/10.1016/j.ejrh.2026.103234
Received 5 November 2025; Received in revised form 3 February 2026; Accepted 5 February 2026

2013). Rising sea levels, more intense precipitation events, and impervious urban surfaces exacerbate flood risks, leading to severe economic losses and threats to public safety and critical coastal infrastructure and assets worldwide (Ezer and Atkinson, 2014; National Academies of Sciences, 2019; Hsiao et al., 2024; Su et al., 2024). By 2100, nuisance flooding is projected to occur on a near-daily basis in U.S. coastal regions, including the Atlantic and Gulf coasts, as well as the Pacific Islands (Sweet et al., 2018). The growing frequency and severity of flooding events in coastal urban areas highlight the urgent need for accurate and timely flood forecasting to support early warning mechanisms and effective flood risk management.

Researchers have focused on developing accurate flood prediction models to enable decision-makers to identify potential flood-prone areas in advance and implement proactive measures to reduce the socioeconomic impacts of urban flooding. Traditional physics-based models, or hydrodynamic models in this context, such as Hydrologic Engineering Center's River Analysis System (HEC-RAS) 2D, MIKE FLOOD, Weather Research and Forecasting-Storm Water Management Model (WRF-SWMM), Two-dimensional Unsteady FLOW (TUFLOW), and Delft3D have been widely used for flood prediction in urban environments (Carr and Smith, 2007; Rangari et al., 2019; Wang et al., 2022; Jeong et al., 2025; Han and Tahvildari, 2024). These models come with the capability to couple one dimensional (1D) hydraulics of the drainage pipe systems with two dimensional (2D) hydrodynamics of tides, surge, and pluvial flow and thus offer high accuracy in simulating flood dynamics in complex urban environments where flow equations need to be solved by accounting for factors such as topography, rainfall, surge, tides, infiltration, and drainage capacity. While these physics-based models provide reliable flood predictions, they require extensive computational resources and long runtimes, which can range from several hours to days for a single flood event, depending on the spatial resolution and domain size (Zahura et al., 2020). The high computational cost of such models makes them impractical for real-time flood forecasting, particularly in rapidly evolving flood scenarios where timely predictions are crucial for emergency response (Luo et al., 2022; Zeng et al., 2025). As a result, there is a growing demand for alternative and complementary approaches that can provide accurate flood predictions with significantly reduced computational requirements.

To address these challenges, Machine Learning (ML) algorithms have been increasingly explored for rapid flood predictions (Mosavi et al., 2018). ML models can either be trained with observed data or trained on output from physics-based models to act as surrogates of these models that can approximate flood simulations with reduced computational costs. Bermúdez et al. (2019) used discharge and tide levels as input data to approximate water depth and velocity at selected control points simulated by the 2-D shallow water model Iber. Zahura and Goodall (2022) developed a Random Forest (RF) surrogate model trained on outputs from a TUFLOW model to enable real-time, street-scale flood prediction under both pluvial and tidal flooding scenarios. Tang et al. (2023) presented a hybrid approach that combined the physics-based Xin'anjiang Model and RF techniques to improve flood forecasting accuracy in the Jingle sub-basin of the Yellow River. These ML surrogate models can learn patterns from historical and synthetic data at a low computational cost and generate predictions much faster than physics-based models alone. However, most of these models are constrained to nowcasting as their relatively simple algorithmic structures prevent them from being effectively applied to long-term, multi-step flood forecasting.

Building upon basic ML, Deep Learning (DL) has emerged as a powerful tool for fast and accurate flood modeling and forecasting (Shen, 2018; Fu et al., 2022). DL typically involves Artificial Neural Networks (ANNs) with multiple layers, which enables it to automatically extract hierarchical features from large and complex datasets (Khan et al., 2020; Kumar et al., 2024). One of the DL algorithms, Convolutional Neural Network (CNN), has been successfully applied in various remote sensing and environmental modeling tasks due to their ability to extract high-level features from spatial or imagery data such as satellite images, terrain maps, static pictures and video frames (Kim and Song, 2020; Kattenborn et al., 2021; Wang et al., 2024). Due to its capability of learning geospatial features such as topography, land use and soil type, CNN-based models have also been developed for forecasting maximum flood inundation and water depth in urban areas (Guo et al., 2021). However, these models are generally less suitable for sequential flood forecasting across multiple time steps, given the limited ability of CNNs to capture temporal dependencies in time series data (Yang et al., 2015; Chen et al., 2023). On the contrary, Recurrent Neural Networks (RNNs) and its variant, Long Short-Term Memory (LSTM) network, have demonstrated strong capabilities in processing time-series hydrological and hydrodynamic data (Waqas and Humphries, 2024), making them mostly applied for flood forecasting (Le et al., 2019; Ding et al., 2020; Shahabi and Tahvildari, 2024; Roy et al., 2025). However, conventional LSTM-based models have two major limitations: (1) they are inherently sequential and struggle to capture spatial dependencies within the input data, and (2) they do not provide the spatial extent of predicted floods, which can be critical information for urban flood warning and risk assessment.

Recent advances in Deep Learning (DL) have introduced hybrid models that combine both spatial and temporal information, improving flood prediction capabilities (Moon et al., 2023). By integrating CNN with LSTM, researchers have developed hybrid deep learning models that can process spatiotemporal data more effectively. These hybrid models leverage the strengths of both architectures - CNN for capturing spatial correlations and LSTM for modeling temporal dependencies - leading to improved predictive performance for flood forecasting (Chen et al., 2023; Situ et al., 2024; Shahabi and Tahvildari, 2024). This fusion of CNN and LSTM enables more precise flood forecasting by effectively capturing both spatial patterns and temporal dynamics, making it a powerful tool for fast flood forecasting and risk assessment. One of the fusion approaches, Convolutional Long Short-Term Memory (ConvLSTM), combines convolution operations with LSTM to extract spatiotemporal features from inputs, which has been applied in hydrological prediction such as large-scale river runoff (Zhu et al., 2023), flood extent (Muckley and Garforth, 2021) and depletion of water resources (Moishin et al., 2021). However, ConvLSTM is limited in long-term forecasting because it introduces high computational load and parameter complexity over long sequences, making it harder to capture long-range temporal dependencies. Malik et al. (2024) developed a Time-Distributed CNN-LSTM (TD-CNN-LSTM) model for watershed-scale fluvial flood forecasting, capable of effectively capturing complex spatial and temporal patterns within hydrological time-series data. Their results demonstrated strong predictive performance, particularly in forecasting flood onset time and peak magnitude. However, the model considered only rainfall and streamflow as inputs, representing a simplified setting compared to the more complex dynamics of coastal urban environments. Hu et al. (2024) proposed a CNN-LSTM hybrid model that incorporated multiple grid-based data sources, including precipitation, vegetation and soil moisture, for daily runoff prediction in the source region of the Yellow River Basin. Although this study demonstrated superior predictive accuracy compared to standalone CNN or LSTM models, its dependence on coarse-resolution, catchment-scale spatial input may constrain its transferability to more heterogeneous coastal urban environments. Shahabi and Tahvildari (2024) developed a CNN-LSTM model for rapid forecasting of tidal flooding for the next 72 h, achieving high accuracy that outperformed the physics-based model. While their model incorporated spatial input variables such as wind fields and atmospheric pressure, its outputs were limited to water levels at selected points, lacking comprehensive spatial coverage. Moreover, rainfall-driven flooding was not accounted for, as the study primarily focused on tidal flooding. Situ et al. (2024) developed an innovative method of CNN-RNN hybrid feature fusion for spatiotemporal forecasting of urban floods, and optimized the hybridization strategy with Bayesian Optimization. While those hybrid models can generate spatiotemporal output, they do not consider the water depth from past time steps, which makes them unavailable for real-time flood forecasting. In addition, the lack of consideration of tide level in their model limits its applicability in coastal areas.

The objective of this study is to develop a deep learning-based spatiotemporal surrogate model for real-time flood forecasting in a coastal urban area. By integrating spatial and temporal features, the proposed model aims to provide spatiotemporal flood forecasting across multiple future time steps in a complicated coastal urban environment while significantly reducing computational time compared to traditional hydrodynamic models. Specifically, this study employs a hybrid model of CNNs and LSTMs to capture both the spatial features from gridded data and the temporal dependencies of time series input. The proposed framework is tested and validated using historical storm events, ensuring its applicability to practical flood risk management scenarios. This approach has the potential to enhance urban flood risk management by providing timely and spatially explicit flood forecasts for coastal cities facing increasing flood hazards. By leveraging deep learning advancements, the study aims to bridge the gap between computational efficiency and prediction accuracy, offering a viable solution for real-time coastal urban flood forecasting.

