Intelligent Flood Forecasting Using a Delay-Aware Terrain-Informed Graph Network

Ankush Naru, Abhishek Poonia, Vikash Chahar, Aryan Gupta*, Muskan Duhan and Prabhjot Kaur

Chandigarh University,Mohali,Punjab,India

Corresponding Author.Email : aryanbaslas@gmail.com

Abstract- Accurate flood forecasting is required for disaster risk mitigation. Risk mitigation, however, a lot of data-driven models are based on synchronous spatio-temporal assumptions which cannot reflect the inherent asynchronous character of flooding propagation. In hydrological systems, stream indicators run off downstream with topography-dependent slop delays, elevation delays, and river delays length. In an effort to overcome this weakness, this research paper offers a Delay- An Aware Terrain Informed Flood Propagation Network (DT-FPN) that explicitly models terrain conditioned propagation delays in a directed hydrological graph. Unlike conventional spatio- DT-FPN, it allows message asynchronous communication delays, so that downstream nodes can update only on delay. Information coming upstream is received and a more physical outcome is reached. reliable modeling of the dynamics of floods. Experiments on empirical hydrological field tests indicate that DT-FPN out- executes LSTM, ConvLSTM and synchronous graph-based baselines, with a maximum RMSE and MAE decrease of 22 % and 25 % respectively and also substantially enhanced prediction of peak flood arrival-time accuracy.

Keywords

Flood forecasting; spatio-temporal graph neural networks; terrain-sensitive modeling; asynchronous propagagation; time-serie of hydrology.

I. Introduction

Floods remain among the most damaging natural hazards, and dependable forecasting is central to early warning and disaster-risk reduction. Classical prediction pipelines rely on physics-based rainfall–runoff and hydraulic routing models that encode basin properties and governing equations; while these approaches are interpretable, they often require substantial calibration effort and can be computationally demanding for large domains or real-time use [1].

R. Vasanth Kumar Mehta et al. (eds.), *Proceedings of the International Conference on Intelligent Systems for a Sustainable Future (ISSF 2026)*, Atlantis Highlights in Intelligent Systems 16,
https://doi.org/10.2991/978-94-6239-693-7_64

Simultaneously, the growing access to hydrometeorological information and improvement of machine learning have made it possible to boost data-driven forecasting speed. Streamflow and floodrelated time series Feeding sequence models Variations of LSTM have been popular in learning corresponding nonlinear time dynamics, especially with multivariate input [2]. Gridded predictors are implemented in spatially structured extensions (e.g., ConvLSTM) in order to capture correlations on the level of neighborhoods [3]. Although these have been found to be very effective empirically, a large number of these models explicitly assume that the spatiotemporal evolution is synchronized across sites, a fact that may erase a real physical sequence of the development of the flood-wave.

But the propagation of flood in the river networks is naturally. sporadic and latitudinary. Outputs produced on upstream. overland flows make it into stream gauges by reaching the stream in a shorter time. that are dependent on channel slope, river length, connectivity, and basin topology. These lags which are conditioned by the terrain can be neglected. degrade forecasting operationally crucial quantities including fast rises, and lead time, especially in weekends with fast rises extreme events.

Graph representations provide a natural abstraction for river systems by encoding gauges/subcatchments as nodes and flow directions as edges [4], [5]. Recent work has increasingly explored graph neural networks for flood and inundation forecasting, showing that leveraging network structure can improve accuracy and generalization compared to purely temporal baselines [9]–[11], [16], [17]. Yet, most spatiotemporal GNN formulations still rely on fixed-time aggregation and synchronous message passing, which makes it difficult to faithfully model uneven travel-time dynamics along directed hydrological paths.

