

Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment



Zelalem Demissie^{a,*}, Prashant Rimal^a, Wondwosen M. Seyoum^b, Atri Dutta^a,
Glen Rimmington^a

^a Wichita State University, 1845 Fairmount St, Wichita, KS, 67260, USA

^b Illinois State University, 100 N University St, Normal, IL, 61761, USA

ARTICLE INFO

Keywords:
Flood susceptibility
Machine learning
Stack generalization
GIS
Kansas

ABSTRACT

Flooding presents a formidable challenge in the United States, endangering lives and causing substantial economic damage, averaging around \$5 billion annually. Addressing this issue and improving community resilience is imperative. This project employed machine learning techniques and publicly available data to explore the factors influencing flooding and to develop flood susceptibility maps at various spatial resolutions. Six machine learning algorithms, including Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN), Adaptive Boosting (Ada Boost), and Extreme Gradient Boosting (XGB) were used. Geospatial datasets comprising thirteen predictor variables and 1528 flood inventory data collected since 1996 were analyzed. The predictor variables are rainfall, elevation, slope, aspect, flow direction, flow accumulation, Topographic Wetness Index (TWI), distance from the nearest stream, evapotranspiration, land cover, impervious surface, land surface temperature, and hydrologic soil group. Five hundred twenty-eight non-flood data points were randomly created using a stream buffer for two scenarios. A total of 2964 data points were classified into flooded (1) and non-flooded (0) categories and used as a target. Overall, testing results showed that the XGB and RF models performed relatively well in both cases over multiple resolutions compared to other models, with an accuracy ranging from 0.82 to 0.97. Variable importance analysis depicted that predictor variables such as distance from the streams, hydrologic soil type, rainfall, elevation, and impervious surfaces significantly affected flood prediction, suggesting a strong association with the underlying driving process. The improved performance and the variation of the susceptible areas across two scenarios showed that considering predictor variables with multiple resolutions and appropriate non-flooding training points is critical for developing flood-susceptibility models. Furthermore, using tree-based ensemble algorithms like RF and XG boost in the stack generalization approach can help achieve robustness in a flood susceptibility model where multiple algorithms are being evaluated.

1. Introduction

Globally, disaster-related events such as weather, climate, or water hazards occur daily, killing 115 people and causing US\$ 202 million in damages (World Meteorological Organization, 2021). Disasters profoundly influence human environments, with floods being the most catastrophic among various types. More than 5000 flood events have been documented since 1980, representing over 40% of all disaster occurrences. These floods have induced an economic impact exceeding 767 billion U.S. dollars and affected approximately 1.71 billion people across the globe, testifying to their severe repercussions (Centre for

Research on the Epidemiology of Disasters, 2022; CRED and UN Office for Disaster Risk Reduction (UNDRR), 2021; CRED and UNDRR, 2020). According to the National Centers for Environmental Information (NCEI), in the United States (U.S.), since 1980, at least 323 weather and climate-related disaster events with financial loss of more than a billion dollars have been recorded with an estimated loss of over 2.1 trillion US Dollars (National Centers for Environmental Information (NCEI), 2022). Among those events, 36 were flooding events that caused over 168.4 billion U.S. Dollars in losses. As a result of anthropogenic climate change, flooding events are more frequent and expected to rise in recent decades (Liu et al., 2022; Meresa et al., 2022; Milly et al., 2002). The

* Corresponding author.

E-mail address: zelalem.demissie@wichita.edu (Z. Demissie).

observed evidence of the rising flood frequency over the Midwest states might cause a more significant impact (Mallakpour and Villarini, 2015). Over the past century, Midwest states have been plagued by deadly and devastating floods like those of 1977, 1993, 2008, 2011, 2013, 2014, and 2019. Those floods have led to adverse societal consequences, including food production decline and displacement of communities, with economic losses reaching billions of dollars (Hauth and Carswell, 1978a,b; Downton et al., 2005; "The, 2019 Floods in the Central U.S.", 2021; Xiao et al., 2013).

The project focuses on the State of Kansas, which suffered a significant impact due to major flood events in the Midwest (National Weather Service, 2022; NOAA et al., 2022; NCEI, 2022). Since the 1980s, six flooding events with an estimated cost of around 2–5 billion U.S. Dollars have occurred in Kansas (NCEI, 2022). Significant efforts have been made to investigate causes and mitigate them in the region. Most of the studies related to flooding in Kansas are either related to post-flood scenarios (Hauth and Carswell, 1978a,b; Clement and Johnson, 1982; Clement, 1987; Sophocleous et al., 1996; Follansbee and Spiegel, 1935; Veatch, 1952) or related to the monitoring of hydrologic conditions (Louen, 2017). Studies associated with flood inundation maps and wetland restoration, along with an assessment of the geomorphic effects of large floods, were published by the United States Geological Survey (USGS) Bowen and Juracek (2011); Heimann et al. (2014); Qaiser et al. (2012) highlighted the harsh impact of urbanization and future scenarios on flooding in their study of the Kansas River basin. In their study, a considerable increase in the peak discharge and inundation areas, along with the increase in water elevation, was predicted for the basin for future scenarios. Yet, studies directly related to flood susceptibility mapping for the entire state are limited. However, inundation mapping has been done in some areas, along with the construction of mitigating structures like dams and levees in various regions of the state (Kansas Department of Agriculture, 2022). Though the mitigation measures are present, the news of reservoirs and levees reaching their capacity in 2019 during extreme rainfall events was reported (Plake, 2019).

Given the noted gap in comprehensive flood susceptibility mapping across the state, technological advancements and predictive modeling may provide a pathway to address this shortfall. Despite significant advancements in flood susceptibility mapping methodologies, traditional approaches often struggle to capture the intricate relationships and complex variables contributing to flood vulnerability. Existing methods may lack the ability to provide accurate and reliable predictions, leading to challenges in proactive flood risk management and disaster preparedness. The need for more sophisticated and data-driven techniques is evident to enhance our understanding of flood-prone areas and improve the effectiveness of mitigation strategies. Therefore, this research aims to address these limitations by developing machine learning-based models that can accurately pinpoint regions susceptible to flooding, leveraging advanced algorithms to analyze factors such as topography, land use, hydrological conditions, and historical flood data. By enhancing the accuracy and reliability of flood susceptibility maps, this study seeks to contribute to more informed decision-making processes and proactive measures for mitigating the adverse impacts of floods.

