Journal of Hydrology 675 (2026) 135601

Contents lists available at ScienceDirect

Journal of Hydrology

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

Convolutional neural network model for rapid prediction of urban flash flood water depth and velocity maps

Muhammad Asif [a], Roberto Bentivoglio [b], Raffaele Albano [c,*]

[a] Department of Engineering, University of Basilicata, Potenza, Italy
[b] Department of Water Management, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, the Netherlands
[c] Department of Health Science, University of Basilicata, Potenza, Italy

ARTICLE INFO

This manuscript was handled by Dan Lu, Editor-in-Chief, with the assistance of Ganesh R. Ghimire, Associate Editor

Keywords:
Convolutional neural network
Urban flash flood
Water depths maps
Flow velocities prediction
Deep learning

ABSTRACT

Urbanization and climate change have intensified the need for rapid and accurate predictions of flash floods, especially in urban areas. Although numerical models produce accurate predictions, their computational cost makes them impractical for real-time simulations. Several machine learning models have been proposed as surrogates for such applications. However, most models focus only on predicting water depths and have little insight into the required variety and amount of data to train such models.

In this study, we present a Convolutional Neural Network (CNN) surrogate model designed to predict both water depths and flow velocities for urban pluvial flooding, using as inputs a rainfall hyetograph and seven hydro-morphological descriptors, such as aspect, curvature, slope, Manning's roughness coefficient, topographic wetness index, flow accumulation, and digital terrain model.

To train and test our approach, we considered a dataset of numerical simulations carried out using a 2D shallow-water hydrodynamic modeling in the city of Matera, Italy. For the physical-based simulations, we considered several rainfall hyetographs, either obtained from real events or derived from design scenarios for different return periods, and simulated the associated pluvial floods propagation extracting as outputs the corresponding maps of the maximum envelope of the water depths and flow velocities. The CNN model obtained a testing 0.37 cm mean absolute error (MAE), 1.7 cm root mean squared error (RMSE), and 0.80 critical success index (CSI) for water depth predictions, and 0.054 m/s MAE, 0.178 m/s RMSE, and 0.84 CSI for flow velocity predictions. The CNN model was 116.5 times faster than the physically-based hydrodynamic model using the same computational hardware.

We also analyzed the effect of different combinations of rainfall events to train and validate the CNN model, showing that it benefits from a balanced dataset in terms of different return periods and presence of both synthetic and real hyetographs. We then employed the model to extrapolate urban flooding for higher return periods (200, 300, 500, and 1000 years), showing that the model can predict well severe extreme events, as highlighted by a high correlation between predicted maximum water volumes and total rainfall.

This study contributes to the practical usability of deep learning models by providing essential support for real-time or low lead-time urban flash flood predictions, which are crucial for effective early warning and emergency management.

1. Introduction

Due to increasingly frequent extreme rainfall events and urban land-use changes, urban flash floods have become one of the most dangerous natural disaster worldwide causing significant causalities and economic losses (Rentschler et al., 2022). Therefore, mitigation and reduction of flood impacts has gained the attention and become worldwide concern (Tanoue et al., 2021). In urban areas, the constant increase in impervious surfaces accelerates the runoff yield and concentration time, leading faster and more intense flash floods (Lou et al., 2022; Yan et al., 2022). These are characterized by a rapid rise in water levels and high flow velocities. This calls for methods that can forecast water depth and flow velocity distributions with short lead time, as they are essential for evaluating flood risk levels and for supporting early warning systems

* Corresponding author.
E-mail addresses: muhammadasif@unibas.it (M. Asif), r.bentivoglio@tudelft.nl (R. Bentivoglio), raffaele.albano@unibas.it (R. Albano).

https://doi.org/10.1016/j.jhydrol.2026.135601
Received 2 February 2026; Received in revised form 13 April 2026; Accepted 30 April 2026
Available online 1 May 2026

Table 1
Comprehensive quantitative comparison of existing CNN/U-Net flood studies.

Model	Inputs	Outputs	No. of rainfall events	Flow velocity considered	Study
U-Net	Rainfall hyetographs + Terrain features	Water depth	18 (Synthetic)	No	Guo et al., (2021)
U-Net	Rainfall hyetographs + Spatial inputs	Water depth	53 (Real)	No	Löwe et al., (2021)
3DCNN, ConvLSTM	Rainfall time series (hyetographs-derived)	Water depth	70 (Synthetic)	No	Liao et al., (2023)
U-Net	Rainfall hyetographs + Topographic data	Water depth	12 (Synthetic)	No	El Baida et al., (2024)
2DCNN	Rainfall hyetographs + Spatial inputs	Water depth	49 (Synthetic)	No	(Gao et al., 2024)
U-Net	Rainfall hyetographs + Terrain features	Water depth	20 (Real)	No	(Farfán-Durán et al., 2025)

(Hosseiny et al., 2020; Piadeh et al., 2022).

Physics-based hydrodynamic models are the state-of-the-art for the assessment of urban flash flood propagation (Chen et al., 2017; Sole et al., 2013; Wu et al., 2017). They are available in a variety of commercial software packages (Bates et al., 2013; DHI, 2017; Hicks and Peacock, 2005) and are based on the mathematical representation of the 2D shallow water equations or one of their simplified forms (Sole et al., 2013). However, they require long computational times, especially when simulating floods for large areas with high spatial resolution (Costabile and Macchione, 2015; Donnelly et al., 2022; Albano et al., 2024). This limits the application of such models for real-time prediction of urban flash flood. To cope with this issue, several alternative approaches were developed: 1) models based on simplified hydraulic concepts (Teng et al., 2022; Albano and Adamowski, 2025; Samela et al., 2018; 2) model that rely on precomputed offline scenarios maps (Leedal et al., 2010); 3) models using GPU acceleration and parallel computation (Dazzi et al., 2021; Shaw et al., 2021); 4) machine learning (ML) based models (Asif et al., 2025b; Bentivoglio et al., 2022; Yang et al., 2022). Among these approaches, ML techniques have received a significant attention due to high computational efficiency and great accuracy. ML models are generally used as surrogates of the physics-based simulations, learning the output regardless of actual physical process (Guo et al., 2021). Many traditional ML models such as random forest (RF) (Hou et al., 2021; Kabir et al., 2021; Liao et al., 2023), gradient boosting decision tree (GBDT) (Wu et al., 2020b), and support vector machine (SVM) (Bermúdez et al., 2019; Li et al., 2016) were used for rapid simulation of urban flash flood inundation depth. In addition, some ensembles model were applied to reduce the uncertainty in predictions (Hou et al., 2021; Wu et al., 2020a). Artificial neural networks (ANNs) have been used more frequently in flood prediction and susceptibility mapping due to non-linear correlation and effectively processed the time series data by stacking the multiple layers. ANNs were utilized to estimate water depth of urban flood at specific coordinates using topographical inputs such as aspect, slope, and curvatures (Bui et al., 2020; Kim and Han, 2020). However, ANNs face the "curse of dimensionality", where the number of required parameters grows exponentially with the inputs' size (LeCun et al., 2015). This is especially relevant in floods, where most inputs have spatially distributed information (Guo et al., 2021).

Convolutional neural networks (CNNs) are a deep learning (DL) method which can leverage spatial biases by feature extraction of patterns. CNNs have outperformed traditional ML models such as SVM, RF, and ANNs in several flood mapping tasks (Isikdogan et al., 2017; Bentivoglio et al., 2022). They can effectively reduce the network's complexity and number of training parameters, thereby improving the efficiency and accuracy in urban flash flood inundation prediction (Liao et al., 2023; Guo et al., 2021; Löwe et al., 2021). Therefore, CNNs have gained increasing attention and are widely adopted for flood studies. Table 1 provides a quantitative comparison of existing studies that employ CNN-based models, summarizing their model inputs, outputs, number of rainfall events, and whether flow velocity is predicted. Guo et al., (2021) proposed a CNN model for the prediction of maximum water depth of urban flood for three catchments using 18 synthetic hyetographs of different return periods. Löwe et al., (2021) also utilized a CNN based U-Net model for predicting the maximum water depth of urban pluvial flood employing 53 observed rain events and generalizing predictions over different parts of an urban catchment. Liao et al., (2023) employed a CNN for the prediction of water depth of urban flood by using 70 design rainstorms based on different return periods, showing that the model reproduced well the trend of water depth in each grid cell and outperformed three classic ML models. El Baida et al., (2024) developed a CNN based U-Net architecture for predicting the pluvial and fluvial floods using 12 rainfall hyetographs of 2-hour duration event, which exhibited high accuracy and reduced computation time. The study by Gao et al., (2024) explored an explainable deep learning approach for urban pluvial flood prediction employing CNN and the explainable artificial intelligence (AI) framework Shapley additive explanation (SHAP). The approach offered a great potential for rapid early warning systems and enhanced the model transparency. Farfán-Durán et al., (2025) proposed a novel deep learning (DL)-based U-Net architecture as a surrogate model that incorporates the net rainfall using 20 events to predict water depth. The outcomes showed that the proposed surrogate model outperformed the benchmark and reduced the computational times. These works focused only on maximum water depths without considering flow velocities. These can be critical as they can make even shallow water depths hazardous, especially in steep slope urban areas (Asif et al., 2025a).

Furthermore, in most existing works, models are solely trained on synthetic or observed rainfall events rather than a combination of both. This limitation of diversity of training data can restrict the model's generalization to unseen or complex events. There is also no analysis on what a suitable set of events to train a model is. This can lead to requiring a larger number of simulations compared to the ones actually needed.

To address these limitations, this study proposes a CNN-based U-Net architecture and introduces three main innovations. First, it enables the simultaneous prediction of water depth and velocity of urban flash flood, providing a comprehensive representation of flood dynamics in urban environments. Second, it evaluates the optimal dataset composition for training this model, by comparing its performance across different combinations of rainfall events, to understand their influence on model training and testing. Third, we also involve a diverse dataset consisting of observed past events and synthetic extreme rainfall scenarios, which improve the robustness and generalizability of the model. We selected as a case study the city of Matera, Italy, since it has a highly urbanized landscape and steep topography, which have led to severe urban flash floods in the past years (Asif et al., 2025a). After comparing the model's performance in terms of mean absolute error (MAE), root mean squared error (RMSE), critical success index (CSI), precision, and recall, we also assessed the model's extrapolation behavior under extreme hydrological conditions. The study provides a novel approach to urban flash flood prediction which can shorten computational time while maintaining high accuracy, needed for the development of rapid early warning systems.

