
AI-Driven Temporal Super-Resolution for Flooding Prediction in Norfolk, Virginia

Chetan Kumar¹ Diana McSpadden^{1,2} Malachi Schram^{1,2} Heather Richter¹
Yidi Wang³ Binata Roy³ Jonathan L. Goodall³

¹Old Dominion University–Thomas Jefferson National Accelerator Facility
Joint Institute on Advanced Computing for Environmental Studies, Portsmouth, VA, USA

²Thomas Jefferson National Accelerator Facility, Newport News, VA, USA

³University of Virginia, Department of Civil and Environmental Engineering, Charlottesville, VA, USA
{ckumar, hrichter}@odu.edu, {dianam, schram}@jlab.org,
{yw5vq, br3xk, goodall}@virginia.edu

Abstract

Accurate and near real-time water depth predictions are essential for supporting transportation and emergency management decisions during flood events. However, traditional physics-based hydrodynamic simulations are computationally expensive and time-consuming, limiting their practicality for real-time response. To address this challenge, we employ the Fourier Neural Operator (FNO) to learn complex spatiotemporal patterns of urban flooding and enable temporal super resolution. Our approach leverages coarse-resolution water depth and rainfall data to predict high-frequency 15-minute resolution water depths. We experiment with varying the temporal resolution during training, from 15 minutes to 1 hour, while always generating predictions at a finer 15-minute temporal resolution during testing. The method is applied to five flooding events between 2017 and 2022 in Norfolk, Virginia, USA. Across different training scenarios, our model achieves an R-squared value higher than 0.79 on test data. These results demonstrate the effectiveness of FNO-based temporal super resolution for accurate and timely water depth predictions.

1 Introduction and Motivation

Coastal urban cities are particularly susceptible to flooding from heavy rainfall, storm surges, river overflows, and high tides. Timely and accurate prediction of flooding is crucial for mitigating these impacts, and enabling effective response measures [1, 2]. This study focuses on Norfolk, Virginia’s second most populous city, where land subsidence and low-lying topography exemplify vulnerable urban communities at risk from high tides and rainfall events [3, 4]. For example, by 2049, there may be over 200 annual nuisance flooding (NF) events in Norfolk [5], with critical transportation corridors flooding during every spring high tide [6]. NF refers to flooding caused by either pluvial, tidal flooding, or both rainfall and tidal surges, i.e., compound events [5]. Additional details of the study area are provided in Appendix A.

Transportation and emergency management decision support require accurate, near real-time water depth predictions with high spatiotemporal resolution. Traditional, physics-based models [7] have been used for flood prediction and hydrological simulation relying on physical laws to understand interactions between subsurface flow, surface water dynamics, and stormwater drainage systems. While robust, accurate, and with high spatiotemporal resolution, these models require significant computation resources and data processing, limiting their practicality for near real-time forecasting.

The hourly rainfall forecasts available from the National Oceanic and Atmospheric Administration [8] are useful for generating hourly flood inundation predictions, but transportation decision support requires higher-frequency predictions to optimize responses. However, the lack of fine-scale environmental inputs makes it challenging to predict water depths at such fine temporal resolutions. To address this challenge, we employ Fourier Neural Operator (FNO) [9], a method that effectively captures the complex spatial and temporal dynamics of flood events. FNO enables super-resolution capabilities in the spatial and temporal dimensions [10, 11, 12]. In this study, the focus is on temporal super resolution, where we conduct multiple training approaches using water depth and rainfall data with gradually varying temporal resolution from fine 15-minute to coarse hourly intervals. The model is trained using coarse-resolution data and learns the underlying spatiotemporal patterns to predict water depths at a 15-minute time resolution during test time.

2 Methodology

2.1 Problem Formulation

Consider $D \subset \mathbb{R}^d$ as the spatial domain. Let the input function $A(x, t)$ and the output function $U(x, t)$ belong to function spaces $\mathcal{X}$ and $\mathcal{Y}$, respectively. Given observations on a coarse temporal grid $\{t_j\}_{j=1}^N \subset [0, T]$ with spacing Δt , the task is to forecast the output at the next coarse step $t_{N+1} = T + \Delta t$, i.e., $\hat{U}(x, t_{N+1})$. In addition, temporal super-resolution is performed within the forecasted interval $[T, T + \Delta t]$ by predicting on a finer grid $\{\tau_k = T + k \delta t\}_{k=1}^K$ with resolution $\delta t = \Delta t/K$, yielding fine-scale outputs $\hat{U}(x, \tau_k)$ for $k = 1, \dots, K$.