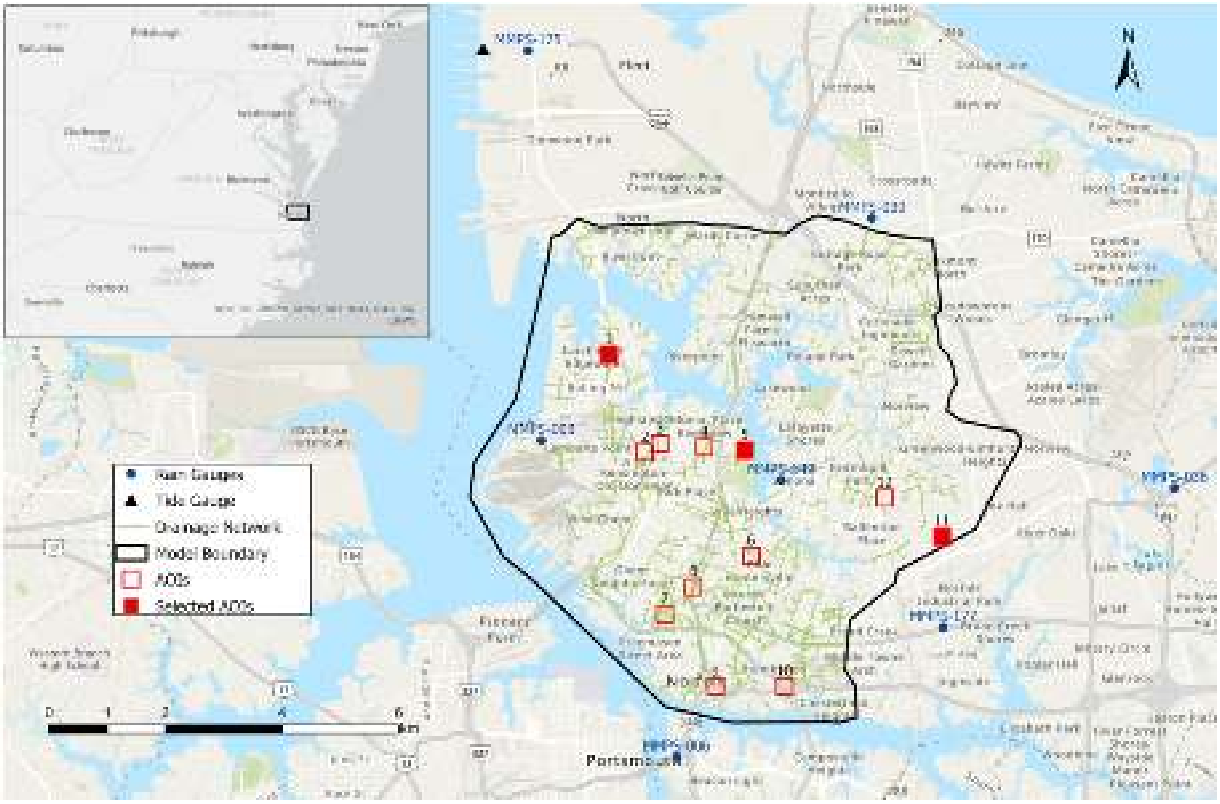

Fig. 1. Map of the study area, including the model domain, drainage network, locations of the seven rain gauges, location of the tide gauge, and locations of the 12 flood prone areas of interest (AOIs). Three of these flood-prone AOIs, Denver Avenue, Myrtle Park and La Valette Avenue highlighted with solid red from left to right, are selected for detailed analysis in Sections 3 and 4.

2. Data and methods

2.1. Study area

Norfolk, Virginia is a coastal city at the confluence of the James River, the Elizabeth River, and Chesapeake Bay. It is in the Hampton Roads metropolitan region in southeastern Virginia, USA, and home to the world's largest naval base. It serves as a vital center for military operations and maritime infrastructure. The city is characterized by predominantly flat terrain, with most areas situated at elevations below 4.5 m relative to the North American Vertical Datum of 1988 (NAVD 88). This low-lying topography, combined with tidal connectivity to Chesapeake Bay, makes Norfolk particularly vulnerable to tidal flooding (Ezer and Atkinson, 2014). This vulnerability is further exacerbated by ongoing sea level rise and land subsidence (Li et al., 2013; Sadler et al., 2017). Along the U.S. Atlantic Coast, Norfolk faces the highest rate of relative sea level rise, which has made the city increasingly susceptible to storm surge and coastal flood hazards (Boon, 2012; Kopp, 2013).

The model domain of this study covers a large portion of Norfolk with a total area of 56.4 km^2, of which 8.7 km^2 is open water (Fig. 1). It is highly urbanized, incorporating dense residential and commercial developments, extensive transportation networks and a complex stormwater system. The interaction between the infrastructures and the inherently flood-prone topography generates complicated hydrodynamic processes in overland flow pathways and subsurface conveyance through the underground drainage network, which poses significant challenges for timely flood modeling.

2.2. Workflow overview

The overall workflow consists of three main steps, as illustrated in Fig. 2. First, the TUFLOW model was set up for the model domain using a range of physical and environmental inputs, including drainage systems, land cover, soil type, imperviousness, Digital Elevation Model (DEM), rainfall and tide level, to simulate flood water depth on a 15-min time step and at a 2.5-m spatial resolution. These outputs were treated as ground truth flood depth data for training and evaluating the surrogate model, as spatially continuous, city-scale flood depth observations are not available for the study area. Second, the TUFLOW water depth outputs were integrated with

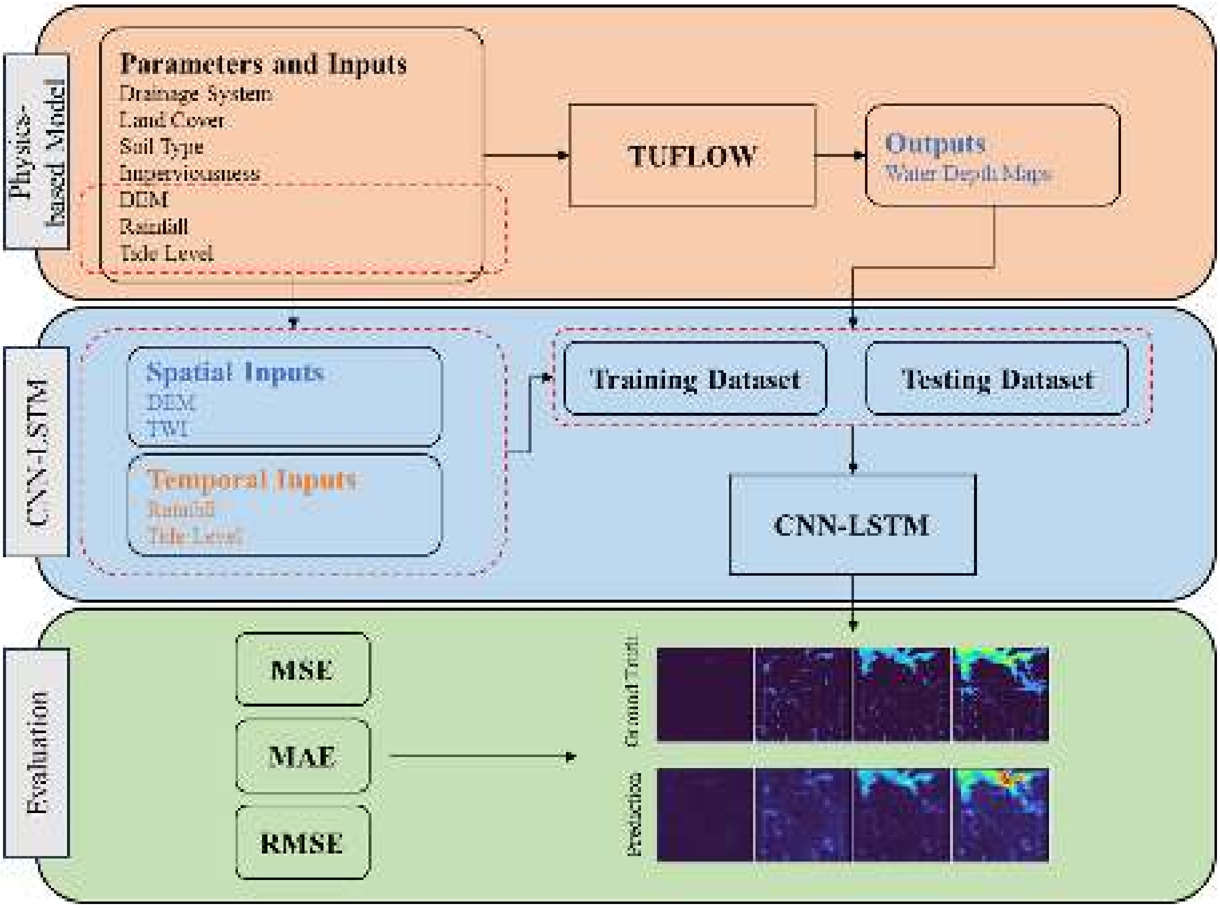

Fig. 2. Workflow for developing a CNN-LSTM surrogate model for spatiotemporal forecasting of coastal urban flooding.

DEM, Topographic Wetness Index (TWI), rainfall and tide level to construct structured dataset. This dataset was then divided into training and testing subsets at a ratio of 80:20 for developing a CNN-LSTM surrogate flood model. Finally, the predictive performance of the surrogate models was assessed using standard error metrics, including Mean Squared Error (MSE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). Further details on each step are provided in the following sections.

2.3. TUFLOW model setup

TUFLOW employs a full set of depth-averaged shallow water equations, solving both momentum and continuity for two-dimensional free-surface flows, and integrates the ESTRY module, a one-dimensional hydraulic engine embedded within TUFLOW that simulates flow through pipe networks, channels, and other conveyance structures, to represent one-dimensional hydrodynamic networks (Syme, 2001). Its capability to couple surface flows on a 2D grid with pipe flows in a 1D system through dynamic linking makes it particularly well-suited for physics-based modeling in low-relief urban settings, where accurate representation of the interaction between overland runoff and subsurface drainage infrastructure is critical. In addition, the High-Performance Computing (HPC) engine of TUFLOW allows parallel execution on multiple GPU units, which considerably enhances computational efficiency. In this study, a coupled 1D/2D TUFLOW model described in Shen et al. (2022) was applied to generate water depth maps as the ground truth data for training and testing the CNN-LSTM framework. The validation of the TUFLOW model was performed using temporary water level observation available for Hurricane Irene in 2011. Further details on the validation procedure and results are provided in Shen et al. (2022) and are beyond the scope of this study. Although comprehensive depth measurements are unavailable during the study period, the TUFLOW model has previously been evaluated using crowdsourced flood occurrence reports from the Waze navigation application for events between September 2017 and August 2018 (Zahura et al., 2020). Table 1 summarizes the datasets used for building the TUFLOW model.

