



OPEN

Unraveling the complexities of urban fluvial flood hydraulics through AI

Md Abdullah Al Mehedi^{1✉}, Virginia Smith¹, Hossein Hosseiny² & Xun Jiao³

As urbanization increases across the globe, urban flooding is an ever-pressing concern. Urban fluvial systems are highly complex, depending on a myriad of interacting variables. Numerous hydraulic models are available for analyzing urban flooding; however, meeting the demand of high spatial extension and finer discretization and solving the physics-based numerical equations are computationally expensive. Computational efforts increase drastically with an increase in model dimension and resolution, preventing current solutions from fully realizing the data revolution. In this research, we demonstrate the effectiveness of artificial intelligence (AI), in particular, machine learning (ML) methods including the emerging deep learning (DL) to quantify urban flooding considering the lower part of Darby Creek, PA, USA. Training datasets comprise multiple geographic and urban hydraulic features (e.g., coordinates, elevation, water depth, flooded locations, discharge, average slope, and the impervious area within the contributing region, downstream distance from stormwater outfalls and dams). ML Classifiers such as logistic regression (LR), decision tree (DT), support vector machine (SVM), and K-nearest neighbors (KNN) are used to identify the flooded locations. A Deep neural network (DNN)-based regression model is used to quantify the water depth. The values of the evaluation matrices indicate satisfactory performance both for the classifiers and DNN model (F-1 scores- 0.975, 0.991, 0.892, and 0.855 for binary classifiers; root mean squared error- 0.027 for DNN regression). In addition, the blocked K-folds Cross Validation (CV) of ML classifiers in detecting flooded locations showed satisfactory performance with the average accuracy of 0.899, which validates the models to generalize to the unseen area. This approach is a significant step towards resolving the complexities of urban fluvial flooding with a large multi-dimensional dataset in a highly computationally efficient manner.

Flooding is globally one of the world's most destructive types of disasters. In the coming years, floods are expected to be more frequent and larger globally¹⁻³. Coupled with the effects of rapid urban growth and climate change, the frequency of large fluvial flooding events is expected to increase, elevating the destructive impact of floods⁴⁻⁹. To untangle this challenge, engineers, planners, and emergency managers must be able to accurately anticipate flood extent and depth¹⁰. Alterations to the depth and time of occurrence of precipitation as a result of climate change are forecasted to reshape the flooding scenarios encountered in many areas in the world shifting flood risk¹¹. These risks are also driven in part by local stormwater management and fluvial infrastructure, making predicting flood events a particularly arduous and critical challenge in the built environments^{12,13}. With the increase in urbanization, more impervious areas are generated resulting in less infiltration and greater flood peaks and runoff¹⁴. Hydrological response time is largely reduced in an urban setting increasing fluvial flood risk where the amount of impervious surface area is high¹⁵. As a result, assessing flood risk in urban areas involves a complex interaction between natural and engineered processes, some of which operate at very local scales, requiring fine-resolution data¹⁶. Numerous investigations have sought to define the pattern of hydrological regime transformation resulting from urban development¹⁷. Alteration in urban river flow regimes is ascribed to the construction of impervious areas which facilitate rapid surface runoff from rainfall, the drainage of surface runoff through sewers to the river, and fluvial infrastructure¹⁸⁻²⁰. The proportion of urban land cover or the proportion of impervious cover within a catchment area provides predictor of changes in hydrograph characteristics, that lacks precision (e.g., the widely used Curve Number method)^{21,22}. Hydraulic models provide more precise results, demand expensive computational results and data.

¹Villanova Centre of Resilient Water System, Villanova University, Villanova, PA, USA. ²Department of Earth and Planetary Sciences, Washington University in St. Louis, St. Louis, MO, USA. ³Department of Electrical and Computer Engineering, Villanova University, Villanova, PA, USA. ✉email: mmehedi@villanova.edu

The complex and ever-changing urban landscape makes urban fluvial flood prediction and modeling computationally expensive and often infeasible due to high-resolution data requirements. Numerical analysis of hydraulic equations across spatio-temporal boundaries can be increasingly expensive depending on the resolution needed. In addition to the complex non-linear relationship among the features of urban fluvial flooding, estimating hydraulic parameters using physics-based equations is computationally expensive, as this process requires a large amount of memory allocation²³. Due to the computational expense, the range of input data is often limited, resulting in models at coarse resolutions, exclusive to different types of hydraulic conditions and potentially relevant parameters. Moreover, the uncertainty of the parameters, defective model calibration and errors in the measurements can serve as an accelerator to the computational expense. Answering the challenge of urban fluvial flood necessitates models that can efficiently and effectively represent flood extent with available data, in a quick and robust manner.

Data-driven prediction with Machine Learning (ML) techniques in the field of artificial intelligence (AI) provides a potential solution. ML is rapidly growing in popularity across many fields. ML methods, including the emerging deep learning (DL) methods, have been successfully applied to the field of water resources for stage-discharge (Q/h) relationships²⁴, rainfall-runoff²⁵, sediment transport^{26,27}, flood prediction²⁸, water quality analysis²⁹. AI models are specifically convenient when the uncertainty in model parameters, complexities in the physics-based equations and computational efforts are significantly high³⁰, such as in urban hydrology. Several previous studies have used ML models for urban fluvial flood detection^{31–33}. For example, a coupled physic-based model and random forest algorithm has been used to detect flood-prone areas in an urban coastal community^{34–37}. The deep convolutional neural network was used to forecast long-term water levels using rainfall intensity with slope and surface curvature³⁸. However, a convolutional network takes a long time and requires tedious hyperparameter optimization for the entire stochastic process, particularly when using large datasets^{39,40}. Neural network models are highly sensitive to the initial randomization of weights, number of layers, number of neurons, activation functions and algorithm to choose (e.g., gradients descent)^{41–43}. In the traditional ML and DL methods, a major challenge lies in developing models that can generalize to unseen case studies and sites⁴⁴. This investigation overcomes this obstacle by leveraging a two-stage approach with a set of ML classifiers and a DNN-based regression model used to predict the flooded extends and magnitudes with a comprehensive set of urban hydraulic features. In several previous studies, data points were randomly divided into a training/testing set for the ML models with satisfactory performance^{45–48}. However, due to the spatial autocorrelation effect, random sampling may not be adequate to validate the models to generalize to perform in the unseen area. Therefore, in this study, the models are tested considering the entire study domain as well as cross-validated spatially to minimize the spatial autocorrelation effect using the blocked K-folds Cross-Validation (CV) technique.

To predict the fluvial flooded locations, a set of classifiers e.g., Logistic Regression (LR), Support Vector Machine (SVM), K-nearest Neighbor (KNN) and Decision Tree (DT) are used. Binary classification generates the output in the form of binary data i.e., 0 s and 1 s, which represent whether a location is flooded or not in the study area. LR model has shown satisfactory performance in the previous study in classifying flooded locations using the matrix of the probability of detection on average for flood events^{49,50}. The DT algorithm showed good performance with Minimum Absolute Error (MAE) and classification accuracy in IoT (Internet of Things) based flood detection and notification system^{51,52}. The SVM attributes the non-linear transformation of geographic and hydraulic features in higher dimensional feature space^{53,54}. Highly satisfactory performance was achieved from the SVM algorithm in detecting the area prone to flood risk in the river basin of Buzau in Roman³⁷. Supervised regression with Deep Neural Network (DNN) is performed to predict the water depth within the model domain. Real-valued regression with an artificial neural network provides a reliable means of predicting the flooding depth with a good performance^{55–57}. In this investigation, a multilayer perceptron-based feed-forward neural network with a back propagation algorithm is used to perform the regression task i.e., predict water depth in the computational space. The DNN regression model is specifically suitable for predicting the flooded depth with representative geographic and hydraulic features⁵⁸. To increase the efficiency of the entire process of flood prediction, an efficient ML workflow plays a vital role by minimizing human involvement and increasing automation through coding^{59,60}. Several previous research works predicted flooding using various features e.g., elevation, slope, aspect, curvatures, topographic wetness index, and hourly rainfall^{61–64}. However, no investigation was performed to incorporate features that are closely related to the flooding in the urban environment e.g., fluvial infrastructure, impervious location within the contributing area. Many previous studies have shown their importance in runoff calculations. Urban streams and rivers are highly complex and particularly sensitive to urban land use and land cover areas^{65–69}, stormwater management^{65,70}, and the presence of fluvial infrastructure^{71,72} in addition to the geology and climate of the watershed. To represent the urban environment, we incorporated the effect of the impervious portion within the contributing area in the models. As a part of urban hydraulics, we introduced the downstream distance of the stormwater outfall and dams within the study area. All the variables under the topography, land covers and fluvial infrastructures linked to the urban flood dynamics were chosen based on an extensive literature search to accurately represent the urban hydrologic environment, without double counting variables^{73–76}. This study delineates a novel data-driven strategy toward unraveling the complexities of the urban flooding environments using multiple AI approaches i.e., a set of binary classifiers to detect flooded locations and DNN regression to predict water depth with blocked K-folds Cross-Validation (CV). The approach incorporates the characteristics of the urban area by introducing urban hydraulic drivers (impervious locations within the contributing area) in training the models for prediction. The outcomes of this research have significant potential to advance the flood preparedness mechanism for urban areas vulnerable to riverine flash floods where devastation due to rapid accumulation of flood water is significantly high. The quick and flexible framework presented here is transferable and can be utilized to prepare large-scale flood maps in an inexpensive and efficient way in the cloud-computing platform across urban areas. The approach outlined in this study has the potential to efficiently predict urban fluvial flooding for a range of scenarios.

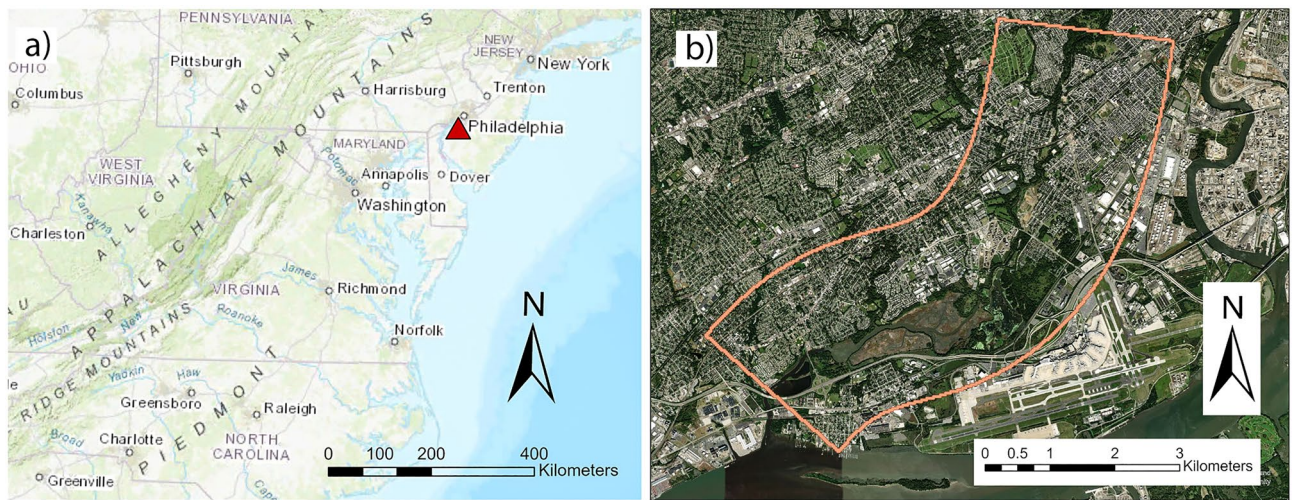


Figure 1. The figure above shows the global (a) and local (b) terrain of the study area. Maps are processed and generated in the ArcGIS Pro platform⁸¹.

Data and methods

Study area. The study area considered in this study is the lower part of Darby Creek (DC), along the southwest border of Philadelphia, PA, USA, shown in Fig. 1⁷⁷. The alluvial channel of the Creek flows through a floodplain with fully urbanized settings which is subject to frequent flooding. The population residing near the creek is subject to flooding significantly⁷⁸. The portion of the river considered in this study flows from the Mt. Moriah Cemetery (upstream) to the confluence with the Delaware River (downstream) and carries alluvial deposits through an urbanized setting⁷⁹, approximately 15 river kilometers (rkm). Darby Creek plays an important role in the adjacent environment and ecology; it is also a flood-prone area⁴⁸. It also offers a unique environment for various plant and animal species⁸⁰.