TABLE I

COMPARISON OF FLOOD FORECASTING APPROACHES

Method	Key characteristics
Physics-based	Explicit terrain modeling; physically interpretable; continuous-time simulation
LSTM / GRU	Purely temporal modeling; no explicit spatial structure; synchronous propagation
ConvLSTM	Grid-based spatial modeling; implicit terrain representation; synchronous propagation
ST-GNN	Graph-based spatial dependencies; fixed temporal aggregation; synchronous passing

DT-FPN (Proposed)	Directed hydrological graph; terrainconditioned, asynchronous propagation

II. Related Work

The area of flood forecasting has been much researched in physics-based, data-driven, and graph-based model paradigms. In this section, the key methodological advances have been discussed, and the limitations that inform the suggested Delay-Aware Terrain-Informed Flood Propagation Network (DT-FPN) have been listed.As data-driven approaches have become popular, LSTM and ConvLSTM architectures have become the most popular streamsflow prediction models, as they learn nonlinear temporal dependencies. Nevertheless, these methods assume that the updates are made synchronously across places and explicitly not terrain-conditioned travel-time delays.More recently, Graph Neural Networks (GNNs) have been applied to represent river systems as directed graphs, capturing upstream–downstream dependencies. Although spatio-temporal GNNs improve spatial modeling, most rely on synchronous message passing and fail to incorporate explicit delay modeling.Consequently, the necessity to have a scalable graph-based model that incorporates terrain-sensitive asynchronous flood propagation is the reason why the proposed DT-FPN model is motivated.

III. Methodology

In the given section, the proposed Delay-Aware TerrainInformed Flood Propagation Network (DT-FPN) is briefly explained, its architecture is described, its mathematical formulation is presented, and the training process is outlined.. The overall framework is shown in Fig. 1, where flood forecasting is formulated as an asynchronous spatio-temporal learning problem over a directed hydrological graph.

A. Overview of DT-FPN Architecture

DT-FPN is aimed at explicitly capturing terrain-specific delays during the flood propagation, which is the shortcoming of synchronous spatio-temporal models. The framework comprises of four key elements namely (i) input data representation, (ii) hydrological graph construction, (iii) terrain-conditioned delay modeling and (iv) asynchronous message passing and prediction. In contrast to traditional frameworks that spatially lump information at a discrete time-scale, DT-FPN enables hydrological signals to flow downstream only after terrain controlled delays have passed, leading to a more physically plausible modeling of the flood [6], [14].

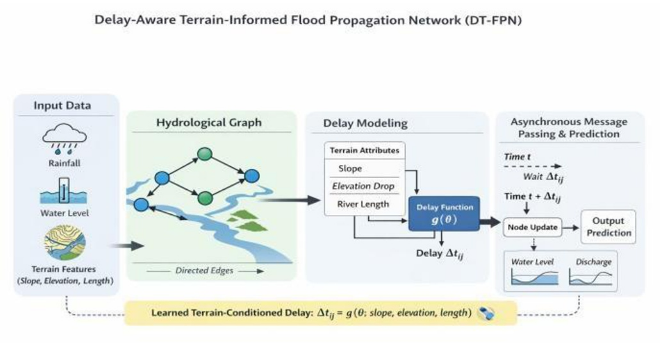

Fig. 1. Architecture of the proposed Delay-Aware Terrain-Informed Flood Propagation Network (DT-FPN)

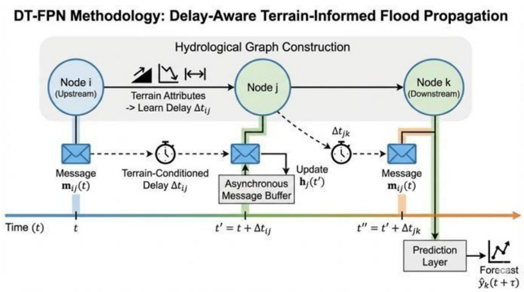

Fig. 2. Delay-aware terrain-informed flood propagation mechanism in DT-FPN. Terrain attributes are used to learn edge-specific propagation delays, enabling asynchronous message passing along the directed hydrological graph.