2.2 Fourier Neural Operator

To solve the temporal super-resolution problem, we use a neural operator framework to learn mappings between function spaces. Specifically, we adopt the Fourier Neural Operator (FNO) [9], which leverages the Fourier domain for efficient global convolution and fast computation. In the FNO, we first project the input data into a higher-dimensional latent space. We then apply the Fourier transform to convert the data from the spatial domain into the frequency domain, where convolutional operations capture non-linear relationships and global patterns. After processing in the frequency domain, we use the inverse Fourier transform to map the predictions back to the spatial domain. We also incorporate linear layers alongside the Fourier layers to model linear, localized patterns that the Fourier-based operations may not fully capture. We combine the outputs from both the Fourier layers and the linear layers, and finally map the resulting data back to its original dimensionality. We provide further details about the framework in Appendix B.

2.3 Dataset

In this study, we examined five Nuisance Flooding (NF) events that occurred between 2017 and 2022. Extreme flooding events, i.e., hurricanes are excluded. Table 1 provides details on the dates and durations of these NF events, along with surge levels, daily precipitation, and associated scenarios. A compound scenario refers to flooding caused by the combination of rainfall and tidal events whereas high surge refers to high tidal surge.

We utilized TUFLOW (Two-dimensional Unsteady FLOW) [13], a hydrodynamic physics model for simulating surface water flow and subsurface flow through stormwater drainage pipes, and subsequently developed a surrogate machine learning model using water depth output from the TUFLOW model and the FNO to predict surface flooding. Using TUFLOW, 15-minute water depth

Table 1: Rainfall events and their duration in days.

Date (mm/dd/yyyy)	Rainfall (days)	Surge (m)	Daily Precipitation (in)	Scenario
08/29/2017	5	1.17	3.93	Compound
09/19/2020	5	1.13	3.60	Compound
11/13/2020	4	0.83	4.96	Heavy rain
01/03/2022	5	1.34	1.68	High Surge
09/30/2022	4	1.25	3.40	High surge

predictions were generated at a spatial resolution of $2.5\text{m} \times 2.5\text{m}$ by discretizing the spatial area into a grid of 3476×3775 cells, resulting in a total of 13,121,900 cells. Spatially interpolated 15-minute rainfall data was also obtained from TUFLOW at a spatial resolution of $10\text{m} \times 10\text{m}$. In addition, we used the U.S. Geological Survey (USGS) digital elevation model (DEM) with a resolution of $2.5\text{m} \times 2.5\text{m}$ [14]. Appendix C provides additional detail about the physics-informed hydrodynamic model.

2.4 Study Area

In this work, we consider a small subset of the fully hydrodynamically simulated area, consisting of 500×500 grid cells, covering a total area of 1.6 km^2 , located in the Hague community of Norfolk, Virginia. The Hague neighborhood, located inland near the Elizabeth River, is one of the most frequently flooded areas in the city and regularly experiences compound flooding from rain and storm tides [15, 16]. For each event, we select a 15-hour period during when flooding was mostly observed.

3 Experiment

3.1 Data Preparation and Model Setup

The input to FNO model includes the TUFLOW simulated water depth, spatially interpolated rainfall, and the USGS DEM. All modalities were resampled to a common spatial resolution of $10\text{m} \times 10\text{m}$ using an averaging-based down sampling approach resulting in a grid size of 125×125 . We divide the 125×125 grid cells into patches of size 50×50 using a sliding window of 25 in both direction, resulting in a total of 16 patches for each flooding event.

We use four events for training and validation, reserving the September 30, 2022 event to test the model’s ability to predict 15-minute water depths for an unseen flooding event. Table 1 lists these five events. For training and validation, we employ k-fold cross-validation [17], where one event is used for validation and the other three for training. Data normalization uses a min–max scaler [18] fitted on the training set. The model is trained with a 120-minute look-back window and predicts a 60-minute look-ahead during inference, with each interval corresponding to 15 minutes.

This work explores five different training approaches. The first approach uses fine temporal resolution data, with both the model input and output during training and testing consisting of 15-minute interval data. This method serves as the benchmark, referred to as the high-resolution method. The next four approaches implement Temporal Super-Resolution (TSR) techniques. In the first TSR method, the model input during both training and testing is at a coarse resolution, using hourly data sampled from the 4th and 8th timesteps, representing the first and second hours. The model is trained to output fine-resolution predictions at 15-minute intervals. This approach is called target-aware. In the second TSR method, the model input during both training and testing is at a coarse resolution, using half-hourly data. The model is trained with half-hourly target data, but at test time, it predicts fine-resolution output at 15-minute intervals, performing interpolation along with the prediction. This method is referred to as the half-hourly method. The third TSR technique uses hourly data as input during both training and testing, with the first timestep provided as an initial condition. The model is trained to predict hourly data during training, but at test time, it predicts fine-resolution data at 15-minute intervals. This is called the hourly with first timestep method. Finally, the fourth method uses hourly resolution for both input and output during training, but at test time, the model predicts 15-minute interval data using hourly input. This method is called the hourly method. These approaches are illustrated in Appendix. D.

