Journal of Hydrology 661 (2025) 133529

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

# FLO-SR: Deep learning-based urban flood super-resolution model

Hyeonjin Choi [a], Hyuna Woo [a], Minyoung Kim [a], Hyungon Ryu [b], Jun-Hak Lee [c,d], Seungsoo Lee [e], Seong Jin Noh [a,*]

[a] Civil Engineering Department, Kumoh National Institute of Technology, Gumi 39177, Republic of Korea
[b] NVIDIA Corporation, Seoul, Republic of Korea
[c] Department of Landscape Architecture, University of Oregon, Eugene, OR 97403, United States
[d] Ojeong Resilience Institute, Korea University, Seoul 02841, Republic of Korea
[e] Water and Land Research Group, Korea Environment Institute, Sejong 30147, Republic of Korea

ARTICLE INFO

This manuscript was handled by Renato Morbidelli, Editor-in-Chief

*Keywords:*
Urban flooding
Inundation mapping
Deep learning
Super-resolution

ABSTRACT

Urban flooding, intensified by both climate change and urbanization, requires high-fidelity and computationally efficient modeling frameworks for effective risk assessment and mitigation. This study presents FLO-SR, a deep learning-based super-resolution (SR) model, to enhance the spatial resolution of urban flood simulations while significantly reducing computational costs. FLO-SR leverages a convolutional neural network (CNN) to convert low-resolution (LR) flood maps into high-resolution (HR) outputs. The model was validated using two distinct flood events: Hurricane Harvey in Houston, Texas (synthetic scenario using bicubic interpolation) and an urban flood event in Portland, Oregon (physics-based simulation scenario). FLO-SR was evaluated in terms of image similarity, flood depth, and inundation extent. FLO-SR achieved accuracy improvements in both cases at scale factors of 2, 4, and 8×, with average RMSE reductions of 56.2, 32.4, and 10.7 % in Houston and 24.5, 33.8, and 44.1 % in Portland. However, performance at the 8× scale was limited due to challenges in reconstructing fine-scale flood features and spatial discontinuities in LR inputs. To address this, future improvements should incorporate hydrodynamic constraints and enhance model generalization. Despite these limitations, FLO-SR combined with physics-based modeling achieved up to 63 and 45.7 % runtime reductions when reconstructing 2 m from 4 m and 4 m from 8 m simulations, respectively, highlighting its potential for real-time urban flood forecasting.

## 1. Introduction

Urban flooding, which is intensifying in both severity and frequency due to climate change and inappropriate urban development, poses a substantial threat to human life and economic sustainability. From 1990 to 2022, floods affected over 3.2 billion people globally, causing 218,353 deaths, and resulting in more than $1.3 trillion in economic losses (Liu et al., 2024). A wide range of natural processes contribute to flooding, including heavy rainfall, spring snowmelt, and rain-on-snow events in cold climate regions (Myers et al., 2023; Zaghloul et al., 2022). As urban areas expand, runoff increases and the natural drainage network changes, affecting hydrological processes (Guan et al., 2015). The implementation of dual drainage systems in urban areas adds further complexity to urban hydrological systems, increases the risk of urban flooding, and complicates simulation and risk management. This growing complexity underscores the need for more accurate flood mapping in urban areas (Liu et al., 2021; Zandsalimi et al., 2025).

Physics modeling based on hydrodynamics or empirical equations has traditionally been employed for simulating the complex rainfall-runoff process leading to flooding (Henonin et al., 2013). However, urban flood predictions remain challenging because these events can last for a short period of time but produce sudden inundation in the dual drainage systems (Chowdhury et al., 2023; Roy et al., 2025). One-dimensional (1D) hydraulic models are typically employed to simulate flow through sewer drainage networks, while two-dimensional (2D) models dynamically simulate the movement of water flow over a two-dimensional surface. These models often incorporate simulation of water flow within the drainage network to represent the interaction between the drainage system and the watershed area, providing an accurate and reliable simulation of urban surface flooding (Gao et al., 2024). However, the complexity of solving physical equations (the full or simplified form of the 2D Saint Venant or Shallow Water Equations

* Corresponding author at: Global Hall Rm 302, 61 Daehak-ro, Gumi-si, Gyeongsangbuk-do, Republic of Korea.
*E-mail address:* seongjin.noh@kumoh.ac.kr (S.J. Noh).

https://doi.org/10.1016/j.jhydrol.2025.133529
Received 28 February 2025; Received in revised form 9 May 2025; Accepted 12 May 2025
Available online 13 May 2025

(SWE)) typically results in time-consuming and high computational costs, limiting their application to real-time or near real-time flood forecasting systems (Bermúdez et al., 2018; Piadeh et al., 2022).

With advances in data-driven technology, surrogate models or alternative methods has emerged as a key strategy for addressing challenges related to computational efficiency (Contreras et al., 2020; Demiray et al., 2021; Ivanov et al., 2021). Deep learning technology can provide an approach to complement physics-based urban flood modeling. It has the potential for high-fidelity simulations by leveraging its capability to learn from data with complex non-linear relationships, while also enabling near-real-time predictions due to significant improvements in computational efficiency. These methods offer significant advantages in terms of computational efficiency and practical applicability to hydrological modeling (Ren et al., 2025). Recent studies have also applied deep learning models such as U-Net and long short-term memory (LSTM) to precipitation nowcasting and time series forecasting, which are essential components of flood forecasting systems (Ghaderpour et al., 2023; Li et al., 2024). These applications highlight the growing role of deep learning in real-time flood prediction and risk mitigation efforts. Convolutional neural networks (CNNs) can extract and reduce the dimensionality of important features by using convolutional and pooling layers. In combination with other neural networks or physics-based methods, several studies have demonstrated the utility of the CNNs technology in flood predictions (Chen et al., 2021; Guo et al., 2021; Kabir et al., 2020). More recently, physics-informed neural networks (PINNs) have emerged as hybrid models that embed governing physical equations into deep learning frameworks to improve generalizability (Qi et al., 2024; Yang et al., 2024).

The application of deep learning techniques to spatial resolution enhancement in hydrological applications is a constantly growing research field. It has been widely applied to spatial data reduction and enhancement tasks in remote sensing, such as downscaling and segmentation of satellite imagery (Cui et al., 2023; Galar et al., 2020; Jian et al., 2025; Li et al., 2023; Mohamadiazar et al., 2024; Poehls et al., 2025; Xu et al., 2022; Zeng et al., 2024). In a similar context, super-resolution (SR) techniques have been actively developed in the fields of computer vision and medical imaging to enhance the resolution of images and videos. SR refers to the process of reconstructing high-resolution (HR) images from low-resolution (LR) inputs. In the context of urban flood mapping, LR flood maps often lack the spatial detail necessary for street-level risk assessment, whereas HR flood maps provide more refined information but are computationally expensive to generate using traditional physics-based models. SR addresses this gap by learning the transformation from LR to HR through data-driven models, thereby enabling the efficient production of HR-like outputs from computationally cheaper LR simulations.

Among various deep learning techniques, CNNs and generative adversarial networks (GANs) have been widely adopted for SR tasks due to their strong capability in extracting spatial features (Wang et al., 2021). CNN-based architectures such as U-Net have shown effectiveness in capturing flood extent and fine-scale hydrodynamic patterns by leveraging encoder–decoder structures (He et al., 2023; Yin et al., 2024). GANs have also been employed to enhance flood inundation maps by learning sharper details from coarse inputs (Demiray et al., 2021). Both approaches have been successful in various applications, including imaging and pixel classification, and enhancement, demonstrating promising results in fields such as satellite image mapping, hydrology, and climatology. For example, the CNN-based SRM (SRMCNN) model, which uses CNN-based SR techniques to map fine-grained land cover from remote sensing images. CNN has proven effective in capturing the spatial characteristics of geographical objects and extrapolating calibrated methods to other research areas. The results showed a performance improvement of 3–5 % over existing SR mapping methods (Jia et al., 2019). In Cheng et al. (2020), they reported that ResLap, a CNN-based method for downscaling in HR, could generate high-quality climate images and improved performance in climate predictions. Their results showed similar or improved levels of bias correction compared with NASA's CMIP6 climate model products for both daily precipitation and temperature. Additionally, the performance of various SR imaging methods in enhancing gridded rainfall data to a higher resolution was compared (Golla et al., 2024). Recently, SR with a U-Net structure was applied to hydrodynamic flood modeling under spatial and temporal dynamic storm conditions (He et al., 2023). In Lombana and Martínez-Graña (2022), a methodology combined with SR algorithms for flood mapping of satellite images in small bodies of water was proposed. They applied image reconstruction in the preprocessing step and used the SEN2RES tool of the Sentinel Application Platform (SNAP) developed by the European Space Agency (ESA). In a similar context, a study combined a UNET-based downscaling model with a two-dimensional rainfall runoff simulation (Jian et al., 2025). The proposed terrain-based attention U-Net model (TA-U-Net) improved the accuracy of regional precipitation estimates by downscaling satellite precipitation data. This approach shares strong similarities with image super-resolution methods in term of algorithmic logic. Another study sought to recover lost physical details and information from LR numerical simulations by generating HR flood maps using CNN-based U-Net and GAN models (Yin et al., 2024). Most existing studies have adopted U-Net-based architectures for SR applications, while CNN-based SR techniques for urban flooding have not been fully explored. Furthermore, current methods lack a thorough analysis of the reliability and applicability of SR techniques in urban flood scenarios. These limitations underscore the need for a more comprehensive approach to improving the spatial resolution and accuracy of urban flood simulations. The objective of this study is to introduce FLO-SR, CNN-based super-resolution (SR) framework developed to enhance the spatial resolution of urban flood simulations. The key contributions of this research are as follows:

- A deep learning-based SR model, FLO-SR, is proposed for upscaling coarse-resolution flood simulation outputs to finer resolutions.
- The performance of FLO-SR is evaluated across multiple scale factors (2, 4, and 8 ×), using both synthetic data and physics-based simulation outputs.
- The accuracy of FLO-SR in reconstructing flood depth and delineating flood extent is quantitatively assessed using multiple evaluation metrics.
- The computational efficiency of the FLO-SR framework, when integrated with physics-based models, is compared against conventional high-resolution simulations.

The remainder of this paper is organized as follows. Section 2 describes the study areas, FLO-SR framework, the data preparation process, and super-resolution modeling. Section 3 presents the experimental setup, results, and evaluation. Section 4 discusses the findings and implications. Finally, Section 5 concludes the study and outlines potential directions for future research.

## 2. Materials and methods

### 2.1. Study region

#### 2.1.1. Houston: Hurricane Harvey flood event

Hurricane Harvey struck southeastern Texas in August 2017. This led to extraordinary flooding of the Houston area from August 26–28, resulting in rainfall totals with staggering statistical return periods of over 9,000 years for the three-day totals in some locations and over 2000 years for the cumulative rainfall of the storm. The flooding caused by Hurricane Harvey was catastrophic, resulting in over 70 deaths. The financial damage caused by Harvey was estimated to be approximately $125 billion, making it the second most financially destructive hurricane in USA history (Noh et al., 2019).

The model domain was approximately 100 × 65 km in the Houston metropolitan area, including Buffalo, Brays, Greens, Hunting, Sims, and White Oak Bayous. Urban flood modeling at 4 km resolution was forced by rainfall data, based on the West Gulf River Forecast Center's radar-based operational multi-sensor precipitation estimator. The amount of precipitation in the Houston metropolitan area ranged from 800 to 1,100 mm from August 25–31, 2017. On August 27, the daily rainfall estimates exceeded 300 mm in most areas and 500 mm in the Brays Bayou lowlands. For more information about inundation modeling of Hurricane Harvey flooding in Houston, refer to (Noh et al., 2019). Fig. 1 shows an enlarged map of the maximum flood depth resulting from a 2D flood simulation of Hurricane Harvey at 10 m resolution. To construct and validate LR and HR pairs of model input data and learn SR processes at various scales, the LR input data were generated by bicubicizing the 10 m resolution Hurricane Harvey 2D flood simulation results to resolutions of approximately 20, 40, and 80 m.

#### *2.1.2. Portland: Urban flood event*

Another study domain includes an urban catchment located in the western region of the Willamette River basin, Portland, OR, USA. Portland is located in a low-lying area between the Oregon Coast and Cascade Ranges, resulting in lower precipitation levels compared with other regions of the Pacific Northwest. It receives an average precipitation of 930 mm annually, based on 1981–2010 data (Chang, 2007; Cooley and Chang, 2017). The majority of precipitation occurs from November to April, primarily in the form of rain. The average annual temperature is 11.3 °C and the mean annual precipitation is approximately 1,095 mm (Franczyk and Chang, 2009). The model domain covers an area of approximately 9 km$^2$.

To construct the model training dataset, rainfall data of urban flooding events for a total of 95 h, from 01:00 on December 6 to 9, 2015, were used. The simulations included five land cover types: water bodies, impervious areas, bare land, grass, and forests. Different runoff coefficients were applied depending on the land cover type. 2D pluvial flood simulations were conducted at different spatial resolutions of 1, 2, 4, and 8 m to construct and validate LR and HR pairs of model input data and to elucidate the SR process at different scales. Fig. 2 shows an enlarged map of the maximum flood depth resulting from a 2D flood simulation of a Portland flood at 1 m resolution.

### *2.2. FLO-SR: Urban flood SR framework*

The proposed framework, FLO-SR, is designed to enhance the spatial resolution of urban flood simulations using SR techniques (Fig. 3). This methodology comprises two major procedures: (1) data preprocessing and (2) SR modeling, as described below.

#### *2.2.1. Data preprocessing*

This procedure prepares the input dataset for utilization in the SR model. The maximum inundation depth images, simulated by physics-based modeling, are preprocessed to create paired LR and HR datasets. In this study, input urban flood data are generated using the H12 physics-based urban flood model (Noh et al., 2018). The input data for the H12 model includes a digital elevation model (DEM) with building footprint data, rainfall, and land cover maps. H12 employs hybrid parallel-computing technologies that combine message-passing interface (MPI) and OpenMP to enable efficient high-resolution simulations. Although H12 can simulate the interactions between 2D surface runoff and 1D sewer network flows, in this study, only the H12 2D flow analysis module was used to assess the effects of maximum inundation depth images at different resolutions for the SR process.