Preparing hydraulic dataset in iRIC. Hydraulic models are simulated in the iRIC platform to generate dataset for ML classifiers and DNN regression model. The iRIC is a numerical tool capable of modelling rainfall runoff generation, flooding, and sediment dynamics. It receives terrain and hydraulic data (e.g., water surface elevation, roughness) for the model calibration purpose. FaSTMECH (Flow and Sediment Transport with Morphological Evolution of Channel) is used as a solver in this study to model flood extent and depth⁸². Notably, this was a fluvial hydraulics model, and did not include a rainfall runoff simulation. Instead, model calibration used elevation data from the floodplain and bathymetry of the channel with water surface elevation at the upstream of USGS Cobb Creek gage at Mt. Moriah Cemetery (USGS gage 01475548) for the flooding event of 30th August of 2009 is utilized to calibrate the hydraulic model⁷⁷. The terrain data is discretized to a size of 5 m² for every computational cell. As the higher discharges from the upstream side of the river are responsible for the morphological changes, higher discharge values from the highest flood event in Darby Creek are chosen to create scenarios for AI models. Multiple scenarios are created using various constant discharge values upstream of DC within a certain range. The discharge data from observed flood events in the time span of 14th July to 16th September is obtained from USGS peak stream flow data (USGS gage 01475548)⁸³. A set of discharge values is chosen to execute ML/DL models used in this study. The discharge values are 37, 42, 45, 50, 52, 61, 83, 95, 99 and 164 m³ per second (cms). The outcomes generated by the iRIC are water surface elevation and flooding depth. A set of urban hydraulic features i.e., the quantity of the impervious areas within the contributing area, and downstream distance from the hydraulic structures e.g., stormwater outfall and dam are introduced in this study to integrate the effect of urban attributes with the flooding extent and magnitude. Furthermore, the average slope of the contributing area is derived through GIS analysis and incorporated to represent the flow accumulation to a specific location.

AI models. The quantification of the flood extent and depth by the ML framework is accomplished in three steps. Firstly, exploratory analysis and feature engineering are performed to study and transform the entire dataset prepared by multiple geographic and hydraulic features that impact the hydrograph, listed in Table 1. After analyzing the dataset and conducting necessary transformation on the features, classifiers, such as Logistic regression (LR), K-nearest neighbors (KNN), decision trees (DT), support vector machines (SVM), are trained using the data prepared in the first step to locate or classify the flooded locations for each scenario of various upstream discharges. Third, a DNN is used to prepare a regression model to predict the depth of water within the computational domain. ML classifiers and DNN models are evaluated using several error matrices, e.g., F1-score, Jaccard similarity score and Root Mean Square Error (RMSE). The algorithms are tuned and optimized by altering the hyperparameters to reduce the error and obtain satisfactory performance. The ML workflow of flood prediction is described in Fig. 2. The entire process can be divided into groups of tasks, i.e., data collection, exploratory data analysis, feature engineering, model training, model evaluation, model deployment, and model

Features	Full descriptions
x_1	x-coordinates of every location in the model domain
x_2	y-coordinates of the same location
x_3	Elevation in meters of same location
x_4/y_2	Depth of water in meters
x_5/y_1	Flooded locations
x_6	Average slope of the contributing area of every point in percentage
x_7	Number of impervious locations of the contributing area
x_8	Downstream distance from the stormwater outfalls
x_9	Downstream distance from the dams
x_{10}	Upstream river discharge in m^3/s

Table 1. Full descriptions of the predictors and target variables used to train/test the ML classifiers and DNN regression model.

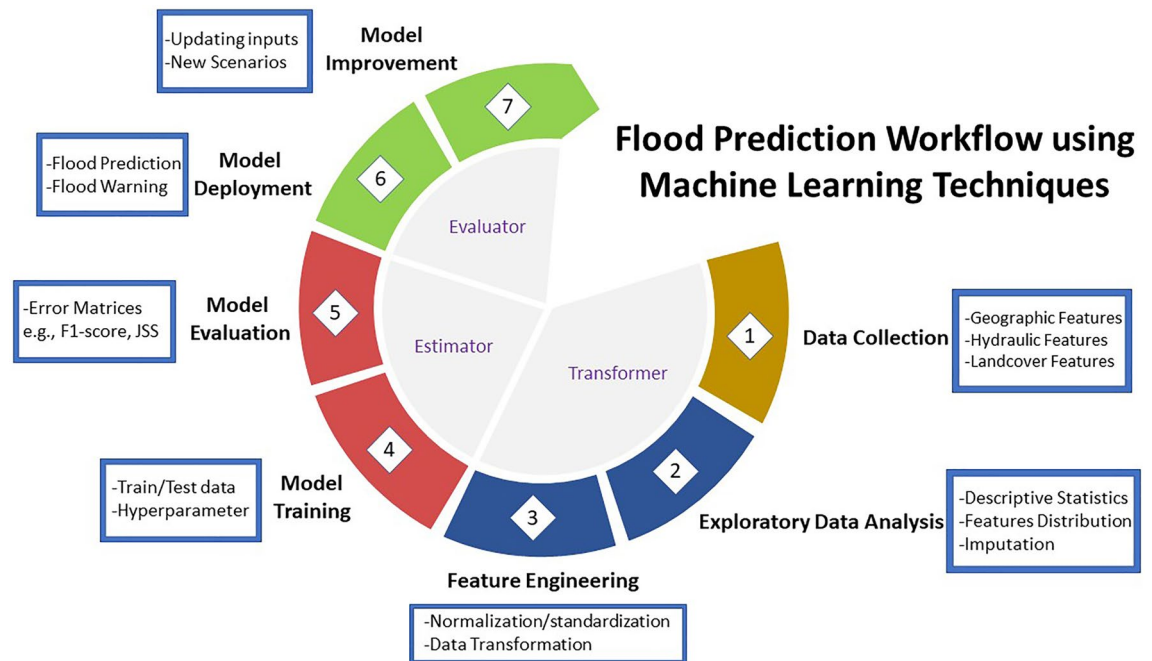


Figure 2. Flood Prediction workflow using machine learning classifiers and neural network regression techniques illustrates how the tasks are linked from the data preprocessing steps to the model deployment and maintenance stage.

improvement. Details are provided in the following sections. The steps are further categorized into distinct groups namely transformer, estimator, and evaluator.

Feature engineering. Within the first group of activities (transformer step), data collection, preprocessing, Exploratory Data Analysis (EDA) and feature engineering are performed (Fig. 2). EDA involves observing descriptive statistics and initial investigation of the variables (Table 1). Scikit-learn is used as the ML library for feature engineering in Python⁸⁴. It offers several classifications, regression and clustering algorithms including LR, KNN, DT and SVM which are used as binary classifiers for identifying flooding locations in this study. Modules needed for ML and Deep Learning algorithms such as optimization, linear algebra, integration, interpolation, and special functions can be accessed through SciPy⁴¹. Independent variables for Binary Classifiers and DNN regression model are listed in the Table 1. Flooded location is used as the target variable denoted by y_1 in case binary classifiers and water depth, y_2 as a target variable in case of DNN model. Spatial information, coordinates and elevation values are obtained from the original Digital Elevation Model (DEM) of the study area using ArcGIS Pro. Water Depth and Discharge values are extracted through simulating multiple hydraulic models in the iRIC platform. Average Slope and number of impervious cells of the contributing area of every point of the DEM are urban hydraulic features, which have not been introduced before as a training feature for AI models. ArcPy, a Python site package that offers an effective and efficient way to perform geographic data analysis, data conversion, data management, and map automation using Python was utilized to generate the contributing

areas of every cell upstream in the model domain⁸⁵. It can be compared with the upstream area contributing to those cells. No modification was needed to alter the data type, as it is generated from iRIC-FaSTMECH simply as binary data type. The main data frame is constructed through concatenating datasets derived from different upstream discharges (Q) scenarios.

Feature Engineering tasks used in this study to prepare the datasets for the ML/DL algorithms include numerical imputation, outlier detection with standard deviation and dropping, splitting training/testing datasets, and scaling with normalization. The proportion of the train-test split is assumed to be 80/20 for both ML classifiers and DNN regression. Dataset is divided into train/test split in a non-reshuffle manner where the datapoints are selected for training purpose without random sampling from original dataset to make the test dataset independent from the train dataset. In Eq. 1 how the normalization of the features performed can be observed. X denotes the feature vector including all the features used to train/test the models. Preparation of dataset for training the DNN is identical to the preparation of the training dataset for ML Classifiers. Eighty percent (80%) of the data is used to train, and the rest of the data is used to test both the ML classifiers and DNN regression model.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

Identifying flooded locations with ML classifiers. *Logistic regression (LR).* In the Activity 4 in Fig. 2, ML classifiers and DNN model are trained using the independent variables (Table 1) to predict the flooded locations and depth. Linear regression searches a function that builds relationships to a continuous dependent feature/variable, y , to some outcome/predictors (independent features x_1, x_2 , etc.). LR is a variation of linear regression, utilized when the existing dependent variable/outcome, y , is categorical. LG uses log loss as the loss/objective function in the classification algorithm. It generates a formula that forecasts the probability of the category as a function of the independent features. Logistic regression fits a special s-shaped curve (sigmoid function) by taking the linear regression and converting the numeric into a probability with the function, which is known as the sigmoid function σ ⁸⁶.

$$h_{\theta}(x) = \sigma(\theta^T X) = \frac{e^{(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots)}}{1 + e^{(\theta_0 + \theta_1 x_1 + \theta_2 x_2 + \dots)}} \quad (2)$$

The probability of a category 1 (a location being flooded) $(Y = 1|X) = \sigma(\theta^T X) = \frac{e^{(\theta^T X)}}{1 + e^{(\theta^T X)}}$. Therefore, LR passes the features (e.g., x_1 = elevation, x_2 = slope of the contributing area, x_3 = water depth, etc.) through the logistic/sigmoid functions; however, considers the outcome as a probability. The goal of LR algorithm is to identify the best parameters θ , for $h(x) = \sigma(\theta^T X)$, in such a way that the algorithm forecasts a cell is being flooded or not in the model domain.

Decision tree (DT). Decision tree learning is one of the predictive modelling approaches used in statistics, data mining and machine learning. It uses a decision tree (as a predictive model) to go from observations about an item e.g., features mentioned in the Table 1 (represented in the branches) to conclusions about the item's target value, e.g., binary decision on a location being flooded or not (represented in the leaves)⁸⁷. From Scikit-Learn, Decision Tree Classifier is used to perform the classification task on flooding location. Gini Impurity is used as a loss function of the DT classifier⁴⁴.

Support vector machine (SVM). SVM works by mapping data to a high-dimensional feature space so that data points can be categorized, even when the data are not otherwise linearly separable. A separator between the categories is found, then the data is transformed in such a way that the separator could be drawn as a hyperplane. Following this, characteristics of new data can be used to predict the group to which a new record should belong. Like the LG classifier, SVM uses the logistic loss function with a piecewise linearization⁸⁸.

K-Nearest Neighbors (KNN). The principle of KNN is based on the concept that the k closest objects or similar cases in the p -dimensional space (the number of dimensions is identical to the number of the features mentioned in the Table 1) determine the class of an unknown variable i.e., flooded locations. KNN aims to partition n observations (number of rows in the flood prediction data frame) into k clusters tagging each observation (rows in the data frame) to a specific cluster with the cluster centers or cluster centroid or the nearest mean serving as a prototype of the cluster. The entire data space is partitioned into Voronoi cells in this approach. As the target variable (flood locations) is predicted by local interpolation of the target associated with the nearest neighbors in the training dataset with the independent variables, no specific loss function is used in the KNN classifier⁴⁶. When features are obtained in different physical units with vastly varying scale, normalizing the training features and outcomes can improve the accuracy of the KNN algorithm as it depends on distance of the data points for the classification⁴⁷.

Predicting flood depth with DNN model. After predicting the flooded locations, DNN regression model is used to predict the water depth (y_2). To do this, a full set of multiple hydraulic variables/features mentioned in Table 1 and flooded locations from ML models are used to train/test the DNN model. Open-source library TensorFlow is used in this study work to construct DNN model as it has an excellent particular focus on the inference and training of DNN⁴⁸. Training a model with TensorFlow typically starts by defining the model architecture.