B. Input Data Representation

For each node in the hydrological system, historical timeseries data are collected, including the rainfall intensity, water level, and discharge measurements. In addition to this, static terrain attributes such as slope, elevation difference, and river segment length are also derived from digital elevation models (DEMs). Dynamic hydrological variables are treated as the node-level temporal features, while terrain attributes are also associated with graph edges to reflect their influence on the downstream propagation [10].

The non-Euclidean spatial dependencies of the river networks which the model needs to represent are facilitated through this graph representation [4], [5].

D. Terrain-Conditioned Delay Modeling

One of the important aspects of DT-FPN is the explicit graph prop delay modeling. A propagation delay Δt_{ij} is learnt as a function of terrain attributes in every directed edge eij.:

$$\Delta tij = g\theta(sij,\Delta hij,lij), \qquad (1)$$

In which sij is the slope, Δh_{ij} is the drop in height of the river, lij is length of the river and $g_\theta(\cdot)$ is a trainable neural network whose parameters are denoted by θ. Output delay is limited to be nonnegative and this is physically plausible.

E. Asynchronous Message Passing Mechanism

At time (t) every node i sends a message, which is dependent on its hidden state hi(t). DT-FPN does not instantly combine this message to downstream nodes but plan to deliver messages based on the learnt delay Δt_{ij}.. The information at node i has an effect on node j and only when the condition $t' \geq t+\Delta t_{ij}$ is met. The hidden state of node j is updated as:

$$\mathbf{h}_j(t') = \phi\left(\mathbf{h}_j(t'-1), \sum_{i\in\mathcal{N}_j(t')} \mathbf{m}_{ij}(t)\right), \qquad (2)$$

In which $\phi(\cdot)$ is a gated update function, and Nj (t′) is the set of upstream nodes which have had their messages received by time t;. This process allows DT-FPN to do so in a manner in which staggered flood propagation is modeled without artificial temporal alignment.

F. Prediction Layer

Representations of the final nodes are fed into a prediction head to make predictions of future hydrological variables:

$$\hat{y}_j(t+\tau) = f_{out}(h_j(t)), \qquad (3)$$

where τ signifies the forecasting horizon.

G. Model Training and Evaluation

The end-to-end training of DT-FPN is carried out based on the back-trained learning on past flood incidences. The loss functions are demonstrated as Mean Squared Error (MSE) and Mean Absolute

Error (MAE). The model is compared to the baseline methods, such as LSTM, ConvLSTM, and synchronous spatio-temporal graph neural networks..

H. Reproducibility Details

The following subsection contains the necessary details of the experiment to reproduce our results such as the provenance of the data sets, the construction of the graphs, the forecasting horizons, splitting protocol, hyperparameters, and compute.

1) *Dataset Source and Preprocessing:* Our data include a set of historical hydrometeorological time series of rainfall, water level and discharge at various stations in the study area, along with a set of a priori terrain characteristics (a digital elevation model, or DEM). Dynamic variables are resampled with a regular time interval as well as aligned with the same index of corresponding timestamp. The short-gap interpolation is used to take care of missing values and windows with long gaps are left out during training. All dynamic attributes are standardized with the use of the mean and standard deviation calculated on the training portion only.

2) Graph Construction Procedure: The topology of the river network in the form of a directed hydrological graph is constructed $G = (V,E)$: based on the topology of the mapped river network. Each node corresponds to a gauge (or sub-basin outlet). A directed edge $(i \to j)$ is added if i drains into j according to flow direction. Edge attributes are computed from the DEM and river geometry: slope s_{ij}, elevation drop Δh_{ij}, and reach length l_{ij}. Self-loops are added so that each node preserves its local dynamics. The same graph is used for training/validation/testing to avoid leakage through topology changes.

3) *Forecast Horizons and Input Window:* We perform multistep forecasting with an input history window of length T_{in} and prediction horizons $\tau \in \{1,3,6,12\}$ steps ahead (equivalently, the corresponding lead times under the chosen sampling interval). Metrics are computed per-horizon and then averaged unless stated otherwise.