Flood maps for an entire simulation area (e.g., maximum inundation depth distributions) often lack sufficient detail for evaluating specific inundation patterns. Therefore, pixel dimensions are selected to enable a detailed examination of the flooding patterns at street level within the simulated area. But, since only a small percentage of pixels in a full-size image correspond to the inundated area, directly processing the full-resolution image would lead to significant class imbalance and computational inefficiency (Li et al., 2023). Therefore, we only retain images where at least 0.2 % of the pixels in each image are labeled as water. For the HR data, we utilized the high-resolution simulated results as synthetic truth. Two methods were employed in the preprocessing stage to construct LR data. First, interpolation techniques were applied to the LR data during the SR process to enhance image quality. Interpolation involves estimating intermediate values between discrete data points and is widely used in digital image processing for resizing images and correcting spatial distortions. Among the various interpolation techniques, bicubic interpolation was used to generate LR images because of

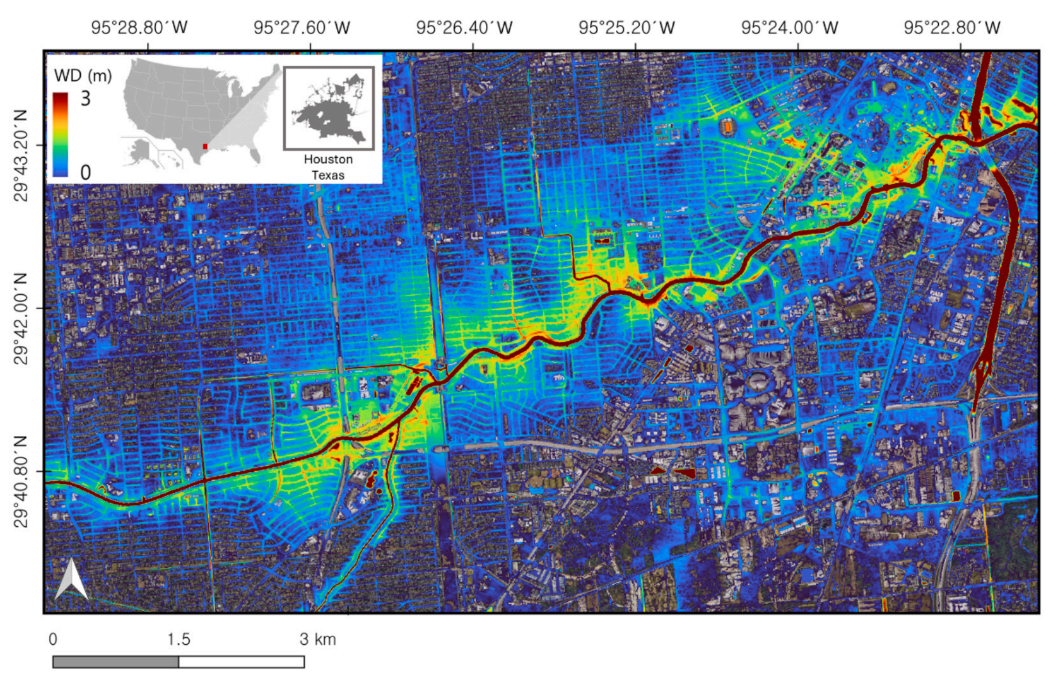


**Fig. 1.** An enlarged maximum inundation depth map for 10 m spatial resolution of the Hurricane Harvey flood, TX, USA.


**Fig. 2.** An enlarged maximum inundation depth map for 1 m spatial resolution in the Portland flood, OR, USA.

superior performance in interpolation accuracy (Khaledyan et al., 2020). Second, the physics-based urban flood simulation results generated at LR were directly used as LR inputs through a resizing process. These LR and HR images are then paired to form input–output datasets for training the SR model. Data augmentation techniques, including flipping and rotation, are applied to the validated data pairs to enhance dataset diversity and reduce the risk of overfitting during model training. This preprocessing pipeline is essential for the effective development of SR models. Table 1 summarizes all datasets utilized in this study, including their types, spatial and temporal characteristics, and the method used to generate LR data from HR data. The generated image datasets for the study can be accessed as described in the "Availability of data and material" section.

#### 2.2.2. SR modeling

The SR model of the FLO-SR framework is based on an enhanced deep SR (EDSR) (Lim et al., 2017) which improves performance and reduces memory usage by removing unnecessary batch normalization layers from the SRResNet architecture. Unlike conventional image classification tasks, in which batch normalization aids in faster and more stable learning, SR tasks can limit the flexibility of the network and produce pixel values that differ from the original data. EDSR is well-suited for processing flood simulation data because it can restore HR images more accurately by removing batch normalization.

The FLO-SR model begins by extracting initial features from LR input images using a convolutional layer with 64 filters and a kernel size of 3 $\times$ 3. These features are then passed through a series of 16 residual blocks. Each residual block has two convolutional layers. The first layer uses the rectified linear unit (ReLU) activation function, and the second layer omits activations to retain raw feature representation. Skip connections within each residual block mitigates vanishing and explosion gradients. After feature extraction, the processed features were converted into HR outputs through a pixel shuffle upsampling process. The number of channels is dynamically increased depending on the scaling factor. For instance, in the 4 $\times$ and 8 $\times$ SR configurations, the number of output channels before pixel shuffle is set to 256 and 512 respectively, corresponding to the spatial resolution increase required. The pixel shuffle operation then rearranges these channels to form the final HR image. To train the FLO-SR model, the mean absolute error (L1 loss) between the predicted HR images and the ground truth reference images was minimized. The model was trained with a batch size of 16 using the Adam optimizer, with default $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\varepsilon = 1 \times 10^{-7}$. A piecewise constant learning rate schedule was applied: the initial learning rate was set to $1 \times 10^{-4}$ and reduced to $5 \times 10^{-5}$ after 5,000 steps to promote convergence. Each model was trained over 400 epochs, with 100 steps per epoch to maintain consistency across experiments. For each case study domain, 70 % of the data were used for training and 30 % for validation. These hyperparameter settings were determined through preliminary tuning experiments and are provided in the supplementary data (Table S1).

### 2.3. Evaluation metrics

We evaluated FLO-SR performance in multiple aspects including image pixel and maximum flood depth reconstructions, and flood area similarities, using all metrics evaluated at the pixel level. Methods for image quality evaluation can be divided into subjective and objective approaches. Subjective evaluation reflects human perception but is time-consuming and lacks clear criteria, whereas objective evaluation employs mathematical algorithms and is widely used. Both methods were used to ensure a comprehensive assessment. Binary classification indicators were used to evaluate the flooded area simulation performance of the FLO-SR. The estimation method measures the proportion of observed and predicted events that are correctly predicted, considering into account hits (correct predictions) and false alarms (incorrect predictions) per pixel based on a minimum acceptable threshold. The minimum allowable threshold was set at 15 cm, considering that it is an urban watershed with high building density (Jasour et al., 2022; Wing et al., 2017).

#### 2.3.1. Peak signal to noise ratio

The peak signal to noise ratio (PSNR) is the most commonly used quality assessment technique to measure the quality of reconstruction of loss image compression codecs (Sara et al., 2019). It evaluates the similarity of images by calculating the loss information for the resulting image quality. In image and video compression quality degradation, the

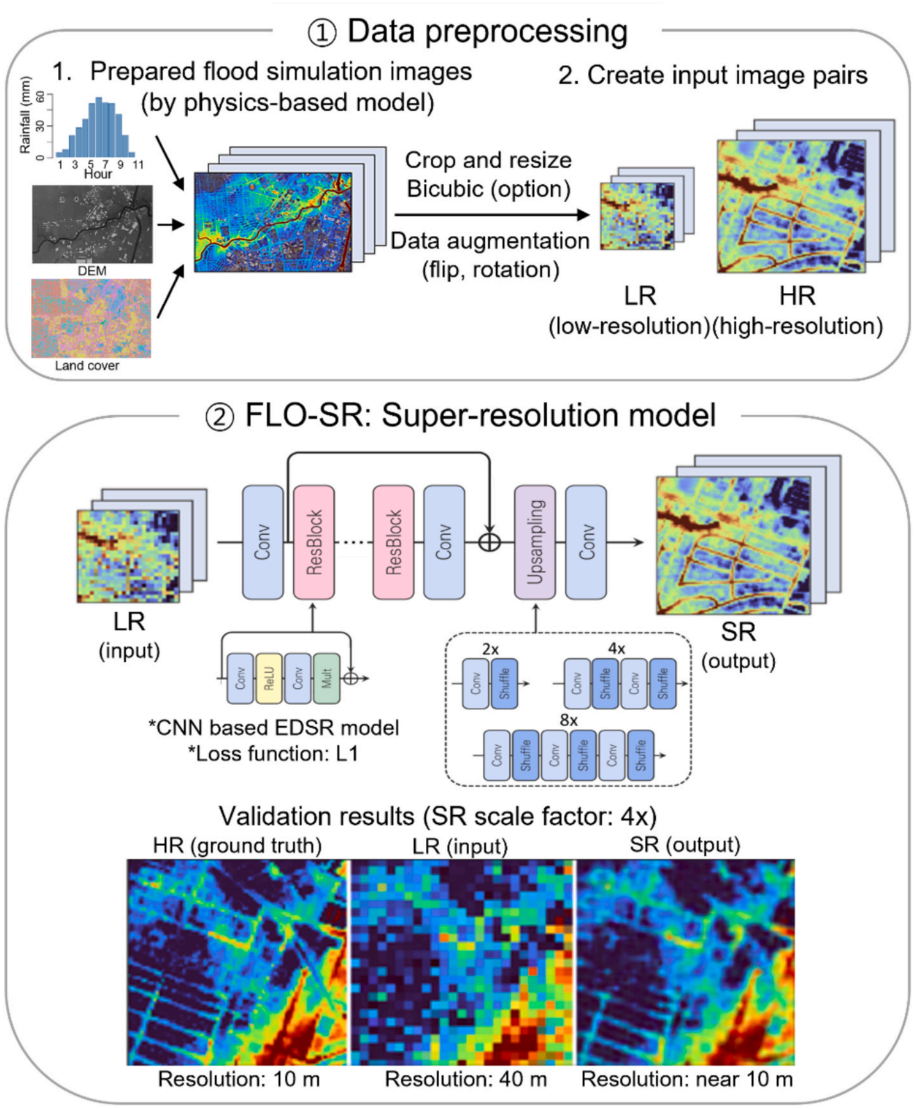


**Fig. 3.** Illustration of urban flood super-resolution framework: FLO-SR.

**Table 1**
Details of datasets utilized in the FLO-SR.

| Type | Dataset | Product type | Spatial resolution | Simulation period | Source/Generation method |
|---|---|---|---|---|---|
| Physics-based flood model (H12) | Houston precipitation | Radar-based precipitation estimates | 4 km | August 25–31, 2017 (hourly) | West Gulf River Forecast Center operational data |
| | Portland precipitation | Ground-based observations | – | December 6–9, 2015 (hourly) | Local meteorological stations |
| | DEM / Land Cover | Elevation and land cover maps | 1 m (Portland), 10 m (Houston) | – | Used as input for H12 model |
| FLO-SR | Houston Hurricane Harvey flood maps | Maximum inundation depth maps | HR: 10 m; LR: 20 m, 40 m, 80 m | August 25–31, 2017 | H12 model (HR) Bicubic interpolation (LR) |
| | Portland urban flood maps | Maximum inundation depth maps | HR: 1 m; LR: 2 m, 4 m, 8 m | December 6–9, 2015 | H12 model (HR, LR) at each resolution |

PSNR value varies from 30 to 50 dB for 8-bit data representation and from 60 to 80 dB for 16-bit data. In wireless transmission, accepted range of quality loss is approximately 20 to 25 dB (Deshpande et al., 2018). PSNR can be defined as:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{M^2}{\mathrm{MSE}}\right) \quad (1)$$

Where M is the maximum value that a pixel value can have regardless of the size of an image. MSE is the mean squared error and calculates the pixel value difference between the super-resolution image and the original image. N is the total number of pixels in the image, $I_y$ is the i-th pixel value of the original image, and $I_{SR}$ is the pixel value of the super-resolution image (Wang et al., 2022). The higher the PSNR value is, the better the result, so the higher the PSNR, the higher the similarity to the original image.

#### 2.3.2. Structure similarity index method

The structure similarity index method (SSIM), proposed by Wang et al. (2004), is based on human perception of structural information within an image. This metric is widely used for comparing the structural similarity between images (Blau and Michaeli, 2018; Sara et al., 2019). The SSIM is more in line with the intuition of the human eye because it comprehensively measures three factors: luminance, contrast, and structure, which are recognized as the main contents of the human visual system (Wang et al., 2022). SSIM is expressed as:

$$I(I_{SR}, I_y) = \frac{2\mu_{I_{SR}}\mu_{I_y} + C_1}{\mu_{I_{SR}}^2 + \mu_{I_y}^2 + C_1} \tag{2}$$

$$c(I_{SR}, I_y) = \frac{2\sigma_{I_{SR}}\sigma_{I_y} + C_2}{\sigma_{I_{SR}}^2 + \sigma_{I_y}^2 + C_2} \tag{3}$$

$$s(I_{SR}, I_y) = \frac{\sigma_{I_{SR}}I_y + C_3}{\sigma_{I_{SR}}\sigma_{I_y} + C_3} \tag{4}$$

$$SSIM(I_{SR}, I_y) = I(I_{SR}, I_y)^{\alpha} \bullet c(I_{SR}, I_y)^{\beta} \bullet s(I_{SR}, I_y)^{\gamma} \tag{5}$$

Eqs. (4)-(6) compare the luminance, contrast, and structure of the two images, respectively. $\mu$ represents the mean, $\sigma$ represents the standard deviation, and $\sigma$ ($I_{SR}$, $I_y$) represents the covariance. $C_1$, $C_2$, and $C_3$ are constants. $\alpha$, $\beta$, and $\gamma$ mean weights. SSIM has values between 0 and 1, and the closer the value is to 1, the more similar the two images are. Therefore, the higher the value, the higher the quality.

#### 2.3.3. Root mean square error

The root mean square error (RMSE) measures the error between predictions flood depth and actual values, emphasizing that smaller RMSE values indicate better performance.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_{SR} - I_y)^2} \tag{6}$$

#### 2.3.4. Hit rate

The Hit rate (H) indicates the proportion of pixels identified as wet in the FLO-SR results that correspond to pixels identified as wet in the benchmark physics-based model (H12), excluding overpredictions. This proportion is expressed as a value that ranges from 0 (indicating that all pixels identified as wet in the FLO-SR are dry in the benchmark physics-based model (H12) results) to 1 (indicating an accurate prediction of all wet pixels in the model).

$$H = \frac{M_1B_1}{M_1B_1 + M_0B_1} \tag{7}$$

#### 2.3.5. False alarm ratio

The false alarm ratio (F) measures the percentage of pixels identified as wet by the FLO-SR that are dry in the reference. This metric assesses the FLO-SR's propensity to overestimate flood coverage, with values ranging from 0 (indicating no false alarms) to 1 (all false alarms).

$$F = \frac{M_1B_0}{M_1B_0 + M_1B_1} \tag{8}$$

#### 2.3.6. Critical success index

The critical success index (C) is a metric of discrepancy between the FLO-SR and the benchmark physics-based model (H12) pixels, encompassing both underprediction and overprediction. The C assumes values within the range of 0 to 1, wherein a C value of 0 signifies a benchmarking result predicting the exact opposite of the FLO-SR, while a C value of 1 denotes a perfect reproduction of the model result.

$$C = \frac{M_1B_1}{M_1B_1 + M_0B_1 + M_1B_0} \tag{9}$$

As illustrated in Table 2, the term $M_1$ denotes to the number of pixels flooded by the benchmark physics-based model (H12) with a water depth of at least 15 cm, while $B_1$ signifies the number of pixels flooded by the FLO-SR at least 15 cm, and $M_0B_0$ represents the number of pixels that remain non-flooded across all models.