We employ an FNO architecture [9] comprising four 3D spectral convolution layers, each followed by GeLU activation layers. The 3D convolution is performed across both spatial and temporal dimensions, allowing the model to simultaneously capture spatial and temporal dependencies within the data. The FNO model leverages six spatial and temporal Fourier modes, chosen empirically. Training is conducted using PyTorch [19] with a batch size of 16, and optimization is performed using the AdamW [20] optimizer with a learning rate of $1\text{e-}5$. Reduce on plateau scheduler with early stopping is employed for optimization. We use relative L_2 error for model training and further assess our model performance using multiple metrics including Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and R-squared (R^2). In addition, we compute precision, recall, and F1 score for water depth categories $< 0.1\text{ m}$, $0.1\text{-}0.3\text{ m}$, and $> 0.3\text{ m}$, based on safety and road closure thresholds [21, 22, 23, 24, 25, 26]. The training is conducted on an NVIDIA H100 GPU.

3.2 Results

Table 2 presents the average test performance across k-fold models for different training data resolution approaches. Both the high-resolution and target-aware models, trained with fine-resolution target data, achieve similar accuracy with an R^2 of 0.97 and the lowest MAE and RMSE values. When trained with half-hourly data, the model attains an R^2 of 0.93, which is slightly lower than the fine-resolution models but still indicates strong predictive capability. Performance decreases further as the training data becomes coarser. Training with hourly data results in an R^2 of 0.79; however, including the initial time step as an additional input in the hourly training data improves performance to an R^2 of 0.90. These results demonstrate that the proposed FNO-based approach maintains high predictive accuracy for temporal super-resolution, even when trained with coarser temporal data.

Figure 1 shows the average F1 score across k-fold models for the test data in three water depth categories under different training approaches. Error bars represent the standard deviation across folds. For depths below 0.1m, all approaches achieve consistently high performance with F1 scores greater than 0.95. In the 0.1–0.3m category, the high-resolution and target-aware models maintain F1 scores around 0.90, with half hourly and hourly with 1st timestep also achieving values above 0.75. For depths greater than 0.3m, all training approaches except hourly achieve F1 scores above 0.80. A decline in performance is observed for both the 0.1–0.3m and > 0.3 m categories when using the hourly approach, reflecting the impact of coarser temporal resolution during training. Overall, these results demonstrate that the proposed approach retains strong predictive capability for deeper flooding levels even with reduced temporal resolution in training.

Additional categorical and qualitative results are shown in Appendices E, F, G and H.

Table 2: Average performance on the test data across k-fold models.

Model	MAE (m)	RMSE (m)	R^2
High-resolution	0.007	0.017	0.971
Target-aware	0.008	0.017	0.971
Half Hourly	0.015	0.026	0.930
Hourly + 1st timestep	0.018	0.031	0.905
Hourly	0.028	0.046	0.791

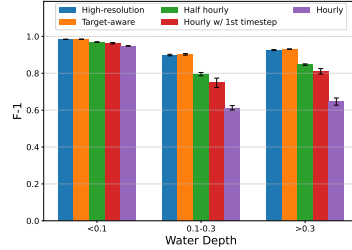


Figure 1: Average F1 score on test data.

4 Discussion

This study employs the FNO to predict water depth with temporal super-resolution. Using coarse resolution input data, the FNO model produces predictions at 15-minute intervals, greatly improving the temporal dimension of the prediction. This is particularly important for emergency management, where timely and accurate information is critical for decision-making. It is also especially valuable given that high-resolution rainfall forecasts at such short intervals are often unavailable.

While the FNO enhances prediction accuracy, this work is limited by the need to scale down the spatial resolution from 2.5m to 10m due to the unavailability of high-dimensional data across different modalities. Another limitation is forecast uncertainty, particularly when input conditions differ significantly from those encountered during model training. Potential future directions include leveraging high-dimensional datasets and improving the model’s robustness and accuracy under varied conditions, with an emphasis on uncertainty quantification.