2. Case study

We selected the downtown of the Matera city (Basilicata region, Italy) as the case study. The city has approximately 60′000 inhabitants and is located in a hilly area in the Bradano river basin, between two

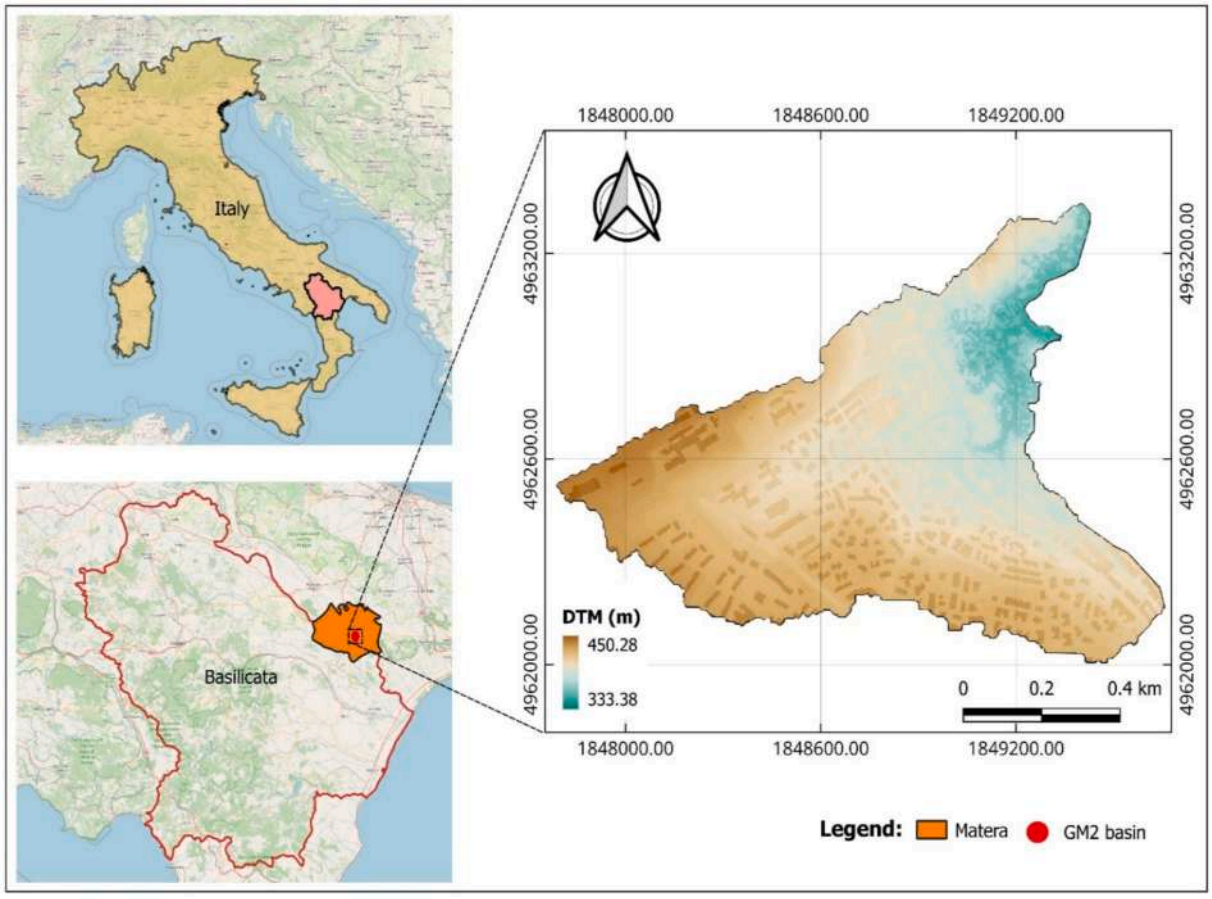

Fig. 1. Geographic location of the city of Matera (Basilicata region, Italy). The figure on the right shows the digital terrain model (DTM) of the GM2 basin, selected as case study.

tributaries, i.e. Gravina di Matera (GM) and Gravina di Picciano (GP) (Perrone et al., 2020). It covers an area of about 390 km^2, with a mean altitude of 400 m above sea level and the yearly average precipitation is around 543 mm/year. Over time, the Matera area has undergone significant transformations (Ermini and Albano, 2023), with population increases and urban development that have altered the hydrological and hydraulic response of the area, leading to more frequent flooding events. The city of Matera is divided into 21 urban basins which extend from the city center to the marginal edges of the transformed areas (Asif at al., 2025a). This study focused on the, so called, GM2 basin (Fig. 1).

The GM2 basin encompasses a very important area of the city that includes Sasso Caveoso, Castle Hill, and the several touristic and commercial activities (Ermini et al., 2024). This urban basin extends over an area of approximately 0.66 km^2, with altitudes ranging from a maximum of 441.52 m above sea level to a minimum of 267.67 m (Mishra and Albano, 2026). The average slope of the basin is 21.7% and its longitudinal extension is about 1.5 km. Because of the highly sloped areas, the city experienced several flash floods, whose severity was further increased by the drastic reduction in permeability and shorter concentration time of the basin (Haastregt, 2021). According to Matera Municipality and Civil Protection, Matera has faced severe urban flash floods in last few years (Coletta et al., 2021). These events caused extensive damage, with people becoming trapped in their cars on flooded roads, losing their balance, and scooters being carried away by the water (Albano et al., 2026).

3. Methodology

We developed a CNN based U-net surrogate model to predict water depths and flow velocities given an extreme rainfall event. We created a dataset of numerical simulations using, as initial conditions, a set of rainfall scenarios collected from past observed events and from synthetic hyetographs characterized by different return periods; the urban flash floods simulations are performed using the physics-based hydrodynamic model LISFLOOD-FP (Bates et al., 2010).

Each simulation corresponds to an input–output pair in which the inputs are given by a set of seven hydro-morphological descriptors and a rainfall time-series, and the outputs are the maps of the maximum envelope of the water depth and flow velocity at each computational cell.

After splitting the dataset into training, validation, and testing datasets, we evaluated the model's performance using both regression and classification metrics, as well as evaluating the computational costs. Then, we evaluated the optimal required dataset composition for training such DL model. Finally, we assessed the model's capability to extrapolate under extreme hydrological conditions.

Fig. 2 summarizes the process to create the datasets, train them, and evaluate them.

3.1. Synthetic and observed hyetographs

Urban flash floods are produced by short term extreme rainfall that produces flooding within few hours or even less (Chen et al., 2018; Wu et al., 2017; Zhang et al., 2022). According to Ravuri et al., 2021, a 2-hour rainfall forecasting is more reliable than higher forecasting lead times. Therefore, we set the duration of each rainfall event to 2 h, with a temporal resolution of 15 min, similarly to other studies (e.g., Löwe et al., 2021). For each event, the net rainfall was computed using the Curve number (CN) method based on land use to estimate the portion of

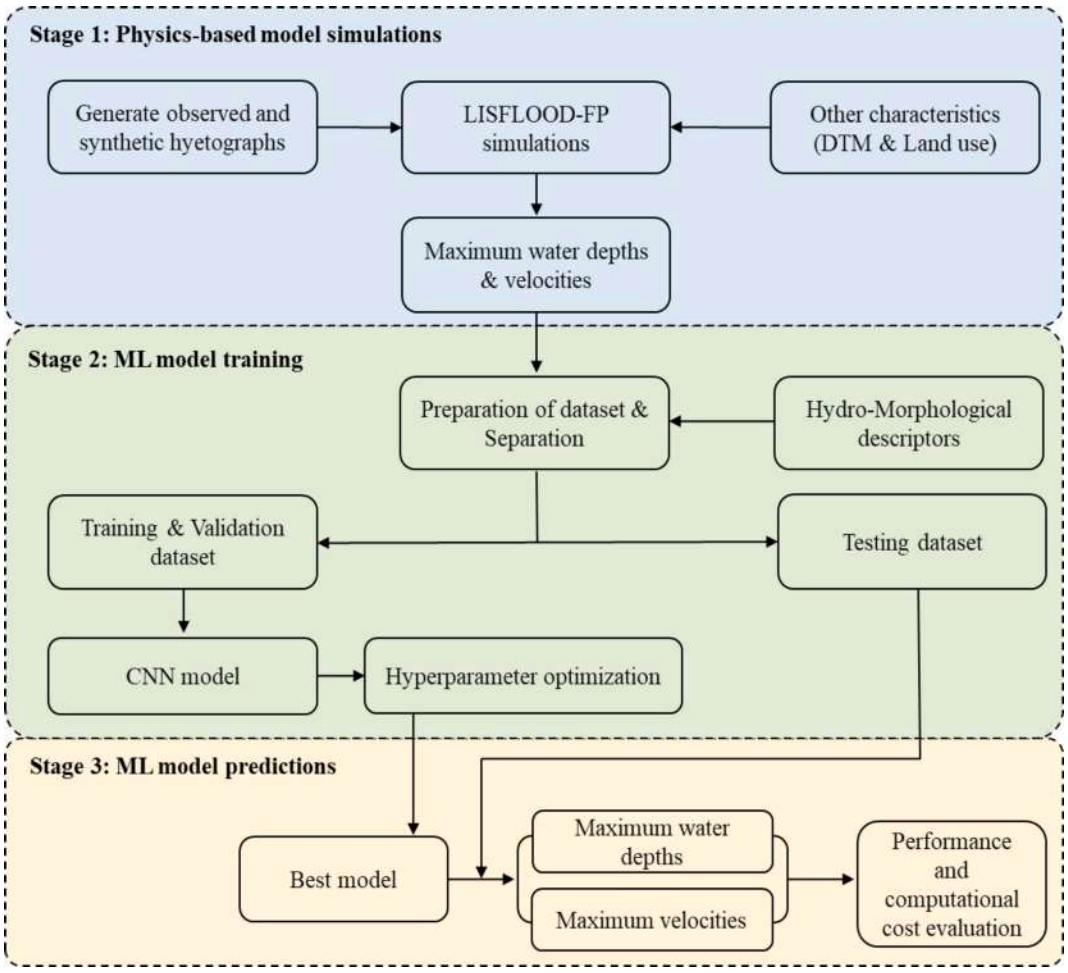

Fig. 2. Flowchart of the urban flash flood prediction using CNN model.

water that does not infiltrate the ground and generates surface runoff (Mishra and Singh, 2003). A total of 20 hyetographs were utilized in this study, comprising 9 observed historical hyetographs and 11 synthetic hyetograph representing different return periods (2, 3, 5, 10, 15, 20, 25, 30, 40, 50, and 100 years). Of these, 14 hyetographs were employed for model training and validation, using an 80/20 split, and the remaining 6 for model testing (Fig. 3).

The synthetic hyetographs were derived from an Intensity-Duration-Frequency (IDF) curve available in the report on the VAPI ("VAlutazione delle Piene in Italia") methodology adapted for Basilicata region (Albano et al., 2024), adopting as time distribution of the precipitation a Chicago hyetograph (Keifer and Chu, 1957). For the observed hyetographs, this study included 9 real past extreme rainfall events that lead to severe flooding in Matera according to Matera Municipality and Civil Protection. These past events occurred on 09 July 2013, 24 August 2018, 12 November 2019, 02 June 2023, 29 February 2024, 28 May 2024, 24 June 2024, 02 July 2024, and 21 July 2024; the dataset are available on the "Centro Funzionale Decentrato della Basilicata" website (https://centrofunzionale.regione.basilicata.it/it/scaricaDati.php).

3.2. Physics-based (LISFLOOF-FP) simulations

We employed the physics-based LISFLOOD-FP model to generate the hydraulic simulations (Bates et al., 2010; De Almeida et al., 2012). For this study, we used version 8.1.0 that implements a Roe finite-volume approximator to solve the full 2D shallow water equations. We used DTM and land use data as inputs, along with the net mean rainfall homogeneously spatially distributed for the whole area. The simulations were carried out for each of the 20 hyetographs described in Section 3.1, generating the maximum envelope of the water depth and flow velocity flood maps as output.