In the settings of the TUFLOW model, the terrain was characterized using a high-resolution LiDAR-derived DEM from the Virginia Geographic Information Network (VGIN), which provides a spatial resolution of 1 m × 1 m (Dewberry, 2014). Because the urban landscape is strongly shaped by built infrastructure, the inclusion of building structures was also necessary. To balance model complexity with numerical stability, only buildings larger than 500 m^2 were incorporated, as prior studies have shown that very small structures exert minimal influence on street-scale overland flow routing while substantially increasing mesh complexity (Shen et al., 2022; Liu et al., 2023). These structures were treated as impermeable surfaces and superimposed onto the bare-earth DEM, thereby creating a more realistic three-dimensional Digital Surface Model (DSM) for simulating overland flow in densely developed areas. The accuracy of the DSM is especially critical in urban flood modeling, as small variations in elevation or building configuration can significantly influence overland flow routing and the spatial distribution of inundation.

The stormwater drainage network comprises over 6200 pipe sections, totaling approximately 179 km, derived from the drainage infrastructure data obtained from the City of Norfolk (City of Norfolk GIS Bureau, 2025). To streamline the model and improve computational efficiency, minor pipes with diameters less than 15 in. were excluded. This threshold follows established practice in previous studies, where minor pipes are often omitted due to limited data availability, higher uncertainty in attribute records, and their relatively localized influence at the neighborhood or street scale (Zahura et al., 2020; Shen et al., 2022; Barbosa et al., 2025). Furthermore, underpasses and tunnels were excluded from the ML surrogate model training because these features have pumps installed to remove ponded water, but the operation of these pumps was unavailable for the TUFLOW modeling. The modeling framework therefore prioritizes primary drainage pathways that govern surface-subsurface interactions during flood events, while acknowledging that smaller pipes may affect localized flow behavior at finer spatial scales. Missing diameters or elevations of major pipes were estimated using a regression-based approach, which correlated the known pipe attributes such as length, slope, upstream and downstream elevations with their corresponding diameters or invert elevations, and then applied the derived relationship to pipes with incomplete data (Shen et al., 2022). Stormwater inlets and grates were modeled as pits, where a depth-discharge relationship can be defined to correlate surface ponding depth to the rate of inflow into the drainage network. This relationship, which was determined by the type and dimension of inlet, controlled the rate of overland runoff entering the subsurface drainage system.

The boundary conditions of the TUFLOW model were defined using 15-minute rainfall data from seven regional rainfall gauges managed by Hampton Roads Sanitation District (HRSD) (HRSD, 2025), and six-minute tide level at National Oceanic and Atmospheric Administration (NOAA)'s Sewells Point tide gauge, referenced to the North American Vertical Datum of 1988 (NAVD 88) (NOAA, 2025). Assuming that the wave propagation speed of tide level is negligible, the tide level at Sewells Point was considered to represent the tide level along the entire shoreline of the study area. The locations of these gauges are shown in Fig. 1.

Infiltration was not considered in the model. Barbosa et al. (2025) found that for major and extreme events (10-, 50-, and 100-year

Table 1
Summary of the datasets used for building the TUFLOW model.

Dataset	Dataset provider	Spatial Resolution (m)	Temporal Resolution (min)
DEM	VGIN	1	-
Land Cover	VGIN	1	-
Drainage System	City of Norfolk	-	-
Rainfall	HRSD	-	15
Tide Level	NOAA	-	0.1
Soil Type	USDA	30	-
Imperviousness	NLCD	30	-

return periods), infiltration processes reduce surface runoff by no more than 12 % for a region in Norfolk. This is because the area is characterized by a high impervious surface coverage of 41 % and shallow groundwater levels that can rise to the surface during storm events. Thus, the model assumes saturated initial conditions and that infiltration has limited influence on significant flood events.

2.4. TUFLOW model simulations

The TUFLOW model was used to simulate water depth maps for 11 selected flood events from 2017 to 2022 (Table 2). These flood events were chosen to capture a range of flood conditions, which span high tide levels up to the 10-year return period and daily rainfall up to the 5-year return period, capturing a diverse set of nuisance to moderate flooding scenarios relevant to urban flood forecasting. This study focuses on the rapid prediction of short-term urban nuisance flooding, for which computational efficiency and timely forecasts are critical. For rarer extreme events or synthetic climate-projection scenarios, timeliness is generally less important, and physics-based hydrodynamic models are more suitable for long-term impact assessment and adaptation planning.

The flood conditions can be classified into three scenarios: rainfall-driven, tide-driven and compound. Rainfall-driven floods are primarily triggered by intense or long-duration rainfall that exceeds the storage and conveyance capacity of the stormwater drainage system, leading to surface water accumulation and localized inundation. Tide-driven floods, in contrast, occur when elevated tidal levels are the dominant factor causing water to overtop or back up through drainage infrastructure, even in the absence of heavy rainfall. Compound floods represent the interaction of these two mechanisms, where the simultaneous occurrence of high tides restricts drainage outflow and heavy rainfall causes increased flood severity. The TUFLOW model generated outputs of water depth every 15 min at a spatial resolution of 2.5 m. All hydrodynamic simulations in this study were based on observed rainfall and tide level time series associated with historical flood events rather than synthetic or idealized hydrographs. Using historical flooding conditions ensures that the simulated flood dynamics reflect realistic interactions between rainfall, tidal influence and urban drainage capacity. These physics-based simulations provide a consistent and physically grounded reference dataset for training and evaluating the CNN-LSTM surrogate model in the absence of spatially continuous, city-scale flood depth observations.

2.5. CNN-LSTM model architecture

This study developed a hybrid CNN-LSTM architecture for the flood surrogate model, which integrated DeepLabv3+ for spatial feature extraction and an LSTM network for temporal sequence modeling, thereby leveraging both spatial and temporal dependencies in flood forecasting (Fig. 3). DeepLabv3+ is an advanced CNN designed for semantic image segmentation, building upon the strengths of the original DeepLab series (Chen et al., 2017, 2018). It integrates atrous convolutions to capture multi-scale contextual information without reducing the spatial resolution of feature maps, enabling precise segmentation of objects at varying scales. Additionally, DeepLabv3+ incorporates an encoder-decoder structure, where the encoder extracts rich contextual features and the decoder refines object boundaries to improve segmentation accuracy. This combination of Atrous Spatial Pyramid Pooling (ASPP) and the encoder-decoder framework allows DeepLabv3+ to perform exceptionally well in complex scene understanding tasks, making it particularly suitable for applications such as urban flood mapping, where accurate delineation of water extents and boundaries is critical. The model has been widely adopted in remote sensing and environmental monitoring due to its balance of accuracy and computational efficiency (Wu et al., 2022; Shi et al., 2025; Sundaresan, Solomon, 2025).

LSTM is a specialized type of RNN designed to model sequential and temporal data while mitigating the vanishing gradient problem common in traditional RNNs (Hochreiter and Schmidhuber, 1997). LSTM achieves this through a gated architecture, consisting of input, forget and output gates, which regulate the flow of information and allow the network to retain long-term dependencies in the data. This capability makes LSTM particularly effective for time-series prediction tasks, such as forecasting water levels or flood dynamics, where past observations strongly influence future outcomes. By capturing temporal correlations in sequential inputs, LSTM networks can generate accurate predictions over multiple time steps even in complex and nonlinear systems. Their robustness in modeling temporal dependencies has led to widespread applications in hydrology, meteorology, and environmental monitoring, particularly when integrated with spatial models like CNN to handle spatiotemporal data (Kratzert et al., 2018; Sabzipour et al., 2023;

Table 2
Summary of the flood events in Norfolk, VA from 2017 to 2022.

Date (mm/dd/yyyy)	High Tide Level		Daily Rainfall		Flood Scenarios
	(m from NAVD 88)	Return Period (year)	(mm)	Return Period (year)	
08/29/2017	1.17	5	99.82	5	Rainfall-driven
11/08/2017	1.07	2	8.67	< 1	Tide-driven
09/06/2019	1.35	10	47.24	1	Compound
10/28/2019	0.93	1	4.83	< 1	Tide-driven
09/19/2020	1.13	2	91.44	2	Rainfall-driven
11/13/2020	0.83	1	125.98	5	Rainfall-driven
05/31/2021	1.21	5	41.66	1	Compound
11/07/2021	1.13	2	0	-	Tide-driven
01/03/2022	1.34	10	42.67	1	Compound
09/30/2022	1.25	5	86.36	2	Rainfall-driven
10/29/2022	0.90	1	0	-	Tide-driven

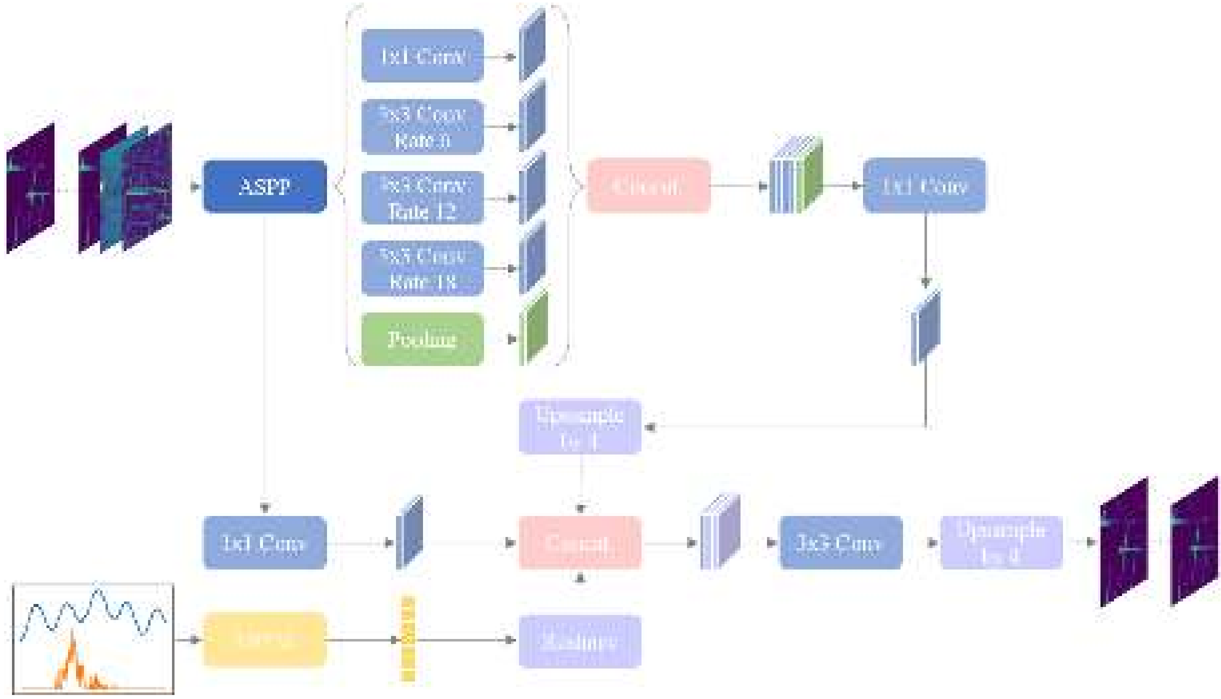

Fig. 3. The architecture of the hybrid CNN-LSTM model.