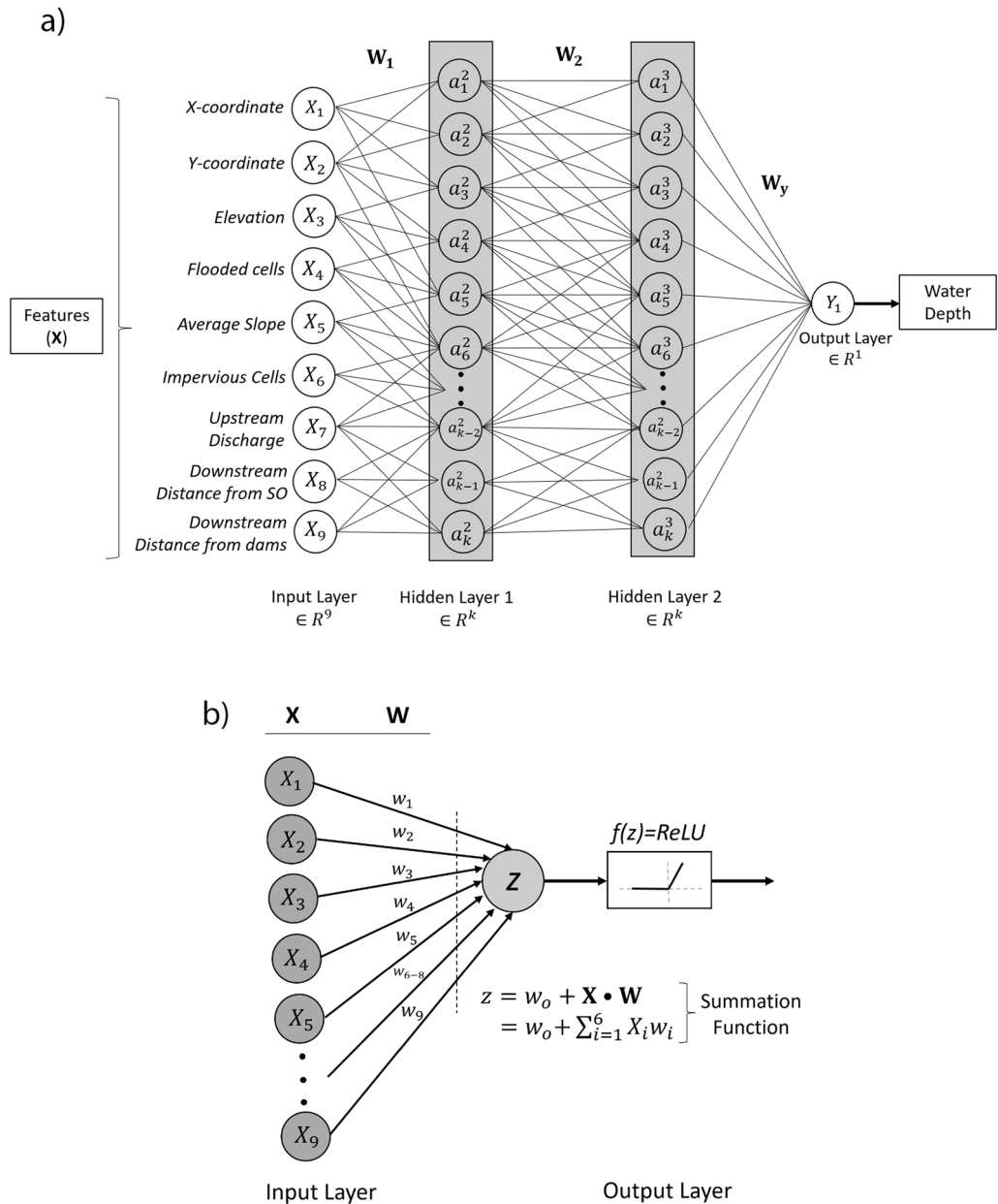


Figure 3. DNN architecture for flood prediction (a). A perceptron with summation of weights, features, and biases with the activation function (b). Input layers consist of all the input features and their corresponding weights. Aggregation of the features, weights and biases constitutes output layer.

Input layer contains features denoted by x_i in general, which is similar to the binary classification problem. The weights imposed on different features, aggregation of multiple features, further weights before the output layer and the activation functions are denoted with W , z , and f respectively. Finally, target variables (water depth) are generated from output layers. In Fig. 3 (a), introducing neural networks improves the prediction performance significantly through the introduction of non-linearity among the input and target features. The activation function used to introduce the non-linearity to the model is ReLU (rectified linear unit) function shown in Fig. 3 (b). This function returns the standard ReLU activation: maximum (X , 0), the element-wise maximum of input tensor (X) and 0 with default values. The total number of layers used to perform DNN is four including a normalized input feature layer, two hidden layers and a linear single-output layer. The total number of weights for each trainable neuron is 4609 where 11 neurons are found to be non-trainable.

Urban hydraulic feature importance is studied by analyzing the sensitivity of the change in feature values over the target variable, water depth and Permutation Feature Importance (PFI) technique in the computational domain^{89,90} (Fig. 4). In PFI, the impact of shuffling the values of a feature, e.g., impervious locations (x_m) within the contributing area over the target variable (\hat{y}_1) is quantified to observe the response in output variables due to the change in input variables. The score of the error matrix (RMSE) derived from the observed and predicted

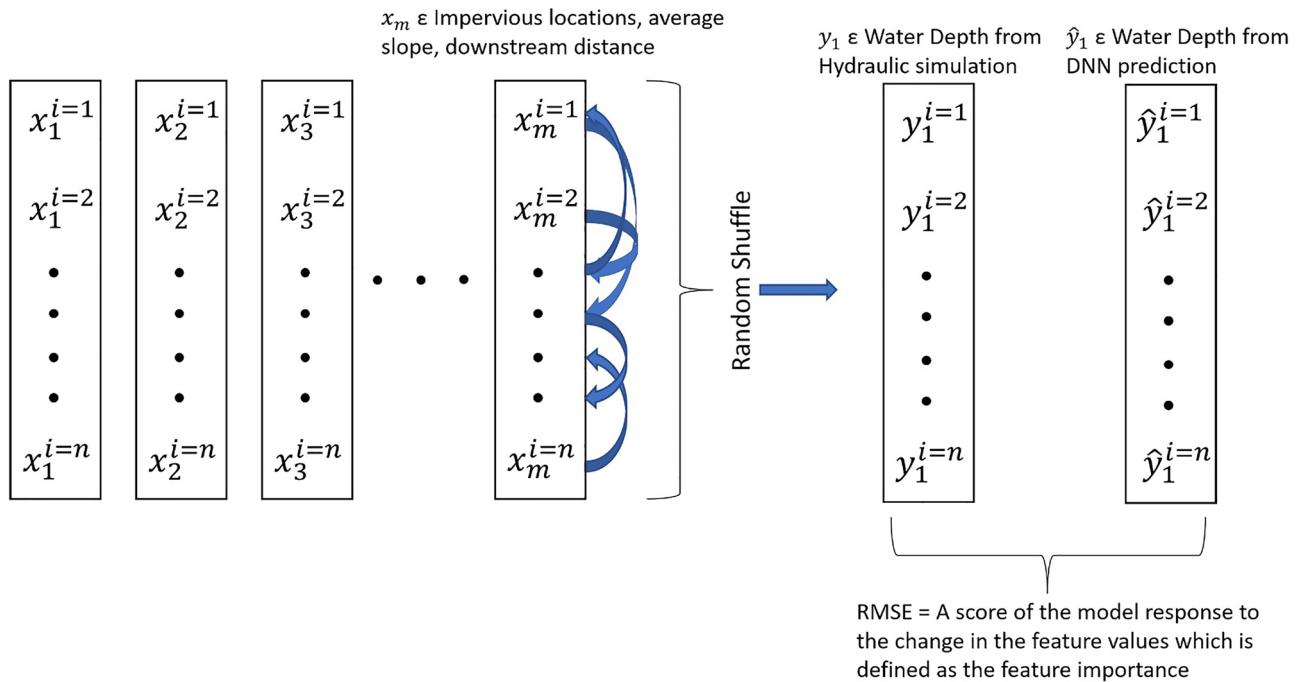


Figure 4. Mechanism of Permutation Feature Importance (PFI).

values of water depth as a result of the shuffle in the independent variable provides the score of feature importance. The values of impervious area, average slope of the contributing, downstream distance (DD) from the Stormwater Outfall (SO) and Dams (DO) are varied (5%, 10% and 20%) to observe the impact on the target variable in the DNN regression Model. The RMSE values are obtained from the difference between the series of the target variable, water depth after running the DNN model with the changed features and the series before running the model. In the PFI technique, DNN model is run with the values of a specific feature, e.g., impervious areas of the contributing area permuted/shuffled keeping the other features constant and the change in the RMSE values are recorded⁴¹. Only the output from DNN model i.e., water depth is used as a target variable in estimating the feature importance of other input variables as the output from ML classifiers are already used as an input variable in the DNN model.

Model evaluation for the study area. In the activity 5, model evaluation in Fig. 2, ML classifiers and DNN model are evaluated for the entire study area using multiple error matrices. Several conventional statistical measures are available to evaluate the performance of the ML classifiers. Mean Absolute Error (MAE), F1-score, True Positive (TP), False Negative (FN), Root Mean Square Error (RMSE) Jaccard similarity score, log-loss are among the popular choices and provide evaluation of the models in quantitative terms⁹¹. In this study, F1-score and Jaccard similarity score are used to evaluate the ML classifiers. The F1-score is used to evaluate binary classification algorithms, e.g., logistic regression, which generates binary outputs of whether a location is flooded or not. The harmonic mean of the model’s precision and recall is calculated to determine the F1-score⁹². The performance of the ML classifiers can also be determined from the confusion matrix, shown in Fig. 8. A confusion matrix is a table that is used to define the performance of a classification algorithm. A confusion matrix visualizes and summarizes the performance of a classification algorithm. It is used to visualize the performance of a classifier, typically a supervised classification algorithm⁹³. Two parameters needed to estimate the F1-score are precision and recall. Precision represents the fraction of the number of instances which the model correctly predicted (T_p) and the sum of all instances that are incorrectly predicted as true (F_p). Recall, sometimes referred to as sensitivity, is the fraction of the number of instances which the model correctly predicted (T_p) and the sum of all instances that are incorrectly predicted as false (F_n)⁹⁴.

$$2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} = \frac{T_p}{T_p + 0.5(F_p + F_n)} \tag{3}$$

The Jaccard coefficient quantifies similarity between finite sample sets and is determined as the size of the intersection divided by the value of the union of the sample sets. Given forecasted values of fluvial flooding occurrence as (\hat{y}) and actual values of flooding occurrence as y , the Jaccard index can be defined as

$$j(y, \hat{y}) = \frac{y \cap \hat{y}}{y \cup \hat{y}} \tag{4}$$

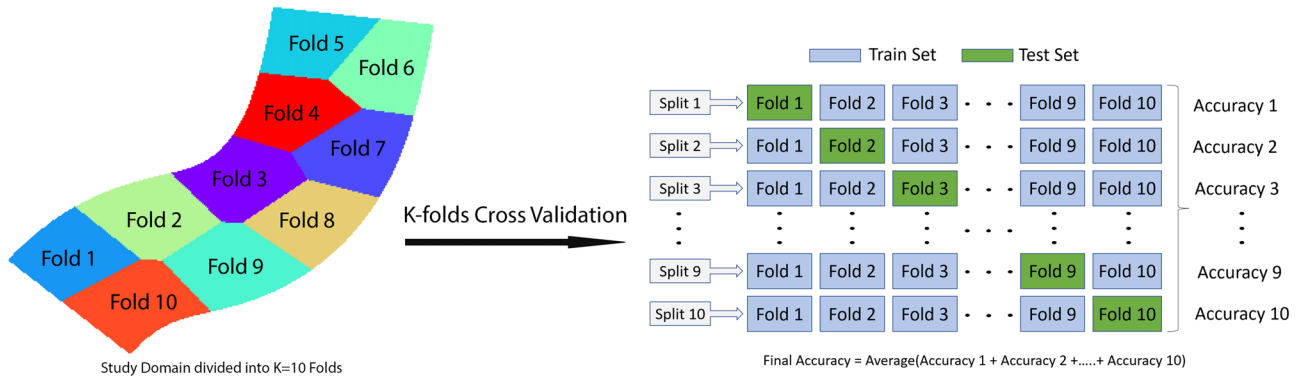


Figure 5. Blocked K-folds CV with a set of ten train/test splits.