3.3. Deep learning (DL) model development

3.3.1. CNN model

We employed a CNN-based U-Net model consisting of a convolutional encoder, a decoder, and a fully connected network in the bottleneck to process the rainfall input similarly to Guo et al., 2021, as shown in Fig. 4. The encoder network consists of three convolutional blocks with channel sizes 32, 64, and 128, respectively. Each block contains two convolutional layers with padding 1, followed by Leaky-ReLU activation functions and one pooling layer. No Batch Normalization/ Dropout layers are employed. The convolutional layers are used to extract the spatial features of inputs, while the pooling layers reduce the dimensionality of inputs to improve computational efficiency. Similarly, the decoder network has three deconvolutional blocks with channel sizes 132, 64, and 32. Each block involves one deconvolution layer with kernel size, stride 2, and followed by two convolutional layers. The deconvolution layers are used for upsampling latent features into a larger shape. The rainfall hyetograph input is processed by a fully connected network composed of two dense layers. Its output is reshaped into 64x89x4 to match the dimension of the encoder's output, allowing for concatenation at the bottleneck of the model. At the end, a final convolutional layer is used to covert 32 channels into the desired 2 outputs. The model contains total 94364770 trainable parameters. The kernel sizes for all convolutional layers are 3x3 and 2x2 for all

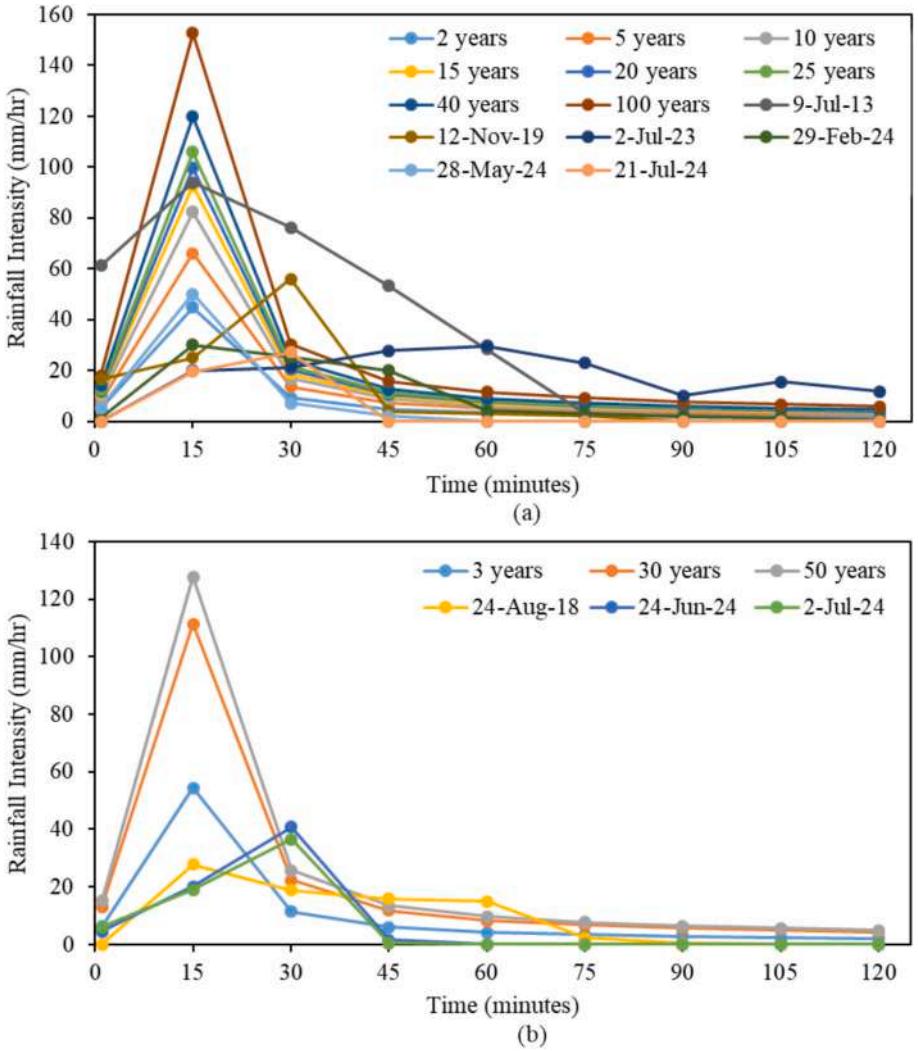

Fig. 3. Hyetographs used for model (a) training & validation and (b) testing.

deconvolution and pooling layers (Guo et al., 2021; Löwe et al., 2021). We utilized a small kernel size to conserve the thin structure of hydro-morphological descriptors and multiple layers to extend the receptive field (Luo et al., 2016). Leaky-ReLU activation functions were applied for all layers with an activation threshold 0.01 (Maas et al., 2013) to ensure the stable gradient flow (Hochreiter, 1998) and avoid the dead neuron problem typical of rectified linear units (Nair and Hinton, 2010). Additionally, we used the maximum pooling in the encoder network as applied in other works (Guo et al., 2021; Ronneberger et al., 2015). When testing, we replaced the final Leaky-ReLU layer with a ReLU layer to enforce a physical positivity of the outputs. The minor difference between training and inference model performance may arise due to use of different activation functions.

Compared to the network proposed by Guo et al., 2021, we have made three changes. First, we applied the skip connections to concatenate the output from each encoding block to the input of the corresponding decoding block, to preserve the fine-grained spatial information that may be lost in deeper parts of the network. Second, we modified the network to predict water velocities too and, third, we added the final ReLU layer for added physical bias.

We trained the model using a weighted mean absolute error loss function $\mathscr{L}$ (Eq. (1), which reads as:

$$\mathscr{L} = \frac{1}{N}\sum_{i}^{N}(\alpha|h_i - \widehat{h}_i| + |v_i - \widehat{v}_i|), \tag{1}$$

where h_i and $\widehat{h}_i$ are the true values and the predicted water depths, v_i and $\widehat{v}_i$ are the true values and the predicted flow velocities, α is a hyper-parameter which balances the different contribution of water depth and flow velocities, and N is the total number of cells across all simulations. Accordingly, α was varied to give higher weights to water depth, emphasizing its importance in the prediction. To select the optimal value of α, we conducted a hyperparameter analysis, as reported in Section S3, which resulted in a value of $\alpha = 7.5$.

3.3.2. *Hydro-Morphological descriptors*

Deep learning models benefit from a wide range of inputs that capture the hydro-morphological characteristics of the basin more exhaustively than raw elevation data only (Löwe et al., 2021). Based on previous studies, we considered seven hydro-morphological descriptors potentially relevant for predicting urban flash floods (Guo et al., 2021; Kabir et al., 2020; Löwe et al., 2021). These were aspect, curvature, slope, Manning's roughness coefficient, topographic wetness index (TWI), flow accumulation, and digital terrain model (DTM).

The slope is defined as the terrain steepness and represents the maximum rate of change in elevation from one cell to its neighboring cells; aspect represents the slope direction at each cell; (plan) curvature identifies the convergence and divergence of the flow; flow accumulation corresponds to the number of upstream cells of a given cell; the TWI quantifies the tendency of an area to accumulate the water based on

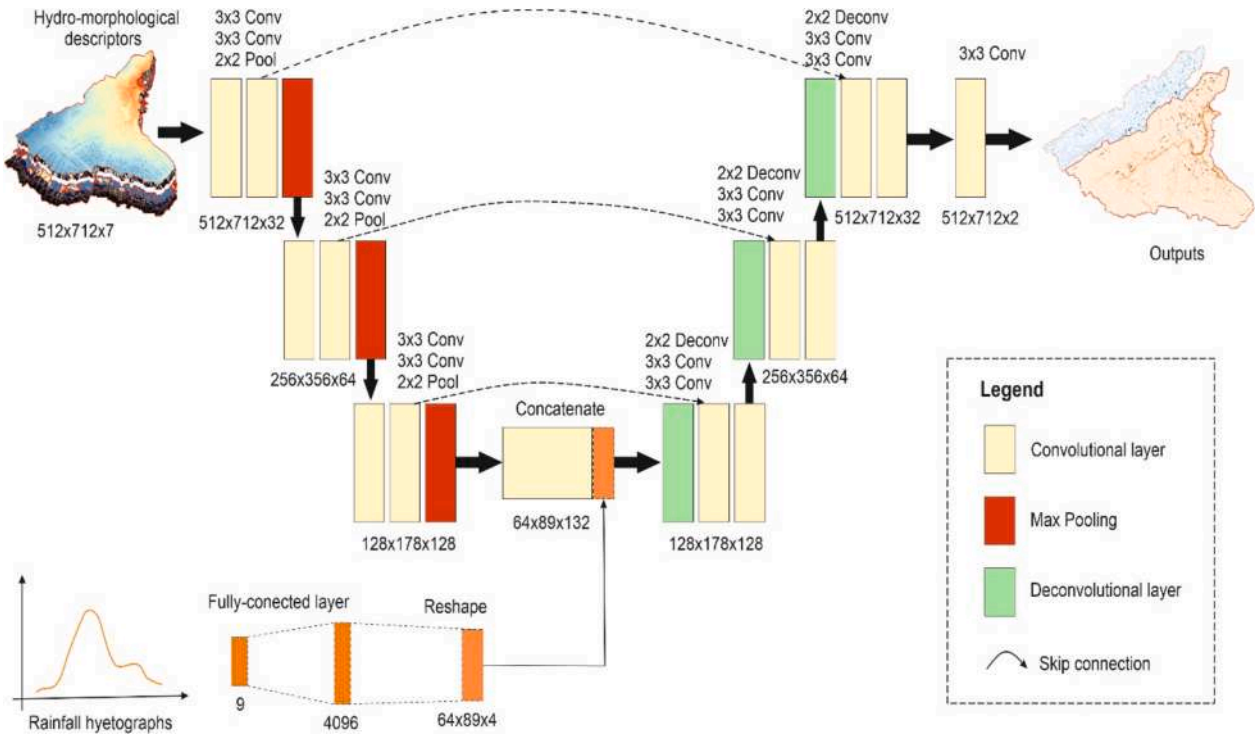

Fig. 4. Schematic of the CNN model utilized in this study. It consists of an encoder, a decoder, and a fully-connected layer that processes rainfall, similarly to Guo et al., (2021).