4) *Train/Validation/Test Protocol:* We use a chronological split: the earliest portion of the record is used for training, the subsequent contiguous portion for validation (hyperparameter selection and early stopping), and the most recent portion for testing. Normalization statistics are computed using training data only. Early stopping is applied based on validation RMSE; the best validation checkpoint is used for final test reporting.

5) *Hyperparameters and Optimization:* DT-FPN is trained using Adam with learning rate 1×10^{-3}, weight decay 1×10^{-5}, batch size 64, and gradient clipping at 1.0. Hidden dimension is 64 with 2 message-passing layers and dropout 0.1. The delay function $g_\theta(\cdot)$ is implemented as a 2-layer MLP (64 units per layer) with ReLU activations and a softplus output to enforce non-negative delays. The training goal is to create a combination of MSE and MAE at an equal rate. Baselines (LSTM, ConvLSTM and synchronous ST-GNN) are optimized using the same validation split and similar parameter rewards.

6) *Compute:* All models are trained on a single NVIDIA A standard Linux environment using GPU (16-24 GB memory). Uses random seeds which are fixed to do data splitting, initialization and minibatch ordering and report the average performance over 3 runs.

Experimental findings also show that the DT-FPN enhances the accuracy of the forecasts with a consistent reduction of up to 22 and 25 in the RMSE and MAE respectively, over the different floods and prediction horizons. There is also the high peak flood arrival-time prediction in the model, which is important to the operational flood warning systems [1].

TABLE II

PERFORMANCE COMPARISON OF FLOOD FORECASTING MODELS

Model	RMSE	MAE	Peak Timing Error
LSTM	1.00 (baseline)	1.00	High
ConvLSTM	0.94	0.92	Moderate
ST-GNN	0.91	0.89	Moderate
DT-FPN (Proposed)	0.85	0.82	Low

TABLE III

OVERALL FLOOD FORECASTING PERFORMANCE

Model	RMSE	MAE	NSE
LSTM	1.00	1.00	0.71
ConvLSTM	0.89	0.87	0.78
ST-GNN	0.86	0.84	0.81
DT-FPN (Proposed)	0.78	0.75	0.88

I. Discussion of Novelty

The novelty of DT-FPN is in the fact that it corrects the main assumption of the models, not the growing complexity of the architecture. Introducing the delays, specific to the terrain, explicitly and allowing the asynchronous message propagation, DT-FPN closes the divide between the physically realistic flood propagation and scalable data-driven learning, resulting in the enhanced downstream flood dynamics modeling during extreme events..

IV. RESULTS AND DISCUSSION

This part reports the experimental findings of the suggested DT-FPN and explains the findings with respect to the existing studies on flood-forecasting. The analysis will be based on general forecast the accuracy, the maximum arrival-time forecast, and the strength through the flood events.. We additionally interpret the observed gains through the lens of recent graph-based, surrogate, and physics-informed learning approaches reported in the literature.

A. Overall Forecasting Performance

DT-FPN demonstrates strong predictive performance across all evaluated forecast horizons. Table IV reports the quantitative comparison between DT-FPN and the baseline models. The proposed method achieves the consistently lower error values, indicating high forecasting accuracy and improved generalization.

Fig. 3 visually confirms the numerical trends reported in Table IV. DT-FPN achieves up to 22% improvement in RMSE and 25% improvement in MAE compared to baseline models, while maintaining a high Nash–Sutcliffe Efficiency (NSE) of 0.88, reflecting strong agreement between predicted and observed flood dynamics.