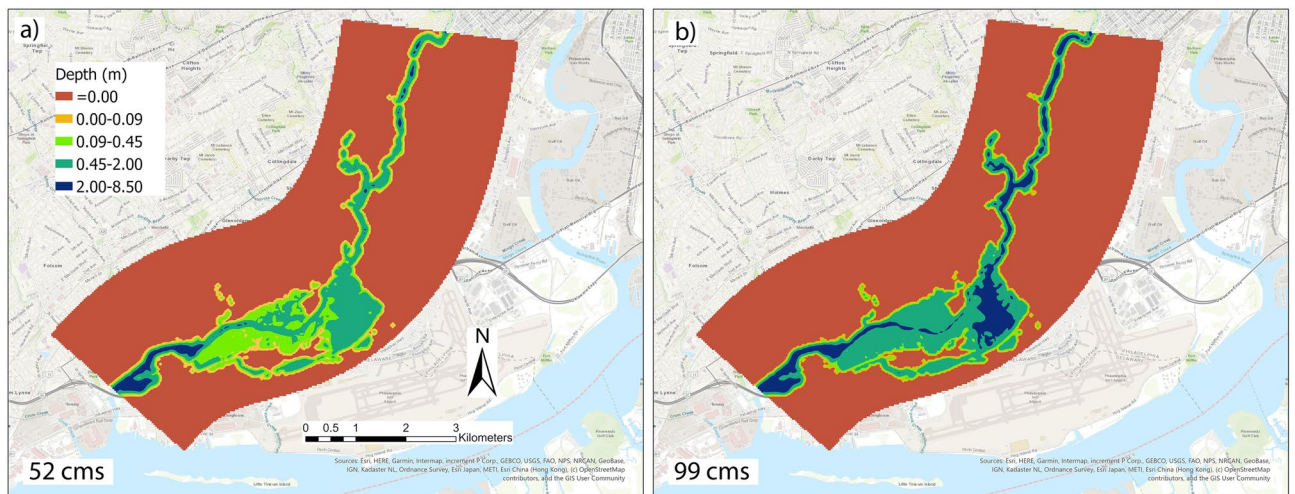


Figure 6. Flooding extent and depth estimated by the FaSTMECH solver in the iRIC platform for the scenarios with discharge values (a) 52 and (b) 99 cms. These scenarios are used to train/test ML classifiers and DNN regression algorithm. Maps are processed and generated in the ArcGIS Pro platform⁸¹.

Root Mean Squared Error (RMSE) value is used for the DNN regression model evaluation. Datapoints are selected randomly from the entire study area to prepare the train/test sets. The train/test split used in this study is 80/20 providing 80% of the total datapoints for train set and 20% for test.

Blocked K-folds cross-validation. Blocked (spatial) K-folds Cross-Validation (CV) is performed for the ML classifiers in classifying the flooded locations^{95,96}. The entire study domain is clustered into 10 folds (zones) as spatial autocorrelation among the nearby cells may lead to bias and wrong model evaluation if the models are evaluated considering the entire study area only^{97,98}. In the random sampling for train/test split, there is a possibility of taking the cells out in the study domain for the training set which are neighbors to the cells taken out for the test set. Consequently, those features in the train and test set are no longer independent, invalidating the evaluation of the ML classifiers. Therefore, the entire study domain is grouped into 10 folds to prepare individual train/test sets to compute the error matrix. Finally, the average of all the values of the error matrix is computed to show the model performance. An illustration of the entire process of the blocked K-folds CV is presented in Fig. 5. Ten different splits are introduced to isolate the test set with an independent fold. In each split the rest of the nine folds are used to train the model.

Results and discussion

Scenarios from hydraulic model. The relationship with the water depth in the computational space and river discharge is highly non-linear. High variation in the geometry of the channel and roughness of both channel and floodplain against the flow can enhance the non-linearity in the system. In this section, water depth variations with their corresponding locations are presented with respect to multiple scenarios with several upstream discharge values, including 37, 42, 45, 50, 52, 61, 83, 95, 99 and 164 m³ per second (cms). The validated hydraulic model from iRIC is used to simulate and create scenarios having the water surface elevation and depth, locations, and binary output regarding a certain location is flooded or not as results. In Fig. 6, two plots of water depth with the locations for the scenarios with discharges 52 and 99 cms are shown, which are obtained from hydraulic simulation in iRIC.

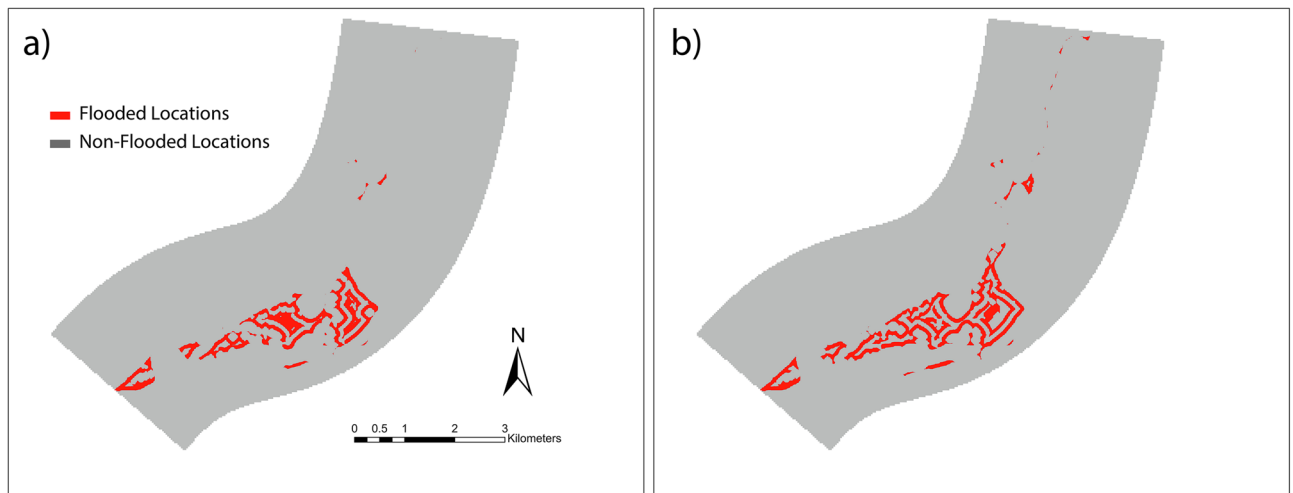


Figure 7. Flooded locations detected by the DT model for representative scenarios with upstream discharge values (a) 52 and (b) 99 cms. Prediction performance of the DT algorithm is the best among all other classifiers. Maps are processed and generated in the ArcGIS Pro platform⁸¹.

Binary classifiers	F1-score	Jaccard similarity score
Logistic Regression (LR)	0.975	0.995
Decision Tree (DT)	0.991	0.966
Support Vector Machine (SVM)	0.892	0.901
K-Nearest Neighbors (KNN)	0.855	0.810

Table 2. Comparison of the performances of ML Classifiers.

Binary flood map. The features used to train the ML classifiers (LR, DT, SVM and KNN) are $X = [x_1: x\text{-coordinate}, x_2: y\text{-coordinate}, x_3: \text{ground elevation}, x_4: \text{water depth}, x_5: \text{average slope}, x_7: \text{number of impervious locations}, x_8: \text{downstream distance from SO}, x_9: \text{downstream distance from dams}, x_{10}: \text{upstream discharge}]$. Locations are classified using their corresponding water depth into two classes, whether a particular location is flooded or not (1/0). Figure 7 illustrates the distribution of flooded location for the scenarios with the upstream discharge values of 52 and 99 cms predicted with the DT algorithm. DT is the top performer among all other ML classifiers with the F1-score and Jaccard Similarity matrix of 0.991 and 0.966.

All ML classifiers conveyed satisfactory performance in isolating flooded locations as the values of the error matrices in Table 2 are closer to unity. From Table 2, it can be observed that LR and DT outperform other classifiers. The F1-score and Jaccard similarity matrix of the LR and DT are 0.975, 0.991 and 0.995, 0.966 respectively which are greater than the values of the SVM and KNN.

Performance of the binary classifiers can also be illustrated in the form of confusion matrix, comparing correctly predicted outcomes with the incorrectly predicted outcomes. In Fig. 8, the confusion matrices showed to illustrate the performance of the ML classifiers. The number of flooded cells predicted correctly by LR algorithm (a) numbers 8332 (96.8%), while the incorrectly predicted cell count is 324 which is significantly lower than the number of correctly predicted locations. Similarly, the number of correctly predicted not-flooded cells count 54,747 (99%) where the number of incorrectly predicted not-flooded cells are 47. A total of 96.8% of predicted cells are correct, suggesting a highly accurate model performance.

The distribution of the individual model accuracy of the ML classifiers in the blocked K-folds CV is presented in Fig. 9 using boxplots. Median values of the model accuracies (F1-score) are presented with the red lines with the values 0.924 (LR), 0.975 (DT), 0.827 (SVM) and 0.871 (KNN). The range of the accuracy scores of the DT algorithm is 0.890–1.000 and is found to be the best ML classifier among all others. The median score of the DT model is 0.975 and is the highest score followed by the LG, KNN and SVM. Overall, LG, DT, KNN, and SVM showed satisfactory performance, as the range of the model accuracy of all classifiers is 0.582–1.000. As the dataset used for the predictive analysis is a tabular dataset with segmented values of the input variables, as well as a categorical-typed target variable (flooded/non-flooded) in this case, tree-based non-parametric algorithm i.e., DT algorithms, outperform other classifiers by capturing the interaction between different features. However, extensive hyperparameter tuning, intermediate feature creation and variation in the size of dataset might lead to alteration in model performance. However, extensive hyperparameter tuning, intermediate feature creation and variation in the size of dataset might lead to alteration in model performance.

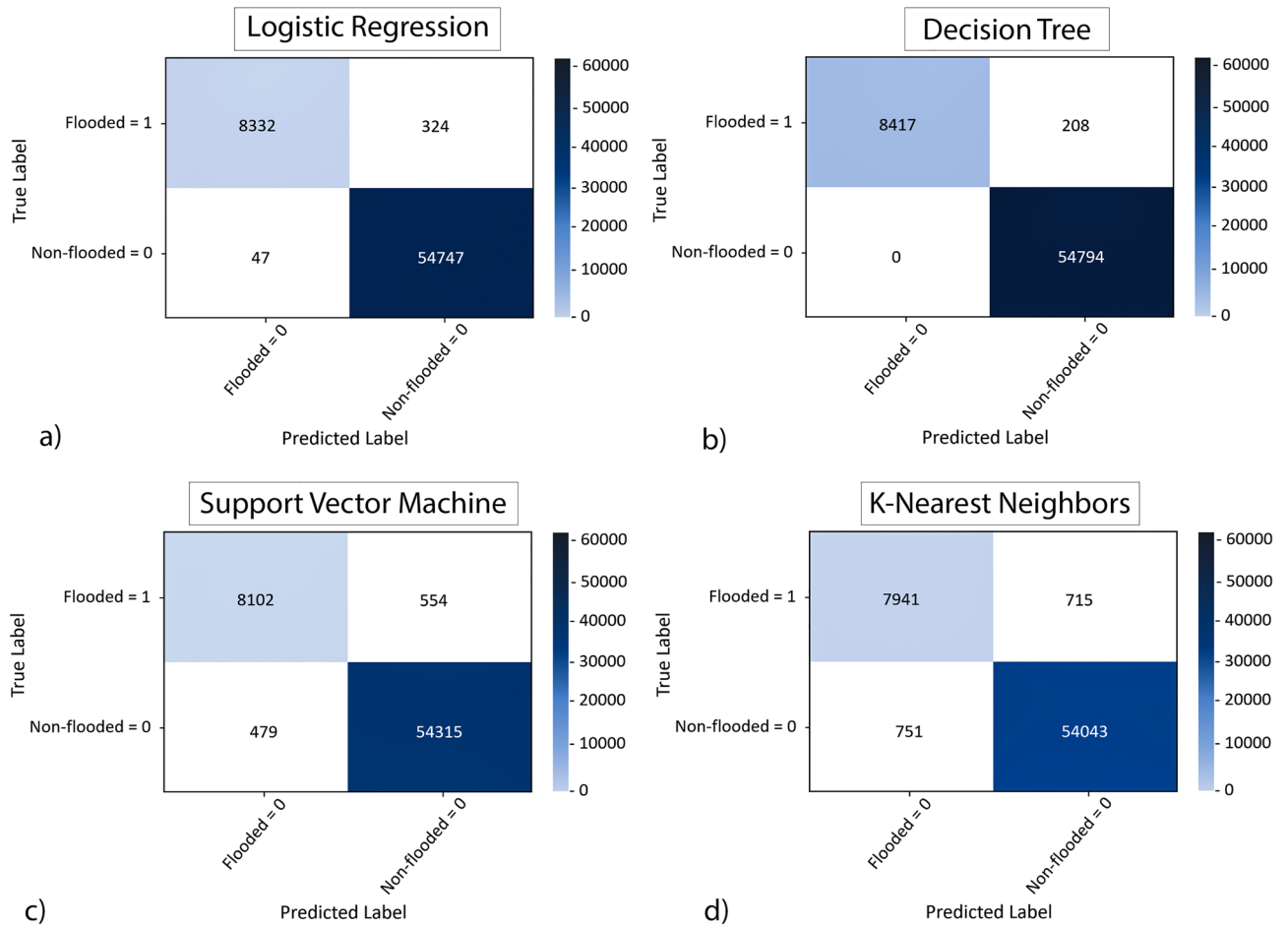


Figure 8. Confusion Matrix is used to observe the performance of ML classifiers i.e., logistic regression (a), decision tree (b), support vector machine (c) and K-nearest neighbors (d). Total number of predicted cells from ML classifiers are compared to the actual values from hydraulic model to extract the number of correctly predicted cells.