**Table 2**
Flood pixel descriptors in a binary classification. States: wet (1) and dry (0) and 15 cm as a threshold water depth to classify the status of each pixel.

| | **Flooded area by FLO-SR** | **Non-flooded area by FLO-SR** |
|---|---|---|
| **Flooded area by H12** | $M_1B_1$ | $M_1B_0$ |
| **Non-flooded area by H12** | $M_0B_1$ | $M_0B_0$ |

### 2.4. Experimental setup

The experimental design intentionally includes two distinct approaches for generating LR inputs (bicubic interpolation for Houston and direct physics-based simulations for Portland) to demonstrate the versatility and limitations of FLO-SR under controlled and realistic conditions. The bicubic interpolation approach in Houston provides a controlled environment where spatial patterns and features from the original HR simulation are systematically downgraded, enabling an isolated assessment of the SR algorithm's capacity to recover lost information. This approach is analogous to traditional SR benchmarking in computer vision and establishes an upper performance bound. In contrast, the physics-based approach in Portland represents realistic operational conditions where different resolution simulations produce inherently different hydrodynamic behaviors, boundary interactions, and flood patterns. These two approaches (bicubic interpolation for Houston and physics-based simulation for Portland) were deliberately chosen to evaluate the performance of the SR model under different conditions. The Houston case provides ideal conditions that show the upper performance limit of the SR algorithm itself, while the Portland case demonstrates its applicability in actual operational conditions. By comparing these approaches, we gain insights into both the theoretical capabilities of SR techniques and their practical applicability in operational flood modeling contexts.

The proposed FLO-SR was evaluated in two regions: Hurricane Harvey in Houston, TX and an urban flood event in Portland, OR. The experimental settings in these two cases differed, depending on the method used to generate.

LR input data that reflect the specific characteristics of flood events and regional modeling requirements. For Houston, LR input images were generated by upsampling HR maximum flood inundation maps (10 m resolution) using bicubic interpolation. To analyze changes in flood patterns in detailed areas at the road level, a pixel size of 240 × 240 was selected. HR maximum flood inundation maps were upsampled to coarser spatial resolutions of 20, 40, and 80 m using bicubic interpolation as the input to the SR process. We used the simulation result with 10 m data as a reference and evaluate the accuracy and performance of other simulation results with different spatial resolutions using accuracy assessment metrics. This method provides a controlled environment for evaluating FLO-SR by isolating the effects of the SR process. This is consistent with the existing SR practice of synthetically generating LR images to ensure a consistent comparison with the HR ground truth. In contrast, for Portland, LR input data were derived directly from physics-based hydrological-simulations. Maximum inundation maps were generated at resolutions of 1, 2, 4, and 8 m, reflecting a more realistic scenario in which LR data were produced using operational flood modeling systems. We used the simulation results with 1 m data as a reference and evaluated the accuracy and performance of other simulation results with different spatial resolutions using accuracy assessment metrics. Unlike the bicubic interpolation approach used in Houston, this method captured the inherent complexities and uncertainties of physical modeling, providing an opportunity to evaluate

FLO-SR under practical urban flood modeling conditions. Considering the size of the simulated area, the pixel size was selected as 200 × 200.

The FLO-SR performance was evaluated using a combination of metrics that assessed different aspects of SR accuracy. The image quality and structural similarity were evaluated using PSNR and SSIM. The accuracy of flood depth reconstruction was estimated using the RMSE, whereas binary classification metrics such as hit rate (H), false alarm ratio (F), and critical success index (C) were used to assess the accuracy of flood extent delineation. To better understand the reconstruction of the flood depths, we systematically mapped the image pixels to their corresponding flood depth values. This was achieved by implementing a color palette that converted pixel RGB values into normalized flood depth values, which were then used to quantitatively compare LR, SR, and HR images. These comprehensive metrics allowed for a thorough evaluation of the capability of FLO-SR to reconstruct both flood depth and extent across diverse scenarios. All the experiments were conducted using TensorFlow version 2.15.0, which was implemented on a system equipped with NVIDIA A100-PCIE-40 GB GPUs and Intel Xeon Gold 6248(R) CPUs.

## 3. Results

### 3.1. FLO-SR model results: visual inspection and overall statistics

#### 3.1.1. Houston: Hurricane harvey flood SR results

To validate the FLO-SR performance, 300 maximum inundation depth images with 240 × 240 pixels from the Hurricane Harvey flooding simulations in Houston were analyzed. FLO-SR was employed to reconstruct HR inundation maps from the LR inputs at various scales (The resolution of the SR magnitude is as follows: 2, 4, and 8 ×). LR input images were generated by upsampling HR maximum inundation maps (10 m resolution) using bicubic interpolation, which is commonly applied in conventional SR studies.

Fig. 4 presents examples of HR (synthetic truth), LR (input), and SR (output) images at various scales. The 2 × scale results, which showed minimal visual difference from the HR images due to the relatively small spatial resolution gap (20 m → 10 m), can be found in the supplementary data (Fig. S1). Higher values of PSNR and SSIM indicate better reconstruction quality and greater similarity to the reference HR image in terms of both pixel-level accuracy and spatial structure. FLO-SR consistently reconstructed more accurate spatial patterns of inundation depth compared to the LR inputs, as evidenced by both visual inspection and enhanced evaluation metrics, including PSNR and SSIM. The structural similarity between the SR and original HR images confirmed that FLO-SR effectively restored the overall inundation distribution at all SR scale factors. When comparing the SR results and HR images, it was found that key information, such as predicted flood contours and depths around road edges and buildings, was restored with minimal distortion. Even in areas with high inundation depths, FLO-SR successfully captured the spatial distribution of maximum inundation depths, similar to the original HR results. Textural details were best preserved at low scale factors, while higher scale factors showed visually consistent inundation patterns, but some blurriness was observed in complex regions. More SR results for Hurricane Harvey flood can be found in the supplementary data (Fig. S2 and S3).

A quantitative evaluation of the FLO-SR performance on the 300 validation images is presented in Table 3 and Fig. 5. The evaluation

**Table 3**
Evaluation metrics for SR performance across different scale factors in the Houston Hurricane Harvey flood event.

| Scale factor | PSNR (dB) | SSIM | RMSE (m) | H | F | C |
|---|---|---|---|---|---|---|
| 2 × | **41.57** | **0.96** | **0.03** | **0.98** | **0.01** | **0.97** |
| 4 × | 32.30 | 0.83 | 0.07 | 0.94 | 0.03 | 0.91 |
| 8 × | 17.75 | 0.61 | 0.14 | 0.80 | 0.06 | 0.73 |

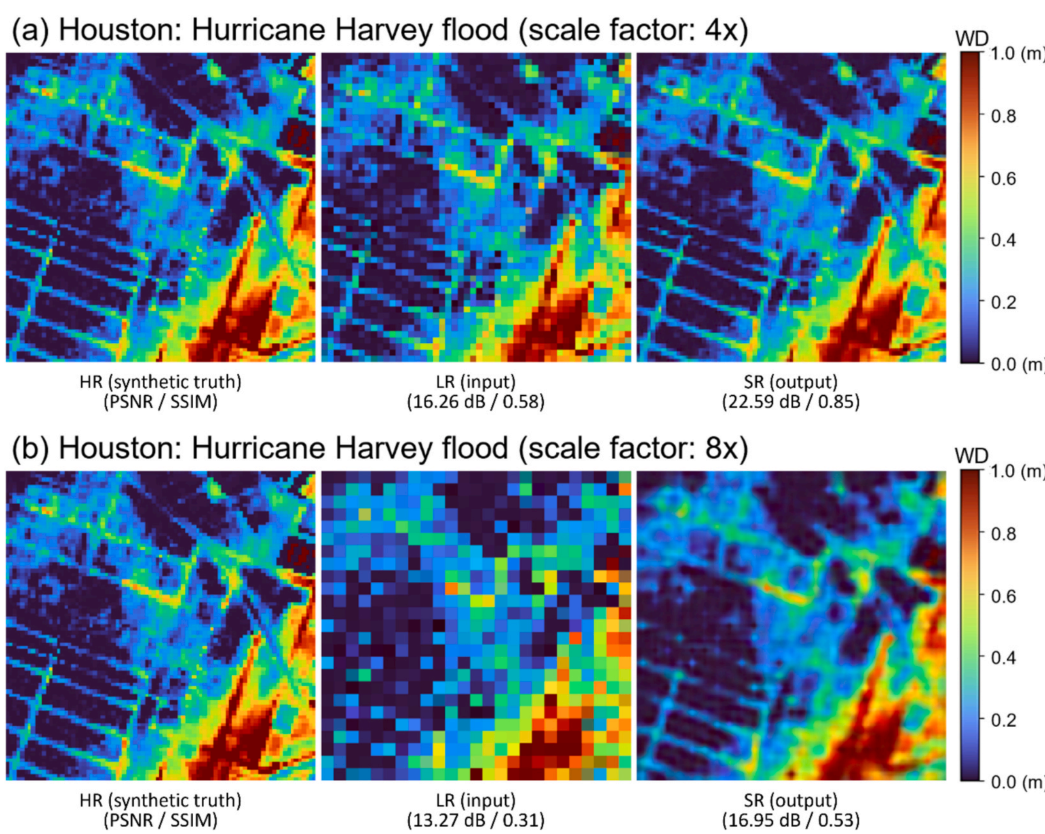


**Fig. 4.** Hurricane Harvey flood SR results by scale factors: (a) 4 × scale factor; (b) 8 × scale factor.

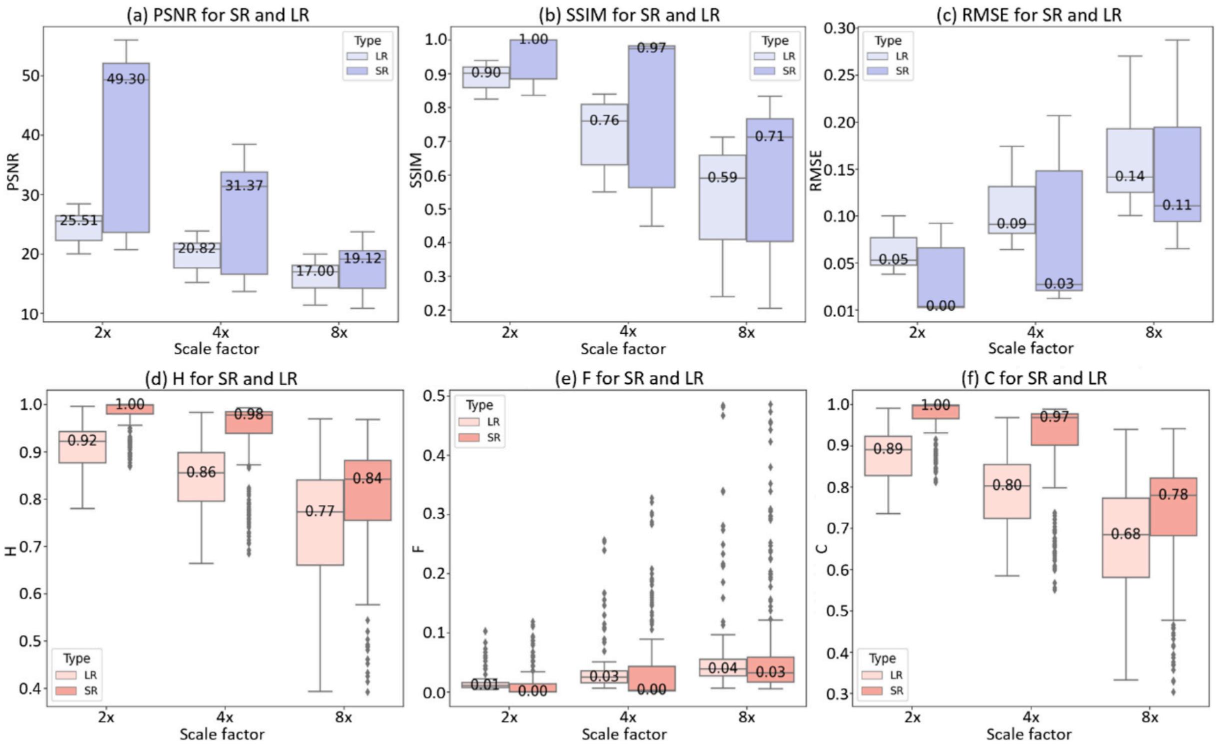


**Fig. 5.** Box plot comparing evaluation metrics of LR and SR performance at different scale factors for the Hurricane Harvey flood event: (a) PSNR; (b) SSIM; (c) RMSE; (d) H; (e) F; (f) C.

focused on three key aspects: image quality (PSNR and SSIM), inundation depth estimation accuracy (RMSE), and flood extent classification (H, F, and C indices). The results showed an average PSNR of 41.6 dB and SSIM of 0.96 at the 2 × scale, indicating relatively accurate reconstruction quality and strong structural similarity to the original HR image (Table 3). For inundation depth estimation, the RMSE was 0.03 m, indicating high accuracy in reconstructing flood depth. In addition, the evaluation results for flood extent classification were H = 0.98, F = 0.01, and C = 0.97, indicating reliable flood area classification. At the 4 × scale, the PSNR and SSIM decreased to 32.30 dB and 0.83, respectively, but still demonstrated strong performance. The RMSE slightly increased to 0.07 m, but showed overall satisfactory results with H = 0.94, F = 0.03, and C = 0.91. However, at the 8 × scale, the PSNR and SSIM further decreased to 17.75 dB and 0.61. The RMSE for flood depth reached 0.14 m, H and C decreased to 0.80 and 0.73, respectively, whereas F increased to 0.06, indicating increasing difficulties in accurately capturing flood depth and spatial details.