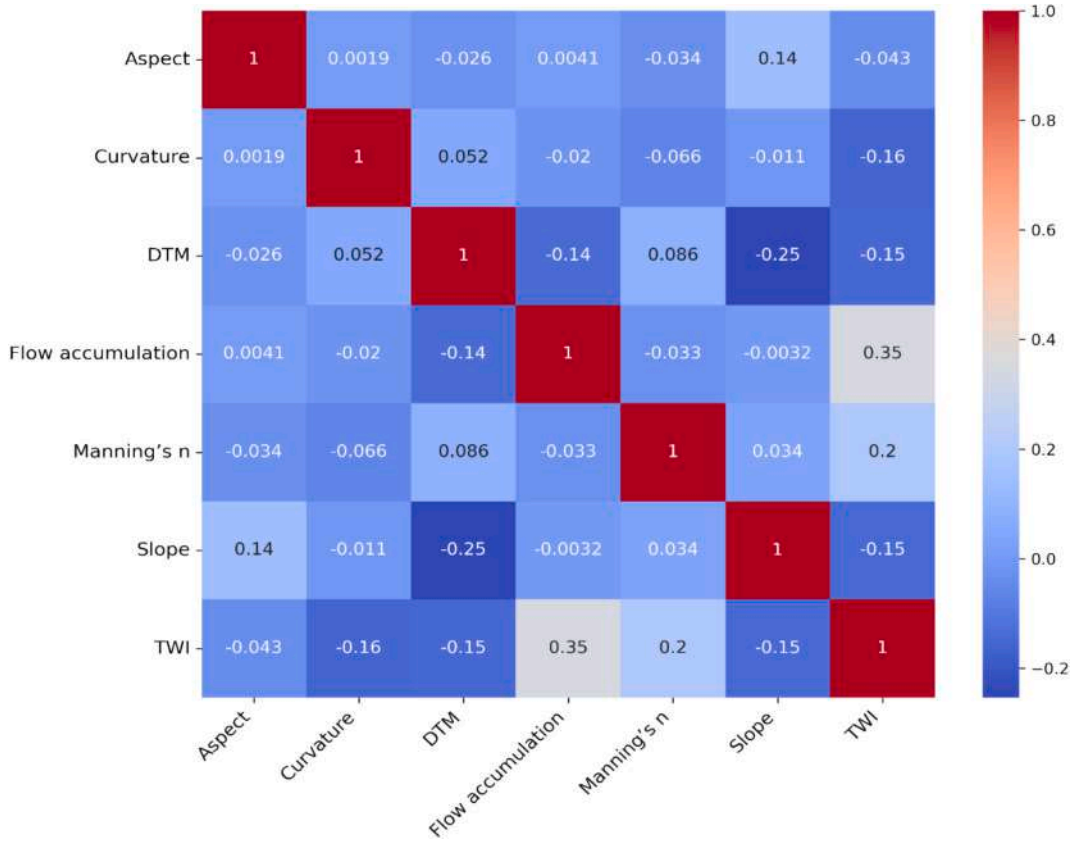

Fig. 5. Pixel-wise Pearson correlation matrix for collinearity analysis among hydro-morphological descriptors.

slope and contributing area. The Manning's roughness coefficient controls the overland flow resistance in flood propagation and was determined based on land use data (Thrysøe et al., 2021). The DTM includes buildings extruded to a 5 m fixed height.

Table 2
Mean and standard deviation of the different evaluation metrics for training, validation, and testing, for both water depths and flow velocities.

Dataset	Event	Water depth			Velocity		
		MAE (cm)	RMSE (cm)	CSI	MAE (m/s)	RMSE (m/s)	CSI
Training	All	0.36 ± 0.13	1.47 ± 0.54	0.85 ± 0.034	0.059 ± 0.012	0.195 ± 0.039	0.86 ± 0.031
Validation	All	0.34 ± 0.06	1.51 ± 0.39	0.74 ± 0.143	0.052 ± 0.003	0.166 ± 0.020	0.80 ± 0.071
Test	3 years	0.33	1.63	0.78	0.049	0.163	0.83
	30 years	0.36	1.62	0.89	0.062	0.202	0.89
	50 years	0.50	2.35	0.87	0.071	0.237	0.89
	24 Aug 2018	0.35	1.49	0.75	0.049	0.165	0.80
	24 June 2024	0.34	1.72	0.76	0.047	0.154	0.82
	02 July 2024	0.31	1.40	0.76	0.044	0.145	0.82
	All	0.37 ± 0.68	1.70 ± 0.34	0.80 ± 0.06	0.054 ± 0.011	0.178 ± 0.035	0.84 ± 0.039

Since these descriptors were obtained with different measurement units and scale ranges, we applied data normalization to ensure the consistency and avoid bias toward a single input. All inputs were rescaled to the intervals [-1,1] if negatives values were present, and [0,1] otherwise. A visualization representation of all the raw hydro-morphological descriptors is provided in the Supporting Material (see S1).

3.3.3. *Feature selection*

We examined the correlations and collinearities between the hydro-morphological descriptors to ensure no bias towards a particular input (Löwe et al., 2021). Based on the selection of relevant inputs, we utilized the Pearson correlation coefficient (PCC), which measures the linear correlation between input variables (Nnanwuba et al., 2022). In this study, PCC is used to quantify the degree of correlation among hydro-morphological descriptors across all valid pixels, helping to identify multi-collinearity or redundancy among them. The resulting correlation matrix highlights the degree of association between each pair of input layers, where – 1 indicates strong inverse correlation and + 1 indicates a strong positive correlation of variables. We assumed to have a high correlation for absolute values above 0.7 (Hong et al., 2020). In addition, we employed mutual information (MI) analysis to compute potential non-linear dependencies among hydro-morphological descriptors. The high value of MI exhibits strong dependency between descriptors, while the lower values close to zero indicate independency.

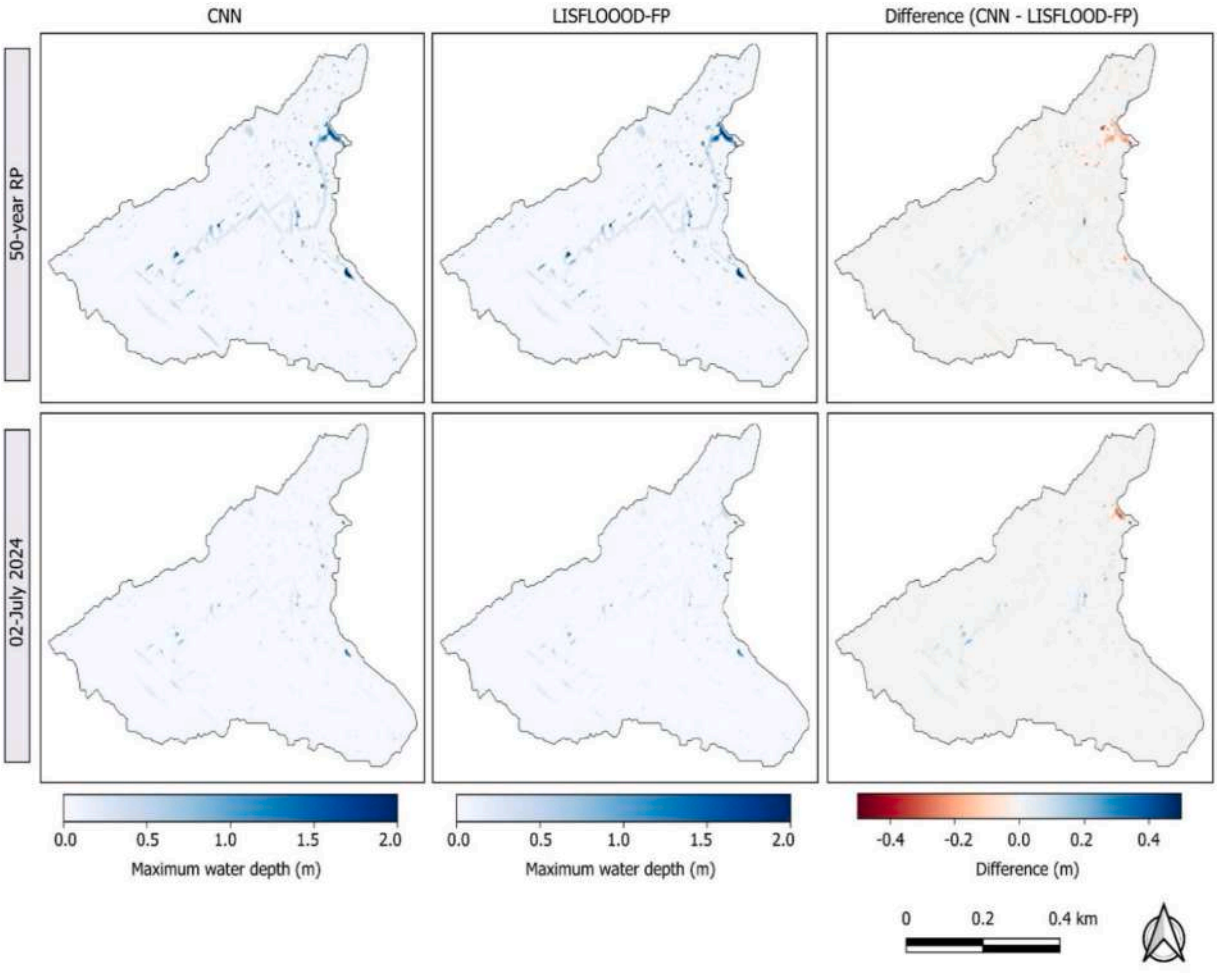

Fig. 6. Water depth maps predicted by CNN model and by the physics-based model for 50 years return period event and the real event of 02 July 2024.

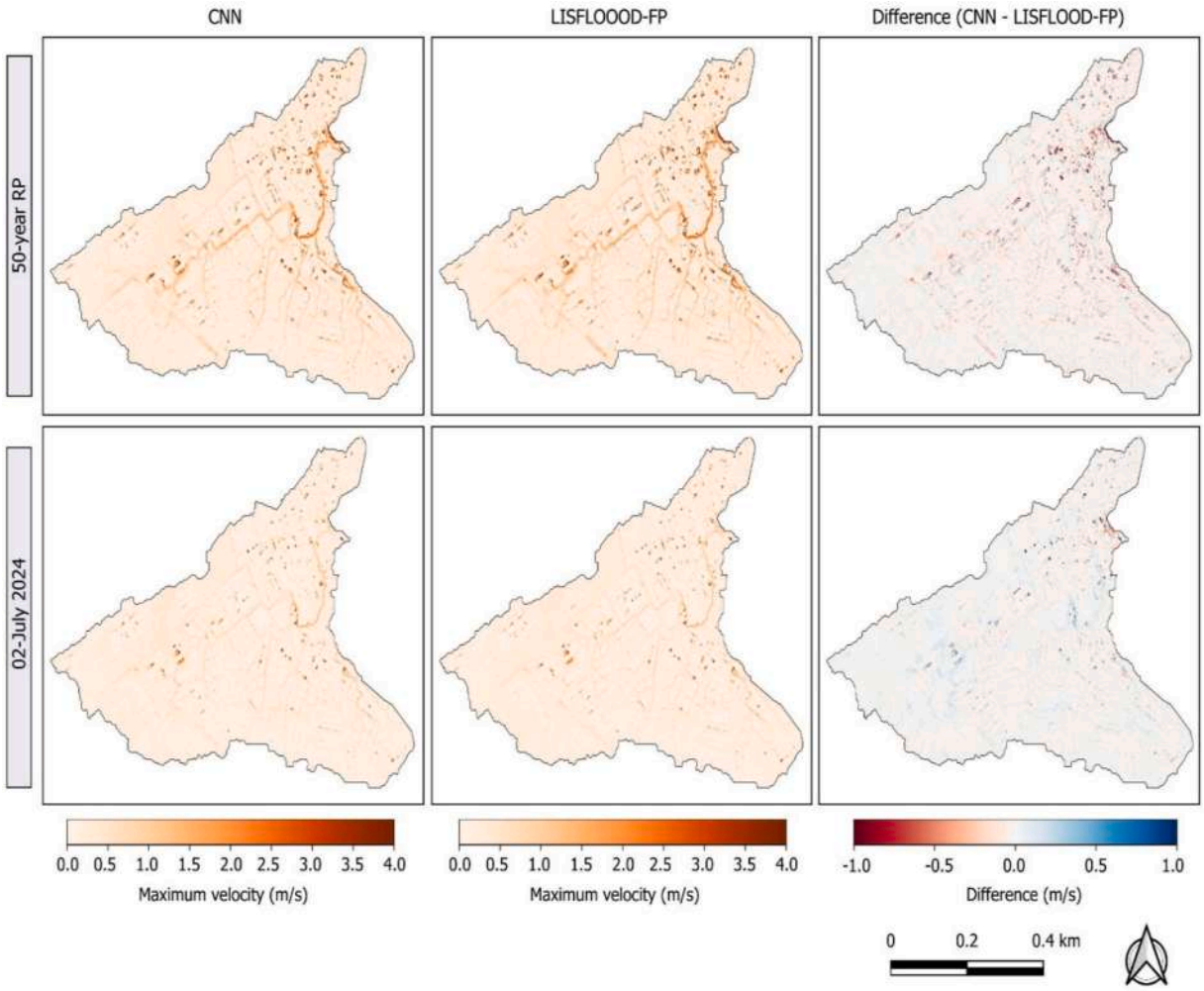

Fig. 7. Flow velocity maps predicted by CNN model and by the physics-based model for 50 years return period event and the real event of 02 July 2024.