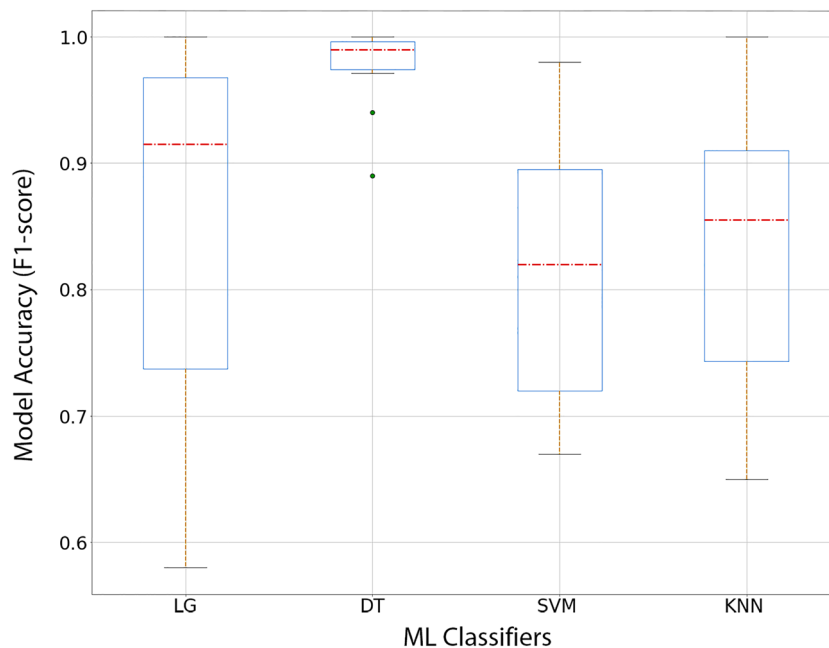


Figure 9. Distribution of the model accuracy (F1-score) in K-folds CV for the ML classifiers.

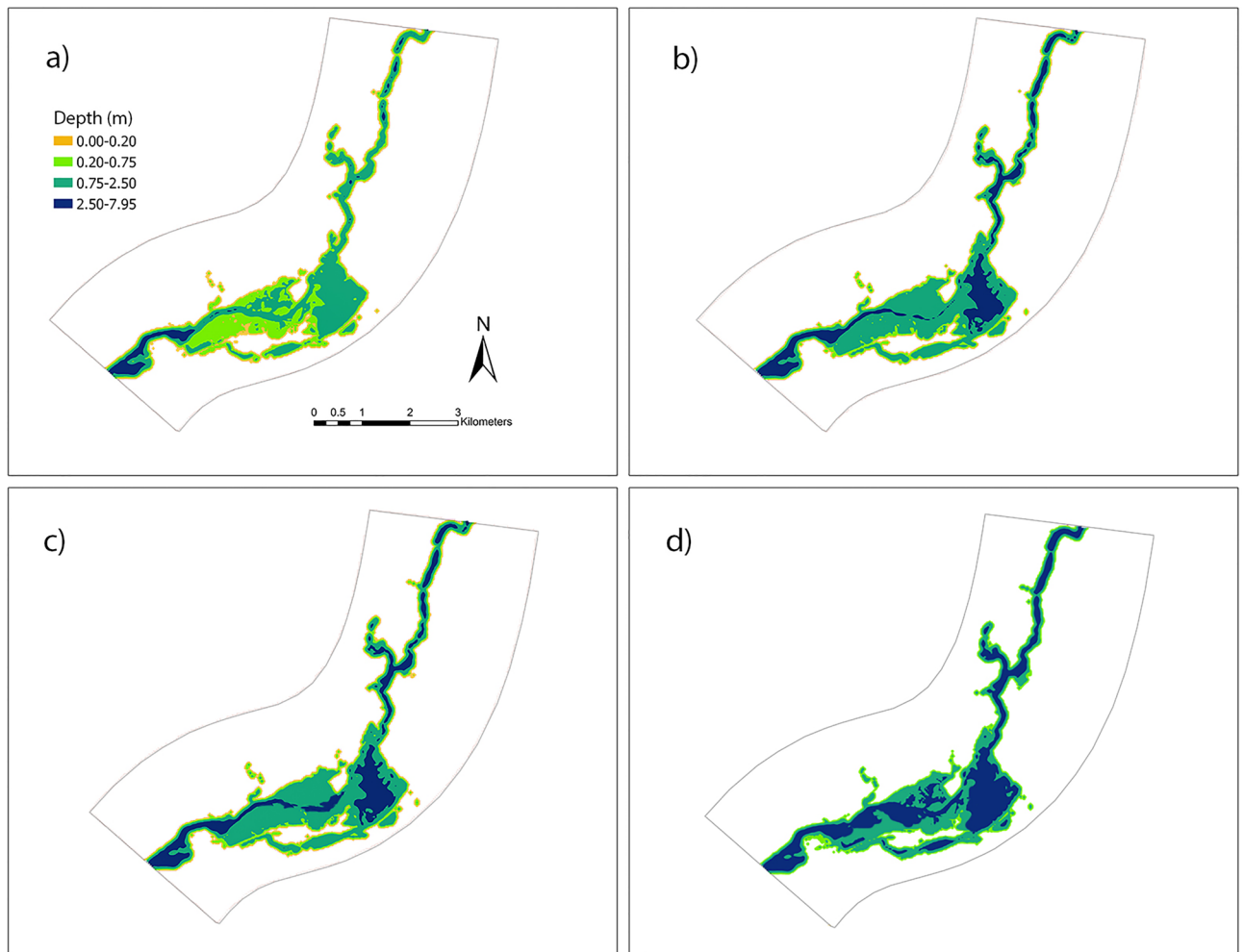


Figure 10. Predicted water depths for the representative scenarios covering the full range of upstream discharge values (a) 42, (b) 52, (c) 99 and (d) 164 cms outputs from the DNN regression model. Maps are processed and generated in the ArcGIS Pro platform⁸¹.

DNN regression to predict flood depth. Artificial neural network with a single hidden layer is not capable of extracting the insights of the non-linearity and complexity of the flood prediction. Therefore, DNN with multiple hidden layers is incorporated in the prediction process of water depth. Adding more hidden layers increases the accuracy of prediction. However, inclusion of a large number of hidden layers requires high computational power and may result in overfitting the model^{99,100}. From the model evaluation, it can be observed that the DNN described in this manuscript reflected the complexities of river flood prediction. To capture the high amount of non-linearity among the geographic and urban hydraulic features mentioned in this paper and establish linkage among them, it is a prerequisite to introduce multiple hidden layers. Hidden layers with the nodes built in them are used to train the model through an iterative optimization process. A total of three hidden layers are used with 64 neurons assigned to each. While the number of epochs found best with a minimum error is 110. 80 percent of the whole dataset was used for model training purposes, while 20 percent was used for testing the performance of the DNN. The activation function for hidden layers used is ReLU. Other popular activation functions such as hyperbolic tangent, sigmoidal or leaky ReLU functions are recommended to introduce the non-linearity in DNN. The model evaluation matrix, RMSE value of 0.027 illustrates the DNN regression algorithm conveyed satisfactory performance in resolving the high non-linearity in the flooding depth prediction process.

Urban hydraulic features, i.e., average slope and the number of impervious cells of the contributing area, are introduced in this process to train the DNN regression model. It is clear from Fig. 10 that the flood depth is highly correlated and sensitive to the upstream discharge. With the increase in the upstream discharge, flood depth also increases. In the DNN regression model training phase, a non-linear correlation is built among these urban hydraulic features and flooding amount and extent. By introducing these urban hydraulic features, connections among the features and target variables i.e., number of impervious locations and water depth are established.

As the output (flooded location) from the ML classifiers is used in the DNN model, CV is not performed for the DNN model evaluation. However, difference mapping (error distribution) of the water depth simulated/predicted in the hydraulic and DNN model is presented in Fig. 11 in addition to the model evaluation for the entire

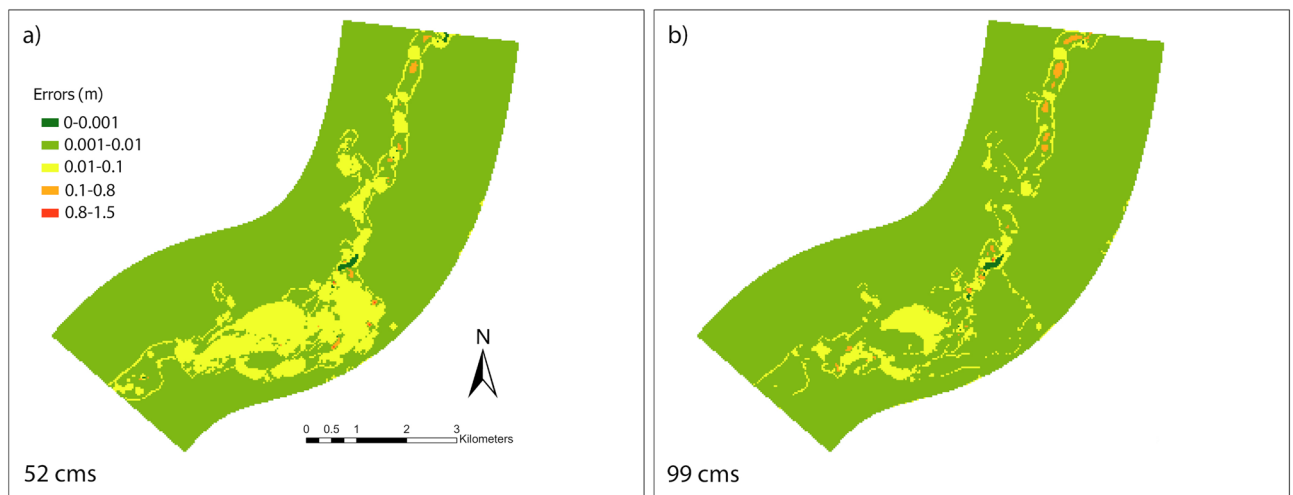


Figure 11. Error distributions calculated by subtracting the water depth from DNN regression model output and iRIC simulation is shown here for two scenarios with upstream discharge of 52 (a) and 99 (b) cms. Maps are processed and generated in the ArcGIS Pro platform⁸¹.

study domain with the error matrix, RMSE. The difference maps in Fig. 11 illustrate the variation in the predicted water depth from DNN regression model and hydraulic model. The spatial distribution of the differences is not significant (in the scale of 0.01 m) for varying upstream discharge condition. Based on the error matrices and difference mapping, it can be concluded that the performance of the DNN regression model is excellent for the urban hydraulic features considered in this study. Orange locations showing high range of difference in errors (predicted and observed flood depth from iRIC) were found mostly in deep water region along the stream obtained from the hydraulic simulation. They are also surrounding the downstream portion of the stream with greater transverse extent with deeper water level. DNN model conveyed a few wider ranges of error in predicting comparatively deeper water along the stream. Imposing more weights on the high water level (extreme values) to reduce the error might lead to a better prediction performance by DNN in the deep-water zones.