The box plots in Fig. 5 provide a detailed comparison of the evaluation metrics between the LR inputs and the FLO-SR outputs across all scale factors, focusing on the resolution enhancement rates achieved by FLO-SR. As shown in Fig. 5 (a), FLO-SR significantly improved the PSNR values compared with the LR inputs at all scales. The resolution enhancement rates for the average PSNR were 68.65, 36.83, and 8.53 % at the 2, 4, and 8 × scales, respectively. FLO-SR achieved the greatest improvement at the 2 × scale, where the median PSNR reached 49.30 dB, compared to 25.51 dB for LR inputs, demonstrating substantial enhancement in image quality. The resolution average enhancement rates for SSIM, which measures the structural similarity, were 7.37, 13.92, and 13.73 % at the 2, 4, and 8 × scales, respectively. FLO-SR maintained consistently high SSIM values at lower scale factors, with medians of 1 and 0.97 at the 2 and 4 × scales, respectively, indicating superior preservation of structural integrity. However, at the 8 × scale, the SSIM decreased to a median of 0.71, reflecting a reduced ability to preserve fine structural details at higher magnifications. Distribution of RMSE for flood depth estimation, demonstrated a clear trend of increasing errors at higher scales. The average RMSE enhancement rates were 56.23, 32.35, and 10.70 % at the 2, 4, and 8 × scales, respectively. At the 2 × scale, FLO-SR achieved a median RMSE of 0 m, representing flawless depth reconstruction compared with the LR input. At the 4 × scale, the median RMSE increased slightly to 0.03 m, and at the 8 × scale, it increased to 0.11 m, reflecting the increasing difficulty in accurately estimating depth at coarser resolutions. Based on the resolution enhancement rates across all metrics, the 2 × scale factor demonstrated the most significant improvements, achieving near-perfect fidelity in terms of image quality, structural similarity, and flood depth estimation. The performance of FLO-SR gradually declined as the scale factor increased, but even at the 8 × scale, the reconstructed images and flood maps retained key spatial patterns and structural coherence.

In summary, Table 3 and Fig. 5 collectively highlight the robustness of FLO-SR in enhancing both the accuracy and consistency of SR reconstruction. At lower scale factors, FLO-SR preserves textural details, structural similarities, and flood extent accuracies, making it highly effective for urban flood analysis. Although the performance decreased slightly at higher scale factors, the reconstructed images and flood maps retained the core spatial patterns and structural integrity, demonstrating the applicability of the model across a range of multiple scaling factors.

### 3.1.2. *Portland: Urban flood SR results*

The Portland urban flood SR image measured 200 × 200 pixels, allowing for a more detailed SR analysis at the road level. A total of 300 validation images were used for evaluation. Similar to the Houston event, the analysis focused on three key aspects. However, unlike in the Houston event, the LR input data for this scenario were generated using a physics-based hydrological simulation model, rather than bicubic upsampling. This approach was taken to provide a more realistic representation of the LR flood data and better reflect real-world conditions.

Fig. 6 provides a visual comparison of 1 m HR maximum flood depth images, LR simulation results, and SR images for each scale factors (4 and 8 ×). As shown in the figure, the lower the scale factor, the more

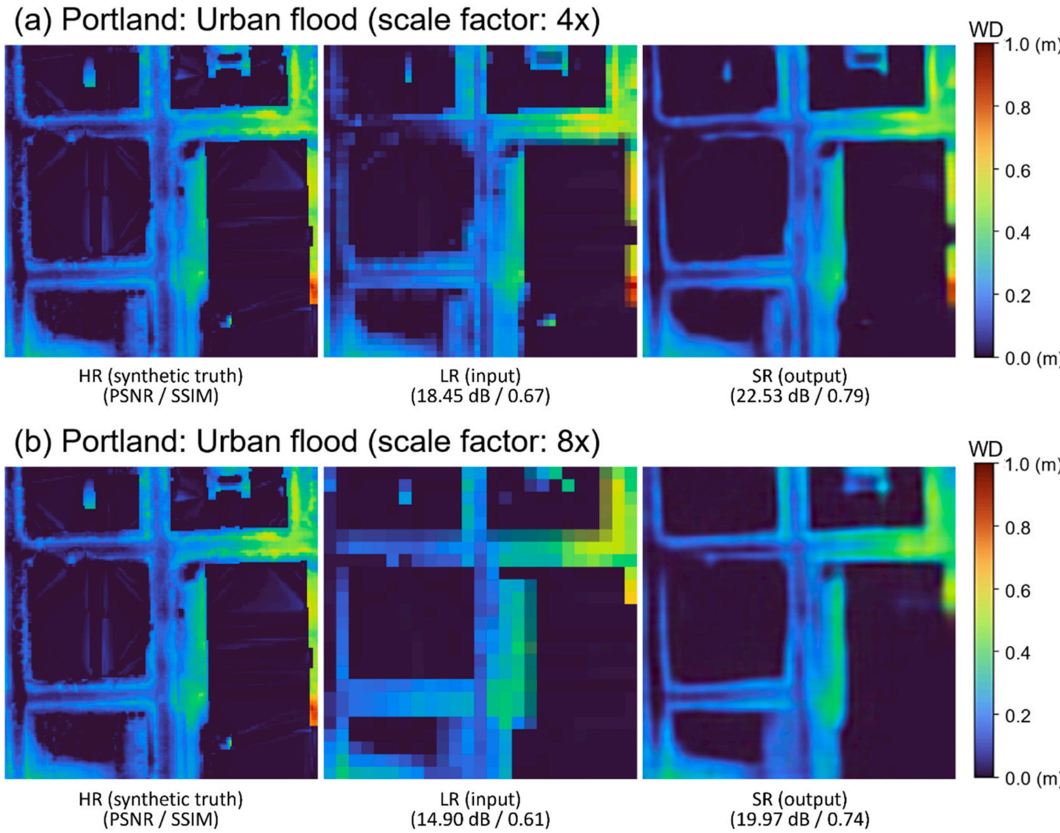


**Fig. 6.** Portland urban flood SR results by scale factors: (a) 4 × scale factor; (b) 8 × scale factor.

accurately FLO-SR restored the detailed features within the flood boundary and flood extent. The 2 × scale results, which showed minimal visual difference from the HR images due to the relatively small spatial resolution gap (2 m → 1 m), can be found in the supplementary data (Fig. S4). At the 4 × scale, the contours of the flooded roads, buildings, and other urban structures closely matched the original HR results, maintaining the integrity of the flood distribution. However, as the scale factor increased, the flood depth boundaries became increasingly blurred, particularly at the 8 × scale, where the differences in the flooded area and boundary details compared with the HR ground truth became more pronounced. More Portland urban flood super-resolutions results can be found in supplementary data (Fig. S5 and S6).

Table 4 presents the averages of various indicators for 300 Portland validation images (physics-based hydrological simulation of LR input data). As shown in the previous example, the accuracy gradually decreased as the scale factor increased. The PSNR indicated a relatively high reconstruction quality ranging from 24.49 dB to 20.59 dB. Similarly, SSIM values decrease from 0.85 at the 2 × scale to 0.76 at the 4 × scale and to 0.69 at the 8 × scale, reflecting reduced structural accuracy at higher scale factors. The RMSE for the flood depth estimation increased with the scale factor, indicating that the accuracy decreased with decreasing resolution. Specifically, the RMSE increased from 0.06 m at the 2 × scale, to 0.08 m at the 4 × scale, and finally to 0.10 m at the 8 × scale. Binary classification metrics such as the H, F, and C showed a more significant decrease as the scale factor increased. The H decreased from 0.82 at the 2 × scale to 0.70 at the 8 × scale, indicating less accuracy in identifying flooded areas. The F remained low, slightly increasing from 0.01 at the 2 × scale to 0.02 at the 8 × scale. The C, which evaluates H and F combined, decreased to 0.78 at the 2 × scale, 0.69 at the 4 × scale, and 0.64 at the 8 × scale, showing that although the composition characteristics of the input data varied, a trade-off occurred as the scale factor (2 to 8 × ) increases, as in the previous example.

**Table 4**
Evaluation metrics for SR performance across different scale factors in the Portland urban flood event.

| Scale factor | PSNR (dB) | SSIM | RMSE (m) | H | F | C |
|---|---|---|---|---|---|---|
| **2 ×** | **24.49** | **0.85** | **0.06** | **0.82** | **0.01** | **0.78** |
| **4 ×** | 22.11 | 0.76 | 0.08 | 0.72 | **0.01** | 0.69 |
| **8 ×** | 20.59 | 0.69 | 0.10 | 0.70 | 0.02 | 0.64 |

Fig. 7 presents the accuracy metrics of the FLO-SR outputs in Portland event compared to the LR physics-based simulations across different scale factors. Same as previous results, FLO-SR significantly improved the PSNR values compared with the LR inputs at all scales. The resolution enhancement rates for the average PSNR were 10.67, 19.03, and 31.83 % at the 2, 4, and 8 × scales, respectively. FLO-SR achieved the greatest improvement at the 2 × scale, where the median PSNR reached 24.59 dB, compared to 22.26 dB for LR inputs, demonstrating substantial enhancement in image quality. Similarly, the SSIM values measure the structural similarity between the reconstructed images and the HR ground truth. The resolution average enhancement rates for SSIM were 5.92, 16.55, and 33.77 % at the 2, 4, and 8 × scales. FLO-SR maintained consistently high SSIM values at lower scale factors, with medians of 0.86, 0.76 and 0.70 at the 2, 4 and 8 × scales. The average RMSE enhancement rates were 24.49, 33.75, and 44.11 % at the 2, 4, and 8 × scales. At the 2 × scale, FLO-SR achieved a median RMSE of 0.06 m. At the 4 × scale, the median RMSE increased slightly to 0.08 m, and at the 8 × scale, it increased to 0.10 m. These results indicate that the FLO-SR performance is highly dependent on the scale factor, with lower scale factors allowing more accurate resolution enhancement and structural similarity preservation. In summary, the 2 × scale factor provided the highest accuracy improvement in both the image quality and flood mapping metrics, making it the most effective scale factor for urban flood analysis using FLO-SR.

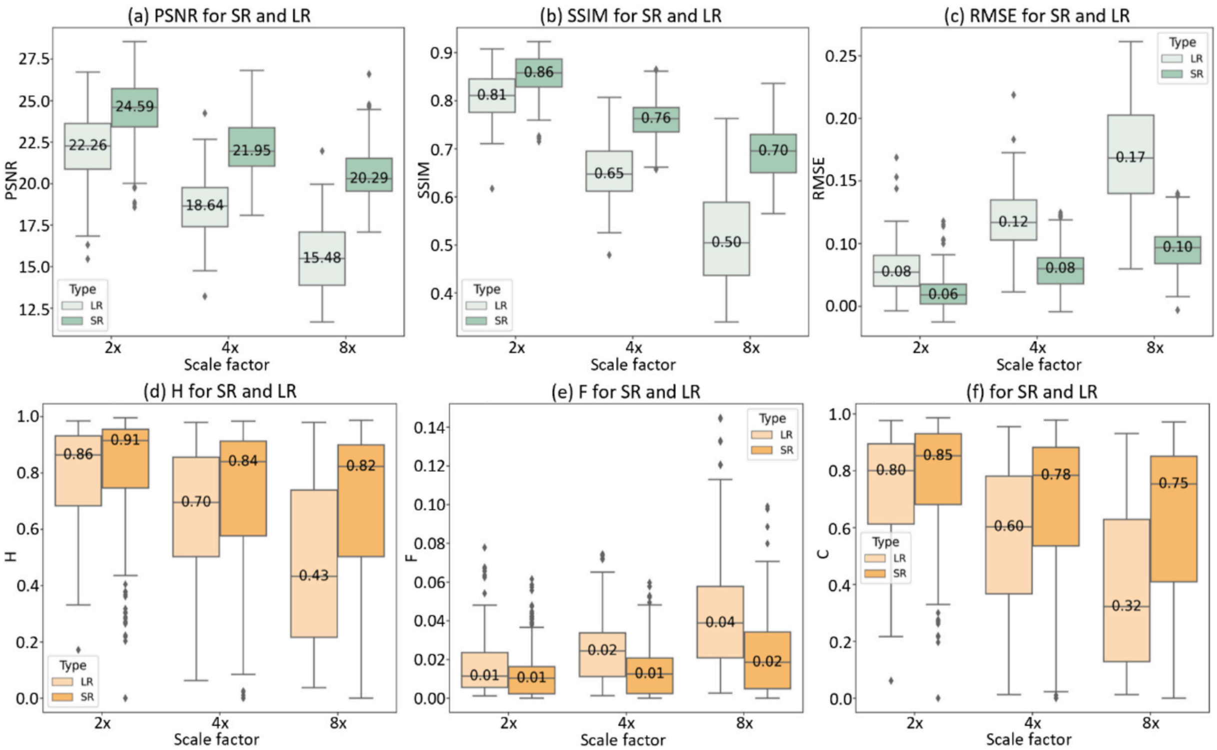


**Fig. 7.** Box plot comparing evaluation metrics of LR and SR performance at different scale factors for the Portland urban flood event: (a) PSNR; (b) SSIM; (c) RMSE; (d) H; (e) F; (f) C.

As shown in the results, FLO-SR demonstrated strong potential in both scenarios. Although some flood information was blurred and some information was lost as the scale of super-resolution increased, the overall pixel accuracy and the evaluation results for flood depth and area were within a reliable range. FLO-SR achieved better performance with bicubic interpolation-generated LR images compared with physics-based modeled inputs. This can be attributed to the uniform and consistent nature of bicubic interpolation, which retains structural similarity to the original HR data. However, physics-based modeled data often introduce spatial artifacts, abrupt transitions, or noise due to hydrodynamic simulation inaccuracies, parameter uncertainties, or non-uniform patterns of inundation. These challenges reduce the ability of the FLO-SR model to reconstruct fine-scale details accurately. Nevertheless, physics-based modeling approaches are important for real-world applications, as LR flood maps are typically generated via fluid dynamics simulations to reflect real-world flood risk scenarios that capture complex urban flood processes and the interactions between rainfall, runoff, and infrastructure. Future work will focus on introducing preprocessing techniques to mitigate the discontinuous inundation patterns and spatial artifacts resulting from the grid size of physics-based modeled LR data, and on improving the robustness of FLO-SR to handle noise and inconsistencies in physics-based modeled data by expanding the training dataset to include a wider range of flood scenarios. By addressing these challenges, FLO-SR can further enhance its utility in urban flood analysis and applications across a variety of data sources and operational contexts.

### 3.2. Configuration sensitivity and in-depth analysis

The sensitivity analysis focused on evaluating the impact of different LR image reconstruction methods on the performance of FLO-SR. The analysis aimed to quantitatively analyze the impact of the two approaches for generating LR input images on the results of FLO-SR, and to investigate the performance and training time of FLO-SR as a function of the image size. This analysis was conducted specifically for the urban flood event in Portland.