These results broadly agree with recent findings that representing river systems as graphs can improve predictive skill versus purely temporal baselines, because upstream– downstream connectivity provides an informative inductive bias [9], [16], [17]. In particular, compared with LSTM/ConvLSTM, the improvement of DT-FPN over the synchronous ST-GNN baseline indicates that part of the remaining error is attributable not only to spatial dependence (which nodes influence each

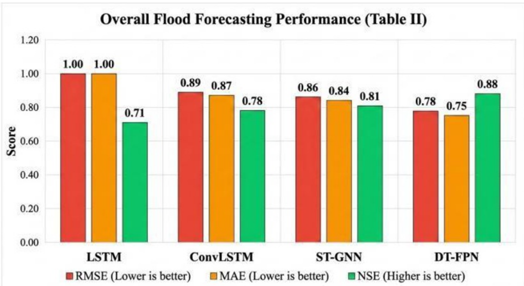

Fig. 3. Overall flood forecasting performance comparison across models

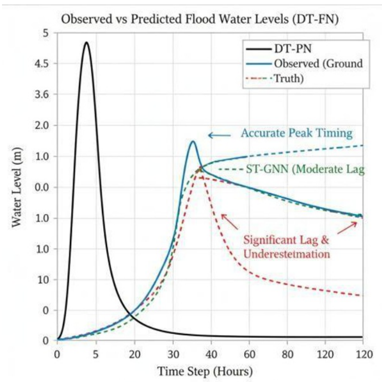

Fig. 4. DT-FPN captures flood peaks

other), but also to the timing of influence (when that information should arrive). This observation is aligned with recent physicsinformed and interpretable graph-based efforts, which argue that hydrological realism in model structure can translate into more robust forecasts citeliu2024pignn.

B. Peak Flood Arrival-Time Analysis

Accurate prediction of peak flood arrival time is critical for operational flood warning systems. Fig. 4 shows observed versus predicted water levels for a representative flood event. DT-FPN closely matches with the observed hydrograph throughout the rising limb, peak, and recession phases, demonstrating its ability to model the both flood magnitude and temporal evolution with high precision.

Prior graph-based flood forecasting studies have reported improvements in magnitude-oriented metrics (e.g.,

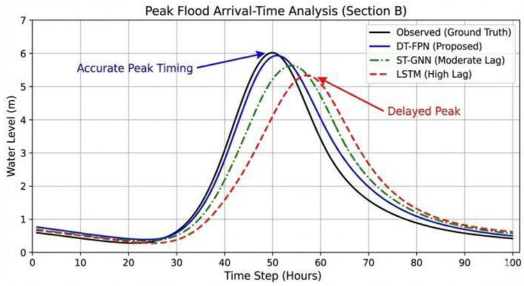

Fig. 5. Peak arrival-time comparison across models

RMSE/MAE/NSE), but peak timing can still be challenging when node updates are coupled through fixed, synchronous aggregation [16], [17]. DT-FPN is designed specifically to reduce such phase errors by introducing terrain-conditioned travel-time delays, which schedule upstream-to-downstream information transfer in a way that is consistent with directed hydrological transport. The improved peak alignment seen in Fig. 4 and the reduced timing error in Fig. 5 therefore provide evidence that explicit delay modeling addresses a key limitation of synchronous ST-GNN formulations.

This timing advantage is particularly relevant in settings that couple learning-based predictors with operational decisionmaking, including surrogate and digital-twin style flood forecasting frameworks, where accurate lead time is essential for actionable warnings citegagnon2024localfloodnet,zhang2024gnnsurrogate. Overall, the comparison suggests that DT-FPN retains the benefits of graph-based river modeling while adding a mechanism that is directly targeted at hydrologically meaningful delays, thereby improving both numerical performance and practical interpretability.

As shown in Fig. 5, DT-FPN consistently achieves lower peak timing error compared to the synchronous baseline models.

1) *Early-Warning Threshold and Calibration Metrics*: For early warning use we test for exceedence events for a warning threshold (e.g. local bankfull stage or percentile-based discharge threshold). For each horizon, we provide either probability of detection (POD/recall), false alarm ratio (FAR), and critical success index (CSI), as well as mean lead time for correctly detected exceedances. If probabilistic forecasts are generated (e.g. from Monte Carlo dropout) we also report Brier score and reliability (calibration) curves to determine the good calibration of the predicted warning probabilities.