Situ et al., 2024).

In this study, the CNN component, based on DeepLabv3+, extracted spatial features from the input variables, including DEM, TWI, and water depth maps. This allowed the model to understand terrain characteristics and flood distribution patterns. The LSTM component processed temporal input features such as rainfall and tide levels, capturing long-term dependencies in flood evolution. By concatenating the bottleneck layer of CNN with the LSTM component, the model effectively fused spatial and temporal features, enabling accurate real-time spatiotemporal flood forecasting.

2.6. Data preparation and preprocessing

The TUFLOW water depth outputs served as the ground truth for training and evaluating the CNN-LSTM model, allowing the surrogate model to learn from physics-based hydrodynamic simulations. The inputs used for training the ML-based surrogate flood forecasting model can be categorized into temporal and spatial inputs, respectively. Their time steps used for training and testing the CNN-LSTM model are summarized in Table 3. Temporal inputs included rainfall and tide level that were used for building the TUFLOW model. The tide level data, originally recorded at six-minute intervals, were interpolated to 15-minute intervals to match the time

Table 3
Summary of the inputs and outputs and the time steps used for training and testing the CNN-LSTM model for each AOI. The input-output window size of 8 previous and 4 future time steps was selected to capture short-term flood dynamics while limiting error accumulation in multi-step prediction. This design enables the LSTM to capture recent hydrodynamic memory effects, such as rising and recession limbs of flood hydrographs, while avoiding excessive sequence lengths that may introduce overfitting or error propagation.

INPUTS					OUTPUT
DEM	TWI	Past Water Depth Map	Rainfall	Tide Level	Water Depth Map Forecast
		t - 7	t - 7	t - 7	
		t - 6	t - 6	t - 6	
		t - 5	t - 5	t - 5	
		t - 4	t - 4	t - 4	
		t - 3	t - 3	t - 3	
Constant	Constant	t - 2	t - 2	t - 2	
		t - 1	t - 1	t - 1	
		t	t	t	
			t + 1	t + 1	t + 1
			t + 2	t + 2	t + 2
			t + 3	t + 3	t + 3
			t + 4	t + 4	t + 4

intervals of the rainfall data. The model was developed using temporal inputs from the previous eight time steps and the subsequent four time steps. During the training process, rainfall and tide levels for the future time steps (t + 1 to t + 4) were taken from historical observations, representing perfect knowledge of future conditions. In real-world applications, however, these inputs are expected to be provided by forecast products generated through available meteorological and tidal forecast models. Spatial data included the 15-minute water depth simulations of the past eight time steps from the TUFLOW model, DSM for the modeling area, and TWI, which were derived from the DEM to measure the tendency of flood accumulation in different areas. TWI is defined as

$$TWI = \ln\frac{a}{\tan S}$$

where a is the contributing area from upslope and S is the topographic slope in radians. TWI provides critical topographical information that influences flood behavior (Pourali et al., 2016), and was therefore incorporated as a spatial input in the CNN-LSTM model. The dataset used for training and testing the model is accessible at Wang (2025).

To focus on areas most affected by flooding, this study selected 12 flood-prone areas of interest (AOIs) from the study area, as shown in Fig. 1. Each AOI was defined as an area with 128 pixels in both length and width, corresponding to 320 m, ensuring that the model had a consistent and comparable spatial context for learning flood patterns. These AOIs were selected based on the crowdsourced flood data from the System to Track, Organize, Record and Map (STORM), an online data portal managed by the City of Norfolk that archives records of reported occurrence of flooding and their locations (STORM, 2023). The selected AOIs covered various urban landscapes including low-lying areas connected to the coastline, highly urbanized inland commercial districts, and infrastructure corridors impacted by both pluvial flooding and storm surge. The variety of the selected flood-prone locations provided a diverse set of training examples for the model.

2.7. Model training and testing

The dataset was split into training and testing subsets using an 80:20 ratio to ensure a balanced evaluation of model performance. The model was trained using Python 3.10.11 on Google Colab, using the NVIDIA Tesla K80 GPU operating at 0.82 GHz with 12 GB VRAM. Training was conducted using the Adam optimizer with a learning rate of 1×10^{-4}, and the MSE was adopted as the loss function. A batch size of 4 was used, with model weights updated iteratively across a maximum of 200 epochs. To prevent overfitting, dropout layers with a rate of 0.2 were incorporated, and an early stopping strategy was applied with a minimum delta of 1×10^{-7} and a patience value of 20 epochs. A sensitivity analysis was conducted to evaluate the robustness of the CNN-LSTM model by tuning key hyperparameters, including the learning rate, batch size, and dropout rate. Three representative AOIs (1, 5, and 11) and two historical events (08/29/2017 and 09/06/2019) were selected for hyperparameter tuning, and model performance was assessed using the MSE on the validation set. The hyperparameter settings for the training process, including hyperparameter tuning, are summarized in Table 4.

To assess the accuracy of the model in predicting the pixel-level water depth, three evaluation metrics were employed: MSE, MAE, and RMSE. They are defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_{i-obs} - y_{i-model}\right|$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_{i-obs} - y_{i-model}\right)^2$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{i-obs} - y_{i-model}\right)^2}$$

where $y_{i\text{-}obs}$ and $y_{i\text{-}model}$ are the physics-based model-generated and the surrogate model-forecasted water depth at the same pixel at time step i, respectively. The TUFLOW-predicted water depths were assumed to be ground ruth for the error assessments, given that a city-wide flood depth observation dataset is unavailable for the modeled flood events. The error metrics were calculated for all AOIs at

Table 4
Hyperparameter settings for training the CNN-LSTM model.

Hyperparameter	Value/Method
Loss Function	MSE
Optimizer	Adam
Learning Rate	1e-4, 5e-3, 1e-3
Batch Size	4, 8, 16
Early Stopping Minimum Delta	1e-7
Early Stopping Patience	20
Dropout Rate	0.5, 0.2, 0.1, 0.05
Maximum Epochs	200

all time steps to provide a comprehensive evaluation of model performance, capturing both absolute and standard deviations between predicted and ground truth flood depths. In addition, the runtime required for forecasting a single flood event was also computed as a quantitative metric for evaluating the real-time applicability of the model.

3. Results

3.1. CNN-LSTM model overall performance

After training, the CNN-LSTM model demonstrated strong predictive performance across all evaluated AOIs. Table 5 presents the quantitative evaluation of the CNN-LSTM model performance across the 12 flood-prone AOIs. The model achieved a high level of accuracy in forecasting water depth maps, with an average MAE of 0.024 m, MSE of 0.0020 m^2, and RMSE of 0.044 m across all AOIs. These low error metrics were observed across various flood-prone locations, suggesting consistent performance in diverse coastal urban flood scenarios. Among the AOIs, the lowest RMSE (0.033 m) was observed at AOI 4, reflecting the strong capability of the model in relatively simple hydrologic conditions. The highest RMSE (0.063 m) occurred at AOI 7, which is associated with more complex topography or mixed pluvial and tidal influences discussed in Shen et al. (2019). Despite the differences in the scenarios of application, the variations in error among AOIs remain small compared to findings from prior studies (Chen et al., 2023; Kumar et al., 2024), as further discussed in Section 4.1. This consistency suggests that the CNN-LSTM model maintains stable predictive capability across diverse flood conditions and urban settings. Overall, the results confirm that the hybrid CNN-LSTM model can effectively capture spatiotemporal flood dynamics with minimal deviation from the reference physics-based model outputs.

The sensitivity analysis of hyperparameters indicates that the CNN-LSTM model performance is generally stable across the tested ranges of learning rate and dropout rate. As shown in Table 6, the MSE varies only slightly, with the lowest value (0.0014 m^2) observed at the combination of learning rate and dropout rate of 1e-4 and 0.2 as well as 5e-3 and 0.5, and the highest value (0.0025 m^2) occurring at a learning rate of 5e-3 and dropout of 0.05. Dropout rates of 0.2 and 0.5 result in similar MSE values, suggesting that the model is relatively insensitive to the choice of dropout within this range. Additional experiments with different batch sizes (4, 8 and 16) produced negligible differences in MSE, indicating that batch size has minimal influence on model performance. This study selected learning rate = 1e-4 and dropout rate = 0.2 for model training primarily because the lower learning rate can reduce the risk of abrupt weight updates during optimization, and generally improve the robustness of the training process.

3.2. Flood inundation mapping

The predicted flood inundation was mapped and visually compared with the physics-based model output as the ground-truth water depth maps (Fig. 4). To further evaluate the mapping performance of the model, three AOIs within the study area were selected for detailed visualization of results: Denver Avenue, Myrtle Park, and La Valette Avenue, as shown in Fig. 1. These locations were selected based on their distinct characteristics of flooding, with Denver Avenue experiencing rainfall-driven flooding, La Valette Avenue primarily affected by high tide levels, and Myrtle Park subject to the combined influence of both rainfall and tidal conditions. For a selected flood event, predicted flood inundation maps at these three locations were generated at different stages of flood evolution, from the beginning of flooding, through its intensification, and up to the peak inundation. Consistent with the error quantification given in the prior section, the map results revealed a high qualitative consistency between the model-predicted inundation and the low-lying, flood-prone areas that closely aligned with areas of high TWI and low elevation, as identified in the topographic data. The CNN-LSTM model consistently identified and captured flood propagation patterns with a high spatial coherence at each stage of flooding.