Acknowledgments and Disclosure of Funding

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

References

- [1] Kun Xie, Kaan Ozbay, Yuan Zhu, and Hong Yang. Evacuation zone modeling under climate change: A data-driven method. *Journal of Infrastructure Systems*, 23(4):04017013, 2017.
- [2] Michael Pitt et al. Learning lessons from the 2007 floods, 2008.
- [3] William Sweet, Joseph Park, John Marra, Chris Zervas, and Stephen Gill. Sea level rise and nuisance flood frequency changes around the United States. 2014.
- [4] John D Boon, John M Brubaker, and David R Forrest. Chesapeake Bay land subsidence and sea level change: An evaluation of past and present trends and future outlook. 2010.
- [5] AG Burgos, BD Hamlington, Philip R Thompson, and Richard D Ray. Future nuisance flooding in Norfolk, VA, from astronomical tides and annual to decadal internal climate variability. *Geophysical Research Letters*, 45(22):12–432, 2018.
- [6] Larry P Atkinson, Tal Ezer, and Elizabeth Smith. Sea level rise and flooding risk in Virginia. *Sea Grant L. & Pol’y J.*, 5:3, 2012.
- [7] Mei Zhao and Harry H Hendon. Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model. *Quarterly Journal of the Royal Meteorological Society*, 135(639):337–352, 2009.
- [8] US Department of Commerce, National Oceanic and Atmospheric Administration. National Weather Service API. 2023.
- [9] Zongyi Li, Nikola Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. *arXiv preprint arXiv:2010.08895*, 2020.
- [10] Peishi Jiang, Zhao Yang, Jiali Wang, Chenfu Huang, Pengfei Xue, TC Chakraborty, Xingyuan Chen, and Yun Qian. Efficient Super-Resolution of Near-Surface Climate Modeling Using the Fourier Neural Operator. *Journal of Advances in Modeling Earth Systems*, 15(7):e2023MS003800, 2023.
- [11] Thorsten Kurth, Shashank Subramanian, Peter Harrington, Jaideep Pathak, Morteza Mardani, David Hall, Andrea Miele, Karthik Kashinath, and Anima Anandkumar. Fourcastnet: Accelerating global high-resolution weather forecasting using adaptive fourier neural operators. In *Proceedings of the platform for advanced scientific computing conference*, pages 1–11, 2023.
- [12] Yuanlong Zhang, Hanyou Zheng, Daiqin Yang, Zhenzhong Chen, Haichuan Ma, and Wenpeng Ding. Space-Time Video Super-resolution with Neural Operator, 2024.
- [13] BMT WBM. *TUFLOW User Manual*, 2016.
- [14] U.S. Geological Survey. TNM download [WWW document]. Natl. Map. <https://viewer.nationalmap.gov/basic/>, 2016. Accessed: 2019-10-01.
- [15] Jeffrey M Sadler, Jonathan L Goodall, Madhur Behl, Benjamin D Bowes, and Mohamed M Morsy. Exploring real-time control of stormwater systems for mitigating flood risk due to sea level rise. *Journal of Hydrology*, 583:124571, 2020.
- [16] Yawen Shen, Mohamed M Morsy, Chris Huxley, Navid Tahvildari, and Jonathan L Goodall. Flood risk assessment and increased resilience for coastal urban watersheds under the combined impact of storm tide and heavy rainfall. *Journal of Hydrology*, 579:124159, 2019.
- [17] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Montreal, Canada, 1995.
- [18] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python journal of machine learning research. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [19] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in Neural Information Processing Systems*, 32, 2019.
- [20] Ilya Loshchilov and Frank Hutter. Decoupled weight decay regularization. In *International Conference on Learning Representations*, 2019.

- [21] TD Shand, RJ Cox, MJ Blacka, and GP Smith. Australian Rainfall and Runoff Project 10: Appropriate Safety Criteria for Vehicles. 2011.
- [22] JL Gattis. *Guide for the Geometric Design of Driveways*, volume 659. Transportation Research Board, 2010.
- [23] AustRoads. Guide to Road Design Part 5: Drainage: General and Hydrology Considerations. Technical report, 2023.
- [24] Jie Yin, Dapeng Yu, Zhane Yin, Min Liu, and Qing He. Evaluating the impact and risk of pluvial flash flood on intra-urban road network: A case study in the city center of Shanghai, China. *Journal of Hydrology*, 537:138–145, 2016.
- [25] Faria T Zahura, Jonathan L Goodall, Jeffrey M Sadler, Yawen Shen, Mohamed M Morsy, and Madhur Behl. Training machine learning surrogate models from a high-fidelity physics-based model: Application for real-time street-scale flood prediction in an urban coastal community. *Water Resources Research*, 56(10):e2019WR027038, 2020.
- [26] VDOT. Virginia Department of Transportation Fact Sheet: Drivcing in the Rain, n.d.
- [27] D Fears. Built on sinking ground, Norfolk tries to hold back tide amid sea-level rise. *Washington Post*, 2012, June 7.
- [28] WJ Syme. TUFLOW-Two & One dimensional unsteady flow Software for rivers, estuaries and coastal waters. In *IEAust Water Panel Seminar and Workshop on 2d Flood Modelling*, Sydney, 2001.
- [29] Donald Shepard. A two-dimensional interpolation function for irregularly-spaced data. In *Proceedings of the 1968 23rd ACM National Conference*, pages 517–524, 1968.
- [30] US Department of Commerce, National Oceanic and Atmospheric Administration. North American Vertical Datum of 1988 (NAVD 88). 1988.