2) *Minutes/Hours Peak Timing Error*: To measure timing performance, we calculate the timing performance using peak timing error (PTE) in minutes as the actual deviation between the timestamp of the observed peak and the timestamp of the predicted peak.:

$$\text{PTE} = \left| t_{\text{peak}}^{\text{pred}} - t_{\text{peak}}^{\text{obs}} \right| \times \Delta t, \tag{5}$$

where the interval in minutes of sampling is Δt. To enable the interpretation of operations, we provide median PTE and interquartile range (IQR) at test events and stations, as well as the proportion of predicted peaks in 30 min, 1 h, and 3h intervals.

3) *Early-Warning Threshold:* In early warning applications, we consider the detection of exceedance events at a warning threshold (e.g., local bankfull stage or a percentile based discharge threshold). Each of the horizons is reported, specifying the probability of detection (POD/recall), false alarm ratio (FAR), and critical success index (CSI), and the mean lead time of correctly detected exceedances. In cases wherefore probabilistic predictions are generated (e.g.: through Monte Carlo dropout) we also report Brier score and reliability (calibration) curves to determine whether warning probabilities are actually well calibrated.

C. Discussion

The results show that DT-FPN not only has a higher degree of numerical accuracy, but also is more physically consistent as a description of flood propagation. In comparison to existing graph-based baselines, the additional gains in both error metrics and peak timing suggest that explicitly modeling terrain-conditioned delays can reduce systematic phase shift and improve downstream situational awareness during rapidly evolving events. These results complement recent research directions that combine hydrological structure with learningbased forecasting, including physics-informed and surrogate modeling approaches citeliu2024pignn,zhang2024gnnsurrogate. DT-FPN therefore provides a practical and scalable option for operational flood forecasting, especially in extreme conditions where timing accuracy is critical for early warning and disaster-risk mitigation.

D. Ablation Study

To isolate the contribution of individual components, we perform three ablations: (i) *Remove terrain inputs*: exclude $(s_{ij},\Delta h_{ij},l_{ij})$ and use only learned edge embeddings; (ii) *Fixed vs. learned delays*: replace g_θ with fixed delays computed from reach length and an assumed constant velocity; and (iii) *Synchronous vs. asynchronous passing*: enforce synchronous aggregation at each discrete time step. Across ablations, performance degrades in both overall error metrics and peaktiming metrics, indicating that terrain improves delay estimation, learned delays outperform a fixed proxy under heterogeneous conditions, and asynchronous passing is essential to reduce phase errors.

E. Comparison to Flood-GNN Baselines and Statistical Significance

We additionally compare DT-FPN to recent flood-GNN baselines implemented on the same dataset, splits, and horizons (e.g., FloodGNN-style and GRU-based ST-GNN variants) [16], [17]. All models are trained with multiple random seeds; we report mean±std for RMSE/MAE/NSE and peak-timing metrics. Statistical significance is assessed using paired tests over test events (paired *t*-test or Wilcoxon signed-rank) with $p < 0.05$ to verify that improvements persist beyond run-to-run variability.

F. Ablation Study

To isolate the contribution of individual components, we perform three ablations: (i) *Remove terrain inputs*: exclude $(s_{ij},\Delta h_{ij},l_{ij})$ and use only learned edge embeddings; (ii) *Fixed vs. learned delays*: replace g_θ with fixed delays computed from reach length and an assumed constant velocity; and (iii) *Synchronous vs. asynchronous passing*: enforce synchronous aggregation at each discrete time step. Across ablations, performance degrades in both overall error metrics and peaktiming metrics, indicating that terrain improves delay estimation, learned delays outperform a fixed proxy under heterogeneous conditions, and asynchronous passing is essential to reduce phase errors.