Over recent decades, significant progress has been made in the prediction sciences due to technological evolution. Various data-driven models, such as bivariate statistical models, Multi-Criteria Decision Making (MCDM), and Artificial Intelligence (AI) approaches, including Machine Learning (ML), computational Intelligence (CI), soft computing (SC), Data Mining (DM), Knowledge of discovery in databases (KDD), and Intelligent Data Analysis (IDA) models have been used in the field of prediction sciences (Solomatine et al., 2008). Along with floods and other natural disasters, several models and their hybrids were used in a wide range of studies (Abedi et al., 2022; Ahmadlou et al., 2019; Bui et al., 2021; Chapi et al., 2017; Costache and Zaharia, 2017; El-Magd et al., 2021; El-Haddad et al., 2021; Farhadi and Najafzadeh, 2021; Madhuri et al., 2021; Meliho et al., 2021; Khosravi et al., 2019; Pradhan,

2010; Tehrany et al., 2015; Shafizadeh-Moghadam et al., 2018; Solomatine et al., 2008; Wang et al., 2015). In recent years, the use of ML techniques in flood risk assessment has been growing because of their ability to capture relationships efficiently (Wang et al., 2015), as the evaluation is dependent solely on detecting flood-prone areas with the use of historical events and topography (Costache and Zaharia, 2017). Such knowledge of the flood-prone regions helps us to know about the areas susceptible to flooding.

Regarding machine learning, various algorithms were used for flood susceptibility mapping across the globe. For example, Tehrany et al. (2015) used an integrated form of Support Vector Machine (SVM) and Frequency Ratio (FR) for flood susceptibility assessment in the Kelantan basin of Malaysia. Khosravi et al. (2019) conducted a comparative evaluation of flood susceptibility modeling using three Multi-Criteria Decision-Making (MCDM) techniques with two ML models, Naïve Bayes Tree (NBT) and Naïve Bayes (NB). Extreme Gradient Boost (XG boost) and K-Nearest Neighborhood (KNN) were used for flood prediction mapping in Egypt, and the model's accuracy was evaluated using receiver operating characteristics (ROC) and Area Under Curve (AUC) (El-Magd et al., 2021). El-Haddad et al. (2021) used four ML models-Boosted Regression Tree (BRT), Functional Data Analysis (FDA), General Linear Model (GLM), and Multivariate Discriminant Analysis (MDA) to generate flood maps in the Wadi Qunea Basin of Egypt. Meliho et al. (2021) employed four supervised models based on ML algorithms: Random Forest (RF), KNN, Artificial Neural Networks (ANN), and XG boost for flood susceptibility in the Souss watershed of Morocco. Madhuri et al., 2021 compared the five ML models: Logistic Regression (LR), SVM, KNN, Ada boost, and XG boost for flood susceptibility assessment and risk management in the Greater Hyderabad Municipal Corporation, India. XG boost performed the best among others and was used to simulate future flood locations.

This research focuses on flood susceptibility mapping using machine learning, which is of paramount importance due to its potential to revolutionize our understanding and prediction of flood-prone areas. Traditional methods often fall short of capturing the complex interplay of diverse variables contributing to flood susceptibility (El-Magd et al., 2021; Meliho et al., 2021; Madhuri et al., 2021). With their ability to discern intricate patterns and relationships within large datasets, machine learning algorithms offer a promising avenue for creating more accurate and reliable flood susceptibility maps. This research aimed to develop machine learning-based models that pinpoint regions susceptible to flooding by assessing factors such as topography, land use, hydrological conditions, and historical flood data. By leveraging advanced modeling techniques, the research aims to enhance our capacity to identify vulnerable regions, enabling proactive and effective flood risk management strategies.

The study employed classification algorithms, including Logistic Regression (LR), Random Forest (RF), Support Vector Machine (SVM), K-nearest neighbor (KNN), Adaptive Boosting (Ada Boost), and Extreme Gradient Boosting (XGB) to model the study area as flood or non-flooded cells using historical flood inventory data for the state of Kansas. In addition, flood susceptibility mapping was conducted by combining the ML outputs with weighted overlay analysis. The novelty of this research lies in using the algorithm's feature importance in weighing predictor variables while establishing flood susceptibility maps. Both feature and permutation-based importance of the thirteen predictor variables were used for that purpose. The findings from this study are poised to make a societal impact by enhancing flood susceptibility mapping. The improved accuracy and reliability of these maps have the potential to play a crucial role in mitigating the adverse consequences of floods, including safeguarding lives, protecting critical infrastructure, and promoting sustainable urban development practices. Using advanced machine learning techniques in flood susceptibility mapping provides valuable insights to inform effective strategies for disaster risk reduction and resilience building in vulnerable regions.

2. Data and methods

This section details the data used and the methodological framework employed to assess flood susceptibility in the study area. We leverage a suite of Machine Learning (ML) algorithms to analyze the relationships between various geospatial factors and historical flood occurrences. Section 2.2 to 2.4 outline the preparation of predictor and predict variables, including data acquisition and pre-processing. Sections 2.5 delves into the specific ML algorithms implemented, providing explanations of the model development process, encompassing training procedures, hyperparameter tuning, and feature importance. Finally, sections 2.6 to 2.8 discuss evaluation metrics and the generation of flood susceptibility maps.

2.1. Study area

The study area covers Kansas, situated in the Great Plains physiographic provinces in the west, and central lowlands in the east. The western and central portion of the state (about 30,800 square miles) is the area of the high plain's aquifer, also called the Ogallala Aquifer (Adams, 1903). Vast acres of fertile farmlands are present in the west, whereas green hills and forests are in the east. Regarding the hydrological soil groups, much of the eastern side of Kansas contains Group D soils (SSUGRO: Soil Survey Geographic database developed by the Natural Resources Conservation Service (NRCS). These soils have high clay content, high shrink-swell potential, and a higher water table, leading to slow infiltration and high runoff (Soil Survey StaffNatural Resources Conservation ServiceUnited States Department of Agriculture, 2019).

