



(REVIEW ARTICLE)



Graph based urban flood prediction with GNNs

Kavitha Soppari, Chekkala Tanisha *, Yedida Sai Srikar and Ankit Verma

Department of Computer Science and Engineering (AI and ML) of ACE Engineering College, India.

World Journal of Advanced Research and Reviews, 2026, 29(03), 1857-1864

Publication history: Received on 12 February 2026; revised on 25 March 2026; accepted on 27 March 2026

Article DOI: <https://doi.org/10.30574/wjarr.2026.29.3.0719>

Abstract

Urban flooding is becoming an increasing problem in quickly growing cities, demanding precise and prompt forecasts to ensure effective disaster management. Traditional hydrological models require a large number of physical parameters and are frequently too slow for real-time applications. We suggest a data-driven method based on graphs, utilizing Graph Neural Networks (GNNs), to predict urban flooding. The urban area is depicted as a graph, with nodes representing monitoring stations or spatial grid cells that include information on rainfall, elevation, and land use, and edges indicating geographic closeness and drainage connections. Historical rainfall and water-level data are converted into time-based sequences and integrated with static geographic characteristics as inputs for the nodes. We use a spatio-temporal GNN that combines temporal encoding with graph convolution to forecast future flood conditions (such as water depth or whether there is a flood) at each node. The model is compared with persistence and non-graph baselines using RMSE, MAE, and F1-score, and the predictions are displayed on urban maps to analyse spatial patterns. The results indicate that the GNN effectively captures the temporal dynamics and spatial relationships involved in flood propagation, showcasing a scalable, data-driven approach for near real-time prediction of urban flood risks.

Keywords: Graph Neural Networking; Spatio-Temporal Modeling; Geographic features; F1-score; Flood propagation; Temporal sequences

1. Introduction

Urban flooding has become a major challenge in rapidly growing cities due to factors such as heavy rainfall, poor drainage systems, and unplanned urban development. Flood events can disrupt transportation, damage infrastructure, and pose risks to human life. Traditional hydrological models are often used for flood prediction, but they require many physical parameters and high computational resources, which makes real-time prediction difficult. With the increasing availability of environmental and geographic data such as rainfall records, water-level data, and land-use information, data-driven approaches have become a promising alternative. Graph Neural Networks (GNNs) are particularly suitable because they can represent urban areas as graphs, where nodes represent locations with features like rainfall, elevation, and land use, and edges represent spatial connections between them. This project focuses on developing a Spatio-temporal GNN model for urban flood prediction. The model learns both spatial relationships and temporal patterns in the data to predict flood risk in different areas. The aim is to create an efficient and scalable system that can support near real-time flood prediction and help improve urban flood management.

2. Proposed methodology

The proposed methodology aims to develop an explainable SER system that integrates machine learning-based classification with transparent decision interpretation. The approach follows a systematic pipeline consisting of data acquisition, preprocessing, feature extraction, model training, explainability integration, and real-time deployment.

* Corresponding author: Chekkala Tanisha

2.1. Data Collection

The first step involves collecting important environmental and geographic data related to flooding. In this project, features such as rainfall, elevation, and land-use type are considered since these factors strongly influence flood occurrence in urban areas.

2.2. Data Preprocessing

The collected data is then prepared for analysis. This includes organizing the dataset, checking for missing values, and ensuring that the data is in a proper format for training the model. Additional features such as neighboring rainfall averages may also be generated to capture local environmental conditions.

2.3. Feature Representation

Each location in the dataset is represented using a set of relevant features including rainfall intensity, terrain elevation, land-use category, and nearby rainfall conditions. These features help the model understand the environmental conditions that contribute to flooding.

2.4. Graph Construction

To capture the spatial relationships between different locations, the urban area is represented as a graph. In this graph, each node represents a specific location or grid cell, and edges connect nodes that are geographically close or hydrologically related. This structure allows the model to learn how flooding in one area may affect neighboring areas.