3.3. Flood forecasting accuracy

To evaluate the capability of the CNN–LSTM model in providing spatiotemporal flood forecasting across a defined time sequence, forecasts were generated for all 12 AOIs at four consecutive 15-minute time steps, and the RMSE was calculated for each time step. As

Table 5
MAE (m), MSE (m^2) and RMSE (m) of the water depth map forecasts across all 12 AOIs using the CNN-LSTM model.

AOI	MAE (m)	MSE (m^2)	RMSE (m)
1	0.016	0.0012	0.035
2	0.023	0.0016	0.041
3	0.028	0.0023	0.048
4	0.014	0.0011	0.033
5	0.019	0.0015	0.039
6	0.022	0.0019	0.043
7	0.040	0.0039	0.063
8	0.017	0.0012	0.035
9	0.025	0.0020	0.045
10	0.031	0.0030	0.055
11	0.020	0.0015	0.039
12	0.029	0.0024	0.049
Average	**0.024**	**0.0020**	**0.044**

Table 6
MSE (m^2) of the CNN-LSTM model for different combinations of learning rate and dropout rate.

		Learning Rate		
		1e-4	5e-3	1e-3
Dropout Rate	**0.5**	0.0016	0.0014	0.0017
	0.2	0.0014	0.0016	0.0016
	0.1	0.0015	0.0018	0.0020
	0.05	0.0022	0.0025	0.0023

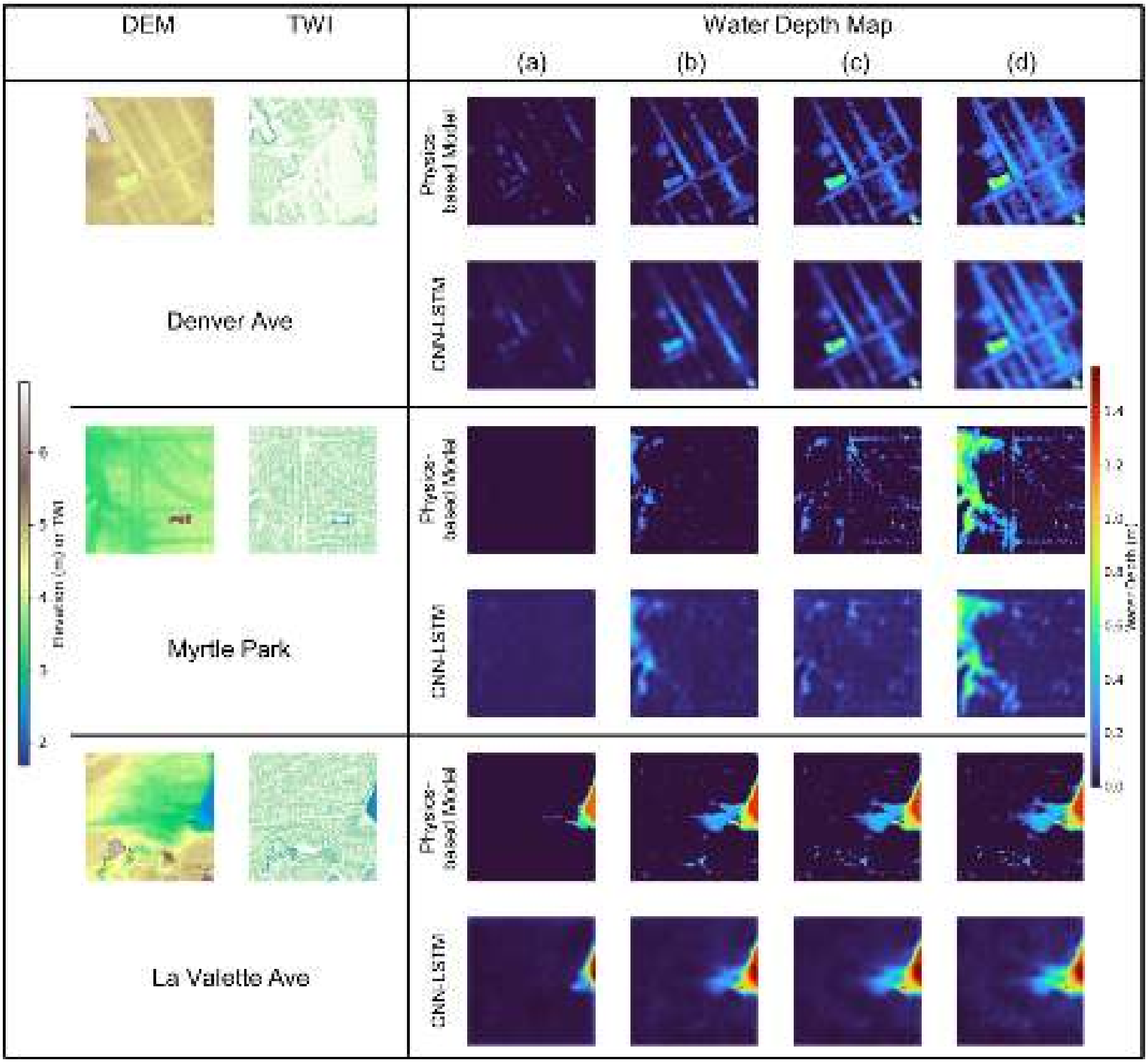

Fig. 4. Water depth maps at different stages of flooding predicted by the CNN-LSTM model in comparison to the physics-based model, where the column headings (a), (b), (c) and (d) represent the beginning of flooding, the formation and spatial expansion of inundation, and the peak flood stage of a storm event.

expected, the results (Fig. 5) reveal a consistent increase in RMSE with longer forecast horizons, indicating that the uncertainty of the forecasts gradually grows from the first time step (15 min) to the fourth time step (60 min). The median RMSE, indicated by the red lines, remains below 0.05 m across all time steps, demonstrating that the model maintains high accuracy throughout the forecasting window considered in this study. The mean RMSE values, represented by black dots connected with a dashed line, exhibit a similar upward trend, further confirming the expected temporal degradation in forecast accuracy. Moreover, the relatively compact

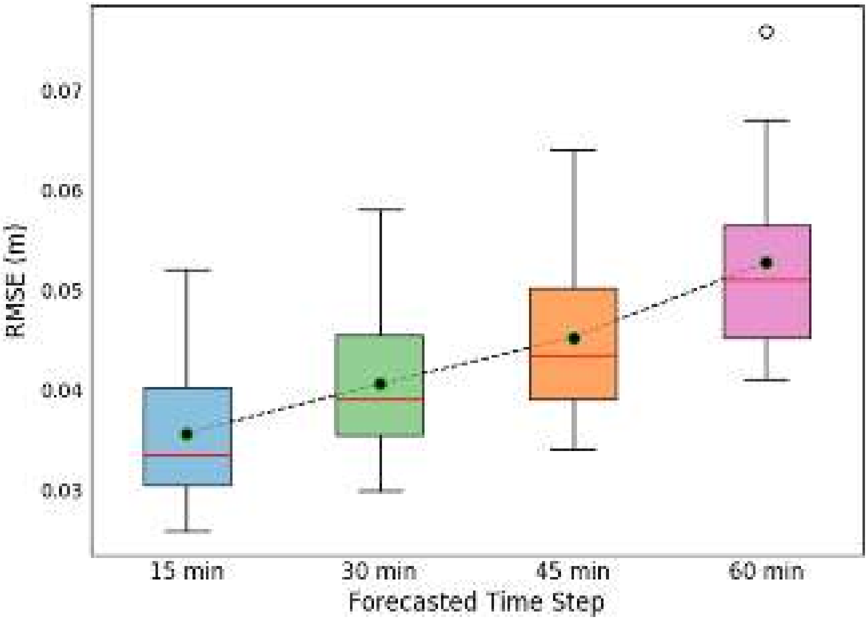

Fig. 5. RMSE in meters of the water depth map forecasts using the CNN-LSTM model across all 12 AOIs for the four future time steps. The red lines indicate the median and the black dots connected with dashed line show the mean RMSE.

interquartile ranges across all time steps suggest that the performance of the model is both stable and spatially consistent among different AOIs, with only a slight broadening of variability at longer prediction intervals. Overall, these findings indicate that the CNN-LSTM model is capable of producing short-term flood forecasts with reliable accuracy.

3.4. Spatial patterns in flood forecasting

The outputs at the three selected AOIs shown in Fig. 1 were visualized with each timestep representing a snapshot of evolving flood conditions and compared against the corresponding ground truth generated by the physics-based model (Fig. 6). Although the relatively short interval between time steps resulted in minimal changes to the overall inundation extent between consecutive frames, variations in water depth can still be observed in the model outputs as evidenced by the shifts in color gradients in the flood maps. This pattern was particularly noticeable in areas where water accumulation was gradual, revealing subtle but consistent increases or decreases in flood depth over time. At Denver Avenue and Myrtle Park, the predicted flood depths gradually increased across the time sequence, corresponding to the relatively steady influence of rainfall in these areas. Similarly, La Valette Avenue exhibited increases in predicted water depth and flood extent across timesteps, aligning with the increased tide level at this location. A time lag in flood prediction was observed at Myrtle Park and La Valette Avenue, where the inundation of some critical spots in the model outputs occurred slightly later than those in the ground truth data. This delay occurred across multiple timesteps, with the predicted flood extent closely matching the true inundation area but offset in timing. Despite the lag, the predicted spatial patterns of flood progression remained consistent, and the areas of increasing water depth aligned well with those identified in the physics-based model.