Kansas's climate varies from east-west, while there is slight variation in a north-south direction. The Eastern part of the state is relatively wet, receiving 900 mm–1150 mm of annual rainfall, while the western portion is relatively drier, with 450 mm–660 mm of rainfall (Daly et al., 1997; Goodin et al., 2004). The rain shadow effect of the Rocky Mountains and the altitude differences have notable impacts on precipitation. Temperature follows a north-south pattern, while rainfall has an east-west gradient. The annual mean temperature ranges from around 11–14.4 °C (NCEI, 2022). Summer is hot and dry, with temperatures crossing 32 °C in the southeast and southwest portions. Temperature exceeds 43 °C due to the Urban Heat Island effect in places like Kansas City and Topeka (Kansas, USA - Yearly & Monthly weather forecast, 2022). The western part is semi-arid, with hot, dry summer months and cold, windy winter months, whereas the eastern part is more humid, with sultry summer and cold winter months (Clement, 1987). Over 80% of the state's land is used for agricultural production. Thus, the study area is highly important for supplying various crops and cattle in the United States.

2.2. Data sources

The historical flooding inventories were extracted from the NOAA's storm event database, reflecting regular flooding occurrences in the study area over the years. The Land Cover dataset and imperviousness were extracted from the National Land Cover Dataset (NLCD) developed by the Multi-Resolution Land Characteristics (MRLC) consortium (Table 1). The hydrologic soils group data, obtained from the Soil Survey Geographic Database (SSUGRO) developed by NRCS, was utilized. The annual average rainfall data from the various stations within the study area was derived from SC-ACIS (Applied Climate Information System). Additionally, datasets from Google Earth Engine (GEE) were utilized for Land Surface Temperature (LST), Evapotranspiration, and Digital Elevation Model (DEM). We used DEM to calculate the other flood-controlling variables: elevation, slope, aspect, flow direction, flow accumulation, and topographic wetness index (TWI). All these datasets were resampled into various resolutions (30 m, 100 m, and 500 m) for further analysis (Table 1).

Table 1

Flood Controlling factors, their respective data sources, minimum and maximum values, and data processing.

Flood Controlling Factors	Value (Min-Max)	Data Sources and Equations
1 Elevation	206.288–1231.511 m	United States Geological Survey; 1/3 Arc-Second (https://earthexplorer.usgs.gov/), (Demissie et al., 2023; Demissie and Rimmington, 2022)
2 Slope	0–61.188°	Derived from item 1-the 1/3 Arc-Second
3 Aspect	0–360 °	Derived using DEM represented in terms of direction (N, NE, E, SE, S, SW, W, NW, N)
4 Flow Direction	1–128	Derived using DEM, used D8 flow direction type, and values are based on neighboring flow cells.
5 Flow Accumulation	0–84125872	Derived using Flow Direction with D8 flow direction type. Represents the total accumulated flow for each cell.
6 TWI (Topographic Wetness Index)	2.942–35.688	$TWI = \ln(A_s)/\tan(\beta)$ Here, A_s is the local upslope area or specific area through a certain point per unit contour length, and β is the slope gradient (in radians). It represents the amount of water contained in every pixel.
7 DNS (Distance from Streams)	0–48019.425 m	Derived using the modified NHD (National Hydrography dataset) (Kansas Biological SurveyUniversity of Kansas, 2022)
8 AET (Actual Evapotranspiration)	282.167–937.916 mm	Yearly averaged map of AET from GEE using the Terra Climate: Monthly Climate and Climatic Water Balance for Global Terrestrial Surfaces (Abatzoglou et al., 2018)
9 LST (Land Surface Temperature)	12.97–28.03 °C	Terra Land Surface Temperature and Emissivity 8-day dataset derived from the Google Earth Engine (GEE) averaged (Wan et al., 2021)
10 Soil Group	1–7	Hydrologic soils group data (from SSUGRO, NRCS). into four major groups (A, B, C, and D) and three dual classes (A/D, B/D, and C/D) according to the rate of water infiltration (Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture., 2019).
11 Land Use and Land Cover	11–95	Derived from the National Land Cover Dataset (NLCD) (Dewitz, 2021). Reclassified into 10 major groups using ArcGIS pro
12 Impervious Surface	0–100%	Derived from National Land Cover Dataset (NLCD) 2019 Impervious Products (Dewitz, 2021)
13 Precipitation	6.226–70.2612 in	Maximum one-year total Precipitation extracted from SC-AIS (NOAA Regional Climate Centers).

2.3. Flood controlling factors

Constructing a susceptible flood model is a complex process, as many topographical and hydrological variables are needed (Arabameri et al., 2020; Abedi et al., 2022; Madhuri et al., 2021; Pradhan, 2010). Recognition of these variables that impact the flood is a critical activity, and these parameters validate the accuracy of the flood susceptibility maps. Based on the available literature on flood susceptibility, for the study area, thirteen flood-influencing factors were selected (Bui et al., 2020; Costache and Zaharia, 2017; El-Haddad et al., 2021; El-Magd et al., 2021; Khosravi et al., 2019; Madhuri et al., 2021; Meliho et al., 2021; Saravanan and Abijith, 2022; Shafizadeh-Moghadam et al., 2018; Tang et al., 2020; Tehrany et al., 2015; Wang et al., 2015). These factors are Elevation, Slope, Aspect, Flow Direction, Flow Accumulation, Topographic Wetness Index (TWI), Distance from Streams (DNS), Precipitation, Evapotranspiration, Land Surface Temperature (LST), Land Cover, and hydrologic soil groups (Fig. 1). All these influencing factors were transformed into a raster format and resampled to the exact resolution. Table 1 shows the minimum and maximum values and the respective data sources for all the controlling factors.

2.4. Flood inventory map

Creating a flood susceptibility map involves preparing a flood inventory map (Bui et al., 2021; Saravanan and Abijith, 2022). In this research, the inventory map was prepared from NOAA's storm database (National Oceanic and Atmospheric Administration and National Centers for Environmental Information, 2022); after cleaning the data, about 1477 historical flooding events occurred between 1996 and 2021 were obtained. These events are categorized into flood points (1), serving as the dependent variable in the models. Further, to avoid bias, a nearly equal number of flooded and non-flood points were introduced (Tang et al., 2020; Towfiqul Islam et al., 2021). While creating the non-flooded points, two scenario maps were developed using a buffer distance from water bodies (Fig. 2). The two scenarios, termed Case I and Case II, were devised for generating non-flooded points using buffer distances from water bodies (Fig. 2). In Case I, non-flood inventories were established beyond a 5 Km buffer from flood points, while in Case II, non-flood points were situated at least 3 Km away from flood points

and 4 Km from streams. These distinct scenarios were introduced to evaluate the impact of non-flood points on the responsiveness of machine learning algorithms. Placing the points outside the buffer zone prevents the overlap of values associated with flooding points across various resolutions. The 1477 non-flood points from each scenario map were randomly selected outside the buffer areas and labeled 0. For each case with various spatial resolutions, the inventory points (flooded and non-flooded) were used to extract the value from flood-conditioning factors.