3.4. Performance evaluation indicators

To evaluate the performance of the DL model against the numerical simulations, we considered five metrics for both regression and classification tasks. For the regression metrics, we used root mean square error (RMSE) and mean absolute error (MAE), comparing the true water depth and velocity maps generated by the physics-based model (indicated in the below sections also as ground truth) against predicted flood maps. These were calculated as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i}^{N}(y_i - \hat{y}_i)^2} \tag{2}$$

$$MAE = \frac{1}{N}\sum_{i}^{N}|y_i - \hat{y}_i| \tag{3}$$

where y_i and $\hat{y}_i$, are the true values and the predicted values, respectively. The RMSE and MAE range from 0 to ∞, with 0 indicating optimal value. These metrics are applied to both water depth and flow velocity.

As classification metric, we considered the critical success index (CSI) to identify flooded cells. We considered an inundation depth threshold of 0.05 m to categorize whether a cell was wet or dry. Similarly, we used a velocity threshold of 0.1 m/s to distinguish between critical flow or not. Furthermore, the precision and recall metrics are also considered to better evaluate spatial flood prediction performance. Precision demonstrates the correctness of predicted flooded grids, while recall evaluates the ability to identify actual flooded grids. The CSI, precision, and recall are defined as:

$$CSI = \frac{TP}{TP + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

where TP (True Positive) refers to the number of grids that are flooded and correctly predicted as flooded; FP (False Positive) indicates the number of grids that are not flooded but incorrectly predicted as flooded; FN (False Negative) indicates the number of grids that are flooded but incorrectly predicted as not flooded by the model. The CSI has range 0 to 1, with 1 being the best value. Likewise, higher values of precision and recall show superior model performance.

3.5. Computational hardware and hyperparameters

We implemented all codes in Python (version 3.10.12), using as deep learning library Pytorch (version 2.5.1). Model training and testing were carried out on an 8-core Intel CPU processor with 10 GB RAM. For training, we employed an Adam optimizer with a learning rate of 0.001 and a batch size of 4, for a maximum of 500 epochs with early stopping to avoid overfitting. Early stopping considered the validation loss at

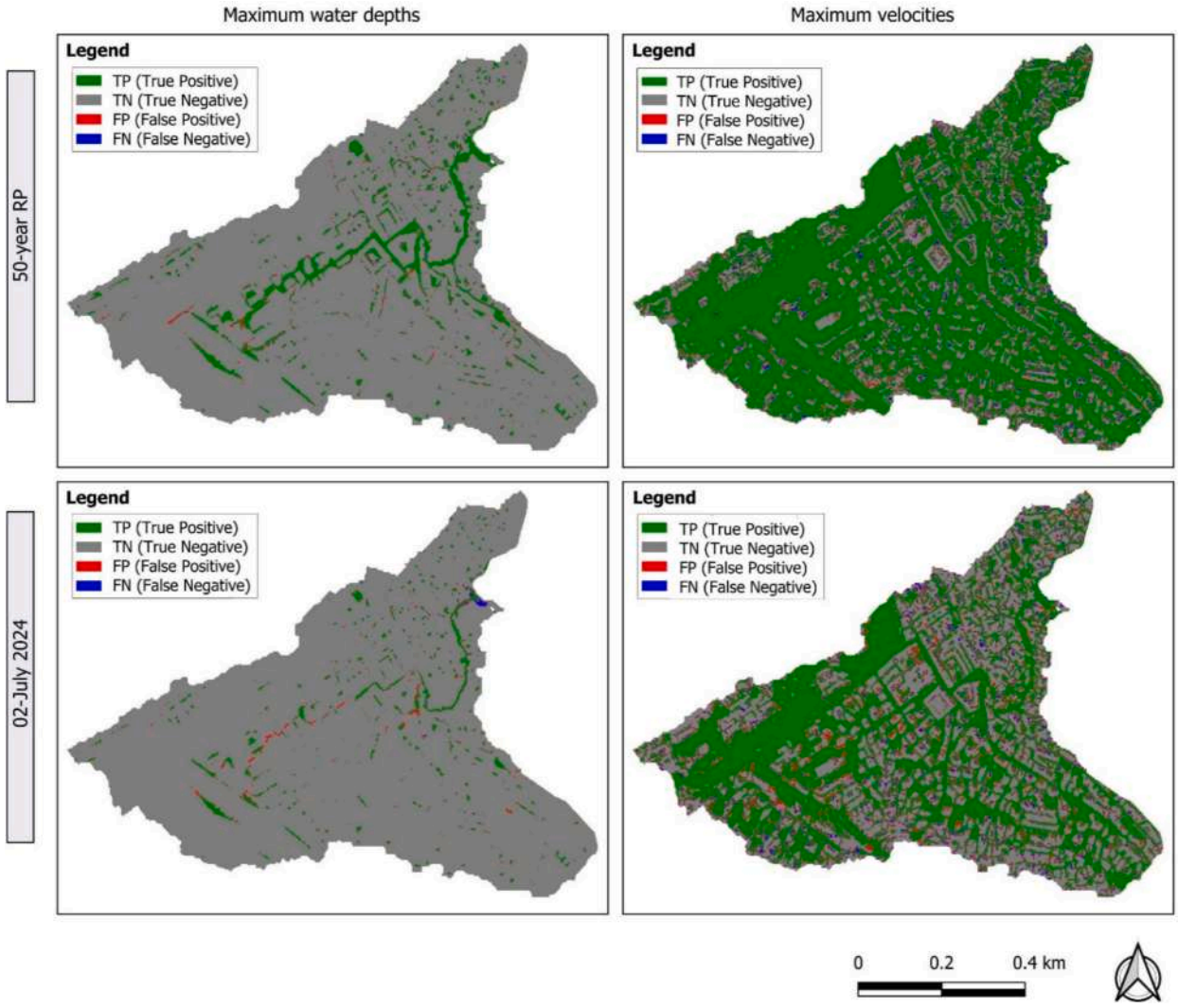

Fig. 8. Binary classification maps of water depths and velocities for the 50 years return period event and the real event of 02 July 2024 (Note that thresholds of 0.05 m for water depth and 0.1 m/s for flow velocity were used to identify flooded areas).

every epoch as monitoring metric and had a patience of 40. The model convergence typically occurred close to the end of training, with best validation loss observed generally around 450–500 epochs.

4. Results

4.1. Feature selection

We first assessed, using Pearson correlation coefficients, the interdependence among the selected hydro-morphological descriptors. Pixel-wise Pearson correlations between descriptors were generally low, with most values below 0.2, indicating limited linear redundancy among the inputs (Fig. 5). The strongest correlation was observed between the Topographic Wetness Index (TWI) and flow accumulation, which is expected given that TWI formulation is explicitly constructed on the base of flow accumulation. Moderate correlations were found between slope and DTM (−0.25) and between Manning's coefficient and TWI (0.20). The observed negative correlation between slope and DTM indicates that steeper areas tend to have larger elevation differences (greater local relief). This pattern is expected in steep catchments like the study basin, where increased slope is associated with stronger vertical gradients and larger elevation drops over short distances. Overall, the obtained low to moderate correlations imply a minimal risk of multicollinearity, justifying inclusion of all selected descriptors in the deep learning model. Second, we computed the non-linear relationship among variables using mutual information analysis. Most of the variables showed relatively low MI values, indicated limited non-linear redundancy. The strong non-linear dependency was observed between flow accumulation and TWI, which is explicitly expected because the derivation of TWI from flow accumulation. Then the moderate non-linear dependency was reported between curvature and slope, illustrating their geomorphological relationships. As a result, this analysis supports the retention of all descriptors due to their distinct physical interpretations and potential contributions to the DL model development.

4.2. DL model prediction performance

Table 2 reports training, validation, and testing performance in terms of MAE, RMSE, and CSI for both water depths and flow velocities. The MAE and RMSE values for water depth remain consistent at approximately 0.36 cm of MAE and 1.5 cm of RMSE for all datasets, with CSI scores averaging around 0.80, indicating a good fit of the predictions. A similar pattern emerges for flow velocities, with MAE around 0.055 m/s, RMSE around 0.17 m/s, and a higher CSI, averaging at 0.84 across all datasets. The higher RMSE, compared with the MAE, indicates that the largest errors occur for the highest water depths and velocities. While the lowest errors are associated with the smallest events, the CSI tends to

Table 3
Testing performance for different combinations (Set) of training and validation events.

Set	Combination Type	Training dataset	Validation dataset	Water depth			Velocity		
				Validation MAE (m)	Test MAE (m)	Test CSI	Validation (m/s)	Test MAE (m/s)	Test CSI
1	Low → Low	2 Low (2019, 29-Feb-2024 real events)	1 Low (28-May-2024 real event)	0.013	0.020	0.30	0.081	0.109	0.73
2	Low → Medium	2 Low (2 years, 5 years return time events)	1 Medium (15 years return time event)	0.018	0.015	0.35	0.102	0.093	0.75
3	Low → High	2 Low (2019, 29-Feb-2024 real events)	1 High (2013 real event)	0.042	0.028	0.22	0.2079	0.1481	0.70
4	High → High	2 High (100 years return time event, 2013 real event)	1 High (40 years return time event)	0.0198	0.0218	0.39	0.0844	0.12	0.74
5	Low + Medium → High	2 Low (2019, 29-Feb-2024 real events) + 1 Medium (2023 real event)	1 High (2013 real event)	0.0384	0.0205	0.29	0.1905	0.1207	0.74
6	Low + High → Medium	2 Low (2 years, 5 years return time events) + 2 High (40 years, 100 years return time events)	1 Medium (25 year return time event)	0.0189	0.0149	0.39	0.0812	0.075	0.76
7	Low + Medium + High → Medium	1 Low (2019 real event) + 1 Medium (2023 real event) + 1 High (2013 real event)	1 Medium (2023 real event)	0.0217	0.017	0.45	0.0824	0.0854	0.76
8	Low + Medium + High → Medium	1 Low (2 years return time event) + 1 Medium (20 years return time event) + 1 High (100 years return time event)	1 Medium (25 years return time event)	0.0175	0.0150	0.42	0.0856	0.0818	0.75
9	Low + Medium + High → Medium	1 Low (2 years return time event) + 1 Medium (20 years return time event) + 1 High (100 years return time event)	1 Medium (2023 real event)	0.027	0.025	0.23	0.17	0.16	0.72
10	Low + Medium + High → Medium	2 Low (21-July-2024 real event, 2 years return time event) + 2 Medium (2023 real event, 20 years return time event) + 2 High (2013 real event, 100 years return time event)	1 Medium (25 years return time event)	0.0138	0.0123	0.52	0.0596	0.0611	0.79
11	Low + Medium → High	4 Low (2 years, 5 years return time events, 2019, 28 May 2024 real events) + 4 Medium (10 years, 15 years, 20 years, return time events, 2023 real event)	I High (2013 real event)	0.035	0.018	0.37	0.170	0.10	0.74
12	Low + Medium + High → Medium	2 Low (2 years, 5 years return time events) + 2 Medium (20 years, 25 years return time events) + 2 High (40 years, 100 years return time events)	1 Medium (2023 real event)	0.022	0.0183	0.35	0.1162	0.1063	0.72
13	Low + Medium + High → Medium	5 Low (2 years, 5 years return time events, 2019, 29-Feb-2024, 28-May-2024 real events) + 3 Medium (15 years, 20 years return time events, 2023 real event) + 3 High (40 years, 100 years return time events, 2013 real event)	1 Medium (10 years, 25 years return time events)	0.014	0.013	0.54	0.0625	0.0635	0.77

improve with the more extreme events, as there are more true positives.