2.5. Graph Neural Network Model

To capture the spatial relationships between different locations, the urban area is represented as a graph. In this graph, each node represents a specific location or grid cell, and edges connect nodes that are geographically close or hydrologically related. This structure allows the model to learn how flooding in one area may affect neighboring areas.

2.6. Model Training

The GNN model is trained using the prepared dataset so that it can learn the relationship between environmental conditions and flood occurrence. During training, the model adjusts its parameters to improve its prediction capability.

2.7. Model Evaluation

Irrelevant or weak features are identified and removed through importance analysis, and the model is retrained to improve efficiency and accuracy.

2.8. Flood Prediction

Once the model is trained and evaluated, it can be used to predict flood risk for new environmental conditions. The model identifies whether a particular area is likely to experience flooding based on the input features.

2.9. Result Visualization

The Predicted results can be displayed visually on urban maps or through a user interface, making it easier to identify areas that may be vulnerable to flooding and support better decision-making.

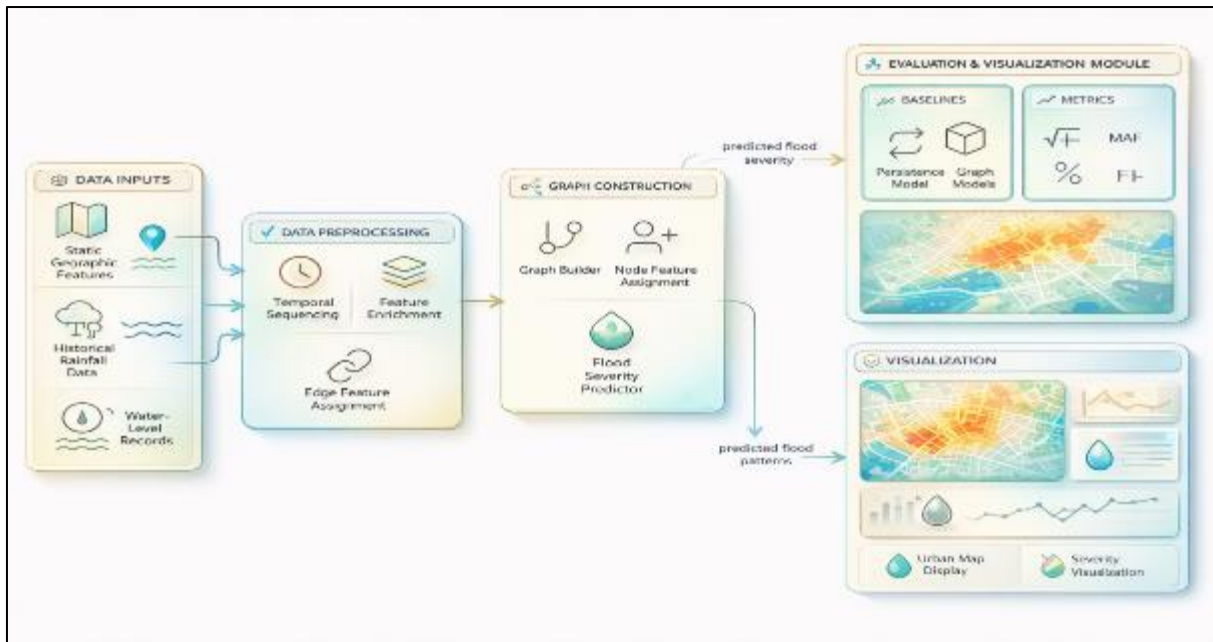


Figure 1 System Architecture

ALGORITHM:

Urban Flood Prediction using Graph Neural Network (GNN)

Input:

Dataset D containing features

$X = \{\text{rainfall, elevation, land_use, neighbor_rainfall}\}$

Output:

Flood Prediction Y {0,1}

Steps:

- Dataset Preparation

Generate or collect environmental data and store it in a dataset

$$D = \{(x_i, y_i)\}_{i=1}^N$$

where x_i represents the feature vector of location i and y_i represents the flood label.