2.5. Machine learning algorithms

Fig. 3 displayed the overall methodology including data processing, preparing the flood inventory map, enhancing the flood controlling factors, training and testing flood susceptibility models using multiple ML algorithms, and evaluating factors using the inbuilt feature and permutation-based importance. We used six ML algorithms: Logistic Regression (LR), Random Forest (RF), K- Nearest Neighbor (KNN), Support Vector Machine (SVM), Adaptive Boosting (Ada Boost), and Extreme Gradient Boosting (XGB), each of which offers a range of techniques to address predictive modeling challenges. LR, a widely used algorithm for binary classification, models the probability of a categorical outcome using a logistic function, making it highly interpretable and efficient for linearly separable data. RF is an ensemble method that constructs multiple decision trees during training and combines their outputs to improve accuracy and robustness while mitigating overfitting (Breiman, 2001). KNN classifies data points based on the majority class of their k-nearest neighbors in the feature space, making it simple yet effective for non-linear and multi-class classification problems. SVM are applicable for both classification and regression problems, finding the optimal hyperplane that maximizes the margin between classes, and performing well with high-dimensional data. Ada Boost combines weak classifiers in a sequential manner to form a strong classifier, focusing more on misclassified instances to improve performance iteratively (Freund and Schapire, 1997). XGB is a highly efficient and scalable implementation of the gradient boosting framework, excelling in handling large-scale data sets and complex prediction tasks with its advanced regularization and parallel processing capabilities (Chen and Guestrin, 2016). The various ML algorithms, their mathematical bases,

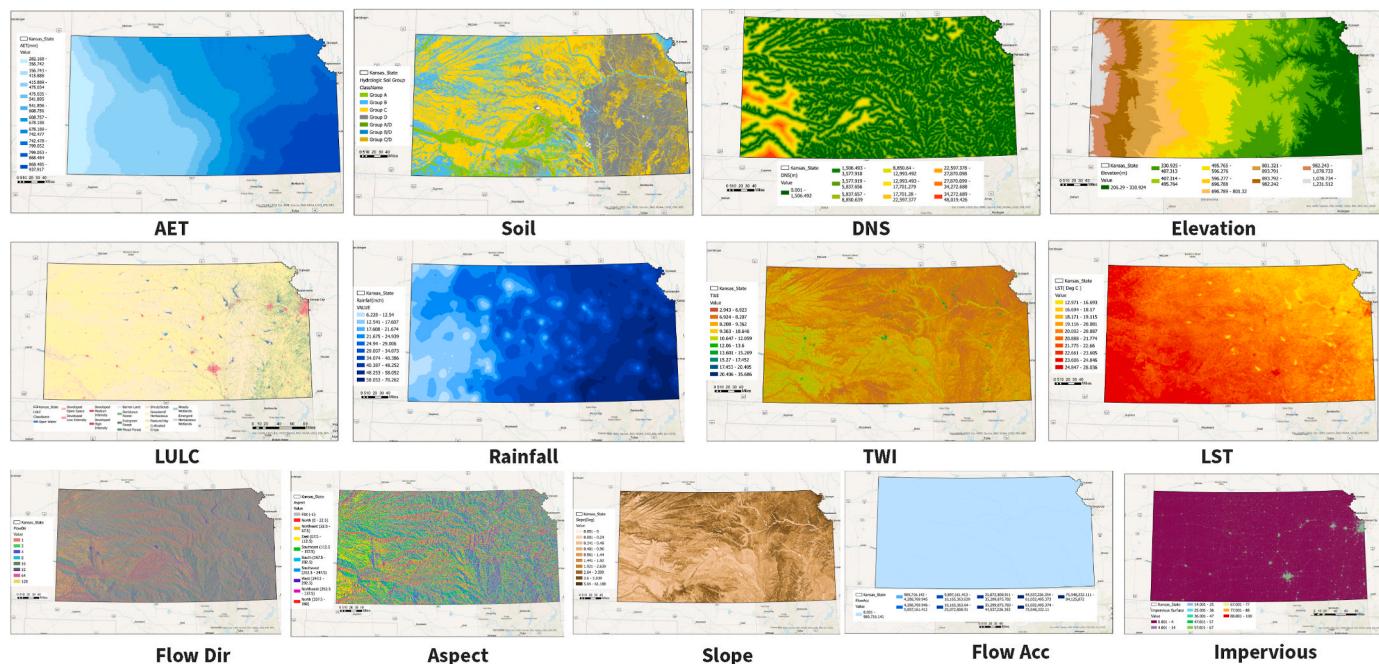


Fig. 1. Thirteen flood-controlling variables were used in the research.