In addition, for the test dataset, the model obtained an average precision of 0.90 and recall of 0.87 for water depth, and 0.92 and 0.90, respectively, for velocity, indicating strong predictive performance. In Figs. 6 and 7, we reported the visual comparison for the largest and smallest events from test dataset, for both water depth and velocity, to evaluate the model's performance under both high and low flood flow conditions. The first event corresponds to 50 years return period rainfall, while the second to the real event of 02 July 2024, which had little rainfall.

For water depths, the CNN model's results closely match the ground truth, effectively capturing the flood both along the main streets along to the steeper outlet and in shallow areas in the upstream part of the study basin (Fig. 6). The biggest errors are observed downstream close to outlet of the basin, where the water depths reach their greatest values due to flow accumulation of whole basin. In this area, the difference reaches a maximum of 30 cm, equivalent to a relative error of approximately 15%, which is reasonable for practical purposes. For flow velocities, the results also accurately match the ground truth for both events (Fig. 7). Here, the error pattern is less clear, with the biggest differences close to buildings, where the flow velocities are the highest. These differences are mainly due to the use of MAE loss function, which treats the errors evenly without prioritizing on large deviations, and therefore allows some outliers while maintaining the overall model performance.

The binary classification maps in Fig. 8 also confirm the performance of CNN model both in terms of water depth and flow velocity. In terms of water depth, for the largest rainfall event, 8.3% of grids are correctly classified as wet and 90.5% exhibited as dry, with only minor misinterpretations observed (0.7% false positives and 0.5% false negatives). For the smallest rainfall event, the wet areas decreased significantly with only 3.2% of grids accurately predicted as wet and 95.9% grids correctly as dry, while the incorrected rates appeared to be slightly lower with 0.7% false positives and 0.3% false negatives. Overall, most errors are false positives and appear upstream of the flooded streets, where water depths are lower.

In terms of flow velocities, for the largest rainfall event, 74.3% of grids are correctly classified as flooded and 17.1% showed as non-flooded areas, with only a small misinterpretation reported (2.9% false positives and 5.7% false negatives). For the smallest rainfall event, flooded areas decreased slightly, with 52.5% of grids accurately predicted as flooded, while non-flooded areas increased, with 36.1% of grids correctly predicted as not flooded. The observed incorrect rates exhibited a slight decrease with 5.1% false positives and 6.3% false negatives. Most misclassifications occurred along the boundaries of buildings, where flow velocities tend to be higher. Part of the reason why velocity maps have more flooded cells compared to the water depth maps is because flow velocity may still exceed 0.1 m/s threshold even in areas where the water depths are very shallow (i.e., <0.05 m).

4.3. *Influence of training data*

To evaluate model robustness under distribution shifts and data scarcity, we trained and validated the model using thirteen different combinations of rainfall events, each with up to only four simulations. The test dataset was kept fixed to ensure a fair comparison across all

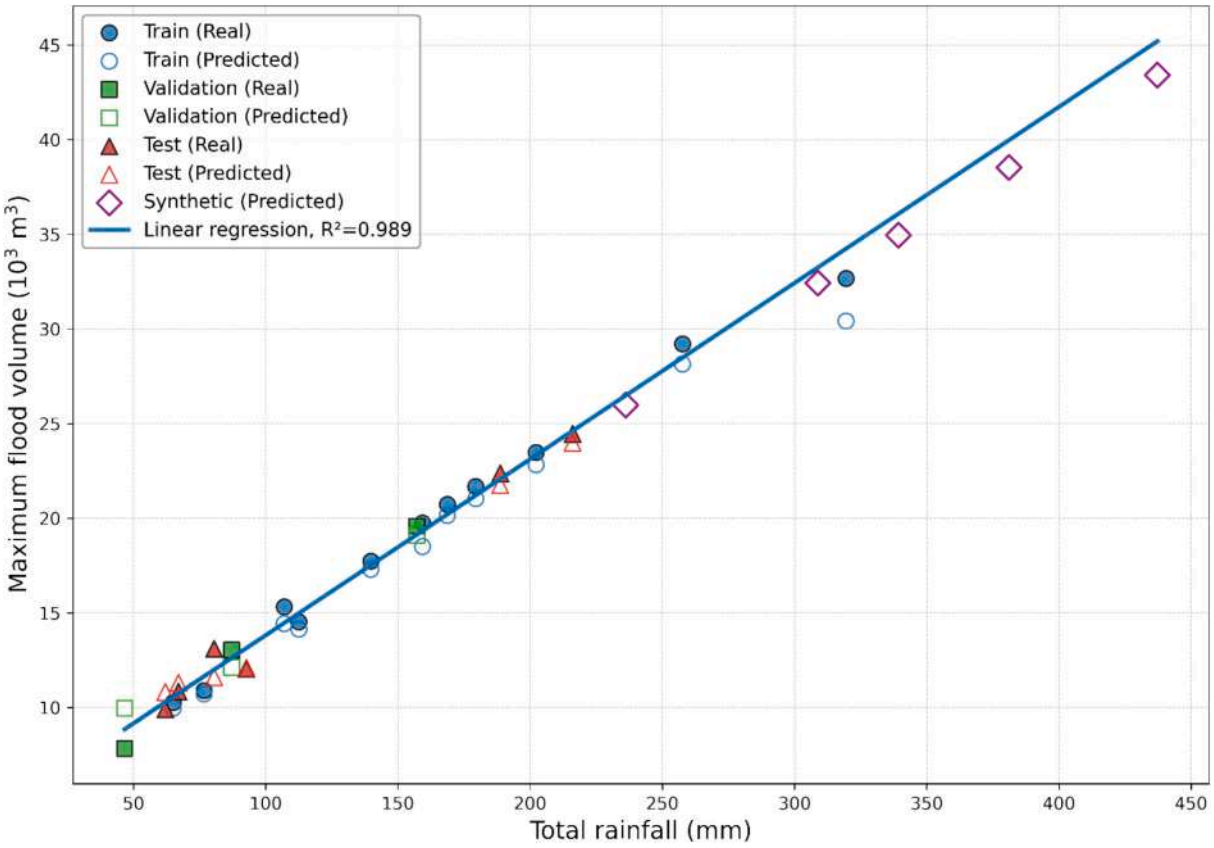

Fig. 9. Relationship between the total rainfall and the total maximum flood volumes for real and predicted datasets using training, validation, and testing events, with synthetic extreme events.

configurations.

Table 3 summarizes the training-validation splits and corresponding validation and test performance. When the model was trained and validated exclusively on low-intensity events (see Set 1 in table 2), it achieved low validation errors (MAE = 0.013 m for depth; 0.081 m/s for velocity) but performed poorly on the test dataset, which included larger events (MAE = 0.020 m; 0.109 m/s). Introducing a larger, but not extreme, validation event (Set 2) substantially improved test performance, particularly for water depth (MAE = 0.015 m) but also velocity (MAE = 0.093 m/s). In contrast, training and validating only on high-intensity events (Sets 3 and 4) resulted in poor generalization, with high errors for both validation and test datasets, especially for velocity.

Increasing the number of training simulations without including high-intensity events (Set 5 & 11) did not resolve this limitation. Models trained on low- and medium-intensity events still exhibited large errors when predicting high events in both validation and test datasets, reflecting the inability to extrapolate beyond the training range.

Performance improved when the training data included events from multiple intensity categories (Sets 6–13). In particular, sets 7 and 8, which combined low-, medium-, and high-intensity events, yielded consistently low validation and test errors for medium-based events. However, a model trained exclusively on synthetic events (Set 9 & 12) performed poorly when predicting real events, indicating limited generalization between synthetic and observed rainfall characteristics. The best performance was achieved with a balanced dataset containing low, medium, and high events from both real and synthetic sources (Set 10 & Set 13), matching the test accuracy obtained when training on the full dataset of 14 simulations.

Overall, these results suggest that under data scarcity, training on low- or medium-intensity events and validating on a medium event provides the most robust performance. When more simulations are available, a balanced mix of low, medium, and high events, including both real and synthetic rainfall, is essential for reliable generalization.

4.4. Generalization to extreme events

To further evaluate the model's generalization capability, we tested its performance to predict four very rare extreme events with return period of 200 years, 300 years, 500 years, and 1000 years, which were not the part of model training nor testing. We considered these scenarios to assess the model's extrapolation behavior under extreme hydrological conditions. To do so, we compared the predicted maximum flood volumes with the total rainfall. The maximum flood volumes were calculated as the sum of the maximum volumes of all grids across the study area, obtained by multiplying the maximum envelope of the water depths at each grid times with their area. Even though this value is not exactly equivalent to the true maximum flood volumes, since it ignores the spatial dynamics, it offers an upper bound that we can use to compare different events.