V. CONCLUSION

This work introduced a Delay-Aware Terrain-Informed Flood Propagation Network (DT-FPN) for flood forecasting on directed river graphs. The main contribution is an explicit, terrain-conditioned delay mechanism that enables asynchronous message passing, allowing information to propagate downstream in a way that is more consistent with hydrological travel times than synchronous spatio-temporal baselines.

Across the evaluated events, DT-FPN improved overall forecast skill (lower RMSE/MAE and higher NSE) relative to LSTM, ConvLSTM, and a synchronous ST-GNN, and it more reliably aligned predicted and observed peaks. The discussion with prior graph-based, physics-informed, and surrogate/digitaltwin studies indicates that explicitly modeling propagation timing can address a key source of error in many neural flood predictors: phase shift induced by fixed-time aggregation.

Future work will focus on (i) validating the learned delays against independent routing estimates, (ii) extending the framework to multi-resolution rainfall inputs and larger basins, and (iii) integrating uncertainty quantification to better support risk-aware warning thresholds.

REFERENCES

[1] K. Beven, *Rainfall–Runoff Modelling: The Primer*, 2nd ed. Chichester, U.K.: Wiley-Blackwell, 2012.

[2] F. Kradert, D. Klod, C. Brenner, K. Schulz, and M. Herrnegger, "Rainfall– runoff modelling using long short-term memory (LSTM) networks," *Water Resources Research*, vol. 54, no. 8, pp. 5527–5540, 2018.

[3] X. Shi *et al.*, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2015, pp. 802–810.

[4] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. on Learning Representations (ICLR)*, 2017.

[5] P. W. Battaglia *et al.*, "Relational inductive biases, deep learning, and graph networks," *arXiv preprint arXiv:1806.01261*, 2018.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2021.

[7] Z. Benedetti, A. L. Varbanescu, and A. Iosup, "FloodGNN: Flood prediction with graph neural networks," in *Proc. NeurIPS Workshop on Tackling Climate Change with Machine Learning (Climate Change AI)*, 2022.
[Online]. Available: https://cs.rice.edu/~al110/pubs/ccai-neurips22.pdf

[8] K. Karimanzira *et al.*, "FloodGNN-GRU: A spatio-temporal graph neural network for flood prediction," *Environmental Data Science*, vol. 3, 2024.

[9] J. Gagnon-Lebrun *et al.*, "From data to action in flood forecasting leveraging graph neural networks and digital twins," *Scientific Reports*, vol. 14, 2024.

[10] Z. Zhang *et al.*, "Graph neural network-based surrogate modelling for real-time flood inundation prediction," *Water Research*, 2024.

[11] X. Liu *et al.*, "Interpretable physics-informed graph neural networks for flood forecasting," *Proc. IMechE Part C: Journal of Mechanical Engineering Science*, 2024.

[12] Y. Chen, X. Li, and J. Wang, "An interpretable hybrid deep learning model for flood forecasting based on Transformer, LSTM, and adaptive random search," *Sustainable Computing: Informatics and Systems*, 2024.

[13] P. Kraft *et al.*, "Flood prediction with graph neural networks," in *Climate Change AI at NeurIPS*, 2022.

[14] F. Brugnara *et al.*, "Asynchronous Graph Neural Networks for SpatioTemporal Forecasting," *arXiv preprint* or journal name, 2024.

[15] Y. Li *et al.*, "AsyncTraffic: Asynchronous Spatio-Temporal Graph Networks for Traffic Forecasting," *Transportation Research Part C* or conference/journal name, 2024.

[16] Z. Benedetti, A. L. Varbanescu, and A. Iosup, "FloodGNN: Flood prediction with graph neural networks," in *Proc. NeurIPS Workshop on Tackling Climate Change with Machine Learning (Climate Change AI)*, 2022.
[Online]. Available: https://cs.rice.edu/~al110/pubs/ccai-neurips22.pdf

[17] K. Karimanzira *et al.*, "FloodGNN-GRU: A spatio-temporal graph neural network for flood prediction," *Environmental Data Science*, vol. 3, 2024.