The results in Figs. 8 and 9 show the sensitivity of FLO-SR to image size based on two distinct input data construction methods: bicubic interpolation (Fig. 8, Fig. S7) and physics-based modeling (Fig. 9, Fig. S8). Each evaluation index represents the average value of the 300 Portland verification images. Input data generated using bicubic interpolation (Fig. 8), FLO-SR showed consistent performance across a variety of metrics regardless of the image size and SR enlargement. Bicubic interpolation produce input images with well-defined spatial features and smooth gradients, thereby minimizing noise and abrupt transitions. These characteristics allowed FLO-SR to accurately reconstruct image quality, flood depth, and spatial patterns with minimal distortion, regardless of the image size or SR scale. In the case of input data generated using physics-based modeling (Fig. 9), similar to in that used in previous results, the average values tended to be slightly lower than those for bicubic interpolation for all image sizes and scales. To investigate the loss in super-resolution performance when using physics-based model inputs instead of bicubic interpolation, we conducted a comparison at a fixed image size of 200 × 200 pixels. As the scale factor increased from 2 to 8×, PSNR decreased (by 7.13–27.50 %) and RMSE increased (by 25.0–66.67 %), reflecting a decline in reconstruction accuracy. C values showed smaller changes (a reduction of 0.14–16.13 %) but still indicated a mild decrease. A notable observation was the decrease in F with increasing image size at higher SR scales, such as 8 × . This phenomenon can be explained by larger image sizes providing additional spatial information, which may allow FLO-SR to better distinguish between flooded and non-flooded areas. Conversely, at reduced image sizes, the paucity of spatial detail may amplify the probability of misclassifications, particularly at magnifications where the reconstruction of finer-scale features becomes more challenging. For the results of various evaluation indicators by image size, refer to supplementary Fig. S9 and Fig. S10.

To better evaluate the accuracy of FLO-SR in reconstructing flood depths, scatter plots were created by systematically comparing predicted depth values to their corresponding reference HR data (Figs. 10

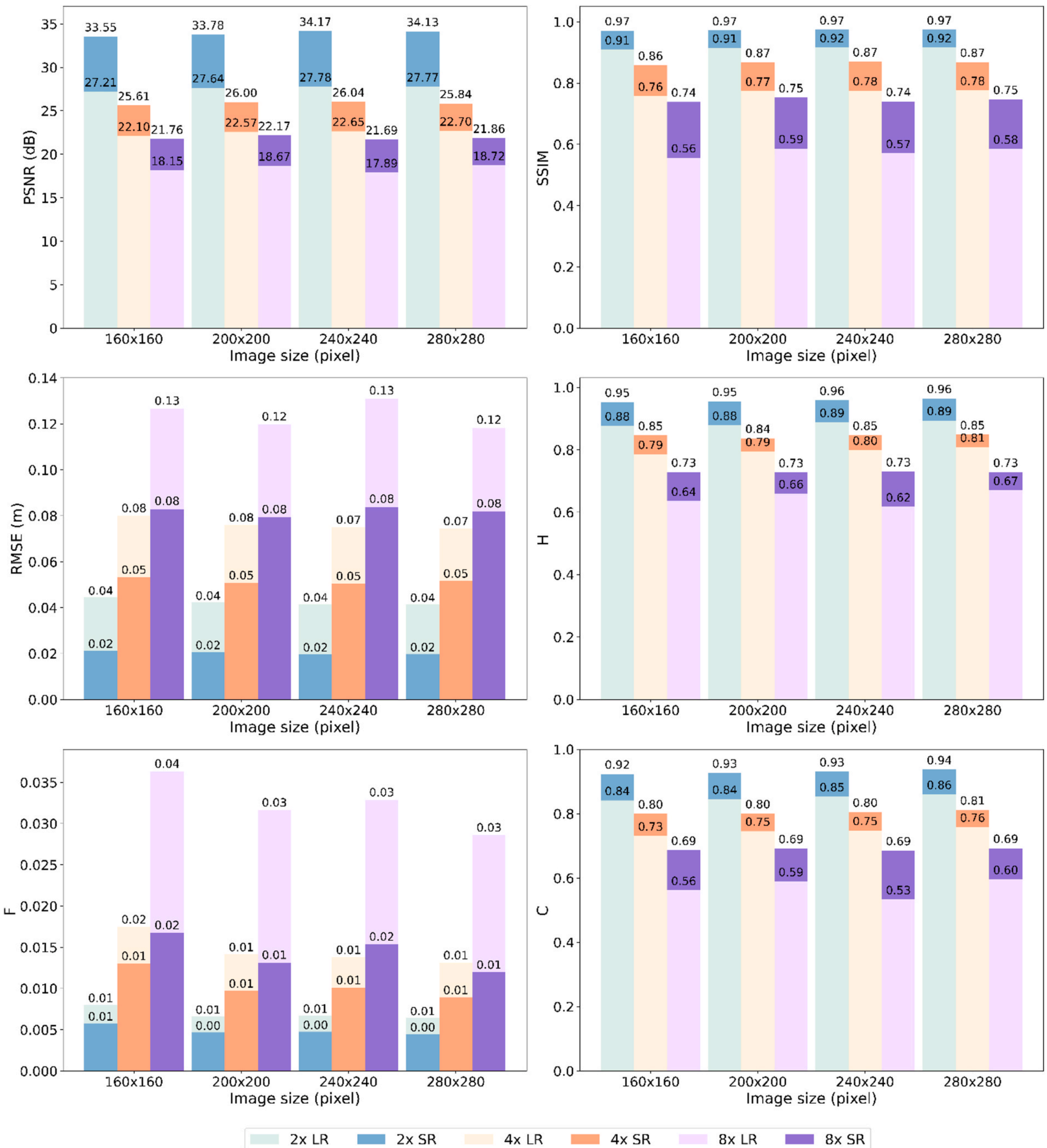


**Fig. 8.** FLO-SR sensitivity to image size using bicubic interpolation for input data generation.

and 11). Additional results at different locations are provided in the supplementary data (Fig. S9 and S10). This was achieved by implementing a color palette that converted pixel RGB values into normalized flood depth values, which were then used to quantitatively compare LR, SR, and HR images. This process involved creating a specified color palette, and normalizing the flood depth value range to an RGB color between 0 and 255. Each image was processed using this color palette to convert pixel colors to their corresponding flood depth values. In Fig. 10, the LR input exhibited significant scatter and variance along the diagonal, reflecting the inaccuracy of the flood depth mapping. However, the FLO-SR output exhibits much denser clustering along the diagonal, indicating a higher level of accuracy in reconstructing the flood depth values. This trend was consistent across various input conditions, demonstrating the robustness of the FLO-SR framework for approximating HR data.

In Fig. 11, the scatter plots illustrate the challenges associated with physically complex flood patterns. Additional results are shown in the supplementary data (Fig. S11). The LR inputs (4 and 8 m) exhibited discontinuous flood patterns and areas, as well as different maximum inundation depths, compared to the HR output (1 m) from the physics-based model. Although LR inputs struggled to capture these complex flood details, resulting in a larger spread in the scatter plots, FLO-SR successfully recovered most of the missing details, as evidenced by the improved alignment along the diagonal. The minor deviations observed in areas of extreme flood depth and the deviations due to missing flood spatial information in the input data highlight the necessity for further improvements in handling variable input data.

Pixel profiles were extracted along specific cross-sections of the flood depth maps, providing a detailed visualization of how each method captures spatial and depth variations in the flood patterns. This analysis

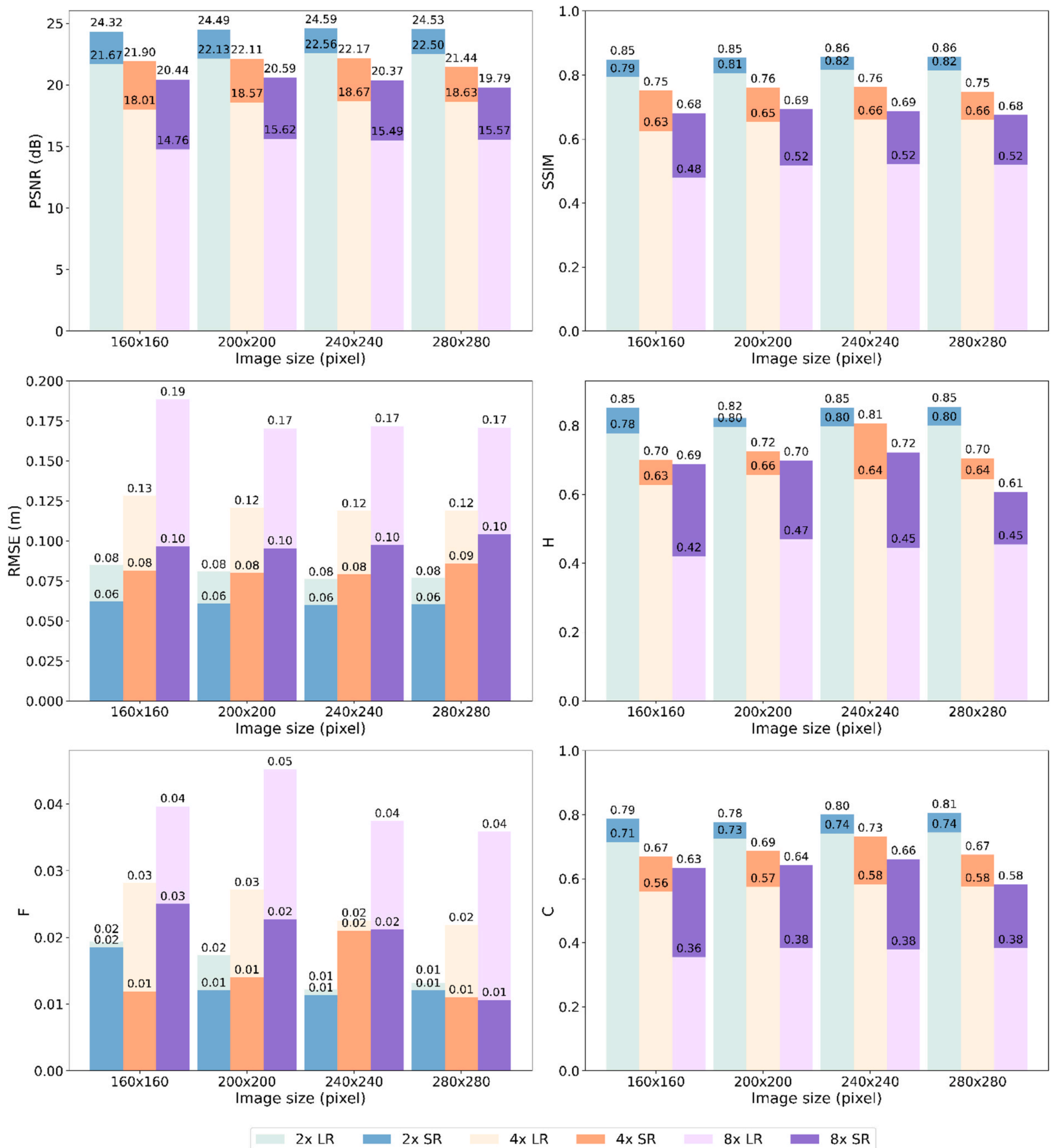


**Fig. 9.** FLO-SR sensitivity to image size using physics-based model for input data generation.

visually demonstrates how FLO-SR bridges the pixel intensity gap between LR and HR. In the context of image analysis, pixel intensity refers to the pixel values that represent specific characteristics depending on the image type. In this analysis, pixel intensity for brightness or luminance was converted into a gray scale image and evaluated. Fig. 12 shows a pixel profile comparison between the HR, LR, and SR results for various cases. The HR profile serves as a reference, whereas the LR profile represents a coarser and less accurate input. The SR profile shows that FLO-SR reconstructs the flood depth with higher fidelity, bridging the pixel gap between LR and HR. Overall, the results show that SR restores the pixel intensity better than LR, but with increasing magnification. It was found that the model did not fully replicate the sharp transitions in the HR reference (implying high inundation depths in small areas). In particular, when LR inputs are used in physics-based models, they exhibit larger inconsistencies as the input information becomes increasingly sparse. This makes it difficult for FLO-SR to accurately interpolate depth changes, resulting in a smoother SR profile.

In summary, information loss issues arise depending on the method used to construct the input data, which has a significant impact on the performance and accuracy of the SR model. In this study, we confirmed that the input data generated from HR data by bicubic interpolation with a smooth gradient enables consistent reconstruction because the information loss is minimized. However, in actual situations, physical modeling using HR information (e.g., DEM or DSM, land cover) requires extensive calculation time and computing resources. Therefore, application of the existing SR methodology, which generates an LR input by applying bicubic interpolation to the results of an HR physical model, has clear limitations. However, as a result of applying SR to the physical model results using LR input in this study, it was possible to reconstruct the inundation depth and area within a reliable range. From a broader