Performance learning curve (PLC) of the ML models conveys the impact of the size of the training examples over the model performance. In the Fig. 12, PLC is shown with the train and cross-validation scores for the ML classifiers in flood detection and DNN model in quantifying flood magnitude. Among the ML classifiers, DT performed best with a top F1-score of 0.994 in training and 0.959 in cross-validation. Model performance improve for all models with the increase in the training phase except LR shown in the Fig. 12 (a). Poor performance of the LR model may result from the overfitting, bias, algorithm incompetence in capturing the complexity in data and overall data quality issues. Future studies could focus on the point of diminishing returns for the size of training data in the context of physics-guided AI models in flood detection. Similar to the ML classifiers, DNN model shows improved performance (lower RMSE) with the increase in size of the trainset. The total number of datapoints (examples) to train the models is 253,800 which is 80% of the total datapoints of 317,250 with 8 m of spatial resolution in the computational domain. In case of DNN model, the performance score, RMSE of 0.027. Training sample size is dependent on the spatial resolution of the computational domain. Increasing the resolution may lead to increase in the size of the dataset (increasing the size of the trainset). However, that may end up with less computationally efficient estimation.

Urban hydraulic feature importance. Permutation Feature Importance (PFI) and sensitivity analysis approach are used to determine the impact of the urban features on the DNN regression prediction output i.e., water depth based on the RMSE value as the indicator. PFI measures the variation in the prediction error of the model after the feature's values are permuted⁷¹. This approach quantifies the change in the RMSE values as the prediction error after a series of feature values of interest is permuted/shuffled breaking the linkage between the feature (e.g., downstream distance from stormwater outfall) and target variable (e.g., water depth). This measure is an indicative of the dependency of the model outcome to a specific feature. A sensitivity analysis for the increase in the feature values (5%, 10% and 20%) is performed to quantify the response of the target variable in the DNN regression model, i.e., water depth to the variation in the urban features using RMSE value. From both analysis, impervious area is found to have the highest importance to predicting flood depth compared to the other features. This is logical, as impervious areas directly contribute to runoff and hence to the accumulation of water. The importance scores in both approaches are shown in Fig. 13 to illustrate the significant response of water depth predicted from the DNN regression model due to the change in the urban hydraulic features. The iRIC model used was a hydraulic fluvial model, and did not include rainfall runoff, but the model was calibrated with USGS gage data and validated with NDWI data, showing a high spatial accuracy for flood inundation⁷⁷. Impervious areas showed the highest influence over the target variable i.e., the water depth predicted from the DNN regression model followed by the downstream distance of the stormwater outfall and dams and average slope to the contributing area in case of PFI and 20% increase in the feature values. For the 5% and 10% increase/

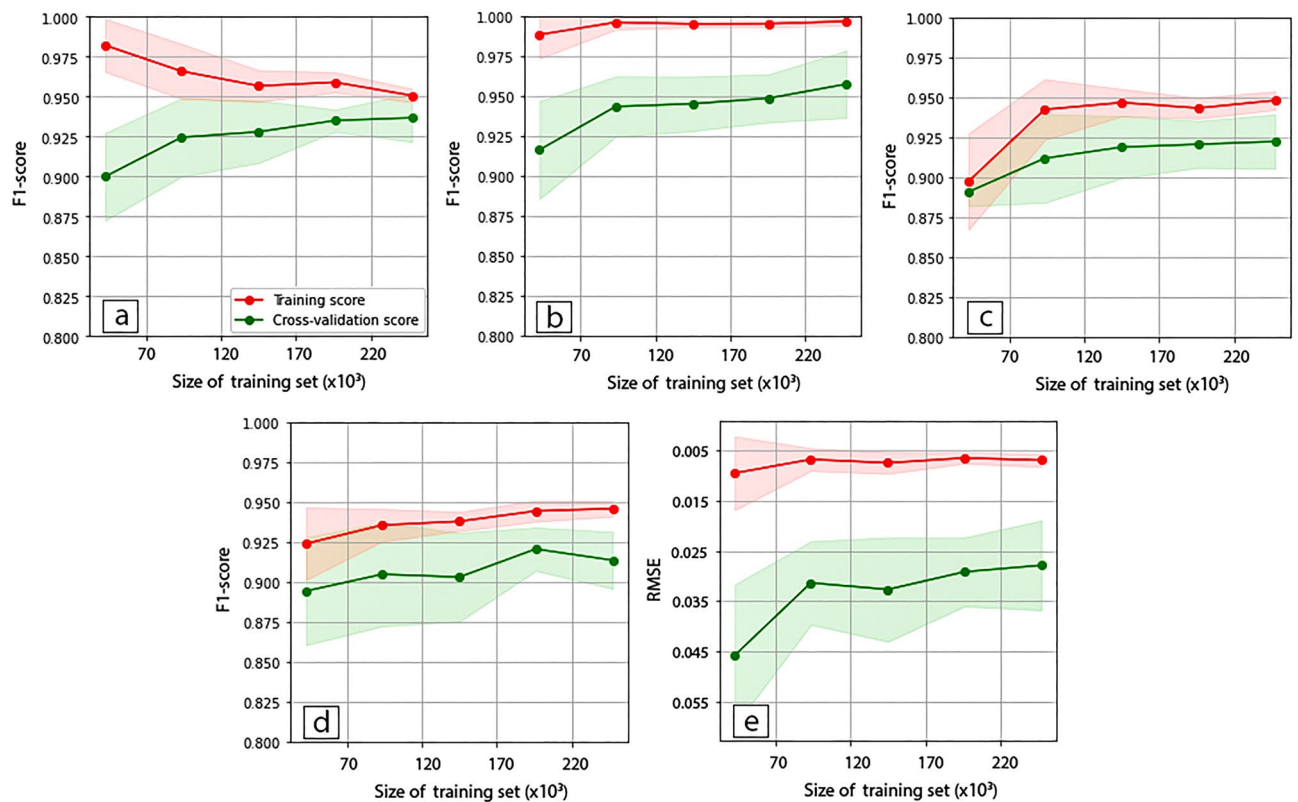


Figure 12. Performance learning curves of the ML classifiers, (a) LG, (b) DT, (c) SVM, (d) KNN and (e) DNN model with the training (red) and cross-validation score (green).

decrease in the feature values, downstream distance from the stormwater outfall showed highest impact. The score of the feature importance, the RMSE value increases for all features with the increase in the change of the feature values from 5 to 20% shown in Fig. 13. Highest response in the RMSE value can be observed in the case of PFI where the feature values are permuted instead of adjusted with a simple sensitivity analysis by percent increase. While these values show the relative importance of features contributing to flood inundation spatially, the disconnection with the physics may create a scenario where the driver(s) of fluvial flooding is limited.

Conclusion

Robust and real-time prediction of flooding is critical to alleviating the growing risk of urban fluvial flooding. Estimation of the water depth of the river and floodplain for various scenarios is of paramount importance in urban flooding planning and management, particularly as many municipalities seek to install or upgrade infrastructure. Data-driven ML approaches provide a path to circumvent the complexities of urban flooding using geographic and urban features outlined in this paper and have the potential to get insights into the flooding attributes. Computationally expensive physics-based numerical models become burdensome at the city-scale and beyond. Traditional hydraulic models depend on solving physics-based differential equations, they require extraordinarily processing power and high memory allocation, specifically for large amount of data and thus perform much slower compared to data-driven methods presented in this paper. In this study, flooding in urban areas, such as the highly urban Darby Creek watershed, is predicted using hybrid physics-informed data-driven techniques. A novel approach to classify and predict the flooded locations and depth using various ML classifiers and DNN-based regression method illustrates a promising ground and potential to entirely shift into the data-driven techniques. Derived urban hydraulic features, i.e., impervious locations and average slope within the contributing area, downstream distance from the stormwater outfall and dams, are introduced in this paper to incorporate the unique impact of urban features on the riverine flooding extent and magnitude which was not present in the previous research works. Future inclusion of additional parameters and resolutions can aid in deepening the understanding of urban hydrology.

A set of binary classifiers (LR, DT, SVM, KNN) is used to identify the flooded locations and a DNN regression model with multiple hidden layers is applied to capture the high non-linearity and quantify the flood magnitude in an urban environment. Both the classification and regression algorithms trained to predict the flooding locations and depth in urban areas with minimum error generated satisfactory outcomes. All error matrices used to evaluate the performance of the binary classifiers are F1-score, Jaccard Similarity matrix and confusion matrix delineate the promising capability of the ML classifiers in isolating flooded locations. The RMSE value used to evaluate the adequacy of the DNN algorithm in predicting water depth. also showed satisfactory performance for the unique datasets with geographic, derived urban and physics-informed hydraulic features. Further, the urban

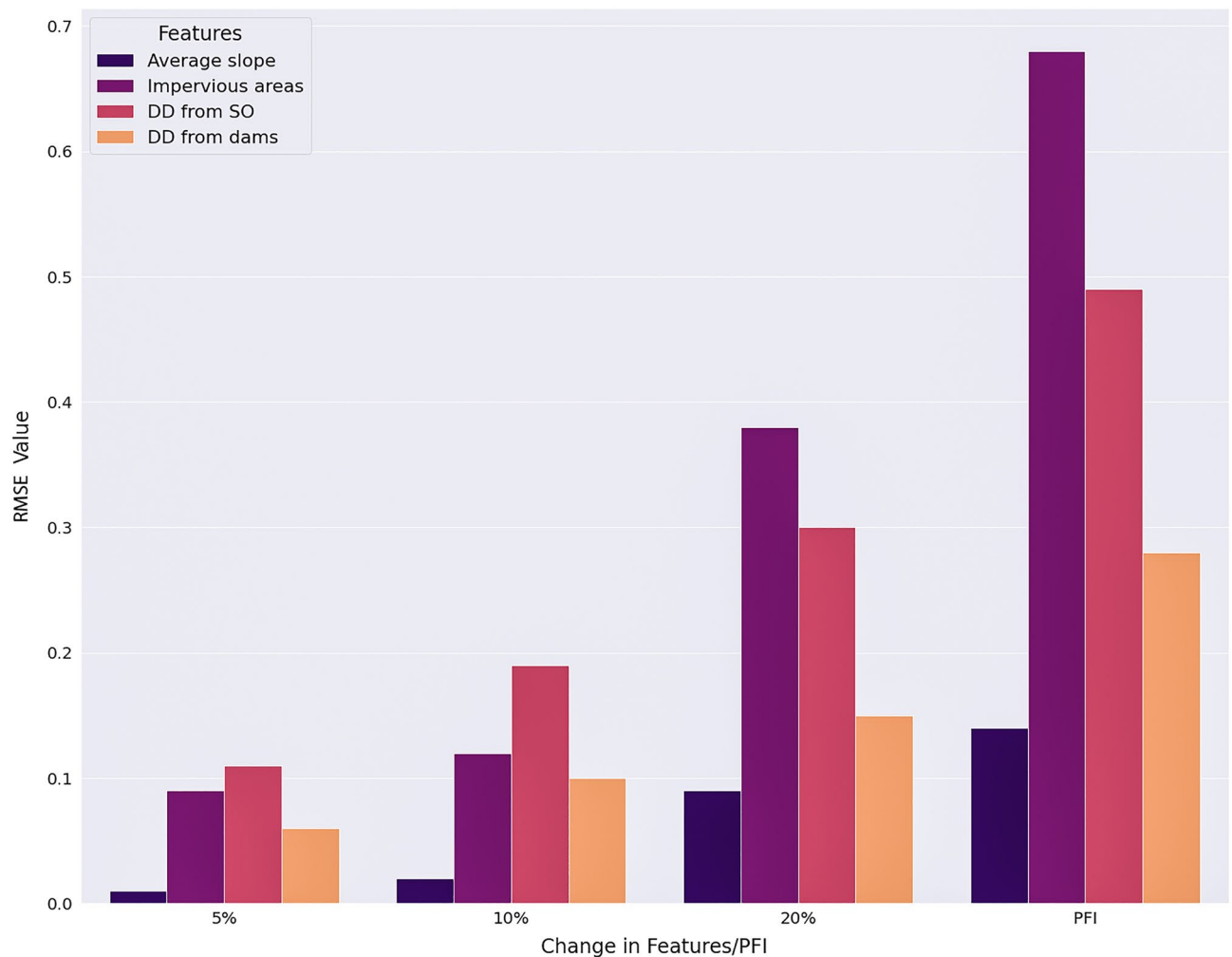


Figure 13. Change in the RMSE value of the DNN regression model due to the % increase/decrease in the feature values and Permutation Feature Importance.

hydraulic feature importance scheme quantified the impact of urban features over the outcome (water depth) of the DNN regression model. Therefore, the satisfactory performance of the proposed framework presented here shows a higher potential for flood prediction in an urban environment, by accounting for the influence of urban features compared to the traditional physics-based hydraulic models. In addition, due to spatial autocorrelation effect, random sampling of data points to prepare the train/test splits from the entire study domain may not yield a satisfactory validation of the model to generalize. In this study, a blocked K-folds CV is performed to further validate the model's performance to generalize to the unseen areas. The performance score for CV showed satisfactory performance for all the spatially segmented blocks as well as the entire study area. Proposed CV framework can be useful in validating ML based flood models to generalize the model performance. For improved flood estimation in complex urban area, a balanced perspective of the proposed framework could serve as a discerning tool for the engineers and decision makers. Data itself cannot be an alternative for physical modeling, however, when combined with the informed and detailed knowledge of the physics-transformed variables from hydrodynamics models, it is highly likely to yield more precise and comprehensive solutions.