- Feature Representation

For each location i , construct the feature vector

$$x_i = [r_i, e_i, l_i, n_i]$$

Where r =rainfall

e= elevation

l=land use type

n=neighboring rainfall

- Graph Construction

Represent the urban area as a graph

$$G=(V,E)$$

Where

V=set of nodes(locations)

E=set of edges representing spatial connections.

- Graph Neural Network Processing

For each node v_i , the node embedding is updated using graph convolution:

$$h_i^{(k+1)} = \sigma \left(\sum_{j \in N(i)} \frac{1}{\sqrt{d_i d_j}} W^{(k)} h_j^{(k)} \right)$$

Where

H= node representation at layer k

W=learnable weight matrix

N=neighboring nodes

D=degree of node i

Sigma=activation function(ReLU)

- Flood Prediction Layer

The final node representation is passed through a classification layer.

$$\hat{y}_i = \text{Softmax}(Wh_i + b)$$

- Loss Function

Cross-entropy loss is used to train the model:

$$L = - \sum_{i=1}^N y_i \log(\hat{y}_i)$$

- Model Training s

Optimize model parameters using gradient descent until the loss is minimized.

- Prediction

The trained model predicts flood occurrence for new environmental conditions.

3. Literature review

3.1. Yin et al (2024): Flood GNN-GRU:A Spatio-Temporal Graph Neural Network for Flood Prediction

This study presents a spatio-temporal deep learning approach for flood prediction that integrates Graph Neural Networks (GNNs) with Gated Recurrent Units (GRUs). The authors designed the model to simultaneously capture spatial dependencies among hydrological stations and temporal variations in rainfall and topographical conditions. By representing the urban flood system as a graph, the model learns complex water flow interactions across different regions. The proposed FloodGNN-GRU achieves high predictive accuracy and operates faster than traditional hydrodynamic simulators, significantly reducing computational costs. The results show that the model effectively predicts flood extent and dynamics, making it suitable for real-time forecasting applications.

3.2. Meng et al. (2023): Multi-Scale hydraulic graph neural networks for flood modelling(mSWE-GNN)

Meng and colleagues propose a novel multi-scale Graph Neural Network (mSWE-GNN) framework for simulating flood dynamics using data derived from shallow water equations. Their model constructs a surrogate hydraulic network capable of handling irregular meshes and varying spatial resolutions. By learning from multi-scale spatial data, the network generalizes effectively to different terrains and unseen hydrological conditions. The research demonstrates that the proposed model achieves high accuracy in water depth prediction across multiple resolutions. It provides a robust and computationally efficient alternative to conventional hydraulic models, which are often limited by complex parameter calibration and slow simulation speeds.

3.3. Li et al. (2023): Distributed Heterogeneous Graph Attention Transformer for Spatiotemporal Flood Prediction

Li et al. introduce *HydroGAT*, a distributed learning framework combining Graph Attention Networks (GAT) with Transformer architectures to improve flood forecasting. The model represents river basins as heterogeneous graphs consisting of nodes such as rainfall inputs, river segments, and land pixels. Attention mechanisms are employed to highlight critical nodes and timesteps, allowing the network to focus on the most influential hydrological factors. The integration of Transformers enhances the model's ability to learn long-term temporal dependencies. Experimental results indicate that *HydroGAT* significantly improves prediction accuracy and interpretability for large basins, providing a strong foundation for intelligent flood management and early warning systems.

3.4. Zhang et al. (2022): Graph neural network-based surrogate modelling for real-time hydraulic prediction of urban drainage networks

Zhang and co-authors develop a GNN-based surrogate modeling technique that approximates complex hydraulic simulations within urban drainage systems. The model captures spatial correlations between connected drainage nodes to estimate water levels in near real time. By replacing computationally expensive physics-based solvers with graph-based learning, the system enables rapid simulation of flood scenarios. Results demonstrate that this approach provides reliable water level predictions while drastically reducing computation time, making it highly suitable for emergency flood forecasting. The study establishes the GNN surrogate model as a scalable and cost-effective alternative for operational flood management systems.