A The Norfolk, Virginia study area

We have selected Norfolk, Virginia as the study area to develop a flood prediction model using FNO. This city exemplifies vulnerable locations on the U.S. East Coast where real-time flooding predictions can provide practical benefits for various stakeholders, including community members, city officials, emergency management agencies, and transportation authorities. Norfolk’s unique combination of sea level rise, land subsidence, and extensive shoreline makes it particularly susceptible to flooding. The region has experienced a significant increase in relative sea level due to vertical land subsidence in the lower Chesapeake Bay and global sea level rise [3]. As one of the most vulnerable communities to coast flooding in the US, behind New Orleans [27], Norfolk faces a pressing risk. The city’s seven miles of beachfront, additional 137 miles of shoreline along rivers and lakes, its naval base, and working waterfronts on major waterways make it an ideal location for testing flood prediction models. The study area and gauges used to collect rainfall and tide measurements are shown in Fig. 2.

B Framework

Fig. 3 illustrates the framework of the proposed study. We combine water depth and rainfall data, both with an coarse temporal resolution, along with DEM data and utilize the FNO to achieve temporal super-resolution, enabling the prediction of water depth at 15 minutes intervals. Within the framework, Fourier and linear layers are applied simultaneously, and their combined output is processed through an activation layer, such as GeLU. The input data is initially passed through fully connected layers to lift the dimensionality of the channel space, and fully connected layers at the end project the output back to the specified dimension.

C TUFLOW: A physics-based hydrodynamic model

TUFLOW is a high-fidelity, physics-based hydrodynamic simulator used here to predict surface flooding with a spatial resolution of $2.5\text{m} \times 2.5\text{m}$. TUFLOW operates by solving the shallow water equations using two-dimensional (2D) numerical methods. Specifically, it employs ESTRY [28] to simulate 1-D flow within linear infrastructures such as pipes and channels, which is then integrated with 2-D numerical methods to simulate overland flow. The model takes into account various factors that influence the spread of water over land, including topography, land use, and surface roughness. By accounting for these factors, TUFLOW predicts flood extents and water depths with high accuracy and at the requested temporal resolution, for this work, 15 minutes.

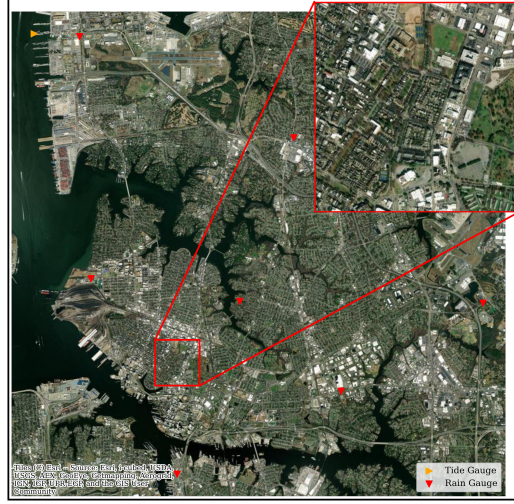


Figure 2: The map displays the Norfolk, Virginia study area and the locations of the gauges used to collect the event data used for the TUFLOW simulations. The locations of seven Hampton Roads Sanitation District rain gauges along with U.S. National Oceanic and Atmospheric Administration Sewell's Point tide station gauge are indicated. The selected study area within the Hague community is highlighted by the red box in the upper right.