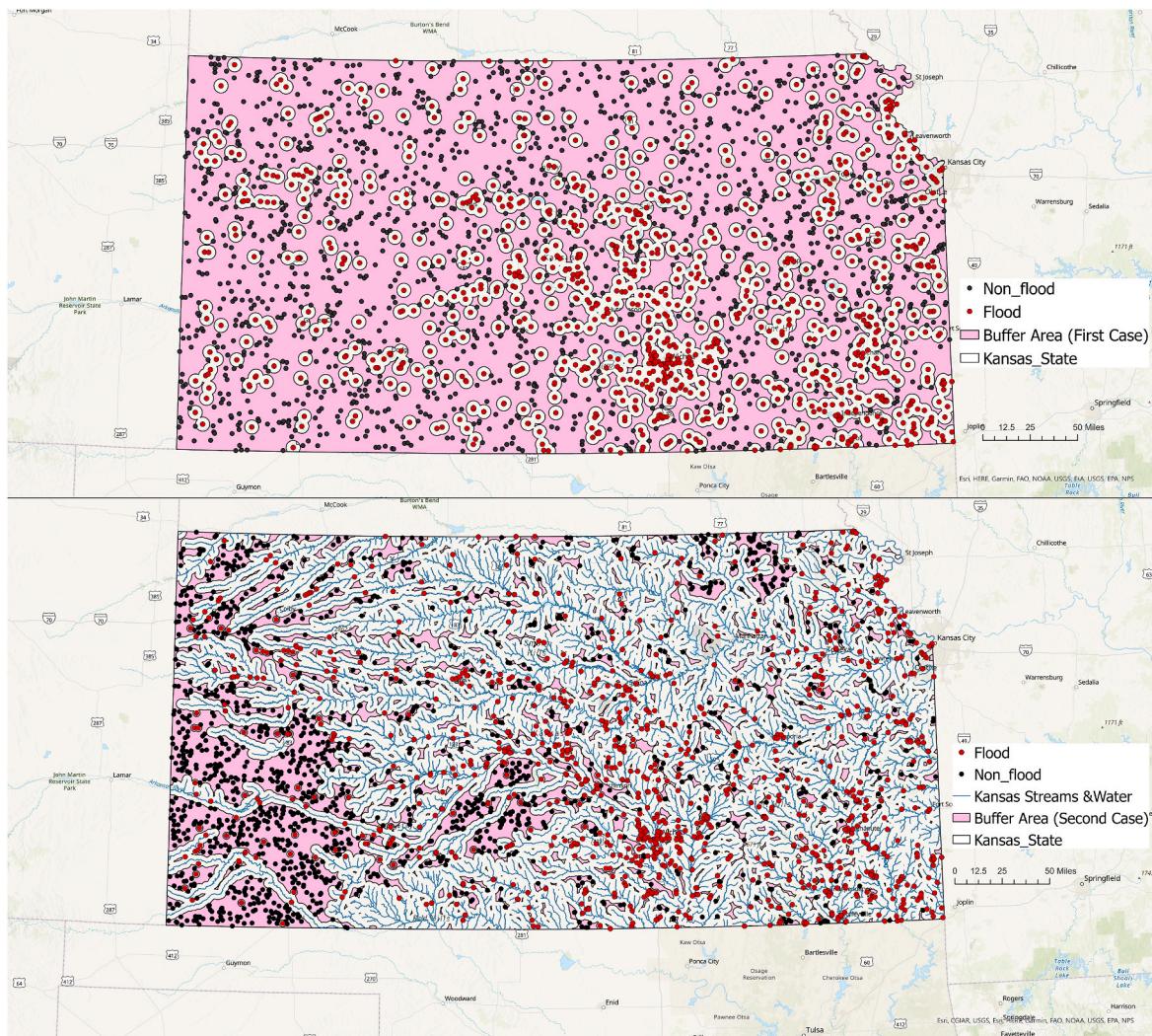


Fig. 2. Map showing the flood inventory points top (first case) and bottom (second case).

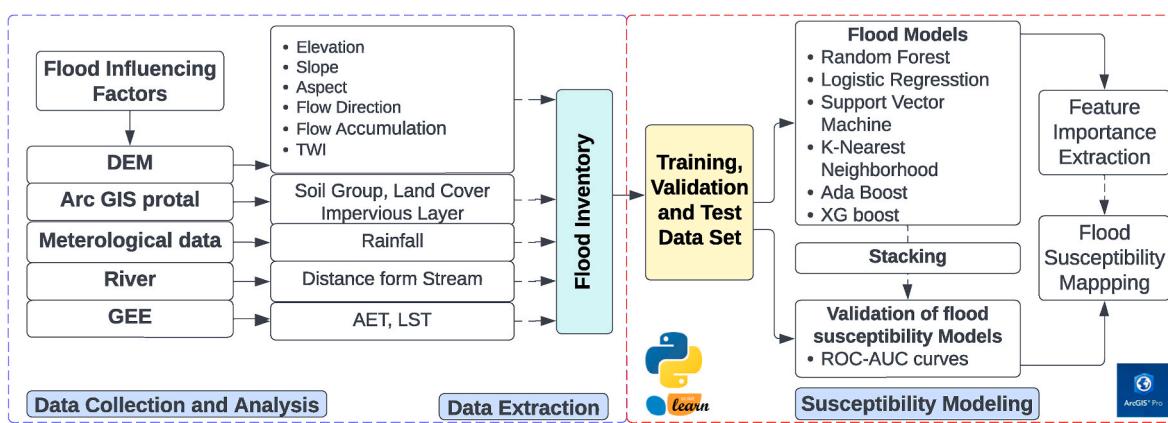


Fig. 3. Overall methodological framework.

and their primary applications are summarized in Table 2. A supervised binary classification method was implemented. First, the data was split into 80% (2363 samples) for ML model training and 20% (591 samples) for testing. Within the training set, 75% served as the final training set and 25% as a validation set for model evaluation and tuning hyperparameters. Thirteen flood-influencing factors were used as independent variables, and their values were derived for flood and non-flood control

points using ArcGIS. The modeling was conducted in a Python environment, refer to supplementary materials (S1 and S2) for the code and data used in the modeling processing.

In such a model, the hyper-parameters should be chosen carefully. Hyperparameters are the variables that are specified while building the ML model and controlling the algorithm's learning process. We used Sklearn's GridSearchCV class on all six machine learning algorithms for

Table 2

Mathematical basis and a brief explanation of the ML algorithms used in the study.

Machine Learning Algorithms	The mathematical basis of the Algorithm	Application of the Algorithm
1 Logistic Regression (LR)	$p = 1/(1 + e^{-z})$ $z = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n$ where p is the probability of occurrence of flooding, w_i ($i = 0, 1, 2, \dots, n$) represents weights and x_i ($i = 0, 1, 2, \dots, n$) are explanatory factors.	Solving binary problems and developed by Cox (1958)
2 Random Forest (RF)	$h(x) = \frac{1}{N} \sum h(x, \theta_k)$ Where x is the controlling factor of flood and k are numbers of the decision tree, θ_k is an independent and identically distributed random variable, and N is the total number of decision trees generated by model.	Uses decision trees on the subset of the data and takes the average to improve the predictive accuracy. Used in Regression and classification tasks and developed by Ho (1995) and Breiman (2001)
3 K- Nearest Neighbor (KNN)	$Distance = (x_{train} - x_{test})^p$ Where p is the Minkowski metric ($p \in [1, 2, 3, 4]$), x_{train} is a training data point and x_{test} is the data point whose class we wish to predict.	Stores all the data and classifies new data points based on the distance function with respect to the stored data (Abu El-Magd et al., 2021)
4 Support Vector Machine (SVM)	$w^T x + b = 0$ Where w is a vector of weights and b is the bias term.	Classifies datasets using hyperplanes in an infinite dimension where each data point is classified as either '1' or '-1' (Cortes and Vapnik, 1995)
5 Adaptive Boosting (Ada Boost)	$H(x) = sign\left(\sum_{t=1}^T \alpha_t h_t(x)\right)$ Where $h_t(x)$ is the prediction by t th weak classifier, α_t is the weight of t th classifier and ϵ_t is the fraction of misclassification by the t th classifier	Developed by Freund and Schapire 1997, combines weak learners to increase performance.
6 Extreme Gradient Boosting (XGB)	$Similarity Score = \frac{(\sum Residual_i)^2}{\sum [p'_i \times (1 - p'_i)] + \lambda}$ where λ is regularization parameter and p'_i is the previous probability computed for the i th training.	Uses tree pruning and handles missing data and uses a minimizing a regularized objective function with a weighting factor for model complexity. Developed by Chen and Guestrin (2016)