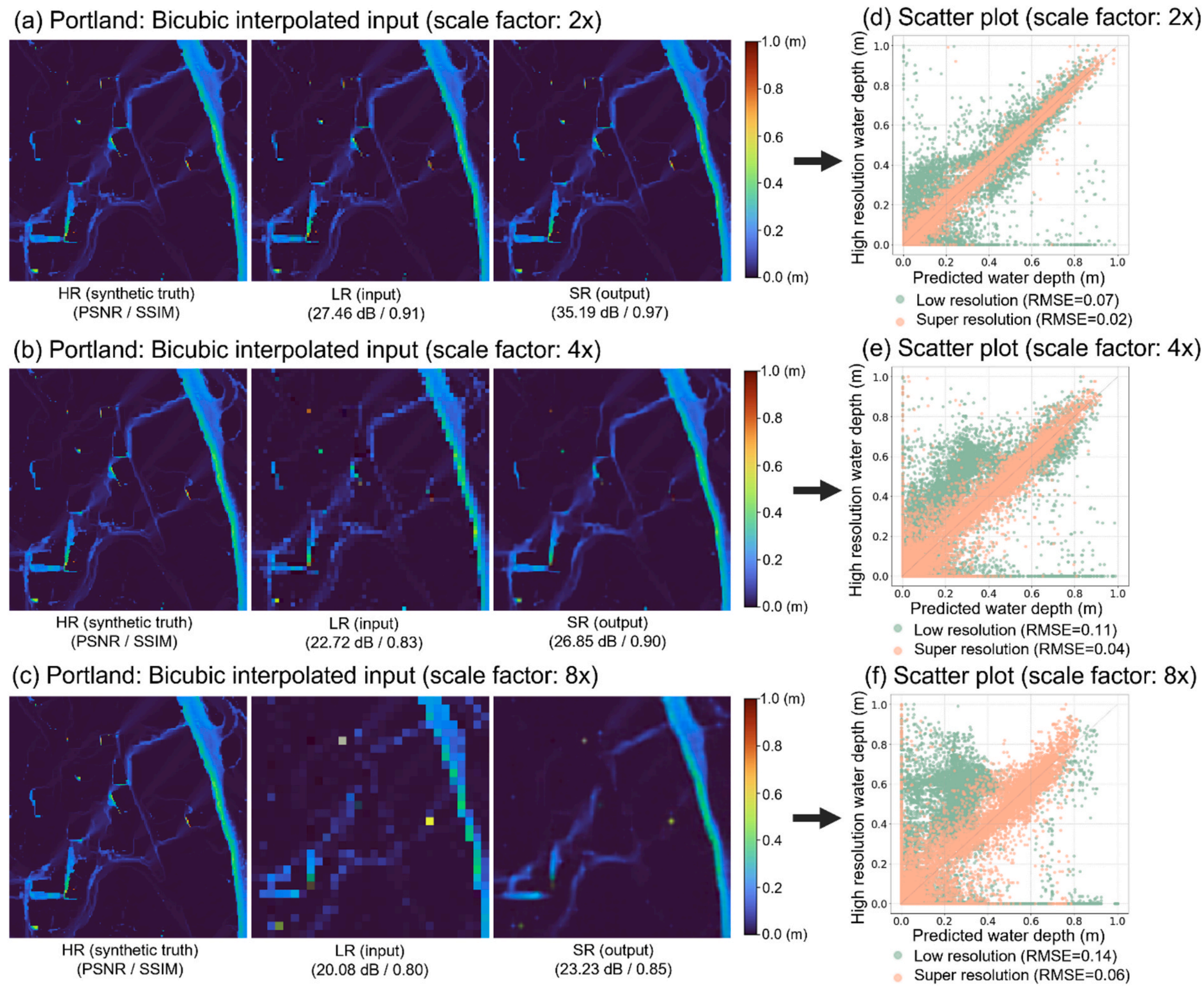


**Fig. 10.** Evaluation of water depth predictions across scale factors using bicubic and SR methods: (a–c) Comparison of HR, LR, and SR water depth maps for Portland at scale factors of 2, 4, and 8 ×; (d–f) Scatter plots comparing predicted SR water depth values against HR references, with RMSE values for LR and SR methods at each scale factor.

perspective, this model represents an important advance in flood modeling, as it allows the efficient generation of HR flood depth maps from LR data. However, similar to existing deep learning-based models, SR models are highly dependent on learned patterns, which means that they perform best when the input data is clean and representative of the target output. Therefore, future studies could focus on improving the robustness of the model by reducing the grid like noise in the input image, preprocessing techniques such as artifact correction, or integrating multi-physics information.

### 3.3. Analysis of FLO-SR computational efficiency

In urban flood modeling that deals with large data sets, assessing computational efficiency is important. It is well-known that for any type of model, the greater the amount of data and information, the longer is the runtime. In this section, we analyze the computational efficiency of the SR model according to the size (number of pixels) of the input image and the scale of SR, and evaluate the computational efficiency compared to the physics-based model.

Fig. 13 illustrates the relationship between the computational efficiency and image size of FLO-SR, showing the correlation between the image size and training time at different SR scale factors. The computational efficiency depends on the image size and scale factor, and as the image size increases, the training time increases exponentially. When the image size was gradually increased from 160 × 160 to 280 × 280, the calculation time increased by 1.33 times on average, requiring approximately 33 % additional calculation time. The computational efficiency of FLO-SR varied at different magnifications (2, 4, and 8 ×). For example, the average training time at each scale was 2,616, 1,163, and 864 s for the 2, 4 and 8 × scales, respectively. The 2 × scale required 124.8 % more calculation time than the 4 × scale and the 4 × scale required 34.6 % more calculation time than the 8 × scale. These results suggest that the image size and computational cost must be balanced for appropriate SR reconstruction performance while maintaining computational efficiency.

To estimate the approximate speedup between the physics-based model and FLO-SR, runtimes for 2D physics-based flood simulations at various spatial resolutions were assessed for the 95 h historical event in the Portland area. The simulation was performed at resolution intervals of 1 m up to 2 m and at 2 m intervals from 2 to 8 m for the 95 h historical event in the Portland area. For physics-based modeling, the number of computational grid cells varied significantly with resolution, ranging from 9,009,045 grids at 1 m resolution to 2,252,235 grids at 2 m resolution, 563,065 grids at 4 m resolution, and 140,773 grids at 8 m resolution. The reduction in the number of grid cells at coarser resolutions highlights the trade-off between computational cost and spatial accuracy. In all resolution cases, the simulation used a time step of 0.1 s and applied the same open boundary conditions at the edge of the study area. The computational run times for each resolution are listed in Table 5. The proposed FLO-SR model has clear advantages over physics-based

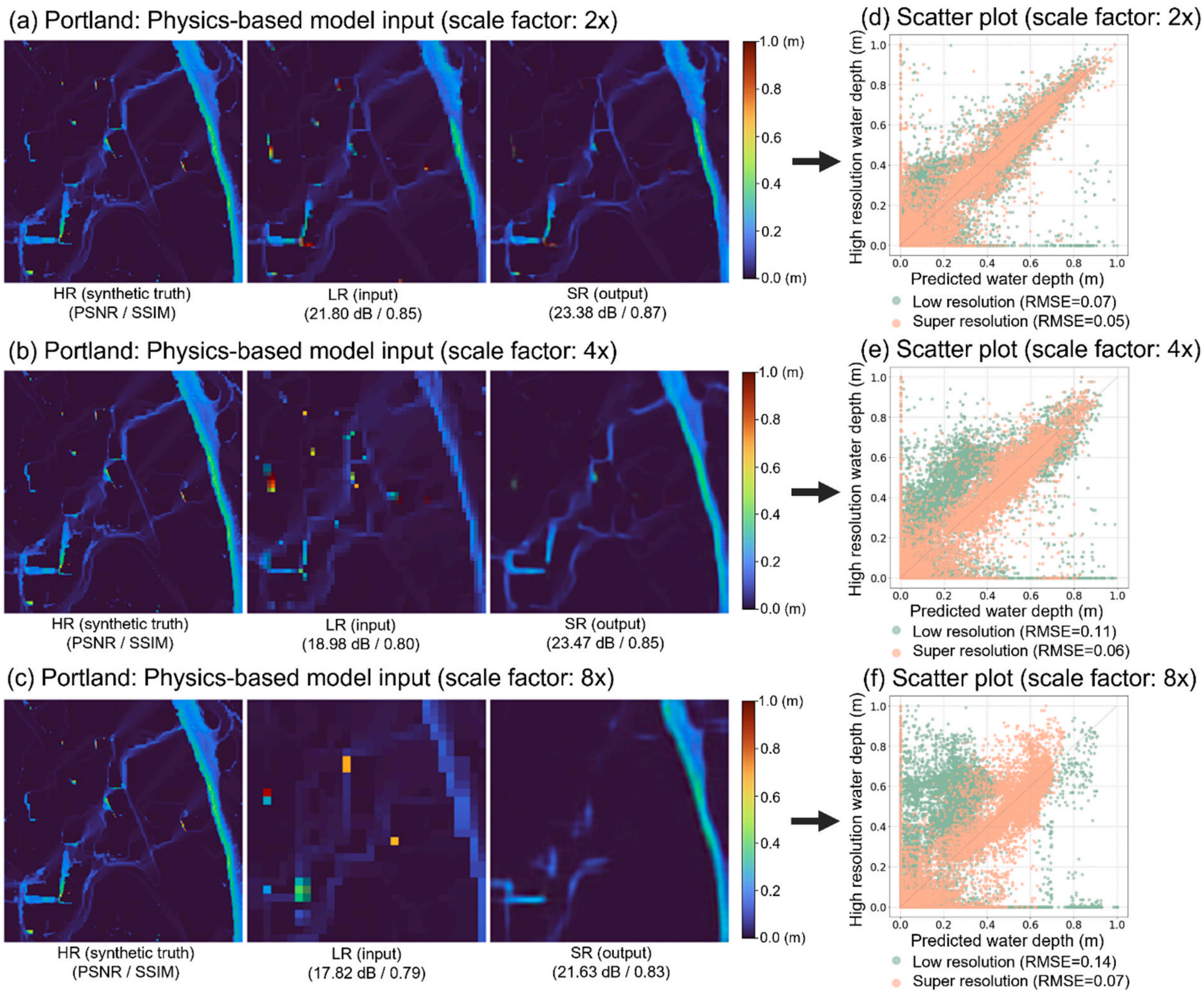


**Fig. 11.** Evaluation of water depth predictions across scale factors using physics-based hydrological simulation and SR methods: (a–c) Comparison of HR, LR, and SR water depth maps for Portland at scale factors of 2, 4, and 8 ×; (d–f) Scatter plots comparing predicted SR water depth values against HR references, with RMSE values for LR and SR methods at each scale factor.

hydrological models in terms of computational cost and scalability. For example, when simulating at 2 m resolution using the physical model, the computation requires 10.73 h with 96 CPUs. In contrast, an alternative approach involves performing the simulation at a coarse 4 m resolution using the physical model, which requires 3.35 h with 48 CPUs, and subsequently applying FLO-SR to reconstruct the 2 m resolution results. The FLO-SR model requires 0.59 h for training and 0.03 h for validation when using a single GPU. The total computation time for this approach was 3.97 h, representing a 63 % reduction in computational time compared with directly running the physical model at 2 m resolution. Similarly, when targeting 4 m resolution, running the physical model at 8 m resolution (1.52 h) and applying FLO-SR (0.27 h for training, 0.03 h for validation) results in a total computation time of 1.82 h, achieving a 45.7 % reduction compared with the directly simulation at 4 m resolution (3.35 h). These results highlight the efficiency of FLO-SR in reducing computational costs while maintaining HR flood predictions.

This significant reduction in computational costs highlights the potential of FLO-SR as an efficient alternative to HR flood simulations. By employing FLO-SR, high-fidelity flood predictions can be obtained while substantially decreasing the computational burden, making large-scale urban flood modeling more feasible and resource-efficient. This drastic reduction in computational cost makes FLO-SR particularly advantageous for time-sensitive applications such as real-time flood forecasting and urban planning. However, it is important to acknowledge the limitations of directly comparing the computation time of the FLO-SR and physics-based models. The methodologies employed by the two approaches differ fundamentally, as FLO-SR focuses on enhancing the resolution of existing flood maps using SR techniques, whereas physics-based models simulate flood dynamics from scratch based on complex hydrological and hydraulic processes. Therefore, the comparison did not reflect an entirely equivalent process. Instead, the computation time reported for the physical model represents the time required to generate the flood simulation results, which is indirectly compared to the FLO-SR's time for super-resolving LR inputs. It is also important to note that FLO-SR relies on the availability of LR input data, which is typically generated by a physical model.

## 4. Discussion

### *4.1. Model performance and scale dependent limitations*

Deep learning has been widely applied to enhance raster-based spatial resolution across various fields. In remote sensing, SR improved satellite imagery for land cover classification and water body detection (Li et al., 2023), while in climate modeling, CNN– and GAN-based methods enhanced precipitation and temperature maps (Cheng et al., 2020). Compared to satellite or climate data, flood dynamics exhibit rapid spatiotemporal changes, posing significant challenges for fine-scale hydrodynamic reconstruction. Overall, the FLO-SR results

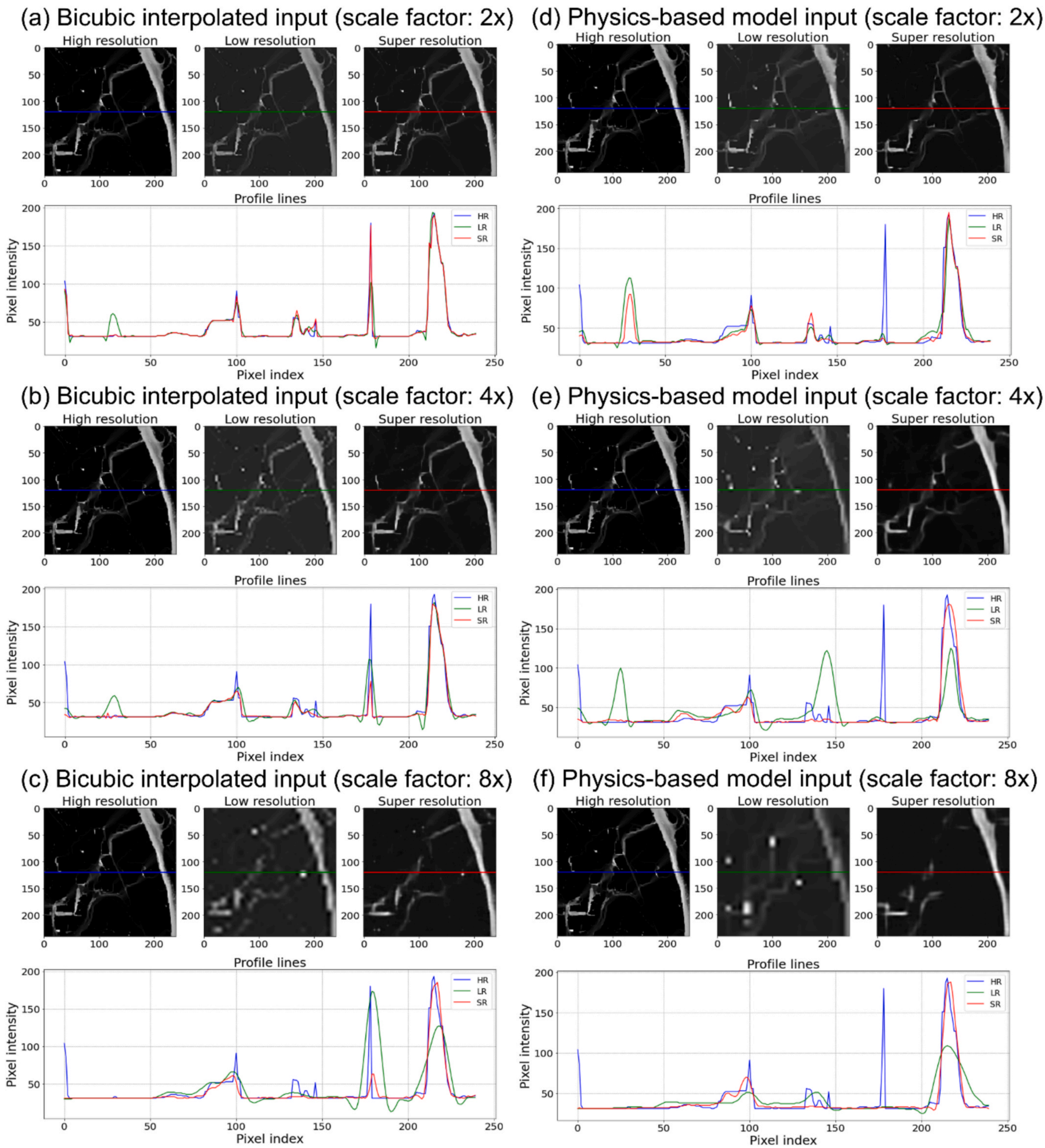


**Fig. 12.** Pixel profile analysis of FLO-SR performance using bicubic interpolation and physics-based hydrological simulation for input data: (a–c) Comparison of HR, LR, and SR profiles for bicubic interpolated inputs at scale factors of 2, 4, and 8 ×; (d–f) Pixel intensity profile comparisons for physics-based model inputs, showing differences in reconstruction accuracy across scale factors.

demonstrate deep learning-based SR technology can effectively improve the spatial details of LR flood simulations and generate HR outputs, aligning with the previous research (e.g. He et al., 2023; Yin et al., 2024).