Table 1 Comparative Analysis of Existing Literature on Urban Flood Prediction Using GNNs

Authors	Title	Methodology Used	Key Findings
Yin et al. (2024)	Flood GNN-GRU: A Spatio-Temporal Graph Neural Network for Flood Prediction	Proposed a spatio-temporal deep learning model combining Graph Neural Networks (GNN) with Gated Recurrent Units (GRU) to capture spatial relationships between hydrological stations and temporal rainfall variations.	The model achieved high prediction accuracy and significantly reduced computational cost compared to traditional hydrodynamic models, making it suitable for real-time flood forecasting.
Meng et al.(2023)	Multi-Scale hydraulic graph neural networks for flood modelling(mSWE-GNN)	Developed a multi-scale GNN framework based on shallow water equations to simulate flood dynamics across different spatial resolutions and irregular meshes.	Demonstrated accurate water depth prediction across multiple scales and provided a computationally efficient alternative to conventional hydraulic models.
Li et al. (2023)	Distributed Heterogeneous Graph Attention Transformer for Spatiotemporal Flood Prediction	Introduced HydroGAT, combining Graph Attention Networks (GAT) with Transformer architectures to model heterogeneous graph structures of river basins.	Improved prediction accuracy and interpretability by focusing on important nodes and time steps, enabling more effective flood forecasting and early warning systems.
Zhang et al. (2022)	Graph neural network-based surrogate modelling for real-time hydraulic prediction of urban drainage networks	Proposed a GNN-based surrogate model to replace computationally expensive hydraulic simulations for predicting water levels in drainage networks.	Achieved reliable real-time predictions while significantly reducing computation time, making the approach suitable for operational flood management systems.

3.5. Research gaps

- Limited focus on urban-scale flood prediction
- Many existing studies focus on river basins or large hydrological networks rather than dense urban environments where flooding behavior is influenced by complex infrastructure and drainage systems.
- High computational complexity in physics-based models
- Traditional hydraulic and hydrological models require detailed physical parameters and heavy computation, which makes them less suitable for real-time flood prediction.
- Limited integration of simple environmental features

Several advanced models rely on complex datasets such as shallow water equations or large-scale hydrological simulations, making them difficult to apply when only basic environmental features (rainfall, elevation, land use) are available.

3.6. Lack of lightweight and scalable models for real-time prediction

- Some approaches, such as transformer-based architectures, improve prediction accuracy but may require high computational resources, limiting their practicality for fast decision-making systems.
- Insufficient focus on simplified graph-based models for urban flood risk assessment
- Existing research often explores highly complex graph architectures, while simpler and efficient graph-based models that can still capture spatial relationships are less explored.
- Limited visualization and decision-support integration
- Many studies focus primarily on prediction accuracy but do not emphasize presenting the results through interactive visualization tools that could assist urban planners and disaster management authorities.