Fig. 9 shows the relationship between total rainfall and maximum flood volumes for the real and predicted volumes across the training, validation, and testing events, along with the synthetic extreme events. For extreme events, we observed a consistent trend between maximum volumes and total rainfall, which confirms a physically rational hydrological response. The predicted volumes are also closely aligned with true flood volumes across all datasets, as demonstrated by a high coefficient of determination (R^2), with value of 0.99. According to the predictions, the 1000 years return period event, corresponding to a total rainfall of 437.25 mm, resulted in the highest maximum volumes of approximately 43000 m^3, which is about 1.4 times higher than the 100 years event (Fig. 9). To further justify the extrapolation ability of our model, we run the numerical model also on the 1000 years return period event, and then the simulated results were compared against the model predictions. The comparison show that the model indeed performs well,

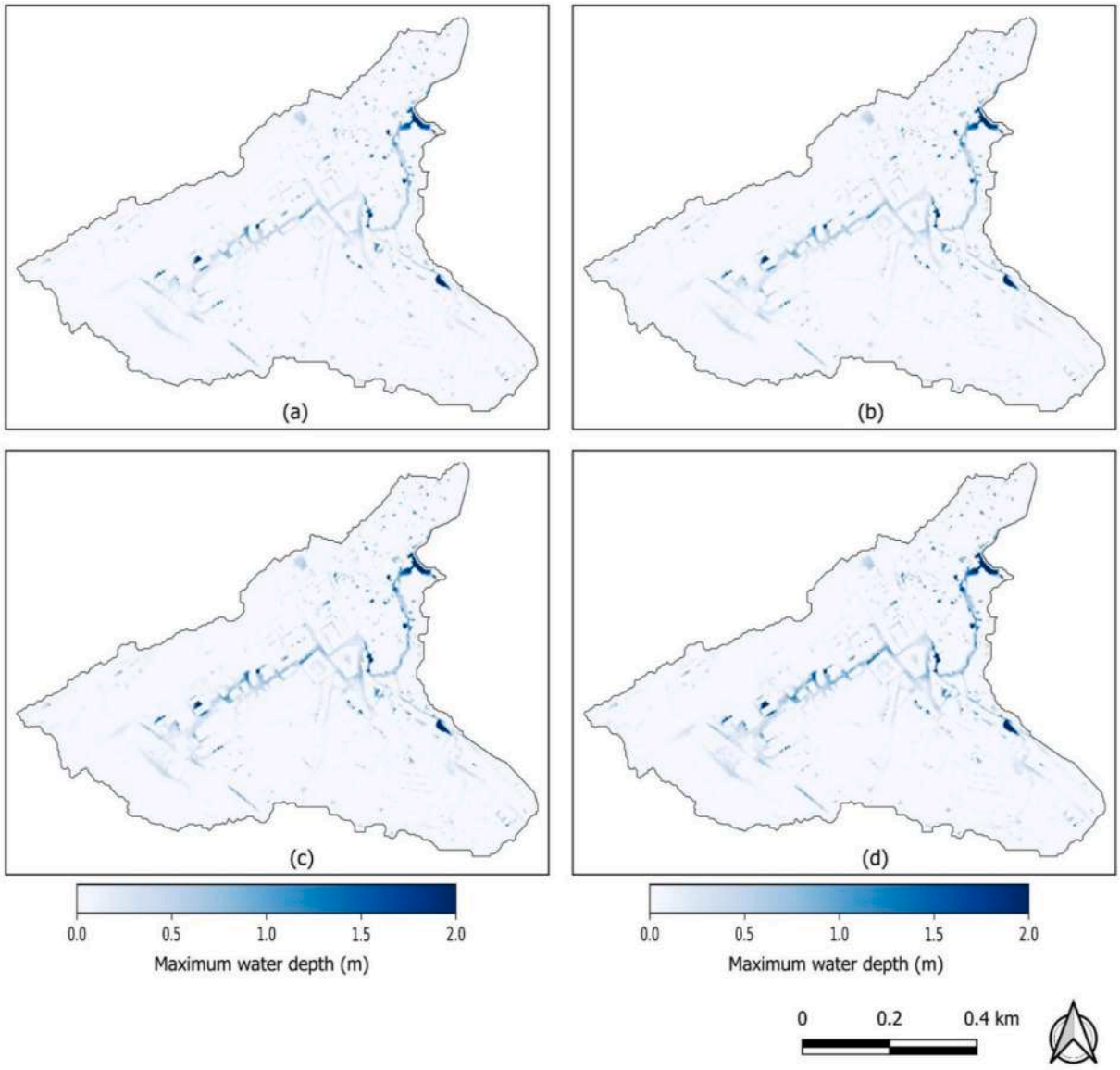

Fig. 10. Water depth maps predicted by CNN model using designed synthetic dataset for different return periods: (a) 200 years, (b) 300 years, (c) 500 years, and (d) 1000 years.

yielding a water depth CSI of 0.71 and a velocity CSI of 0.88, which confirms its extrapolation behavior under unseen extreme events. This is further highlighted by binary classification maps reported in the Supporting Material (see S2), which exhibit strong spatial consistency between predicted and reference maps with only limited misclassified areas.

Fig. 10 illustrated the CNN model predictions of unseen extreme events. The CNN model effectively captured the underlying flood patterns and are consistent with physically rational hydrological response. The model predicts high water depths in low-lying areas (main street) and in the downstream along the outlet of the basin. At the upstream of the basin, the model predicts low water depths, particularly in shallow-areas. This similar capturing trend was observed in all predictions of extreme events, which indicates that the model can accurately reproduce the flood response using different rainfall magnitudes.

4.5. Computational cost

The CNN model required approximately 200 mins for training and 17 s for inference. The simulation time of physics-based model was on average 45 min with the same hardware, leading to an average speed-up of 116.5 times.

5. Discussion and Conclusion

In this study, we proposed a CNN-based approach to rapidly predict spatial maps of the flood water depths and velocities in case of urban flash flood, using the city of Matera, Italy, as a case study. A U-Net architecture was developed and trained using outputs from the physics-based LISFLOOD-FP model. A key contribution of this work is the simultaneous prediction of both flood depth and flow velocity for urban flash flood, which must be combined to assess hazard to pedestrians and vehicles. The CNN predictions closely matched those of the physics-based hydrodynamic model. For water depths maps, the model achieved average test-set MAE, RMSE, CSI, precision, and recall values of 0.37 cm, 1.7 cm, 0.80, 0.90, and 0.87, respectively. For flow velocities maps, the corresponding values were 0.054 m/s, 0.178 m/s, 0.84, 0.92, and 0.90. In terms of computational efficiency, the CNN predicted flood maps at 2×2 m spatial resolution over a 2-hour event in approximately 17 s, making it about 116 times faster than the LISFLOOD-FP model.

Beyond model development, this study assessed the impact of the training dataset composition on the deep learning model's performance. By evaluating different training and validation combinations of rainfall events, we demonstrated that a balanced training dataset including both real and synthetic events across low, medium, and high intensities is critical for achieving robust generalization. The results show that training data selection strongly influences model accuracy, and we recommend including at least one real and one synthetic event per intensity class.

We also evaluated the model's sensitivity and extrapolation capability under extreme rainfall conditions. The CNN successfully extrapolated to extreme flood scenarios, producing maximum flood volumes consistent with total rainfall. When tested on synthetic rainfall events with return periods of 200, 300, 500, and 1000 years, the model realistically captured the evolving spatial patterns and severity of flood maps, indicating stable and reliable extrapolation behavior.

The proposed CNN was trained using 2D shallow water equations-based simulations for Matera case study and therefore cannot be directly transferred to other cities without retraining. The same methodology could, however, be directly applied to other case studies, with similar expected results, as demonstrated by other studies (Guo et al., 2021; Liao et al., 2023; Löwe et al., 2021). Future work could explore transferability and domain adaptation, building on recent advances in this area (e.g., Bentivoglio et al., 2025, 2023; Guo et al., 2022). To enhance transferability, transfer learning approach has shown great potential, especially in pluvial flood modelling. Recent investigations highlighted that the data-driven and neural-operator models can effectively transfer learned features to new spatial environments and enhance the predictive performance in data-scarce areas (Seleem et al., 2022; Xu et al., 2025). However, their effectiveness strongly depends on the similarity between source and target domain. In this regard, domain adaptation enables flood prediction to transfer knowledge with promising performance and can be widely adopted in these cross-city problems (Zhou and Pan, 2022). Additionally, physics-informed deep learning provides an additional approach that embed physical laws into neural networks to improve generalization and robustness (Bentivoglio et al., 2025). Overall, these approaches are promising strategies and direction for improving cross-city urban flash flood prediction.

In addition, the physics-based model assumed spatially-uniform effective rainfall, which was justified by the limited size of the case study. Larger case studies should include the effect of spatially-distributed rainfall and sewer system information, which was also not accounted for in our experiments. Incorporating these features would improve the model's applicability for real-time rainfall forecasts but would likely require a larger training dataset to cover a bigger variability in flooding patterns (Hofmann and Schüttrumpf, 2021; Burrichter et al., 2024).

Nonetheless, these results demonstrate that the proposed CNN provides accurate and computationally efficient flood predictions, highlighting its strong potential for real-time urban flood forecasting and monitoring. Based on this, the CNN model can be readily utilized for real-time flood early warning systems, where near real-time rainfall forecasts are used as inputs to produce rapid predictions of flood water depth and velocity maps at local scale. These maps can be employed to identify the critical zones across the basin, and predefined threshold based-alerts related to flood instability for people (Albano et al., 2026) and vehicles can be issued to support decision-making and community resilience.

CRediT authorship contribution statement

Muhammad Asif: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Roberto Bentivoglio:** Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis. **Raffaele Albano:** Writing – review & editing, Writing – original draft, Validation, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Raffaele Albano reports financial support for this work was provided within the scope of the "Casa delle Tecnologie Emergenti di Matera" project. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work contributes to the outputs of the "Casa delle Tecnologie Emergenti di Matera" project. The contribution of the first author was financed under the Italian PhD Grant with D.M. 352/2022 at University of Basilicata.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2026.135601.

Data availability

Data will be made available on request.

References

Asif, M., Dal Sasso, S.F., Lacava, T., Albano, R., 2025a. Estimating flash flood surface flow velocity in heritage city using citizen-recorded videos. Prog. Disaster Sci. 100474.

Albano, R., Adamowski, J., 2025. Use of digital elevation models for flood susceptibility assessment via a hydrogeomorphic approach: A case study of the Basento River in Italy. Nat. Hazards 121, 9021–9042. https://doi.org/10.1007/s11069-025-07144-z.

Albano, R., Ermini, R., Moramarco, T., Sole, A., 2026. A physically-based conceptual model calibrated on experimental data for esti-mating the flood instability hazard for pedestrians in urban areas. Model. Earth Syst. Environ. 12, 132. https://doi.org/10.1007/s40808-026-02779-8.

Albano, R., Limongi, C., Dal Sasso, S.F., Mancusi, L., Adamowski, J., 2024. Flood scenario spatio-temporal mapping via hydrological and hydrodynamic modelling and a remote sensing dataset: A case study of the Basento river (Southern Italy). Int. J. Disaster Risk Reduct. 111, 104758. ISSN 2212-4209.

Asif, M., Kuglitsch, M.M., Pelivan, I., Albano, R., 2025b. Review and Intercomparison of Machine Learning applications for Short-term Flood forecasting. Water Resour. Manag. 1–21.

Bates, P., Trigg, M., Neal, J., Dabrowa, A., 2013. LISFLOOD-FP user manual. Code Release 5.

Bates, P.D., Horritt, M.S., Fewtrell, T.J., 2010. A simple inertial formulation of the shallow water equations for efficient two-dimensional flood inundation modelling. J. Hydrol. 387, 33–45.

Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2022. Deep learning methods for flood mapping: a review of existing applications and future research directions. Hydrol. Earth Syst. Sci. 26, 4345–4378.

Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2023. Rapid spatio-temporal flood modelling via hydraulics-based graph neural networks. Hydrol. Earth Syst. Sci. 27, 4227–4246.

Bentivoglio, R., Isufi, E., Jonkman, S.N., Taormina, R., 2025. Multi-scale hydraulic graph neural networks for flood modelling. Nat. Hazards Earth Syst. Sci. 25, 335–351.

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

Bui, D.T., Hoang, N.-D., Martínez-Álvarez, F., Ngo, P.-T.-T., Hoa, P.V., Pham, T.D., Samui, P., Costache, R., 2020. A novel deep learning neural network approach for predicting flash flood susceptibility: a case study at a high frequency tropical storm area. Sci. Total Environ. 701, 134413.