optimal fit. It applies a grid search to an array of hyperparameters and cross-validates the model using k-fold cross-validation. The hyperparameters chosen by GridSearchCV for each of the algorithms are reported in Table 3.

We conducted feature importance analysis to understand the physical processes driving flood prediction. When a model calculates feature

importance, it identifies which variables significantly influence its output or predictions. This estimation varies by the model's architecture and hyperparameters. In LR, feature importance is determined by the magnitude of the coefficients: larger absolute values indicate stronger influence (Cox, 1958; Pradhan, 2010). For KNN, feature importance is assessed indirectly through feature scaling and distance calculations, with larger scales having more influence (Abu El-Magd et al., 2021). In SVM, feature importance is linked to the magnitude of weights assigned to support vectors; non-zero coefficients indicate important features for the decision boundary (Cortes and Vapnik, 1995). RF calculates feature importance based on the average decrease in impurity (e.g., Gini impurity) when a feature is used for data splitting, with higher reductions indicating more important features (Breiman, 2001). Ada Boost determines feature importance by tracking how often features are used in weak learners (decision trees) and their contribution to classification accuracy; frequently used features with better classification results are more critical (Freund and Schapire, 1997). Similarly, XGB estimates feature importance using methods like Gini impurity or gain, with features that provide higher impurity reduction or gain being deemed more important (Chen and Guestrin, 2016).

2.6. Accuracy assessment

The ROC (Receiver Operating Characteristic) Curve is a visual tool used to evaluate the performance of a model. The ROC curve and related metrics are commonly used in flood evaluations (Band et al., 2020; Madhuri et al., 2021; Pham and Prakash, 2019; Saravanan and Abijith). The curve represents sensitivity on the y-axis, and specificity on the x-axis. A specific decision threshold can be obtained for each point on the curve to predict the model's accuracy. The area under the ROC curve (AUC) is a quantitative measure used to evaluate the model's performance, where a higher value of AUC (closer to 1) indicates better performance (Band et al., 2020; Pham and Prakash, 2019). The sensitivity of a flooded cell is the probability of it being correctly classified, while the false-negative rate is the complement of sensitivity. Specificity refers to the probability of a non-flooded cell being correctly classified, while the false-positive rate complements specificity. Statistical indices such as sensitivity (SS), specificity (SP), positive predictive (PP), and negative predictive (NP) values are used to assess the performance of flood models. A higher value of these indices indicates better machine learning model performance and vice versa.

2.7. Stack generalization

Stack Generalization or Stacking is a type of ensemble machine learning algorithm that can improve the accuracy of weaker models by combining them through a higher-level model (Wolpert, 1992). This method involves two stages or levels: level-0 and level-1, where base models are trained in level-0, and their outputs are combined into a single score using a meta-model to create the output of level-1 (Sesmero et al., 2015). The goal of stacking is to reduce bias and variance and improve the model's predictive power by integrating multiple models. The two-layer stacking process used in this study is shown in Fig. 4.

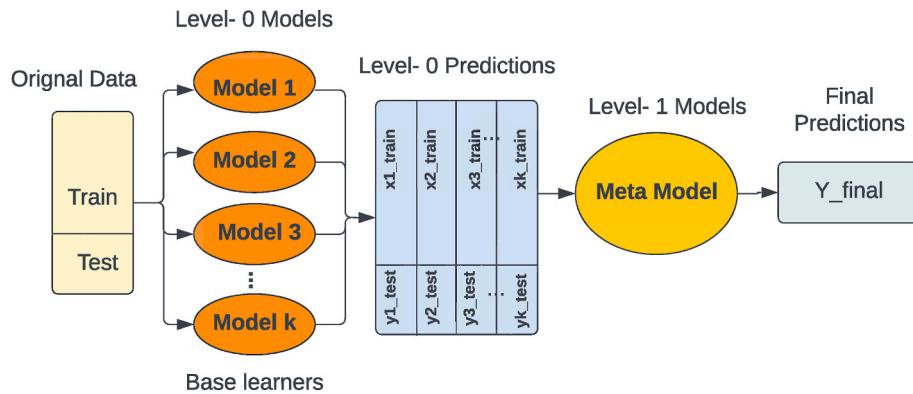
2.8. Flood susceptibility maps

Using the ArcGIS pro weighted overlay tool, flood susceptibility maps were generated. As the weighted overlay is based on the common measurement scale, the raster maps representing the flood-controlling factors were reclassified using values ranging from 1 to 10. The significance values in the weighted overlay are represented by the feature-based importance of each flood-controlling factor extracted from the models. Feature-based and permutation-based importance are vital for extracting the importance of each flood controlling factor as these values represent the relative importance of each feature when predicting each model. Among the various approaches for reclassifying the flood