Our analysis reveals important insights about SR methodology through the comparative evaluation of the Houston and Portland cases. The controlled approach in Houston, where bicubic interpolation was used to generate LR inputs, effectively isolated the intrinsic limitations of the SR algorithm itself. This isolation allowed us to establish theoretical performance boundaries of FLO-SR without the confounding variables present in operational settings. Conversely, the Portland case demonstrated additional challenges that emerge when applying SR to outputs from physics-based models at different resolutions, where each resolution produces inherently different hydrodynamic behaviors and

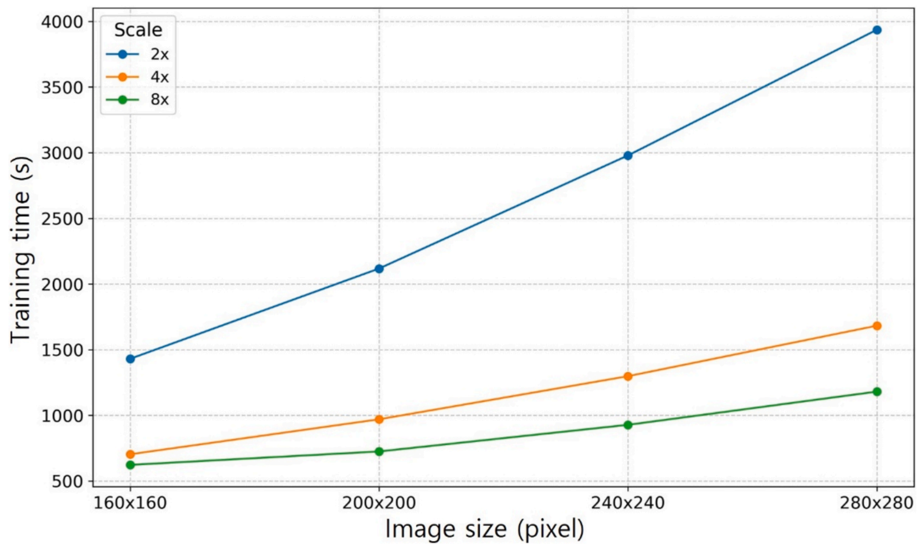


**Fig. 13.** Relationship between image size and FLO-SR training time across different scale factors.

**Table 5**
Comparison of computation time between physics-based urban flood model and FLO-SR.

| Spatial resolution | Physics-based model | | FLO-SR | | |
|---|---|---|---|---|---|
| | Calculation time (hr) | No. CPU | Training time (hr) | Validation time (hr) | No. GPU |
| **2 m** | 10.73 | 96 | 0.59 | 0.03 | 1 |
| **4 m** | 3.35 | 48 | 0.27 | 0.03 | 1 |
| **8 m** | 1.52 | 48 | 0.20 | 0.03 | 1 |

boundary interactions. While the physics-based approach offers greater operational relevance, our results indicate that understanding the baseline performance established through controlled evaluation provides critical context for interpreting SR performance in complex urban flood modeling applications. This dual-approach evaluation becomes particularly valuable given the limited previous research on SR applications in urban flood modeling.

Furthermore, the suitability of SR models in flood modeling may vary depending on spatial resolution, geographical characteristics, and hydrological conditions. For example, He et al. (2023) reported that the grid sizes ranging from 450 to 1,800 m were appropriate for downscaling into a 30 m resolution on large scale SR flood modeling with a U-Net-based model. Yin et al. (2024) applied SR techniques to two-dimensional fluvial flood simulations of river inundation events using U-Net and GAN models. In their study, the LR grid size was defined as 150 feet (45.72 m), while the HR grid size was set to 20 feet (6.096 m), resulting in the number of HR cells being approximately 56 times greater than that of LR cells. Although the study did not specify an optimal grid size for SR application, it demonstrated that the model performance tended to degrade as the input resolution became coarser, indicating the sensitivity of SR accuracy to input grid resolution. In contrast to previous studies that primarily addressed large-scale or fluvial flood scenarios, FLO-SR is specifically designed for urban flood modeling applications where finer spatial resolution is essential due to the complexity of built environments. Unlike He et al. (2023), who focused on rural watersheds with grid refinement from 450–1800 m to 30 m, and Yin et al. (2024), who demonstrated efficiency gains in fluvial flood simulations, FLO-SR aims to capture the intricate interactions of urban infrastructure, such as buildings and road networks, with flood dynamics at sub-10 m scales. While differences in datasets and evaluation protocols hinder direct metric-based comparisons between studies, FLO-SR shows strong capabilities in reconstructing high-resolution flood patterns in complex urban settings, particularly in areas where both pluvial and fluvial processes interact. In urban environments, the complex layout of buildings and roads can lead to significant prediction errors when using coarse grids. Therefore, for accurate urban flood prediction, SR should be applied at a finer grid scale. In the Portland domain, we downscaled urban flood simulation results, covering a range of resolutions from 40 to 2 m, to a fine-scale 1 m resolution. However, the enhancement of SR was also found only up to an 8 × scaling factor. This limitation suggests a need for further strategies to improve SR performance at higher magnification levels. Although digital elevation models (DEMs), land cover data, and building footprints were not explicitly included as separate input channels in FLO-SR, these geospatial features were inherently embedded within the high-resolution simulation outputs used for model training, which were generated through physics-based hydraulic modeling. Consequently, the model was able to implicitly learn patterns associated with urban topography and infrastructure. To overcome the performance degradation observed at higher scaling factors, future versions of FLO-SR could incorporate these geospatial datasets explicitly as additional input channels. By leveraging multi-channel architectures, the model may achieve more accurate reconstructions of complex urban flood dynamics, particularly in environments characterized by dense infrastructure and heterogeneous surface conditions. Furthermore, incorporating remote sensing data such as Synthetic Aperture Radar (SAR)-based flood maps may offer valuable complementary information, enhancing the model's capability to generalize across diverse flooding scenarios. Recent studies have demonstrated that U-Net-based architectures can effectively handle SAR data for urban flood mapping by capturing complex spatial relationships and mitigating the effects of low resolution and speckle noise (Yadav et al., 2022; Zhao et al., 2022). U-Net's encoder-decoder structure, combined with skip connections, enables the model to learn and retain both global context and local spatial details, which is crucial for accurately delineating flood extents influenced by complex topography and infrastructure (He et al., 2023; Yin et al., 2024). Building on these advances, coupling multi-channel geospatial inputs, including SAR-based flood maps, with a U-Net structure may further improve FLO-SR's ability to reconstruct fine-scale flood dynamics across heterogeneous urban environments.

### 4.2. Transferability, input uncertainty, and future directions

This section discusses the model's transferability across domains, the uncertainties involved in flood super-resolution modeling, and future directions for improving the generalizability and robustness of FLO-SR. Although developing a generalized structure of SR flood models is beyond the scope of this research, the model transferability remains an important issue for further improving the applicability of SR in urban flood modeling. While this study is based on a single historical flood event in each case study area, the training data were constructed to include diverse local inundation patterns through spatial sampling. This approach allowed the model to learn from varying flood dynamics, despite being limited to a single event. To evaluate its transferability, the model trained on the Hurricane Harvey flood event in Houston was validated on the Portland urban flood event, which exhibits different hydrological and geographical characteristics. Cross-validation results related to the model's generalization are presented in the supplemental data. A marginal improvement was observed at a lower magnification (2 ×), but the SR output did not show clear enhancement at higher scales (4 and 8 ×) (Table S2, Fig. S12). These results suggest that while the model exhibits limited transferability under certain conditions, its generalization capability across distinct hydrological contexts remains a challenge. To address this, multi-event and multi-region training will be conducted in future research. Moreover, the model's sensitivity to input characteristics was further evident when the model trained using LR data generated by bicubic interpolation was validated on the Portland dataset constructed from physics-based flood simulation results, where the SR performance further declined (Table S3, Fig. S13). In addition to training data diversity, the model's performance is also affected by the configuration of its SR pipeline, particularly the preprocessing strategies applied to LR inputs. However, the qualitative analysis of the preprocessing strategies in SR is not well discussed in the existing literature. One possible approach is the use of adaptive downscaling methods, such as wavelet-based decomposition (Ren et al., 2017) and deep learning-based downscale method (Mayya et al., 2023). These methods can minimize the loss of key physical information while refining low-resolution input images. Even in super-resolution models, including these techniques in the preprocessing process may reduce noise and improve grid patterns. Additionally, by leveraging multi-modal data fusion, FLO-SR could integrate additional hydrological variables, such as soil moisture or land use, to enhance flood prediction accuracy. These preprocessing advancements would not only improve super-resolution fidelity but also expand FLO-SR's scalability for real-time applications. Future research can explore more advanced AI-driven preprocessing techniques and hybrid modeling approaches to further bridge the gap between physics-based and deep learning based super-resolution flood simulations.

## 5. Conclusion

In this study, a deep learning-based urban flood SR model (FLO-SR) was presented to improve the output of grid-based flood analysis models driven by physical processes. This model was evaluated for two flood events in two different geographical regions, Houston, Texas, and Portland, Oregon, with a focus on analyzing the performance and stability at different spatial resolutions. FLO-SR's performance was assessed using image pixel reconstruction, maximum flood depth prediction, and flood area similarity metrics. Based on the results, the following conclusions can be drawn:

1. Comparison of bicubic interpolation and physics-based modeling approaches: FLO-SR demonstrated superior performance when trained with bicubic interpolated low-resolution data, achieving consistent structural fidelity and computational efficiency. However, performance declined when physics-based low-resolution flood simulation inputs were used, due to spatial inconsistencies and hydrodynamic complexities inherent in physically modeled data. Specifically, when compared to bicubic interpolation, FLO-SR showed performance reductions characterized by a 7.13–27.50 % decrease in PSNR, a 0.14–16.13 % decrease in C, and a 25.0–66.67 % increase in RMSE, reflecting greater reconstruction error. Nevertheless, FLO-SR still significantly enhanced coarse resolution simulations, demonstrating its capability to effectively capture key flood patterns and improve flood depth reconstruction accuracy. These results emphasize both the computational efficiency benefits and the practical limitations of applying SR to physically modeled flood simulations, highlighting its potential for enhancing real-time flood modeling applications.
2. Performance of FLO-SR at different resolutions: FLO-SR exhibited the highest accuracy at lower scale factors (2 and 4 ×), with substantial improvements in PSNR, SSIM, RMSE, and flood extent classification metrics (H, F, and C). At higher magnifications (8 ×), performance declined due to the increased difficulty in reconstructing fine-scale flood dynamics, particularly for physics-based input data. These results suggest that FLO-SR is most effective when applied at moderate scale factors to balance resolution enhancement and accuracy.
3. Balance between computational efficiency and accuracy: Compared to conventional physics-based hydrodynamic models, FLO-SR significantly reduced computation time by up to 63 % when downscaling from 4 to 2 m resolution and 45.7 % when downscaling from 8 to 4 m resolution. This efficiency makes FLO-SR a viable tool for large-scale urban flood modeling and real-time applications. However, the model's dependence on input data quality highlights the need for further optimization, particularly when working with physics-based low-resolution inputs.

These findings demonstrate that FLO-SR effectively enhances urban flood simulations while maintaining computational efficiency. Future research should focus on improving its transferability by integrating additional hydrodynamic constraints and exploring adaptive learning techniques for varying flood conditions.

**Funding statement**

This work was supported by Korea Environment Industry & Technology Institute (KEITI) grant, funded by Korea Ministry of Environment (MOE) (RS-2023-00218973), Korea Environmental Industry & Technology Institute (KEITI) through R&D Program for Innovative Flood Protection Technologies against Climate Crisis, funded by Korea Ministry of Environment (MOE) (2022003470001), and the Technology Innovation Program (RS202400398858, Development of AI-based urban flood damage risk prediction and evaluation technology for practical use), funded by the Ministry of the Interior and Safety (MOIS, Republic of Korea).

**CRediT authorship contribution statement**

**Hyeonjin Choi:** Writing – original draft, Visualization, Formal analysis, Conceptualization. **Hyuna Woo:** Writing – review & editing, Visualization, Methodology. **Minyoung Kim:** Writing – review & editing, Visualization, Methodology. **Hyungon Ryu:** Writing – review & editing, Methodology. **Jun-Hak Lee:** Writing – review & editing. **Seungsoo Lee:** Writing – review & editing. **Seong Jin Noh:** Writing – review & editing, Supervision, Project administration, Conceptualization.

**Informed consent**

All Authors consent to the article's publication after acceptance.

**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.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2025.133529.

## Data availability

The Hurricane Harvey flood simulation datasets used in this study are available in the 'data' folder (Houston.zip) of the FLO-SR GitHub repository (https://github.com/cyber-hydrology/FLO-SR-flood-super-resolution-model/tree/main). This includes the Houston HR data at 10 m resolution and LR data at 20 m, 40 m, and 80 m resolutions. The Portland urban flood datasets, including HR data at 1 m resolution and LR data at 2 m, 4 m, and 8 m resolutions, are available upon request to the corresponding author. The source code for implementing the FLO-SR model is available in the 'code' folder of the FLO-SR GitHub repository (https://github.com/cyber-hydrology/FLO-SR-flood-super-resolution-model/tree/main).

## References


Bermúdez, M., Ntegeka, V., Wolfs, V., Willems, P., 2018. Development and comparison of two fast surrogate models for urban pluvial flood simulations. Water Resour. Manag. 32, 2801–2815. https://doi.org/10.1007/s11269-018-1959-8.

Blau, Y., Michaeli, T., 2018. The Perception-Distortion Tradeoff. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the Salt Lake City, UT, USA, pp. 6228–6237.