The computational time required to converge to the solution by the physic-based model iRIC has been reported approximately one and half times higher (1 h 7 min for iRIC; 32 min for ML models) than the full pipeline of training ML models for flood detection, estimation and feature importance with the ML classifiers and DNN regression model. However, average time required by the trained ML classifiers was 4.7 s in estimating the flood extent where it was 9.3 s for the DNN model in quantifying the flood magnitude. Trained AI models can also be highly efficient in reproducing a range of scenarios which may aid to the decision-making process in a faster and more efficacious way compared to the hydrodynamic model. In addition, several influencing factors (e.g., average slope and impervious locations of the contributing area, downstream distance from the stormwater outfall and dams) used in the ML classifiers and DNN model, were not taken into consideration in the physic-based model equations to quantify the flood extent and magnitude. Therefore, in both, the computational time required and inclusion of the number of influencing factors to the target variable, ML models outperformed the physics based iRIC simulation. The performance of the ML classifiers and DNN regression models can be improved with the increase in the discretization of the computational domain creating more training and testing

data. Geographic and hydraulic features can be stored in the web where the entire training or testing workflow is possible to be executed in the cloud-computing platform. Further, other machine learning and deep learning classifiers and regression models such as Gaussian Process classification, Bayesian classification, Histogram-based gradient boosting, and Long Short-Term Memory regression can be studied with the river hydraulic dataset. Notably the transferability of this method is data-limited. Linking to physical models has the potential to advance model capabilities, as well as allow for deeper insight into urban hydrologic processes. Future work in this area is highly recommended as the data availability and computational power are increasing rapidly. The approach outlined in this study has the potential to be combined with the weather forecast models paving the way of feasible and inexpensive quantification of real-time riverine flooding scenarios.

Data availability

Data collected for the study can be made available upon request from the corresponding author.

Received: 27 April 2022; Accepted: 26 October 2022

Published online: 04 November 2022

References

- Buchanan, M. K., Oppenheimer, M. & Kopp, R. E. Amplification of flood frequencies with local sea level rise and emerging flood regimes. *Environ. Res. Lett.* **12**, 064009 (2017).
- Managing the risks of extreme events and disasters to advance climate change adaptation — IPCC. <https://www.ipcc.ch/report/managing-the-risks-of-extreme-events-and-disasters-to-advance-climate-change-adaptation/>.
- Pielke, R. A. & Downton, M. W. Precipitation and damaging floods: Trends in the United States, 1932–97. *J. Clim.* **13**, 3625–3637 (2000).
- Hirabayashi, Y. *et al.* Global flood risk under climate change. *Nat. Clim. Change* **3**, 816–821 (2013).
- Tanoue, M., Hirabayashi, Y. & Ikeuchi, H. Global-scale River flood vulnerability in the last 50 years. *Sci. Rep.* **6**, 36021 (2016).
- Arnell, N. W. & Lloyd-Hughes, B. The global-scale impacts of climate change on water resources and flooding under new climate and socio-economic scenarios. *Clim. Change* **122**, 127–140 (2014).
- Miller, J. D. & Hutchins, M. The impacts of urbanization and climate change on urban flooding and urban water quality: A review of the evidence concerning the United Kingdom. *J. Hydrol. Reg. Stud.* **12**, 345–362 (2017).
- Ashley, R. M., Balmforth, D. J., Saul, A. J. & Blanksby, J. D. Flooding in the future—predicting climate change, risks and responses in urban areas. *Water Sci. Technol.* **52**(5), 265–273 (2005).
- Wheater, H. & Evans, E. Land use, water management and future flood risk. *Land Use Policy* **26**, S251–S264 (2009).
- Tingsanchali, T. Urban flood disaster management. *Proc. Eng.* **32**, 25–37 (2012).
- Feng, B., Zhang, Y. & Bourke, R. Urbanization impacts on flood risks based on urban growth data and coupled flood models. *Nat. Hazards* **106**, 613–627 (2021).
- Zhou, Q., Leng, G., Su, J. & Ren, Y. Comparison of urbanization and climate change impacts on urban flood volumes: Importance of urban planning and drainage adaptation. *Sci. Total Environ.* **658**, 24–33 (2019).
- Wilby, R. L. & Keenan, R. Adapting to flood risk under climate change. *Prog. Phys. Geogr. Earth Environ.* **36**, 348–378 (2012).
- Dawson, R. J. *et al.* Attribution of flood risk in urban areas. *J. Hydroinf.* **10**(4), 275–288 (2008).
- Rogers, G. O. & Defee li, B. B. Long-term impact of development on a watershed: Early indicators of future problems. *Landsc. Urban Plan.* **73**(2–3), 215–233 (2005).
- Walsh, C. J., Fletcher, T. D. & Burns, M. J. Urban stormwater runoff: A new class of environmental flow problem. *PLoS ONE* **7**, e45814 (2012).
- Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E. & Smith, D. R. Impacts of impervious surface on watershed hydrology: A review. *Urban Water J.* **2**, 263–275 (2005).
- Sahu, R. K., Mishra, S. K. & Eldho, T. I. An improved AMC-coupled runoff curve number model. *Hydrol. Process.* **24**, 2834–2839 (2010).
- Nazif, S., Soleimani, P. & Eslamian, S. *Dynamic curve numbers: Concept and Application* (In Flood Handbook CRC Press, 2022).
- Mishra, S. K., Singh, V. P. & Singh, P. K. Revisiting the Soil Conservation Service Curve Number Method. In Singh, V. P., Yadav, S. & Yadava, R. N. (eds) *Hydrologic Modeling* (Springer, 2018) pp 667–693 https://doi.org/10.1007/978-981-10-5801-1_46.
- Wilkerson, G. & Parker, G. Physical basis for quasi-universal relationships describing bankfull hydraulic geometry of sand-bed rivers. *J. Hydraul. Eng.* **137**, 739–753 (2011).
- Hosseiny, H. & Smith, V. Two dimensional model for backwater geomorphology: Darby creek PA. *Water* **11**(11), 2204 (2019).
- Baghalian, S. & Ghodsian, M. Experimental analysis and prediction of velocity profiles of turbidity current in a channel with abrupt slope using artificial neural network. *J. Braz. Soc. Mech. Sci. Eng.* **39**(11), 4503–4517 (2017).
- Yitian, L. & Gu, R. R. Modeling flow and sediment transport in a river system using an artificial neural network. *Environ. Manag.* **31**, 122–134 (2003).
- Chu, K.-S., Oh, C.-H., Choi, J.-R. & Kim, B.-S. Estimation of threshold rainfall in ungauged areas using machine learning. *Water* **14**, 859 (2022).
- Bhattacharya, B., Price, R. K. & Solomatine, D. P. Machine learning approach to modeling sediment transport. *J. Hydraul. Eng.* **133**, 440–450 (2007).
- Abeshu, G. W., Li, H.-Y., Zhu, Z., Tan, Z. & Leung, L. R. Median bed-material sediment particle size across rivers in the contiguous US. *Earth Syst. Sci. Data* **14**, 929–942 (2022).
- Mosavi, A., Ozturk, P. & Chau, K. Flood prediction using machine learning models: Literature review. *Water* **10**, 1536 (2018).
- Asadollah, S. B. H. S., Sharafati, A., Motta, D. & Yaseen, Z. M. River water quality index prediction and uncertainty analysis: A comparative study of machine learning models. *J. Environ. Chem. Eng.* **9**, 104599 (2021).
- Bhattacharya, B., Price, R. & Solomatine, D. A machine learning approach to modeling sediment transport. *J. Hydraul. Eng.* **133**, 440–450 (2007).
- Moy de Vitry, M. & Leitão, J. P. The potential of proxy water level measurements for calibrating urban pluvial flood models. *Water Res.* **175**, 115669 (2020).
- Sun, A. Y. & Scanlon, B. R. How can big data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. *Environ. Res. Lett.* **14**, 073001 (2019).
- Ke, Q. *et al.* Urban pluvial flooding prediction by machine learning approaches—A case study of Shenzhen city China. *Adv. Water Resour.* **145**, 103719 (2020).
- Dikshit, A., Pradhan, B. & Alamri, A. M. Pathways and challenges of the application of artificial intelligence to geohazards modelling. *Gondwana Res.* **100**, 290–301 (2021).
- Hosseiny, H. A deep learning model for predicting river flood depth and extent. *Environ. Model. Softw.* **145**, 105186 (2021).