Burrichter, B., Koltermann da Silva, J., Niemann, A., Quirmbach, M., 2024. A temporal fusion transformer model to forecast overflow from sewer manholes during pluvial flash flood events. Hydrology 11, 41.

Chen, W., Huang, G., Zhang, H., 2017. Urban stormwater inundation simulation based on SWMM and diffusive overland-flow model. Water Sci. Technol. 76, 3392–3403.

Chen, W., Huang, G., Zhang, H., Wang, W., 2018. Urban inundation response to rainstorm patterns with a coupled hydrodynamic model: a case study in Haidian Island. China. J. Hydrol. 564, 1022–1035.

Coletta, V., Mascitelli, A., Bonazza, A., Ciarravano, A., Federico, S., Prestileo, F., Torcasio, R.C., Dietrich, S., 2021. Multi-instrumental analysis of the extreme meteorological event occurred in matera (Italy). In: On November 2019. in: International Conference on Computational Science and Its Applications. Springer, pp. 140–154.
Costabile, P., Macchione, F., 2015. Enhancing river model set-up for 2-D dynamic flood modelling. Environ Model Softw. 67, 89–107.
Dazzi, S., Vacondio, R., Mignosa, P., 2021. Flood stage forecasting using machine-learning methods: a case study on the Parma River (Italy). Water 13, 1612.
De Almeida, G.A.M., Bates, P., Freer, J.E., Souvignet, M., 2012. Improving the stability of a simple formulation of the shallow water equations for 2-D flood modeling. Water Resour. Res. 48.
DHI, 2017. MIKE 21 flow model & MIKE 21 flood screening tool hydrodynamic module scientific documentation.
Donnelly, J., Abolfathi, S., Pearson, J., Chatrabgoun, O., Daneshkhah, A., 2022. Gaussian process emulation of spatio-temporal outputs of a 2D inland flood model. Water Res. 225, 119100.
El Baida, M., Boushaba, F., Chourak, M., Hosni, M., 2024. Real-time urban flood depth mapping: convolutional neural networks for pluvial and fluvial flood emulation. Water Resour. Manag. 38, 4763–4782.
Ermini, R., Albano, R., 2023. Hydromorphic analysis of urban areas transformations: the case study of the Matera city. In: In: IOP Conference Series: Earth and Environmental Science. IOP Publishing, p. 12026.
Ermini, R., Fattore, C., Zoubi, A.A., 2024. Hydromorphic Impact of Matera's Urban Area. Geographies 4, 152–167.
Farfán-Durán, J.F., Montalvo, C., Cea, L., Leitão, J.P., 2025. Integrating net rainfall calculation in deep learning-based surrogate modeling frameworks for 2D flood prediction. J. Hydrol. 661, 133632.
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.
Guo, Z., Leitao, J.P., Simões, N.E., Moosavi, V., 2021. Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. J. Flood Risk Manag. 14, e12684.
Guo, Z., Moosavi, V., Leitão, J.P., 2022. Data-driven rapid flood prediction mapping with catchment generalizability. J. Hydrol. 609, 127726.
Haastregt, A. V, 2021. Selecting a flood mitigation measure for Matera, Italy, and determining its effectiveness in reducing physical flood characteristics.
Hicks, F.E., Peacock, T., 2005. Suitability of HEC-RAS for flood forecasting. Can. Water Resour. J. 30, 159–174.
Hochreiter, S., 1998. The vanishing gradient problem during learning recurrent neural nets and problem solutions. Int. J. Uncertainty, Fuzziness Knowledge-Based Syst. 6, 107–116.
Hofmann, J., Schüttrumpf, H., 2021. Floodgan: using deep adversarial learning to predict pluvial flooding in real time. Water 13, 2255.
Hong, H., Liu, J., Zhu, A.-X., 2020. Modeling landslide susceptibility using LogitBoost alternating decision trees and forest by penalizing attributes with the bagging ensemble. Sci. Total Environ. 718, 137231.
Hosseiny, H., Nazari, F., Smith, V., Nataraj, C., 2020. A framework for modeling flood depth using a hybrid of hydraulics and machine learning. Sci. Rep. 10, 8222.
Hou, J., Zhou, N., Chen, G., Huang, M., Bai, G., 2021. Rapid forecasting of urban flood inundation using multiple machine learning models. Nat. Hazards 108, 2335–2356.
Isikdogan, F., Bovik, A.C., Passalacqua, P., 2017. Surface water mapping by deep learning. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 4909–4918.
Kabir, S., Patidar, S., Pender, G., 2021. A machine learning approach for forecasting and visualising flood inundation information. In: Proceedings of the Institution of Civil Engineers-Water Management. Thomas Telford Ltd, pp. 27–41.
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.
Keifer, C.J., Chu, H.H., 1957. Synthetic storm pattern for drainage design. J. Hydraul. Div. 83 (4), 1–1332.
Kim, H.I., Han, K.Y., 2020. Urban flood prediction using deep neural network with data augmentation. Water 12, 899.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444.
Leedal, D., Neal, J., Beven, K., Young, P., Bates, P., 2010. Visualization approaches for communicating real-time flood forecasting level and inundation information. J. Flood Risk Manag. 3, 140–150.
Li, S., Ma, K., Jin, Z., Zhu, Y., 2016. A new flood forecasting model based on SVM and boosting learning algorithms. In: In: 2016 IEEE Congress on Evolutionary Computation (CEC). IEEE, pp. 1343–1348.
Liao, Y., Wang, Z., Chen, X., Lai, C., 2023. Fast simulation and prediction of urban pluvial floods using a deep convolutional neural network model. J. Hydrol. 624, 129945.
Lou, H., Zhang, Y., Yang, S., Wang, X., Pan, Z., Luo, Y., 2022. A New Method for Long-Term River Discharge Estimation of Small-and Medium-Scale Rivers by using Multisource Remote Sensing and RSHS: Application and Validation. Remote Sens. 14, 1798.
Löwe, R., Böhm, J., Jensen, D.G., Leandro, J., Rasmussen, S.H., 2021. U-FLOOD–Topographic deep learning for predicting urban pluvial flood water depth. J. Hydrol. 603, 126898.
Luo, W., Li, Y., Urtasun, R., Zemel, R., 2016. Understanding the effective receptive field in deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 29.
Maas, A.L., Hannun, A.Y., Ng, A.Y., 2013. Rectifier nonlinearities improve neural network acoustic models. In: Proc. Icml. Atlanta, GA, p. 3.
Mishra, S.K., Singh, V.P., 2003. SCS-CN Method. In: Soil Conservation Service Curve Number (SCS-CN) Methodology. Springer, pp. 84–146.
Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10). pp. 807–814.
Mishra, M., Albano, R., 2026. FLOOD-DEPTH-ML: Machine Learning-Driven Python Application for Estimation of Urban Flood Depths through Submerged Vehicles Detection. Results Eng. 29, 109495.
Nnanwuba, U.E., Qin, S., Adeyeye, O.A., Cosmas, N.C., Yao, J., Qiao, S., Jingbo, S., Egwuonwu, E.M., 2022. Prediction of spatial likelihood of shallow landslide using GIS-based machine learning in awgu, southeast/Nigeria. Sustainability 14, 12000.
Perrone, A., Inam, A., Albano, R., Adamowski, J., Sole, A., 2020. A participatory system dynamics modeling approach to facilitate collaborative flood risk management: a case study in the Bradano River (Italy). J. Hydrol. 580, 124354.
Piadeh, F., Behzadian, K., Alani, A., 2022. A critical review of real-time modelling of flood forecasting in urban drainage systems. J. Hydrol. 127476.
Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., 2021. Skilful precipitation nowcasting using deep generative models of radar. Nature 597, 672–677.
Rentschler, J., Salhab, M., Jafino, B.A., 2022. Flood exposure and poverty in 188 countries. Nat. Commun. 13, 3527.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer, pp. 234–241.
Samela, C., Albano, R., Sole, A., Manfreda, S., 2018. A GIS tool for cost-effective delineation of flood-prone areas. Computers, Environment and Urban. Systems 70, 43–52. https://doi.org/10.1016/j.compenvurbsys.2018.01.013.
Seleem, O., Ayzel, G., Bronstert, A., Heistermann, M., 2022. Transferability of data-driven models to predict urban pluvial flood water depth in Berlin, Germany. Nat. Hazards Earth Syst. Sci. Discuss. 2022, 1–23.
Shaw, J., Kesserwani, G., Neal, J., Bates, P., Sharifian, M.K., 2021. LISFLOOD-FP 8.0: the new discontinuous Galerkin shallow-water solver for multi-core CPUs and GPUs. Geosci. Model Dev. 14, 3577–3602.
Sole, A., Giosa, L., Albano, R., Cantisani, A., 2013. The laser scan data as a key element in the hydraulic flood modelling in urban areas. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 40, 65–70.
Tanoue, M., Taguchi, R., Alifu, H., Hirabayashi, Y., 2021. Residual flood damage under intensive adaptation. Nat. Clim. Chang. 11, 823–826.
Teng, J., Penton, D.J., Ticehurst, C., Sengupta, A., Freebairn, A., Marvanek, S., Vaze, J., Gibbs, M., Streeton, N., Karim, F., 2022. A comprehensive assessment of floodwater depth estimation models in semiarid regions. Water Resour. Res. 58, e2022WR032031.
Thrysøe, C., Balstrøm, T., Borup, M., Löwe, R., Jamali, B., Arnbjerg-Nielsen, K., 2021. FloodStroem: a fast dynamic GIS-based urban flood and damage model. J. Hydrol. 600, 126521.
Wu, W., Emerton, R., Duan, Q., Wood, A.W., Wetterhall, F., Robertson, D.E., 2020a. Ensemble flood forecasting: current status and future opportunities. Wiley Interdiscip. Rev. Water 7, e1432.
Wu, X., Wang, Z., Guo, S., Liao, W., Zeng, Z., Chen, X., 2017. Scenario-based projections of future urban inundation within a coupled hydrodynamic model framework: a case study in Dongguan City. China. J. Hydrol. 547, 428–442.
Wu, Z., Zhou, Y., Wang, H., Jiang, Z., 2020b. Depth prediction of urban flood under different rainfall return periods based on deep learning and data warehouse. Sci. Total Environ. 716, 137077.
Xu, Q., De Vos, L.F., Shi, Y., Rüther, N., Bronstert, A., Zhu, X.X., 2025. Urban flood modeling and forecasting with deep neural operator and transfer learning. J. Hydrol. 661, 133705.
Yan, B., Mu, R., Guo, J., Liu, Y., Tang, J., Wang, H., 2022. Flood risk analysis of reservoirs based on full-series ARIMA model under climate change. J. Hydrol. 610, 127979.
Yang, Q., Wu, W., Wang, Q.J., Vaze, J., 2022. A 2D hydrodynamic model-based method for efficient flood inundation modelling. J. Hydroinformatics 24, 1004–1019.
Zhang, Y., Ragettli, S., Molnar, P., Fink, O., Peleg, N., 2022. Generalization of an Encoder-Decoder LSTM model for flood prediction in ungauged catchments. J. Hydrol. 614, 128577.
Zhou, R., Pan, Y., 2022. Flooddan: Unsupervised flood forecasting based on adversarial domain adaptation. In: In: 2022 IEEE 5th International Conference on Big Data and Artificial Intelligence (BDAI). IEEE, pp. 6–12.