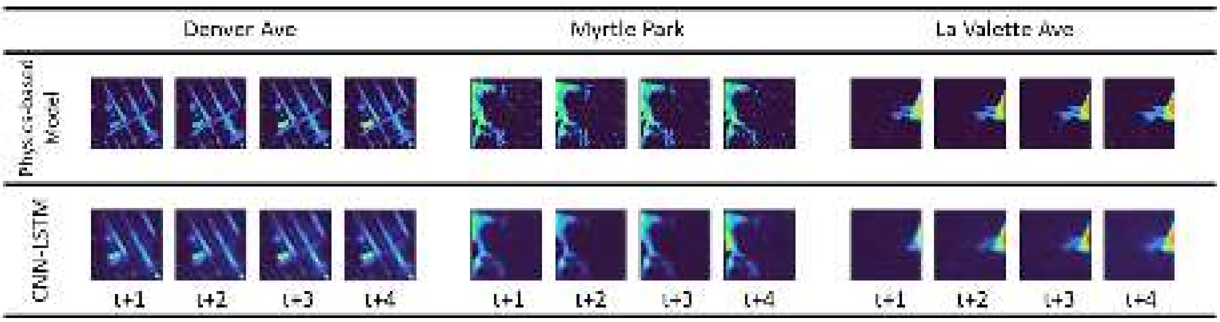

Fig. 6. Water depth maps at four consecutive time steps forecast by the CNN-LSTM model in comparison to the physics-based model.

3.5. Computational cost

In terms of computational efficiency, the CNN-LSTM model demonstrated a substantial advantage in simulation speed. While conventional physics-based hydrodynamic simulations using the TUFLOW model typically require four to six hours to complete a single flood event scenario, the CNN-LSTM model accomplished the same simulation task within a substantially shorter time frame. Given an identical dimension of 128 × 128 pixels for each AOI in this study, the model exhibited minimal variation in runtime across AOIs, with an average of 3.2 min per flood event per AOI. This drastic reduction in computational time was consistent across all 11 selected flood events, demonstrating the scalability and stability of the model when applied to different spatial domains. Such efficiency suggests the potential for near-real-time flood forecasting, especially under conditions requiring frequent updates or large-scale spatial coverage. Compared to the physics-based model that requires GPU-based computation, the trained CNN-LSTM model can be executed on a standard CPU, which offers a considerable advantage in terms of accessibility and computational resource requirements. Furthermore, the rapid processing capability was maintained without compromising the accuracy of flood depth outputs, as confirmed by statistical comparisons conducted over the full testing dataset.

4. Discussions

4.1. Applicability of the CNN-LSTM model in urban flood forecasting

The results from this study demonstrate the strong applicability of the CNN-LSTM model in the context of coastal urban flood forecasting. Across all evaluated flood-prone locations, the model consistently produced accurate predictions of flood depth with low error metrics. These results affirm the reliability of the model in capturing the magnitude of flooding under diverse coastal urban conditions. Furthermore, the spatial consistency between predicted flood extents and low-elevation, high-TWI regions indicates that the model effectively learned key spatial features influencing flood susceptibility. This is critical for urban flood modeling, where topographic and hydrologic variability play a central role in flood propagation. The model also proved capable of capturing temporal flood dynamics. Predictions generated at consecutive time steps showed coherent changes in water depth and spatial extent, reflecting the evolving nature of urban flood events. Although slight time lags were observed in some cases, the spatial accuracy and overall progression of flood extent remained consistent with the ground truth. These results support the ability of the model to produce both spatially and temporally relevant forecasts, which is crucial for applications requiring time-sensitive flood information.

Recent studies have emphasized that spatial characteristics play a fundamental physical role in shaping flood dynamics. Li et al. (2022) demonstrated that watershed slope and land cover significantly affect hydraulic geometry responses under tropical cyclone forcing, highlighting how terrain and vegetation regulate flood propagation and adjustment processes. These findings reinforce the importance of incorporating topography-derived features, such as elevation and topographic wetness, into flood forecasting models to capture physically meaningful spatial controls. In coastal urban settings, flood susceptibility is further governed by the interaction between geomorphology, shoreline configuration, and patterns of urban development. Giannakidou et al. (2020) showed that foreshore slope, coastline length, and industrial exposure are key determinants of coastal flood vulnerability, while Xu et al. (2021) demonstrated that local sea level rise and spatially explicit urban growth jointly drive the evolution of coastal flood risk. These studies underscore that flood response emerges from the coupling of hydrodynamic forcing with spatially heterogeneous urban and coastal features. The strong alignment observed in this study between predicted inundation patterns and low-elevation, high-TWI regions suggests that the CNN component effectively learns these physically grounded spatial relationships, supporting the applicability of the proposed framework to complex coastal urban environments.

When compared to existing ML-based models, the proposed CNN-LSTM model in this study demonstrates competitive capabilities for urban flood forecasting. With an average MAE of 0.024 m, MSE of 0.0020 m^2 and RMSE of 0.044 m, the model consistently produced accurate water depth predictions under varying coastal urban conditions. In comparison to existing studies, such as Kumar et al. (2024), who utilized an FNO-based model yielding a lower average MAE of 0.021 m but a higher MSE of 0.0037 m^2, and Chen et al. (2023), who reported an average MAE of 0.06 m and MSE of 0.03 m^2, the proposed model performs competitively. These findings support the effectiveness of the proposed CNN-LSTM model as a reliable and efficient surrogate model for real-time flood forecasting in complex urban coastal environments.

In addition to its predictive accuracy, the CNN-LSTM model demonstrated substantial computational advantages over physics-based models. While previous studies achieved similar speeds, their flood forecasts were limited to point-based locations (Zahura et al., 2020; Roy et al., 2025). The proposed model in this study provides more detailed information in real-time flood forecasts while maintaining a good computational efficiency, which makes the model especially ideal for operational deployment, where near real-time forecasting is essential for emergency response, early warning systems, and proactive decision-making in coastal urban flood management.

4.2. Roles of major components in the CNN-LSTM model

The CNN-LSTM architecture integrates two key components that serve different yet complementary purposes in spatiotemporal flood forecasting. The CNN component is primarily used for extracting spatial features from input data, while the LSTM component is used for capturing temporal features. The combined architecture allows the model to learn spatial dependencies and temporal sequences simultaneously, which is essential in coastal urban flood scenarios.

The CNN component demonstrated a strong capacity to interpret spatially heterogeneous input features. This is evident from the consistent alignment between predicted flood extents and areas of low elevation and high TWI. These areas are generally more prone to surface water accumulation. The model's capacity to capture these patterns, which arise from diverse flooding mechanisms such as rainfall-driven, tide-driven, and compound scenarios, suggests that the CNN component can learn the spatial relationships embedded in the input. This also emphasizes the importance of incorporating terrain-aware features in data-driven flood modeling, while showing how CNNs can be leveraged to detect and prioritize spatially sensitive flood-prone zones. The LSTM component, on the other hand, enabled the model to track and forecast changes in flood conditions over time. Through evaluation across four consecutive time steps, the model demonstrated an ability to follow the progression of water depth, with outputs showing clear temporal transitions in flood conditions. While changes between successive time steps were sometimes subtle due to short intervals, the model can consistently capture trends in flood intensification.

The effectiveness of the CNN-LSTM architecture observed in this study is consistent with recent applications in urban flood prediction. Despite differences in spatial resolution, input variables and flooding mechanisms, the findings of Chen et al. (2023) support the conclusion that CNN-LSTM models can efficiently capture coupled spatial and temporal flood processes, which aligns with the results presented in this study. Huang et al. (2024) integrated SHAP-based explainability into a CNN-LSTM framework, providing insights into the relative importance of input features and addressing the “black-box” nature of DL models. While this study primarily focuses on predictive performance, incorporating explainability techniques represents a promising direction for future work to improve transparency and confidence in practical flood forecasting applications.

4.3. Shortcomings of the CNN-LSTM model and prospects of future work

While the CNN-LSTM model shows promising performance in terms of accuracy and computational efficiency, there are several limitations in the context of real-world applicability and generalizability. Unlike physics-based models, ML-based models do not inherently encode physical laws such as the conservation of mass, momentum, or energy. As a result, predictions made by ML models may violate fundamental principles like mass balance, leading to unrealistic or physically implausible outcomes in certain scenarios. To address this limitation, recent research has explored the integration of physical constraints into the training process of ML models. For example, incorporating a water balance equation into the loss function allows the network to penalize physically inconsistent predictions, thereby encouraging the learning of hydrologically sound behavior (Feng et al., 2023; Donnelly et al., 2024; Taghizadeh et al., 2025). This hybrid approach, often referred to as physics-informed machine learning, offers a promising pathway toward building more robust and interpretable models that adhere to known physical processes while retaining the flexibility and efficiency of data-driven methods. Other studies have shown that embedding physical constraints into data-driven flood models does not necessarily improve predictive performance and may, in some cases, reduce accuracy when the imposed constraints over-limit the learning process or are imperfectly specified (Frame et al., 2022, 2023). In this study, the explicit incorporation of physical constraints is beyond the intended scope, as the focus is on evaluating the feasibility and efficiency of a purely data-driven surrogate for rapid urban flood forecasting; however, the development of physics-informed frameworks that balance physical consistency and predictive performance remains a potential direction for future research.

An important consideration in interpreting the performance of the proposed CNN-LSTM model is that its training and evaluation are based on outputs from a physics-based model rather than direct flood depth observations. Due to the lack of spatially continuous, city-scale flood depth measurements in Norfolk, the TUFLOW simulations, driven by historical rainfall and tide observations and previously validated against limited water level data, were treated as the ground truth data. Consequently, the reported error metrics (e.g., MAE, MSE, and RMSE) quantify the surrogate model’s ability to reproduce the spatiotemporal flood dynamics simulated by the hydrodynamic model, rather than direct agreement with real-world flood depths. This limitation is common in surrogate modeling studies for urban flooding and should be considered when interpreting model accuracy and transferring the framework to operational forecasting applications.