Table 3

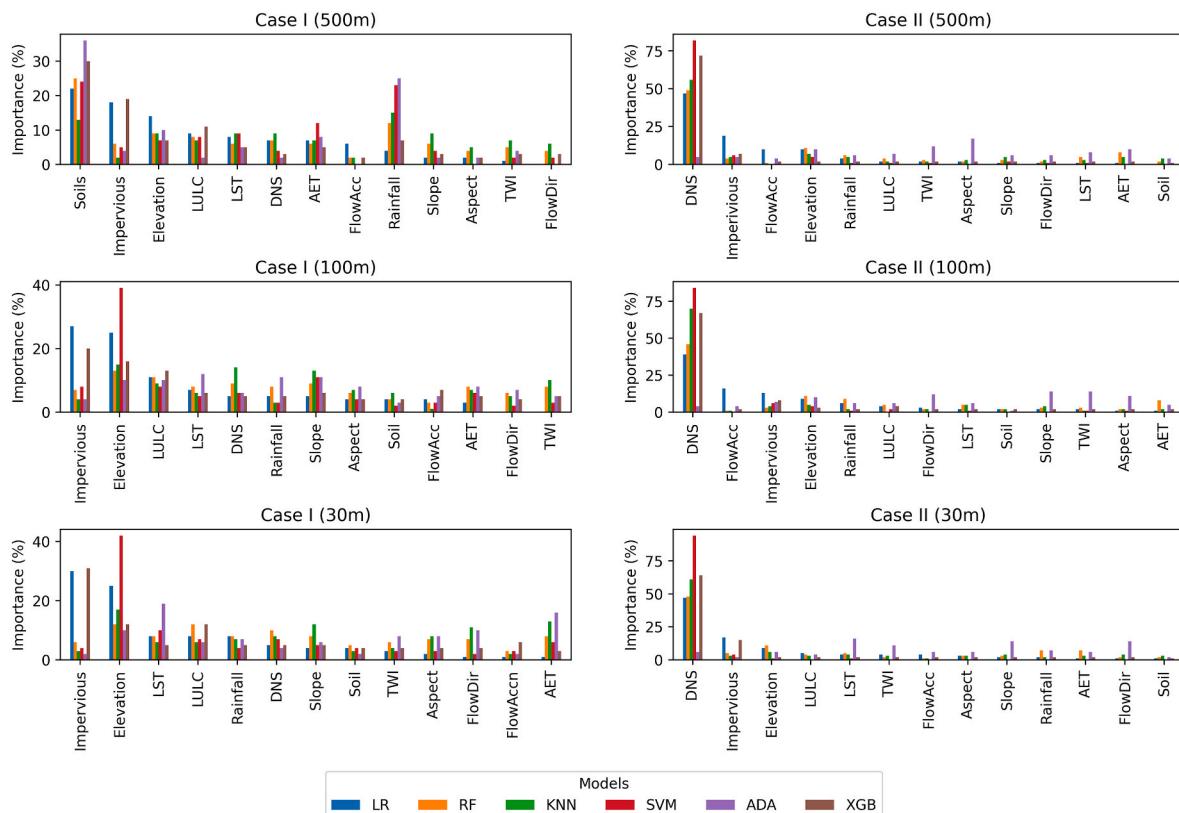
ML Algorithms response for test dataset and associated hyperparameters.

Algorithm	Case I (F1 Score)			Case II (F1-Score)			Considered Hyperparameters
	500 m	100 m	30 m	500 m	100 m	30 m	
LR	0.73	0.7	0.76	0.94	0.91	0.92	{'C': 10, 'penalty': 'l1'}
RF	0.86	0.72	0.79	0.96	0.95	0.94	{'max_depth': 20, 'min_samples_leaf': 1, 'min_samples_split': 2, 'n_estimators': 500}
KNN	0.77	0.73	0.76	0.92	0.92	0.92	{'algorithm': 'kd_tree', 'n_neighbors': 9, 'weights': 'distance'}
SVM	0.78	0.72	0.77	0.94	0.93	0.92	{'C': 10, 'gamma': 'scale', 'kernel': 'rbf'}
Ada Boost	0.83	0.71	0.76	0.94	0.92	0.92	{'algorithm': 'SAMME.R', 'learning_rate': 0.5, 'n_estimators': 200, 'random_state': 42}
XGB	0.88	0.72	0.78	0.96	0.94	0.94	{'colsample_bytree': 1, 'eval_metric': 'auc', 'gamma': 0.2, 'learning_rate': 0.2, 'max_depth': 5, 'min_child_weight': 1}

**Fig. 4.** Process of stack generalization.

susceptible models, Quantile and natural breaks methods are two of the strategies that have been frequently described in flood susceptibility studies (Al-Aizari et al., 2022; Tehrany et al., 2015; Towfiqul Islam et al., 2021). The final flood susceptibility map, created using Jenks' Natural

Breaks method in ArcGIS, was divided into five classes: very high, high, moderate, low, and very low susceptibility levels. (Al-Aizari et al., 2022 Madhuri et al., 2021).

**Fig. 5.** Feature Importance of six algorithms across multiple resolutions in two cases.

3. Results

3.1. Models response

Based on the observed F1 scores, the responses of all the models varied in terms of resolutions in both cases. In both cases, the scores for the 500 m spatial resolutions were better for testing and training sets. Based on the response of the testing set, XGB and RF performance was relatively higher than other algorithms in all spatial resolutions for case I. However, in Case II scenario, all the algorithms performed well with F1 scores above 0.90. The F1 scores of all the models, along with the associated hyperparameters considered for building the models for both cases and various spatial resolutions, are listed in the table below (Table 3).

These results indicate that most of the models performed consistently across both cases, regardless of the resolution level. The use of the same set of hyperparameters allowed for a fair comparison between the models, showcasing their relative strengths and weaknesses. The differences observed in the models' response between the two cases can be attributed to the variation in input values for non-flood inventory control points. The relatively higher accuracy values for all models in Case II were expected, as the non-flooding training and testing samples were created outside buffered areas from water bodies, leading to a more favorable scenario for classification.

3.2. Feature importance

Fig. 5 presents feature importance for each model and scenarios. For Case I, at 500 m resolution, the hydrologic soil group, followed by rainfall, impervious surfaces, elevation, land use/land cover (LULC), and land surface temperature, are the most important variables. However, at 100 m and 30 m resolutions, elevation emerged as the most

important factor, followed by impervious surfaces, distance from streams, rainfall, slope, and so on as shown in Fig. 5. In the Case II scenario, a similar feature importance characteristic was observed. The percentage importance of flood-controlling factors varied across the six ML algorithms. However, regarding the resolution, distance from water bodies emerged as the most important factor for prediction, followed by impervious surfaces, elevation, rainfall, and LULC.