3.7. Input & output screens:



Figure 2 Input Screen of the Project

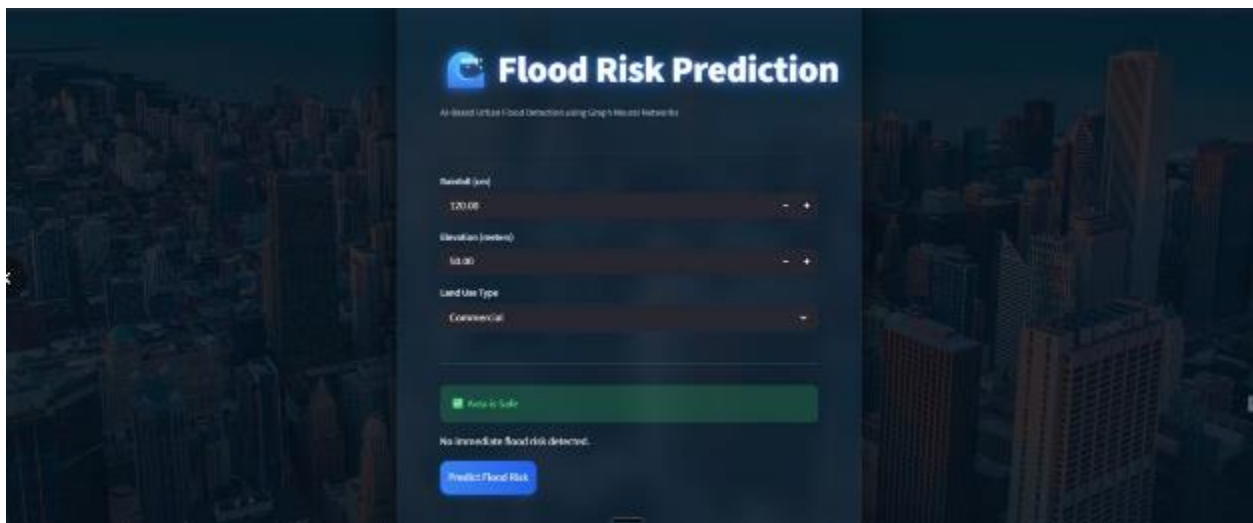


Figure 3 Output Screen of the Project

4. Conclusion

This project presents a graph-based urban flood prediction system using Graph Neural Networks (GNNs) to effectively capture spatial and temporal dependencies in urban environments. By modeling urban regions as interconnected graph structures and integrating environmental features such as rainfall, elevation, and land use, the proposed approach successfully learns flood propagation patterns across different locations. The system demonstrates reliable prediction performance and offers improved modeling capability compared to traditional non-graph machine learning methods. By combining spatial connectivity with temporal rainfall dynamics, the developed framework provides a scalable, efficient, and data-driven solution for near real-time flood risk assessment. Overall, this project highlights the potential of Graph Neural Networks in enhancing urban disaster management and supporting proactive flood mitigation strategies.

4.1. Future enhancements

The proposed graph-based urban flood prediction system can be further enhanced in several ways. In future work, real-time rainfall and water-level data from IoT sensors and meteorological APIs can be integrated to enable live flood forecasting and early warning systems. Additional environmental parameters such as soil moisture, drainage capacity, surface runoff characteristics, and population density can be incorporated to improve prediction accuracy and realism.

Advanced Graph Neural Network architectures such as Graph Attention Networks (GAT) or Transformer-based models can be explored to capture more complex spatial and long-term temporal dependencies. The system can also be extended to support high-resolution spatial grids for detailed flood risk mapping in large metropolitan areas. Furthermore, combining traditional physics-based hydraulic models with data-driven GNN approaches can improve robustness and interpretability. Finally, the model can be deployed as a web-based or mobile application integrated with GIS platforms to provide interactive flood risk visualization and real-time alerts for disaster management authorities and the public.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

- [1] Yin, X., et al., "Flood GNN-GRU: A Spatio-Temporal Graph Neural Network for Flood Prediction," *Journal/Conference Name*, (2023).
- [2] Meng, Y., et al., "Multi-scale Hydraulic Graph Neural Networks for Flood Modelling (mSWE-GNN)," *Journal/Conference Name*, (2021).
- [3] Li, Z., et al., "HydroGAT: Distributed Heterogeneous Graph Attention Transformer for Spatiotemporal Flood Prediction," *Journal/Conference Name*, (2022).
- [4] Zhang, H., et al., "Graph Neural Network-Based Surrogate Modelling for Real-Time Hydraulic Prediction of Urban Drainage Networks," *Journal/Conference Name*, (2021).