The LSTM component exhibited minor inaccuracy in temporal alignment. A time lag in the predicted flood extents was observed at Myrtle Park and La Valette Avenue, where the model outputs slightly delayed the inundation in critical spots in comparison to the ground truth. This lag suggests that while the LSTM effectively captures near-term temporal dynamics, it may not fully account for complex or delayed hydrological responses without further optimization. Additionally, the accuracy of the model in flood forecasting declines as the time horizon window is extended. This limitation is common among sequence-based models, where accumulated errors in recurrent forecasting can impair long-term fidelity.

A key assumption in this study is the use of historical rainfall and tide levels at future timesteps ($t + 1$ to $t + 4$) as model inputs. While this approach is suitable for training and evaluation, in real-world forecasting these values would not be directly available. Instead, they would be obtained from external forecasting systems such as numerical weather prediction models or tide prediction platforms. Therefore, the results presented here should be interpreted as demonstrating the potential performance of the CNN-LSTM surrogate under the assumption of accurate rainfall and tide forecasts. Although uncertainties in rainfall and tide forecasts may propagate through the CNN-LSTM model, quantifying its impact on flood forecasting is beyond the scope of this study. Future work should explore the integration of actual forecast products to evaluate the robustness of the model under realistic operational conditions.

Another key limitation concerns the geographic scope of the model's applicability. The CNN-LSTM was trained on data from specific locations, and its predictions are reliable only within the spatial domain covered by the training dataset. This restriction limits its transferability to other urban areas with different topographic or hydrologic conditions or built environments. The lack of generalization poses a barrier to developing a universal model for urban flood forecasting. Addressing this will require the incorporation of techniques such as transfer learning, domain adaptation, or multi-site training frameworks, which can help expand the applicability of the model beyond the original training domain.

4.4. Integration with urban flood monitoring systems

One practical consideration of the CNN-LSTM model developed in this study is its reliance on past water depth maps as input for forecasting future flood conditions. While this temporal input is essential for capturing the dynamic evolution of flood events, it raises the question of how such information can be reliably obtained in real-world applications. Recent advancements in camera-based urban flood monitoring systems provide a promising solution to this challenge. Wang et al. (2024) demonstrated the use of CNN to segment flood extents from real-world surveillance camera images, effectively extracting water depth information from visual data. Integrating such camera-based monitoring systems with ML-based forecasting models like the CNN-LSTM framework can enable real-time generation of input data needed for flood prediction. This integration has the potential to create a robust end-to-end urban flood forecasting pipeline, where surveillance imagery feeds directly into a deep learning model for continuous, automated flood extent prediction. This integration can enhance situational awareness and decision-making capabilities in urban flood management.

5. Conclusion

The CNN-LSTM real-time surrogate model represents a practical alternative to physics-based hydrodynamic simulations for real-time coastal urban flood forecasting. Its ability to rapidly generate accurate flood predictions with minimal computational demand makes it well-suited for integration into early warning systems and adaptive flood risk management frameworks. As cities continue to face the growing challenge of urban flooding due to increasing rainfall extremes, sea level rise and land-use changes, data-driven models such as the one presented in this study offer a promising path forward for enhancing urban resilience. A key advantage of this approach is its significantly reduced runtime without losing much prediction accuracy, compared to traditional physics-based models, making it a viable option for real-time flood forecasting and decision support in coastal urban areas. As coastal cities continue to face the growing challenge of flooding, this study highlights the potential of ML-based surrogate models such as CNN-LSTM for enhancing flood risk management and emergency response planning.

CRediT authorship contribution statement

Chetan Kumar: Writing – review & editing, Visualization, Resources, Methodology, Conceptualization. **Diana McSpadden:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Conceptualization. **Sergio A. Barbosa:** Writing – review & editing, Writing – original draft, Resources, Methodology, Data curation. **Binata Roy:** Writing – review & editing, Visualization, Software, Resources. **Ali Shahabi:** Writing – review & editing, Visualization, Validation, Resources, Methodology, Conceptualization. **Navid Tahvildari:** Writing – review & editing, Resources, Methodology, Conceptualization. **Yidi Wang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Methodology, Formal analysis, Data curation, Conceptualization. **Jonathan L. Goodall:** Writing – review & editing, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used ChatGPT and M365 Copilot in order to improve readability. After using the tools/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

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.

Acknowledgement

This study was supported by the U.S. National Science Foundation under the awards No. 1951745 and 2209139. The Hampton Roads Biomedical Research Consortium provided funding for this work as part of the efforts associated with the Joint Institute for Advanced Computing on Environmental Studies between Old Dominion University and Jefferson Laboratory. This manuscript has been authored by Jefferson Science Associates (JSA), operating the Thomas Jefferson National Accelerator Facility for the U.S. Department of Energy under Contract No. DE-AC05-06OR23177.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.ejrh.2026.103234.

Data availability

Data is available in Hydroshare at the following link: http://www.hydroshare.org/resource/43244f815e7947e6bac6b6705a9f7941.

References

Barbosa, S.A., Wang, Y., Goodall, J.L., 2025. Exploring infiltration effects on coastal urban flooding: Insights from nuisance to extreme events using 2D/1D hydrodynamic modeling and crowdsourced flood reports. Sci. Total Environ. 968, 178908. https://doi.org/10.1016/j.scitotenv.2025.178908.

Bermúdez, M., Cea, L., Puertas, J., 2019. A rapid flood inundation model for hazard mapping based on least squares support vector machine regression. J. Flood Risk Manag. 12, e12522.

Boon, J.D., 2012. Evidence of sea level acceleration at US and Canadian tide stations, Atlantic Coast, North America. J. Coast. Res. 28 (6), 1437–1445.

Carr, R.S., Smith, G.P., 2007. Linking of 2D and Pipe hydraulic models at fine spatial scales. Water Pract. Technol. 2 (2), wpt2007038.

Chen, J., Li, Y., Zhang, S., 2023. Fast prediction of urban flooding water depth based on CNN− LSTM. Water 15 (7), 1397.

Chen, L.C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation. Proc. Eur. Conf. Comput. Vis. 801–818.

City of Norfolk GIS Bureau, (2025). Interactive Norfolk | City of Norfolk, Virginia - Official Website [WWW Document]. URL ⟨https://www.norfolk.gov/1605/Interactive-Norfolk⟩ (accessed 1.3.26).

Dewberry, 2014. USGS Norfolk, VA LiDAR: Report Produced for U.S. Geological Survey, Norfolk, VA, USA.

Ding, Y., Zhu, Y., Feng, J., Zhang, P., Cheng, Z., 2020. Interpretable spatio-temporal attention LSTM model for flood forecasting. Neurocomputing 403, 348–359.

Donnelly, J., Daneshkhah, A., Abolfathi, S., 2024. Physics-informed neural networks as surrogate models of hydrodynamic simulators. Sci. Total Environ. 912, 168814.

Ezer, T., Atkinson, L.P., 2014. Accelerated flooding along the US East Coast: on the impact of sea-level rise, tides, storms, the Gulf Stream, and the North Atlantic Oscillations. Earth's Future 2 (8), 362–382.

Feng, D., Tan, Z., He, Q., 2023. Physics-informed neural networks of the Saint-Venant equations for downscaling a large-scale river model. Water Resour. Res. 59 (2), e2022WR033168.

Frame, J.M., Kratzert, F., Klotz, D., Gauch, M., Shalev, G., Gilon, O., Nearing, G.S., 2022. Deep learning rainfall–runoff predictions of extreme events. Hydrol. Earth Syst. Sci. 26 (13), 3377–3392.

Frame, J.M., Kratzert, F., Gupta, H.V., Ullrich, P., Nearing, G.S., 2023. On strictly enforced mass conservation constraints for modelling the Rainfall-Runoff process. Hydrol. Process. 37 (3), e14847.

Fu, G., Jin, Y., Sun, S., Yuan, Z., Butler, D., 2022. The role of deep learning in urban water management: a critical review. Water Res. 223, 118973.

Giannakidou, C., Diakoulaki, D., Memos, C.D., 2020. Vulnerability to coastal flooding of industrial urban areas in Greece. Environ. Process. 7 (3), 749–766.

Guo, Z., Leitao, J.P., Simões, N.E., Moosavi, V., 2021. Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. J. Flood Risk Manag. 14 (1), e12684.

Hallegatte, S., Green, C., Nicholls, R.J., Corfee-Morlot, J., 2013. Future flood losses in major coastal cities. Nat. Clim. Chang 3, 802–806. https://doi.org/10.1038/nclimate1979.

Han, S., Tahvildari, N., 2024. Compound flooding hazards due to storm surge and pluvial flow in a low-gradient coastal region. Water Resour. Res. 60 (11), e2023WR037014.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.

HRSD, (2025). HRSD | HRSD [WWW Document]. URL ⟨https://www.hrsd.com/⟩ (accessed 1.3.26).

Hsiao, S.C., Fu, H.S., Wu, H.L., Liang, T.Y., Chang, C.H., Chen, Y.M., Chen, W.B., 2024. Impact assessment of sea level rise-induced high tide flooding and socioeconomic losses in a highly vulnerable coastal region. J. Hydrol. Reg. Stud. 55, 101921.

Hu, F., Yang, Q., Yang, J., Luo, Z., Shao, J., Wang, G., 2024. Incorporating multiple grid-based data in CNN-LSTM hybrid model for daily runoff prediction in the source region of the Yellow River Basin. J. Hydrol. Reg. Stud. 51, 101652.

Huang, H., Wang, Z., Liao, Y., Gao, W., Lai, C., Wu, X., Zeng, Z., 2024. Improving the explainability of CNN-LSTM-based flood prediction with integrating SHAP technique. Ecol. Inform. 84, 102904.

Jeong, J., Barbosa, S.A., Goodall, J.L., 2025. Quantifying the impact of nuisance flooding on urban coastal communities under present and future climatic conditions: Norfolk, Virginia as a case study. J. Hydrol. Reg. Stud. 59, 102304.

Kattenborn, T., Leitloff, J., Schiefer, F., Hinz, S., 2021. Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. ISPRS J. Photogramm. Remote Sens. 173, 24–49.

Khan, A., Sohail, A., Zahoora, U., Qureshi, A.S., 2020. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 53 (8), 5455–5516.