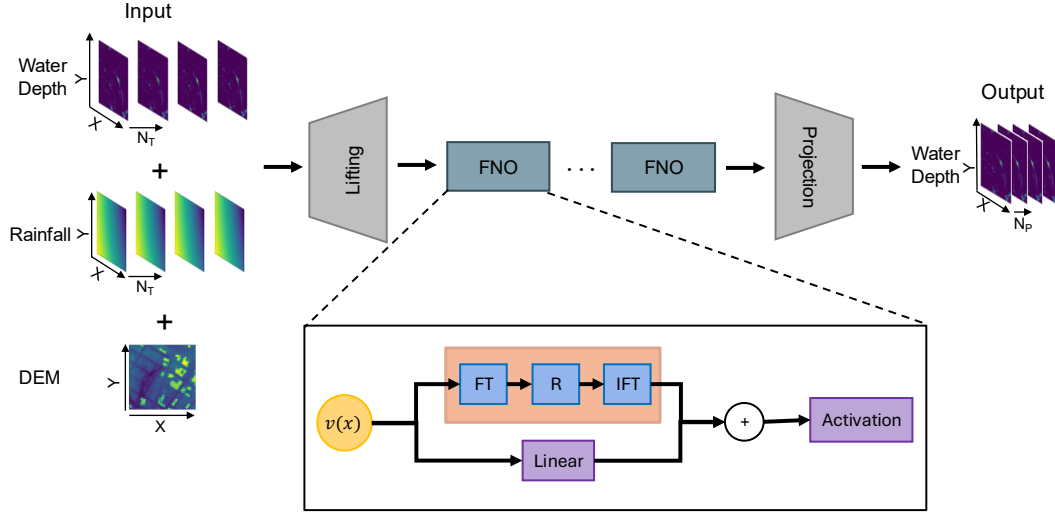


Figure 3: Framework of our work and architecture of FNO. X and Y denote the width and height of the patch, each consisting of 50 grid cells with a spatial resolution of 10m. FT and IFT represents Fourier and Inverse Fourier Transforms respectively. R denotes the linear transform emphasizing the lower Fourier modes while filtering the higher modes. N_T and N_P represent input and output temporal resolution respectively.

The TUFLOW model utilized a modified bare-earth digital elevation model (DEM) that incorporated building footprints and heights. To reduce the computational time and instability in these simulations, only buildings within the study area with an area greater than 500m² were considered. This exclusion criterion was applied to focus on larger structures that would have a significant impact on flood simulation results. Additionally, approximately 3% of the data, specifically Norfolk underpasses, were excluded from the simulations due to the lack of information regarding underpass pump stations during this time period.

The hydrodynamic model used environmental observations, including rainfall and tide level measurements, to generate predictions. Rainfall data was collected from seven gauges managed by the Hampton Roads Sanitation District (HRSD) with a 15-minute temporal resolution. TUFLOW carried out inverse-distance weighted [29]

interpolation to estimate rainfall for entire study area. Tide level measurements were obtained from the National Oceanic and Atmospheric Administration's (NOAA) Sewell's Point station, referenced to the North American Vertical Datum (NAVD 88) [30], with a temporal resolution of 6-minutes.

D Training Approaches

Fig. 4 illustrates the five temporal resolution approaches used to train the model, ranging from fine 15-minute resolution to coarse hourly resolution. The figure shows the timesteps considered for the input and target data during both training and testing, highlighting differences in how temporal information is provided to the model across approaches.

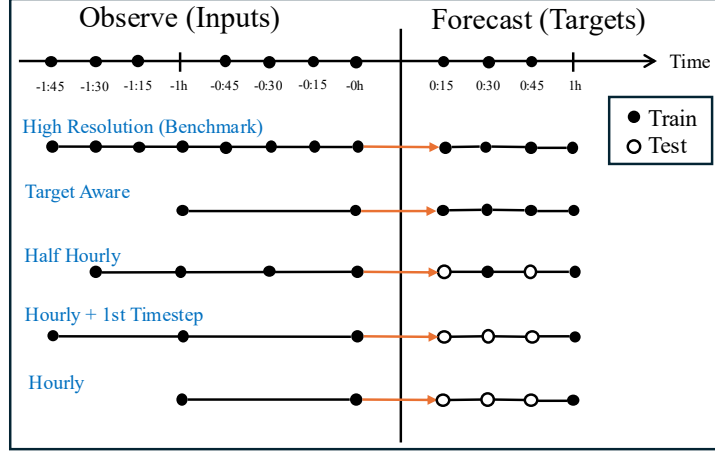


Figure 4: Training approaches with different temporal resolution settings. Solid circles indicate the timesteps considered in both the input and target data during training, while open circles indicate the timesteps considered in the target data during testing.