36. Yan, X., Mohammadian, A. & Khelifa, A. Modeling spatial distribution of flow depth in fluvial systems using a hybrid two-dimensional hydraulic-multigene genetic programming approach. *J. Hydrol.* **600**, 126517 (2021).
37. Zahura, F. T. *et al.* 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 Resour. Res.* **56**, e2019WR027038 (2020).
38. Guo, Z., Leitão, J. P., Simões, N. E. & Moosavi, V. Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. *J. Flood Risk Manag.* **14**, e12684 (2021).
39. Zhu, G. & Zhu, R. Accelerating Hyperparameter Optimization of Deep Neural Network via Progressive Multi-Fidelity Evaluation. In: Lauw, H W *et al.* (eds) *Advances in knowledge discovery and data mining* (Springer International Publishing, 2020) pp 752–763
40. Passos, D. & Mishra, P. A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemom. Intell. Lab. Syst.* **223**, 104520 (2022).
41. Cao, W., Wang, X., Ming, Z. & Gao, J. A review on neural networks with random weights. *Neurocomputing* **275**, 278–287 (2018).
42. Ding, B., Qian, H. & Zhou, J. Activation functions and their characteristics in deep neural networks. In *2018 Chinese Control And Decision Conference (CCDC)* 1836–1841 (2018). <https://doi.org/10.1109/CCDC.2018.8407425>.
43. Tsimenidis, S. Limitations of deep neural networks: A discussion of G Marcus' critical appraisal of deep learning. *ArXiv Preprint arXiv* <https://doi.org/10.48550/arXiv.2012.15754> (2020).
44. Bentivoglio, R., Isufi, E., Jonkman, S. N. & Taormina, R. Deep learning methods for flood mapping: A review of existing applications and future research directions. *Hydrol. Earth Syst. Sci. Discuss* <https://doi.org/10.5194/hess-2022-83> (2022).
45. Ebtehaj, I. & Bonakdari, H. Performance evaluation of adaptive neural fuzzy inference system for sediment transport in sewers. *Water Resour. Manage* **28**(13), 4765–4779 (2014).
46. Bui, D. T. *et al.* Novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci. Rep.* **8**(1), 1–4 (2018).
47. Dodangeh, E. *et al.* Integrated machine learning methods with resampling algorithms for flood susceptibility prediction. *Sci. Total Environ.* **25**(705), 135983 (2020).
48. Hosseiny, H., Nazari, F., Smith, V. & Nataraj, C. A framework for modeling flood depth using a hybrid of hydraulics and machine learning. *Sci. Rep.* **10**(1), 1–4 (2020).
49. Viteri López, A. S. & Morales Rodriguez, C. A. Flash flood forecasting in são paulo using a binary logistic regression model. *Atmosphere* **11**, 473 (2020).
50. Flood susceptibility modelling using novel hybrid approach of reduced-error pruning trees with bagging and random subspace ensembles—ScienceDirect. <https://www.sciencedirect.com/science/article/pii/S0022169419305347>.
51. Vinothini, K. & Jayanthi, S. IoT Based Flood Detection and Notification System using Decision Tree Algorithm. in *2019 International Conference on Intelligent Computing and Control Systems (ICCS)* 1481–1486 (2019) :<https://doi.org/10.1109/ICCS45141.2019.9065799>.
52. Yariyan, P. *et al.* Improvement of best first decision trees using bagging and dagging ensembles for flood probability mapping. *Water Resour. Manag.* **34**, 3037–3053 (2020).
53. Yilmaz, I. Comparison of landslide susceptibility mapping methodologies for Koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine. *Environ. Earth Sci.* **61**, 821–836 (2010).
54. Bermúdez, M., Cea, L. & Puertas, J. A rapid flood inundation model for hazard mapping based on least squares support vector machine regression. *J. Flood Risk Manag.* **12**, e12522 (2019).
55. Tsakiri, K., Marsellos, A. & Kapetanakis, S. Artificial neural network and multiple linear regression for flood prediction in Mohawk River New York. *Water* **10**, 1158 (2018).
56. Campolo, M., Andreussi, P. & Soldati, A. River flood forecasting with a neural network model. *Water Resour. Res.* **35**, 1191–1197 (1999).
57. Kim, H. I. & Han, K. Y. Urban flood prediction using deep neural network with data augmentation. *Water* **12**, 899 (2020).
58. Full article: Detection of areas prone to flood risk using state-of-the-art machine learning models. <https://www.tandfonline.com/doi/full/https://doi.org/10.1080/19475705.2021.1920480>.
59. Dtissibe, F. Y., Ari, A. A. A., Titouana, C., Thiare, O. & Gueroui, A. M. Flood forecasting based on an artificial neural network scheme. *Nat. Hazards J. Int. Soc. Prev. Mitig. Nat. Hazards* **104**, 1211–1237 (2020).
60. Gessang, O. M. & Lasminto, U. The flood prediction model using Artificial Neural Network (ANN) and weather application programming interface (API) as an alternative effort to flood mitigation in the Jenelata Sub-watershed. *IOP Conf. Ser. Mater. Sci. Eng.* **930**, 012080 (2020).
61. Zahura, F. T. *et al.* 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 Resour. Res.* **56**, e2019WR027038 (2020).
62. Moosavi, V. Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks. *J. Flood Risk Manag.* **14**, e12684 (2021).
63. Ma, M. *et al.* Flash flood risk analysis based on machine learning techniques in the Yunnan Province, China. *Remote Sensing* **11**(2), 170 (2019).
64. Nachappa, T. G. *et al.* Flood susceptibility mapping with machine learning, multi-criteria decision analysis and ensemble using Dempster Shafer Theory. *J. Hydrol.* **590**, 125275 (2020).
65. Russell, K. L., Vietz, G. J. & Fletcher, T. D. Urban sediment supply to streams from hillslope sources. *Sci. Total Environ.* **653**, 684–697 (2019).
66. Gellis, A. C. *et al.* Storms, channel changes, and a sediment budget for an urban-suburban stream, difficult run, Virginia, USA. *Geomorphology* **278**, 128–148. <https://doi.org/10.1016/j.geomorph.2016.10.031> (2017).
67. Smith, S. & Wilcock, P. Upland sediment supply and its relation to watershed sediment delivery in the contemporary mid-Atlantic Piedmont (USA). *Geomorphology* **232**, 33–46. <https://doi.org/10.1016/j.geomorph.2014.12.036> (2015).
68. Taylor, K. Urban environments. In *Environmental Sedimentology* (eds Perry, C. & Taylor, K.) 190–222 (Blackwell Publishing, 2007).
69. Wolman, M. G. A cycle of sedimentation and erosion in urban river channels *Geografiska Annaler. Series A Phys. Geogr.* **49**(2/4), 385–395 (1967).
70. Taylor, K. G. & Owens, P. N. Sediments in urban river basins: A review of sediment–contaminant dynamics in an environmental system conditioned by human activities. *J. Soils Sedim.* **9**(4), 281–303 (2009).
71. Ampomah, R. *et al.* 3-Dimensional numerical modeling of urban green stormwater infrastructure curb inlets. *J. Irrig. Drain.* **147**(9), 04021039 (2020).
72. Smith, V. B. & Mohrig, D. Geomorphic signature of a dammed sandy river: The lower trinity river downstream of livingston dam in Texas, USA. *Geomorphology* **297**, 122–136 (2017).
73. Bruwier, M. *et al.* Influence of urban forms on surface flow in urban pluvial flooding. *J. Hydrol.* **582**, 124493 (2020).
74. Ozkan, S. P. & Tarhan, C. Detection of flood hazard in urban areas using GIS: Izmir case. *Procedia Technol.* **22**, 373–381 (2016).
75. Brody, S., Blessing, R., Sebastian, A. & Bedient, P. Examining the impact of land use/land cover characteristics on flood losses. *J. Environ. Planning Manag.* **57**(8), 1252–1265 (2014).
76. Kim, T. J. Modeling for mitigating storm water urban flooding and water quality issues by using small serial dams: A case study of the city of San Angelo. *Global NEST J.* **23**(2), 288–296 (2021).
77. Hosseiny, H. & Smith, V. Two-dimensional model for backwater geomorphology: Darby creek PA. *Water* **11**, 2204 (2019).

78. Sudheer, K., Nayak, P. C. & Ramasastri, K. Improving peak flow estimates in artificial neural network river flow models. *Hydrol. Process* **17**, 677–686 (2003).
79. Zarzar, C. M. *et al.* A hydraulic multimodel ensemble framework for visualizing flood inundation uncertainty. *JAWRA J. Am. Water Resour. Assoc.* **54**, 807–819 (2018).
80. City Council Testimony on Environmental Health Disparities and Environmental Racism in Philadelphia | The Public Interest Law Center. <https://www.pubintlaw.org/cases-and-projects/city-council-testimony-on-environmental-health-disparities-and-environmental-racism-in-philadelphia/>.
81. 2D, 3D & 4D GIS Mapping Software | ArcGIS Pro. <https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview>.
82. Solvers | iRIC Software. <https://i-ric.org/en/solvers/>.
83. USGS Current Conditions for USGS 01475548 Cobbs Creek at Mt. Moriah Cemetery, Philadelphia. https://nwis.waterdata.usgs.gov/pa/nwis/uv?cb_00065=on&cb_00060=on&format=gif_default&site_no=01475548&period=&begin_date=2014-04-29&end_date=2014-04-30.
84. Scikit-learn: machine learning in Python—scikit-learn 1.0.2 documentation. <https://scikit-learn.org/stable/>.
85. ArcGIS Pro | Documentation. <https://pro.arcgis.com/en/pro-app/2.8/arcpy/get-started/what-is-arcpy-.htm>.
86. Zou, X., Hu, Y., Tian, Z. & Shen, K. Logistic Regression Model Optimization and Case Analysis. in *2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT)* 135–139 (2019). <https://doi.org/10.1109/ICCSNT47585.2019.8962457>.
87. Charbuty, B. & Abdulazeez, A. Classification based on decision tree algorithm for machine learning. *J. Appl. Sci. Technol. Trends* **2**, 20–28 (2021).
88. Suthaharan, S. Support Vector Machine. in *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning* (ed. Suthaharan, S.) 207–235 (Springer, Boston, 2016)
89. Altmann, A., Tološi, L., Sander, O. & Lengauer, T. Permutation importance: A corrected feature importance measure. *Bioinformatics* **26**, 1340–1347 (2010).
90. König, G., Molnar, C., Bischl, B. & Grosse-Wentrup, M. Relative feature importance. in *2020 25th International Conference on Pattern Recognition (ICPR)* 9318–9325. <https://doi.org/10.1109/ICPR48806.2021.9413090> (2021).
91. Choudhary, R. & Gianey, H. K. Comprehensive review on supervised machine learning algorithms. in *2017 International Conference on Machine Learning and Data Science (MLDS)* 37–43. <https://doi.org/10.1109/MLDS.2017.11> (2017).
92. Lipton, Z. C., Elkan, C. & Narayanaswamy, B. Thresholding Classifiers to Maximize F1 Score. *ArXiv14021892 Cs Stat* (2014).
93. Al-jabery, K. K., Obafemi-Ajayi, T., Olbricht, G. R. & Wunsch II, D. C. 9 - Data analysis and machine learning tools in MATLAB and Python. in *Computational Learning Approaches to Data Analytics in Biomedical Applications* (eds. Al-jabery, K. K., Obafemi-Ajayi, T., Olbricht, G. R. & Wunsch II, D. C.) 231–290 (Academic Press, 2020). <https://doi.org/10.1016/B978-0-12-814482-4.00009-7>
94. Goutte, C. & Gaussier, E. A Probabilistic interpretation of precision, recall and F-score, with implication for evaluation. in *Advances in Information Retrieval* (eds. Losada, D. E. & Fernández-Luna, J. M.) 345–359 (Springer, 2005). https://doi.org/10.1007/978-3-540-31865-1_25
95. Lai, K. *et al.* A natural language processing approach to understanding context in the extraction and geocoding of historical floods, storms, and adaptation measures. *Inf. Process. Manag.* **59**(1), 102735 (2022).
96. Arabameri, A., Chandra Pal, S., Rezaie, F., Chakraborty, R., Saha, A., Blaschke, T., & Thi Ngo, P. T. Decision tree based ensemble machine learning approaches for landslide susceptibility mapping. *Geocarto Int.* 1–35. (2021).
97. Chung, C. J. & Fabbri, A. G. Predicting landslides for risk analysis—Spatial models tested by a cross-validation technique. *Geomorphology* **94**(3–4), 438–452 (2008).
98. Jung, C., Ahn, S., Sheng, Z., Ayana, E. K., Srinivasan, R., & Yeganantham, D. (2021). Evaluate river water salinity in a semi-arid agricultural watershed by coupling ensemble machine learning technique with SWAT model. *JAWRA J. Am. Water Resour. Assoc.*
99. Shafi, I., Ahmad, J., Shah, S. I. & Kashif, F. M. (2006) Impact of varying neurons and hidden layers in neural network architecture for a time frequency application. in *2006 IEEE International Multitopic Conference* <https://doi.org/10.1109/INMIC.2006.358160>.
100. Panchal, G., Ganatra, A., Shah, P. & Panchal, D. Determination of over-learning and over-fitting problem in back propagation neural network. *Int. J. Sofi Comput.* **2**, 40–51 (2011).

Acknowledgements

This work was supported by the Villanova Center of Resilient Water System (VCRWS) of Villanova University.

Author contributions

A.M.A.A.M.—conceived of the presented idea and developed the algorithm to predict flooding extent and magnitude in the urban environment, lead the investigation of urban hydraulic features in the predictive analysis and supervised the findings of the work, carried out the modeling experiment and analysis, wrote the manuscript. B.V.S.—conceived of the presented idea and developed the algorithm to predict flooding extent and magnitude in the urban environment, oversaw the investigation of urban hydraulic features in the predictive analysis and supervised the findings of the work, reviewed the data-driven approach and AI models and verification of results, oversaw the writing and development of the manuscript. C.H.H.—conceived of the presented idea and developed the algorithm to predict flooding extent and magnitude in the urban environment, contributed ideas to the investigation of urban hydraulic features in the predictive analysis and supervised the findings of the work, developed, and produced data from the ArcGIS and hydraulic model, reviewed the data-driven approach and AI models and verification of results, provided feedback on the manuscript. D.X.J.—reviewed the data-driven approach and AI models and verification of results provided feedback on the manuscript.

Funding

Department of Civil and Environmental Engineering.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-23214-9>.

Correspondence and requests for materials should be addressed to M.A.A.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022