Kim, D.Y., Song, C.M., 2020. Developing a discharge estimation model for ungauged watershed using CNN and hydrological image. Water 12 (12), 3534.

Kopp, R.E., 2013. Does the mid-Atlantic United States sea level acceleration hot spot reflect ocean dynamic variability? Geophys. Res. Lett. 40 (15), 3981–3985.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., Herrnegger, M., 2018. Rainfall–runoff modelling using long short-term memory (LSTM) networks. Hydrol. Earth Syst. Sci. 22 (11), 6005–6022.

Kumar, C., McSpadden, D., Goldenberg, S., Schram, M., Richter, H., Wang, Y., Goodall, J.L., 2024. Location Invariant Flood Prediction using Fourier Neural Operator. 2024 International Conference on Machine Learning and Applications (ICMLA). IEEE, pp. 1105–1110 (December).

Le, X.H., Ho, H.V., Lee, G., Jung, S., 2019. Application of long short-term memory (LSTM) neural network for flood forecasting. Water 11 (7), 1387.

Li, H., Lin, L., & Burks-Copes, K.A. (2013). Modeling of coastal inundation, storm surge, and relative sea-level rise at Naval Station Norfolk, Norfolk, Virginia, USA. *Journal of coastal research*, *29*(1), 18-30.

Li, Y., Wright, D.B., Bledsoe, B.P., 2022. Watershed controls and tropical cyclone-induced changes in river hydraulic geometry in Puerto Rico. J. Hydrol. Reg. Stud. 44, 101268.

Liu, Y., Bates, P.D., Neal, J.C., 2023. Bare-earth DEM generation from ArcticDEM and its use in flood simulation. Nat. Hazards Earth Syst. Sci. 23 (1), 375–391.

Luo, P., Luo, M., Li, F., Qi, X., Huo, A., Wang, Z., Wang, Y., 2022. Urban flood numerical simulation: research, methods and future perspectives. Environ. Model. Softw. 156, 105478.

Malik, H., Feng, J., Shao, P., Abduljabbar, Z.A., 2024. Improving flood forecasting using time-distributed CNN-LSTM model: a time-distributed spatiotemporal method. Earth Sci. Inform. 17 (4), 3455–3474.

Moishin, M., Deo, R.C., Prasad, R., Raj, N., Abdulla, S., 2021. Designing deep-based learning flood forecast model with ConvLSTM hybrid algorithm. Ieee Access 9, 50982–50993.

Moon, H., Yoon, S., Moon, Y., 2023. Urban flood forecasting using a hybrid modeling approach based on a deep learning technique. J. Hydroinformatics 25 (2), 593–610.

Mosavi, A., Ozturk, P., Chau, K.W., 2018. Flood prediction using machine learning models: Literature review. Water 10 (11), 1536.

Muckley, L., Garforth, J., 2021. Multi-input convlstm for flood extent prediction. Proceeding of International Conference on Pattern Recognition. Springer International Publishing, Cham, pp. 75–85.

National Academies of Sciences, 2019. Medicine, Division on Earth, Life Studies, Water Science, Technology Board, & Committee on Urban Flooding in the United States. Framing the challenge of urban flooding in the United States. National Academies Press.

NOAA, 2025. Sewells Point - Station Home Page - NOAA Tides & Currents [WWW Document]. URL⟨⟨https://tidesandcurrents.noaa.gov/waterlevels.html?id=8638610⟩⟩ (accessed 1.3.26).
Pourali, S.H., Arrowsmith, C., Chrisman, N., Matkan, A.A., Mitchell, D., 2016. Topography wetness index application in flood-risk-based land use planning. Appl. Spat. Anal. Policy 9, 39–54.
Rangari, V.A., Umamahesh, N.V., Bhatt, C.M., 2019. Assessment of inundation risk in urban floods using HEC RAS 2D. Model. Earth Syst. Environ. 5 (4), 1839–1851.
Roy, B., Goodall, J.L., McSpadden, D., Goldenberg, S., Schram, M., 2025. Forecasting multi-step-ahead street-scale nuisance flooding using seq2seq LSTM surrogate model for real-time applications in a Coastal-Urban city. J. Hydrol., 132697
Sabzipour, B., Arsenault, R., Troin, M., Martel, J.L., Brissette, F., Brunet, F., Mai, J., 2023. Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. J. Hydrol. 627, 130380.
Sadler, J.M., Haselden, N., Mellon, K., Hackel, A., Son, V., Mayfield, J., Goodall, J.L., 2017. Impact of sea-level rise on roadway flooding in the Hampton Roads region, Virginia. J. Infrastruct. Syst. 23 (4), 05017006.
Shahabi, A., Tahvildari, N., 2024. A deep-learning model for rapid spatiotemporal prediction of coastal water levels. Coast. Eng. 190, 104504.
Shen, C., 2018. A transdisciplinary review of deep learning research and its relevance for water resources scientists. Water Resour. Res. 54 (11), 8558–8593.
Shen, Y., Morsy, M.M., Huxley, C., Tahvildari, N., Goodall, J.L., 2019. Flood risk assessment and increased resilience for coastal urban watersheds under the combined impact of storm tide and heavy rainfall. J. Hydrol. 579, 124159.
Shen, Y., Tahvildari, N., Morsy, M.M., Huxley, C., Chen, T.D., Goodall, J.L., 2022. Dynamic modeling of inland flooding and storm surge on coastal cities under climate change scenarios: Transportation infrastructure impacts in Norfolk, Virginia USA as a case study. Geosciences 12 (6), 224.
Shi, L., Lan, X., Duan, X., Liu, X., 2025. FV-DLV3+: a light-weight flooded vegetation extraction method with attention-based DeepLabv3. Int. J. Remote Sens. 46 (1), 366–391.
Situ, Z., Wang, Q., Teng, S., Feng, W., Chen, G., Zhou, Q., Fu, G., 2024. Improving urban flood prediction using LSTM-DeepLabv3+ and Bayesian optimization with spatiotemporal feature fusion. J. Hydrol. 630, 130743.
STORM. (2023). Storm Map: Norfolk Open Data. ⟨https://data.norfolk.gov/Maps/STORM-MAP/fefd-kf5s⟩.
Su, W.R., Chen, Y.H., Fu, H.S., Chang, T.Y., Chen, W.B., 2024. Assessing the inundation risk of cultural heritages along the southwestern coast of Taiwan: present and future. Reg. Environ. Change 24 (2), 56.
Sundaresan, A.A., Solomon, A.A., 2025. Post-disaster flooded region segmentation using DeepLabv3+ and unmanned aerial system imagery. Nat. Hazards Res. 5 (2), 363–371.
Sweet, W.W.V., Dusek, G., Obeysekera, J.T.B., & Marra, J.J. (2018). Patterns and projections of high tide flooding along the US coastline using a common impact threshold.
Syme, W.J. (2001, February). TUFLOW-Two & Onedimensional unsteady flow Software for rivers, estuaries and coastal waters. In *IEAust Water Panel Seminar and Workshop on 2d Flood Modelling, Sydney*.
Taghizadeh, M., Zandsalimi, Z., Nabian, M.A., Shafiee-Jood, M., Alemazkoor, N., 2025. Interpretable physics-informed graph neural networks for flood forecasting. Comput. -Aided Civ. Infrastruct. Eng.
Tang, Y., Sun, Y., Han, Z., Wu, Q., Tan, B., Hu, C., 2023. Flood forecasting based on machine learning pattern recognition and dynamic migration of parameters. J. Hydrol. Reg. Stud. 47, 101406.
Wang, Y., 2025. CNN-LSTM coastal urban flood dataset and source code. HydroShare. ⟨http://www.hydroshare.org/resource/43244f815e7947e6bac6b6705a9f7941⟩.
Wang, H., Hu, Y., Guo, Y., Wu, Z., Yan, D., 2022. Urban flood forecasting based on the coupling of numerical weather model and stormwater model: a case study of Zhengzhou city. J. Hydrol. Reg. Stud. 39, 100985.
Wang, Y., Shen, Y., Salahshour, B., Cetin, M., Iftekharuddin, K., Tahvildari, N., Goodall, J.L., 2024. Urban flood extent segmentation and evaluation from real-world surveillance camera images using deep convolutional neural network. Environ. Model. Softw. 173, 105939.
Waqas, M., Humphries, U.W., 2024. A critical review of RNN and LSTM variants in hydrological time series predictions. MethodsX, 102946.
Wu, H., Song, H., Huang, J., Zhong, H., Zhan, R., Teng, X., Cao, J., 2022. Flood detection in dual-polarization SAR images based on multi-scale deeplab model. Remote Sens. 14 (20), 5181.
Xu, L., Cui, S., Wang, X., Tang, J., Nitivattananon, V., Ding, S., Nguyen, M.N., 2021. Dynamic risk of coastal flood and driving factors: Integrating local sea level rise and spatially explicit urban growth. J. Clean. Prod. 321, 129039.
Yang, J., Nguyen, M.N., San, P.P., Li, X., Krishnaswamy, S., 2015. Deep convolutional neural networks on multichannel time series for human activity recognition (July). Ijcai 15, 3995–4001.
Zahura, F.T., Goodall, J.L., 2022. Predicting combined tidal and pluvial flood inundation using a machine learning surrogate model. J. Hydrol. Reg. Stud. 41, 101087.
Zahura, F.T., Goodall, J.L., Sadler, J.M., Shen, Y., Morsy, M.M., Behl, M., 2020. Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction in an urban coastal community. Water Resour. Res. 56 (10) e2019WR027038.
Zeng, B., Huang, G., Chen, W., 2025. Research progress and prospects of urban flooding simulation: From traditional numerical models to deep learning approaches. Environ. Model. Softw. 183, 106213.
Zhu, S., Wei, J., Zhang, H., Xu, Y., Qin, H., 2023. Spatiotemporal deep learning rainfall-runoff forecasting combined with remote sensing precipitation products in large scale basins. J. Hydrol. 616, 128727.