E Scatter Plot

Fig. 5 shows scatter plots of true versus predicted water depth for different training resolution approaches. The results are illustrated for one of the k-fold cross-validation experiment in which the validation event was September 19, 2020. Across approaches from high-resolution to hourly with 1st timestep, the density of points is aligned along the diagonal, which is the ideal case. For the hourly approach, more points appear below the diagonal, indicating underprediction; however, the majority of points still lie close to the diagonal. Similar performance is observed for different k-fold cross-validation training-validation splits.

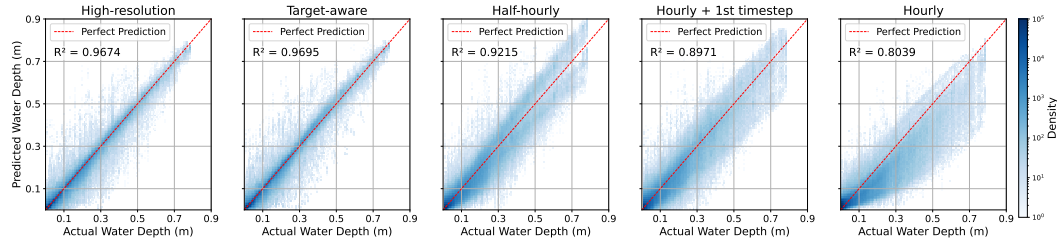


Figure 5: Scatter plots for true water depth vs. predicted water depth for different training resolution approaches.

F Categorical Results

Fig. 6a and 6b present the average precision and recall results across k-fold models for the test data. Precision for water depths below 0.1m and above 0.3m remains consistently higher than 0.95 across all training approaches, indicating the model's strong ability to correctly identify these depth categories. However, precision decreases

for the 0.1–0.3m category as the temporal resolution of the training data becomes coarser. Recall for water depths below 0.1m remains high across all training approaches, exceeding 0.98. In contrast, recall for the other two depth categories declines as the temporal resolution of the training data is reduced from fine to coarse.

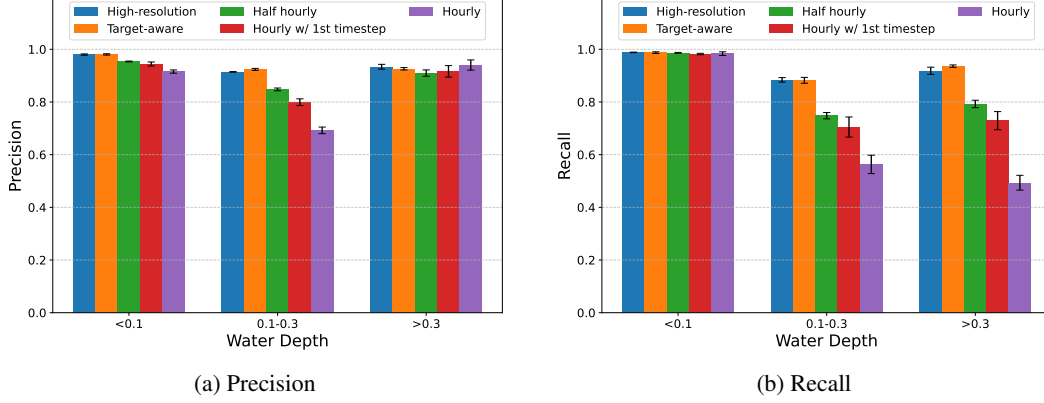


Figure 6: Average precision and recall score on test data across k-fold models for different water depth categories.

G High Water Depth

Fig. 7 shows the normalized density distribution of predicted water depths greater than 0.3m for different training approaches, compared with the ground truth distribution. The high-resolution and target-aware models closely match the ground truth. The half-hourly and hourly with 1st timestep approaches follow a similar overall pattern to the ground truth but with slight deviations in magnitude. The hourly approach shows a shift toward lower depths, indicating an underpredict deeper water levels.