Chang, H., 2007. Comparative streamflow characteristics in urbanizing basins in the Portland Metropolitan Area, Oregon, USA. Hydrol. Process. 21, 211–222. https://doi.org/10.1002/hyp.6233.

Chen, C., Hui, Q., Xie, W., Wan, S., Zhou, Y., Pei, Q., 2021. Convolutional Neural Networks for forecasting flood process in Internet-of-Things enabled smart city. Comput. Netw. 186, 107744. https://doi.org/10.1016/j.comnet.2020.107744.

Cheng, J., Kuang, Q., Shen, C., Liu, J., Tan, X., Liu, W., 2020. ResLap: generating high-resolution climate prediction through image super-resolution. IEEE Access 8, 39623–39634. https://doi.org/10.1109/ACCESS.2020.2974785.

Chowdhury, M.E., Islam, A.S., Lemans, M., Hegnauer, M., Sajib, A.R., Pieu, N.M., Das, M. K., Shadia, N., Haque, A., Roy, B., Billah, M., Abdullah, F., Mamoon, W.B., Kaiser, S., Bala, S.K., Islam, G.M.T., Sarker, G.C., Rahman, S., Bhuyan, A., 2023. An efficient flash flood forecasting system for the un-gaged Meghna basin using open source platform Delft-FEWS. Environ. Model. Softw. 161, 105614. https://doi.org/10.1016/j.envsoft.2022.105614.

Contreras, M.T., Gironás, J., Escauriaza, C., 2020. Forecasting flood hazards in real time: a surrogate model for hydrometeorological events in an Andean watershed. Nat. Hazards Earth Syst. Sci. 20, 3261–3277. https://doi.org/10.5194/nhess-20-3261-2020.

Cooley, A., Chang, H., 2017. Precipitation intensity trend detection using hourly and daily observations in Portland Oregon. Climate 5, 10. https://doi.org/10.3390/cli5010010.

Cui, Z., Chen, Q., Liu, G., 2023. A two-stage downscaling hydrological modeling approach via convolutional conditional neural process and geostatistical bias correction. J. Hydrol. 620, 129498. https://doi.org/10.1016/j.jhydrol.2023.129498.

Demiray, B.Z., Sit, M., Demir, I., 2021. D-SRGAN: DEM Super-resolution with generative adversarial networks. SN Comput. Sci. 2, 48. https://doi.org/10.1007/s42979-020-00442-2.

Deshpande, R., Ragha, L., Sharma, S., 2018. Video quality assessment through PSNR estimation for different compression standards. Indones. J. Electr. Eng. Comput. Sci. 11, 918–924. https://doi.org/10.11591/ijeecs.v11.i3.pp918-924.

Franczyk, J., Chang, H., 2009. The effects of climate change and urbanization on the runoff of the Rock Creek basin in the Portland metropolitan area, Oregon, USA. Hydrol. Process. 23, 805–815. https://doi.org/10.1002/hyp.7176.

Galar, M., Sesma, R., Ayala, C., Albizua, L., Aranda, C., 2020. Super-resolution of sentinel-2 images using convolutional neural networks and real ground truth data. Remote Sens. (Basel) 12, 2941. https://doi.org/10.3390/rs12182941.

Gao, W., Liao, Y., Chen, Y., Lai, C., He, S., Wang, Z., 2024. Enhancing transparency in data-driven urban pluvial flood prediction using an explainable CNN model. J. Hydrol. 645, 132228. https://doi.org/10.1016/j.jhydrol.2024.132228.

Ghaderpour, E., Dadkhah, H., Dabiri, H., Bozzano, F., Scarascia Mugnozza, G., Mazzanti, P., 2023. Precipitation time series analysis and forecasting for italian regions. Eng. Proc. 39, 23. https://doi.org/10.3390/engproc2023039023.

Golla, S., Murukesh, M., Kumar, P., 2024. Comparative assessment of image super-resolution techniques for spatial downscaling of gridded rainfall data. SN Comput. Sci. 5, 312. https://doi.org/10.1007/s42979-024-02653-3.

Guan, M., Sillanpää, N., Koivusalo, H., 2015. Modelling and assessment of hydrological changes in a developing urban catchment. Hydrol. Process. 29, 2880–2894. https://doi.org/10.1002/hyp.10410.

Guo, Z., Leitão, 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 Manage. 14, e12684. https://doi.org/10.1111/jfr3.12684.

He, J., Zhang, L., Xiao, T., Wang, H., Luo, H., 2023. Deep learning enables super-resolution hydrodynamic flooding process modeling under spatiotemporally varying rainstorms. Water Res. 239, 120057. https://doi.org/10.1016/j.watres.2023.120057.

Henonin, J., Russo, B., Mark, O., Gourbesville, P., 2013. Real-time urban flood forecasting and modelling - a state of the art. J. Hydroinform. 15, 717–736. https://doi.org/10.2166/hydro.2013.132.

Ivanov, V.Y., Xu, D., Dwelle, M.C., Sargsyan, K., Wright, D.B., Katopodes, N., Kim, J., Tran, V.N., Warnock, A., Fatichi, S., Burlando, P., Caporali, E., Restrepo, P., Sanders, B.F., Chaney, M.M., Nunes, A.M.B., Nardi, F., Vivoni, E.R., Istanbulluoglu, E., Bisht, G., Bras, R.L., 2021. Breaking down the computational barriers to real-time urban flood forecasting. Geophys. Res. Lett. 48, e2021GL093585. https://doi.org/10.1029/2021GL093585.

Jasour, Z.Y., Reilly, A.C., Tonn, G.L., Ferreira, C.M., 2022. Roadway flooding as a bellwether for household retreat in rural, coastal regions vulnerable to sea-level rise. Clim. Risk Manag. 36, 100425. https://doi.org/10.1016/j.crm.2022.100425.

Jia, Y., Ge, Y., Chen, Y., Li, S., Heuvelink, G.B.M., Ling, F., 2019. Super-resolution land cover mapping based on the convolutional neural network. Remote Sens. (Basel) 11, 1815. https://doi.org/10.3390/rs11151815.

Jian, J., He, S., Liu, W., Liu, S., Guo, L., 2025. A refined method for the simulation of catchment rainfall–runoff based on satellite–precipitation downscaling. J. Hydrol. 653, 132795. https://doi.org/10.1016/j.jhydrol.2025.132795.

Kabir, S., Patidar, S., Xia, X., Liang, Q., Neal, J., Pender, G., 2020. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. J. Hydrol. 590, 125481. https://doi.org/10.1016/j.jhydrol.2020.125481.

Khaledyan, D., Amirany, A., Jafari, K., Moaiyeri, M.H., Khuzani, A.Z., Mashhadi, N., 2020. Low-Cost Implementation of Bilinear and Bicubic Image Interpolation for Real-Time Image Super-Resolution. In: In: 2020 IEEE Global Humanitarian Technology Conference (GHTC). Presented at the 2020 IEEE Global Humanitarian Technology Conference (GHTC), pp. 1–5. https://doi.org/10.1109/GHTC46280.2020.9342625.

Li, J., Li, L., Zhang, T., Xing, H., Shi, Y., Li, Z., Wang, C., Liu, J., 2024. Flood forecasting based on radar precipitation nowcasting using U-net and its improved models. J. Hydrol. 632, 130871. https://doi.org/10.1016/j.jhydrol.2024.130871.

Li, Z., Leong, W.J., Durand, M., Howat, I., Wadkowski, K., Yadav, B., Moortgat, J., 2023. Super-resolution deep neural networks for water classification from free multispectral satellite imagery. J. Hydrol. 626, 130248. https://doi.org/10.1016/j.jhydrol.2023.130248.

Lim, B., Son, S., Kim, H., Nah, S., Lee, K.M., 2017. Enhanced Deep Residual Networks for Single Image Super-Resolution. In: In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Presented at the Honolulu, pp. 1132–1140. https://doi.org/10.1109/CVPRW.2017.151.

Liu, Q., Du, M., Wang, Y., Deng, J., Yan, W., Qin, C., Liu, M., Liu, J., 2024. Global, regional and national trends and impacts of natural floods, 1990-2022. Bull. World Health Organ. 102, 410–420. https://doi.org/10.2471/BLT.23.290243.

Liu, Y., Bates, P.D., Neal, J.C., Yamazaki, D., 2021. Bare-earth DEM generation in urban areas for flood inundation simulation using global digital elevation models. Water Resour. Res. 57, e2020WR028516. https://doi.org/10.1029/2020WR028516.

Lombana, L., Martínez-Graña, A., 2022. A flood mapping method for land use management in small-size water bodies: validation of spectral indexes and a machine learning technique. Agronomy 12, 1280. https://doi.org/10.3390/agronomy12061280.

Mayya, V., Sowmya, S.K., Kulkarni, U., Surya, D.K., Acharya, U.R., 2023. An empirical study of preprocessing techniques with convolutional neural networks for accurate detection of chronic ocular diseases using fundus images. Appl. Intell. 53, 1548–1566. https://doi.org/10.1007/s10489-022-03490-8.

Mohamadiazar, N., Ebrahimian, A., Hosseiny, H., 2024. Integrating deep learning, satellite image processing, and spatial-temporal analysis for urban flood prediction. J. Hydrol. 639, 131508. https://doi.org/10.1016/j.jhydrol.2024.131508.

Myers, D.T., Ficklin, D.L., Robeson, S.M., 2023. Hydrologic implications of projected changes in rain-on-snow melt for Great Lakes Basin watersheds. Hydrol. Earth Syst. Sci. 27, 1755–1770. https://doi.org/10.5194/hess-27-1755-2023.

Noh, S.J., Lee, J.-H., Lee, S., Kawaike, K., Seo, D.-J., 2018. Hyper-resolution 1D-2D urban flood modelling using LiDAR data and hybrid parallelization. Environ. Model. Softw. 103, 131–145. https://doi.org/10.1016/j.envsoft.2018.02.008.

Noh, S.J., Lee, J.-H., Lee, S., Seo, D.-J., 2019. Retrospective dynamic inundation mapping of hurricane harvey flooding in the houston metropolitan area using high-resolution modeling and high-performance computing. Water 11, 597. https://doi.org/10.3390/w11030597.

Piadeh, F., Behzadian, K., Alani, A.M., 2022. A critical review of real-time modelling of flood forecasting in urban drainage systems. J. Hydrol. 607, 127476. https://doi.org/10.1016/j.jhydrol.2022.127476.

Poehls, J., Alonso, L., Koirala, S., Reichstein, M., Carvalhais, N., 2025. Downscaling soil moisture to sub-km resolutions with simple machine learning ensembles. J. Hydrol. 652, 132624. https://doi.org/10.1016/j.jhydrol.2024.132624.

Qi, X., de Almeida, G.A.M., Maldonado, S., 2024. Physics-informed neural networks for solving flow problems modeled by the 2D Shallow Water Equations without labeled data. J. Hydrol. 636, 131263. https://doi.org/10.1016/j.jhydrol.2024.131263.

Ren, H., Pang, B., Zhao, G., Yu, H., Tian, P., Xie, C., 2025. Incorporating dynamic drainage supervision into deep learning for accurate real-time flood simulation in urban areas. Water Res. 270, 122816. https://doi.org/10.1016/j.watres.2024.122816.

Ren, R., Gu, L., Fu, H., Sun, C., 2017. Super-resolution algorithm based on sparse representation and wavelet preprocessing for remote sensing imagery. JARS 11, 026014. https://doi.org/10.1117/1.JRS.11.026014.

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. 656, 132697. https://doi.org/10.1016/j.jhydrol.2025.132697.

Sara, U., Akter, M., Uddin, M.S., 2019. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. J. Comput. Comm. 7, 8–18. https://doi.org/10.4236/jcc.2019.73002.

Wang, X., Yi, J., Guo, J., Song, Y., Lyu, J., Xu, J., Yan, W., Zhao, J., Cai, Q., Min, H., 2022. A review of image super-resolution approaches based on deep learning and applications in remote sensing. Remote Sens. (Basel) 14, 5423. https://doi.org/10.3390/rs14215423.

Wang, Z., Chen, J., Hoi, S.C.H., 2021. Deep learning for image super-resolution: a survey. IEEE Trans. Pattern Anal. Mach. Intell. 43, 3365–3387. https://doi.org/10.1109/TPAMI.2020.2982166.

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

Xu, M., Yao, N., Yang, H., Xu, J., Hu, A., Goncalves, G., de Goncalves, L., Liu, G., 2022. Downscaling SMAP soil moisture using a wide & deep learning method over the Continental United States. J. Hydrol. 609, 127784. https://doi.org/10.1016/j.jhydrol.2022.127784.

Yadav, R., Nascetti, A., Ban, Y., 2022. Attentive Dual Stream Siamese U-Net for Flood Detection on Multi-Temporal Sentinel-1 Data, in. In: IGARSS 2022–2022 IEEE International Geoscience and Remote Sensing Symposium. Presented at the Kuala Lumpur, pp. 5222–5225. https://doi.org/10.1109/IGARSS46834.2022.9883132.

Yang, F., Ding, W., Zhao, J., Song, L., Yang, D., Li, X., 2024. Rapid urban flood inundation forecasting using a physics-informed deep learning approach. J. Hydrol. 643, 131998. https://doi.org/10.1016/j.jhydrol.2024.131998.

Yin, Z., Saadati, Y., Hu, B., Leon, A.S., Amini, M.H., McDaniel, D., 2024. Fast high-fidelity flood inundation map generation by super-resolution techniques. J. Hydroinform. https://doi.org/10.2166/hydro.2024.228 jh2024228.

Zaghloul, M.S., Ghaderpour, E., Dastour, H., Farjad, B., Gupta, A., Eum, H., Achari, G., Hassan, Q.K., 2022. Long term trend analysis of river flow and climate in Northern Canada. Hydrology 9, 197. https://doi.org/10.3390/hydrology9110197.

Zandsalimi, Z., Barbosa, S.A., Alemazkoor, N., Goodall, J.L., Shafiee-Jood, M., 2025. Deep learning-based downscaling of global digital elevation models for enhanced urban flood modeling. J. Hydrol. 653, 132687. https://doi.org/10.1016/j.jhydrol.2025.132687.

Zeng, Y.-F., Chang, M.-J., Lin, G.-F., 2024. A novel AI-based model for real-time flooding image recognition using super-resolution generative adversarial network. J. Hydrol. 638, 131475. https://doi.org/10.1016/j.jhydrol.2024.131475.

Zhao, J., Li, Y., Matgen, P., Pelich, R., Hostache, R., Wagner, W., Chini, M., 2022. Urban-aware U-net for large-scale urban flood mapping using multitemporal sentinel-1 intensity and interferometric coherence. IEEE Trans. Geosci. Remote Sens. 60, 1–21. https://doi.org/10.1109/TGRS.2022.3199036.