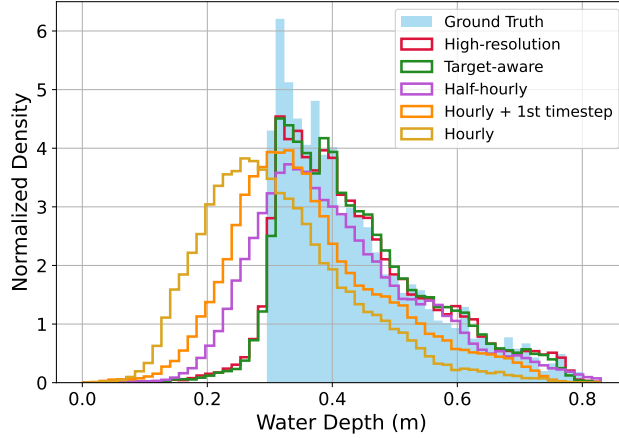
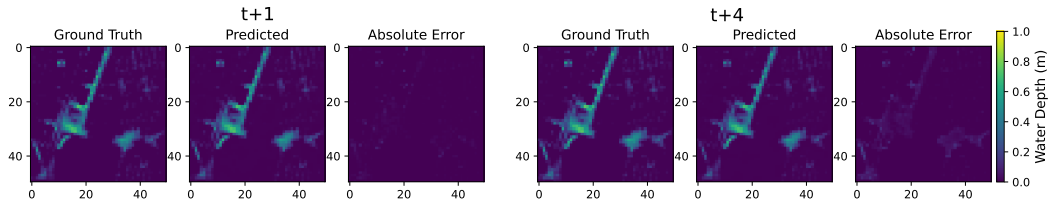


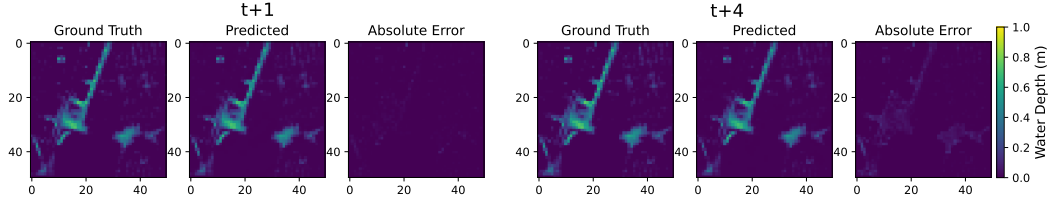
Figure 7: Test data distribution plot of water depth greater than 0.3m for different training approaches.

H Qualitative Results

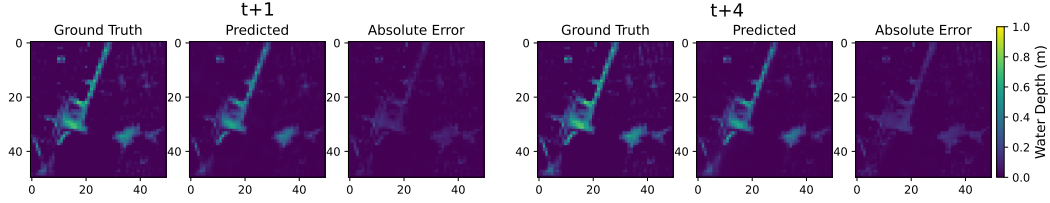
We present visual results of our model in Fig. 8 for different training approaches on one of the test patches. The figure displays the ground truth, predicted outputs, and the absolute differences between them for the first and fourth timesteps in the target data of the test set. Consistent results are observed across all training approaches and for both target timesteps, further validating the effectiveness of our approach.



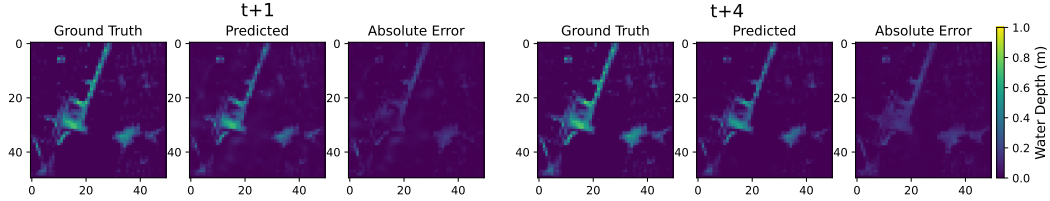
(a) High-resolution



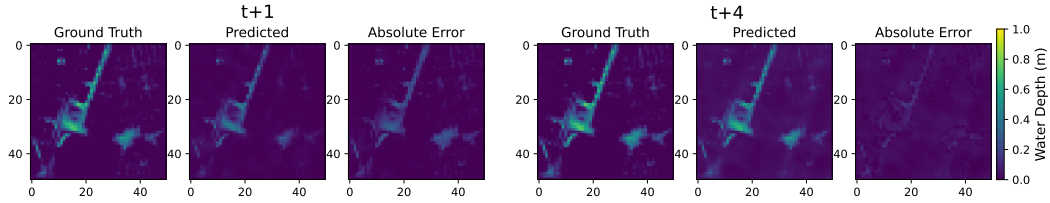
(b) Target-aware



(c) Half hourly



(d) Hourly with 1st timestep



(e) Hourly

Figure 8: Visual Performance with different approaches.