In the case of the LR, from our study, the Hydrologic Soil groups, Elevation, and DNS were the major factors influencing the predictions. Hydrologic Soil groups, Elevation, and DNS were major features affecting the predictions in the KNN model. While in SVM, The Hydrologic Soil group, Elevation, Rainfall, and DNS were the major features. In RF, Hydrologic Soil Groups, DNS, Rainfall, Elevation, and LULC were considered important factors. The features that were deemed important were the Hydrologic Soil group, Rainfall, Elevation, LULC, LST, and AET. If we compare the importance of AdaBoost features with other algorithms for the second scenario, it doesn't classify the DNS as the major important ones; instead, features like slope and flow direction were given more importance. Like others, Hydrologic soil groups, Impervious Surface, Elevation, LULC, and DNS were the significant factors across the multiple resolutions.

3.3. Flood susceptible maps and areas

Flood Susceptibility Maps (FSM) were created in five classes or levels: very low, low, moderate, high, and very high Susceptibility. Significant differences in terms of area within algorithms and various resolutions were observed (Fig. 6). In the first scenario at 500 m resolution, areas ranged from 179.54 to 29,311.72 square miles across susceptibility levels. Logistic Regression (LR) showed a decrease in areas, especially for very low susceptibility. At 100 m resolution, very high susceptibility areas decreased, but at 30 m, they increased.

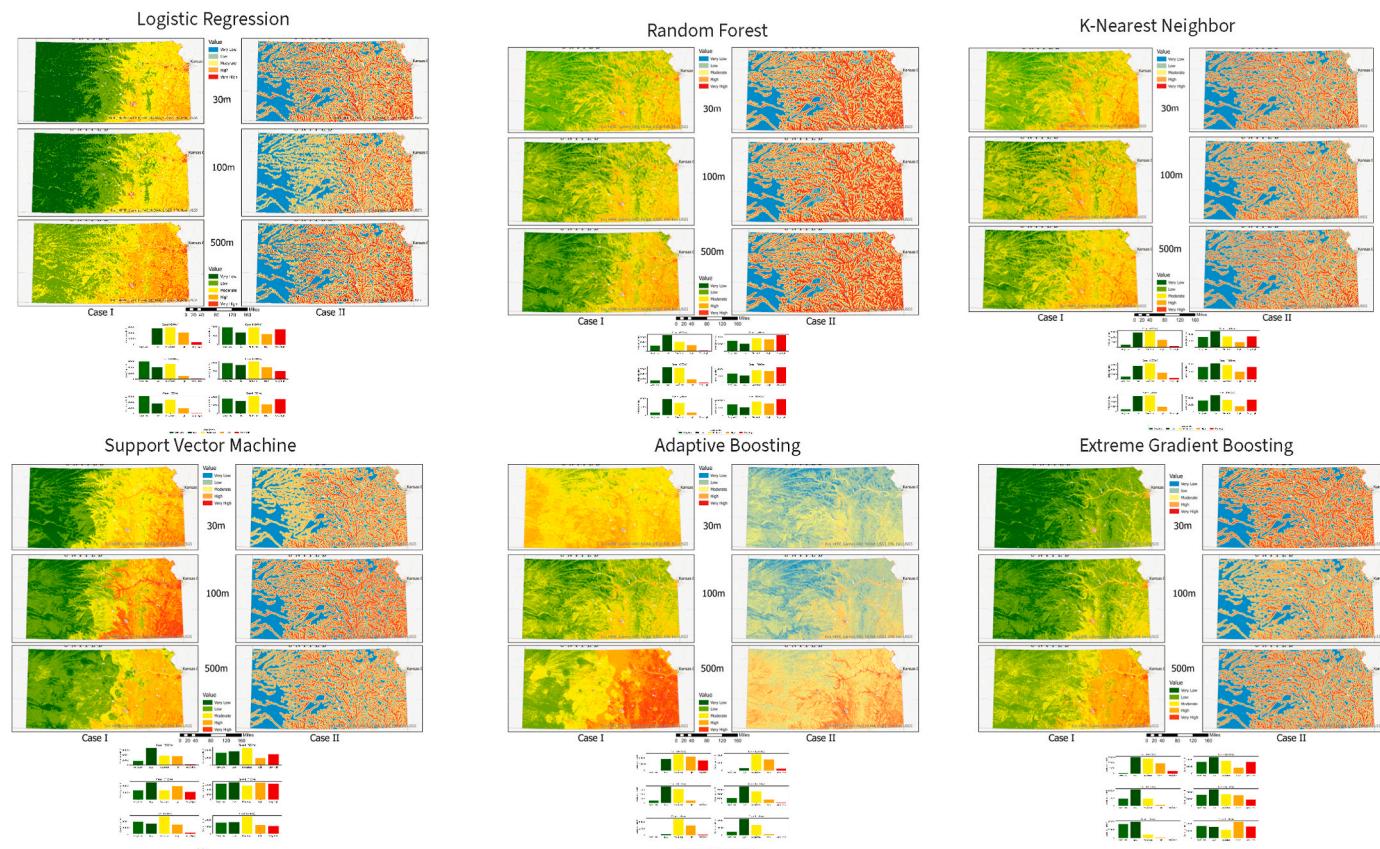


Fig. 6. Flood Susceptibility Maps and Area comparison of all six ML Algorithms over multiple resolutions in two scenarios.

Random Forest (RF) at 500 m exhibited varying areas across susceptibility levels (1677 to 34,809 sq. miles). Unlike LR, RF showed a decrease in very high susceptibility areas with increasing resolution. For the second case, high and very high susceptibility areas were notably higher (Fig. 6). At 500 m, K-Nearest Neighbor (KNN) showed areas ranging from 1858 to 31,244 sq. miles for different susceptibility levels. Similar to RF, high and very high susceptibility classes decreased in the first scenario. In the second case, susceptibility classes first increased and then decreased with finer resolutions. Support Vector Machine (SVM) at 500 m displayed varying areas (1581 to 34,471 sq. miles) across susceptibility levels. Like KNN, the increase and decrease in areas

were observed. Case II scenario showed significantly higher areas for very highly susceptible regions (Fig. 6). Adaptive Boosting (Ada Boost) at 500 m had areas ranging from 93 to 25,328 sq. miles. Like LR, it struggled with very low susceptibility classes. Finer resolutions showed a decline in very high susceptibility areas in the first case, while the second scenario exhibited significant increases in the low susceptibility class. In Case I, moderate, high, and very high susceptibility areas decreased with finer resolutions using Extreme Gradient Boosting (XGB). At 500 m, areas ranged from 291.70 to 30,390.78 sq. miles. Case II scenario exhibited varying areas, with higher coverage for all susceptibility areas except very low susceptibility.

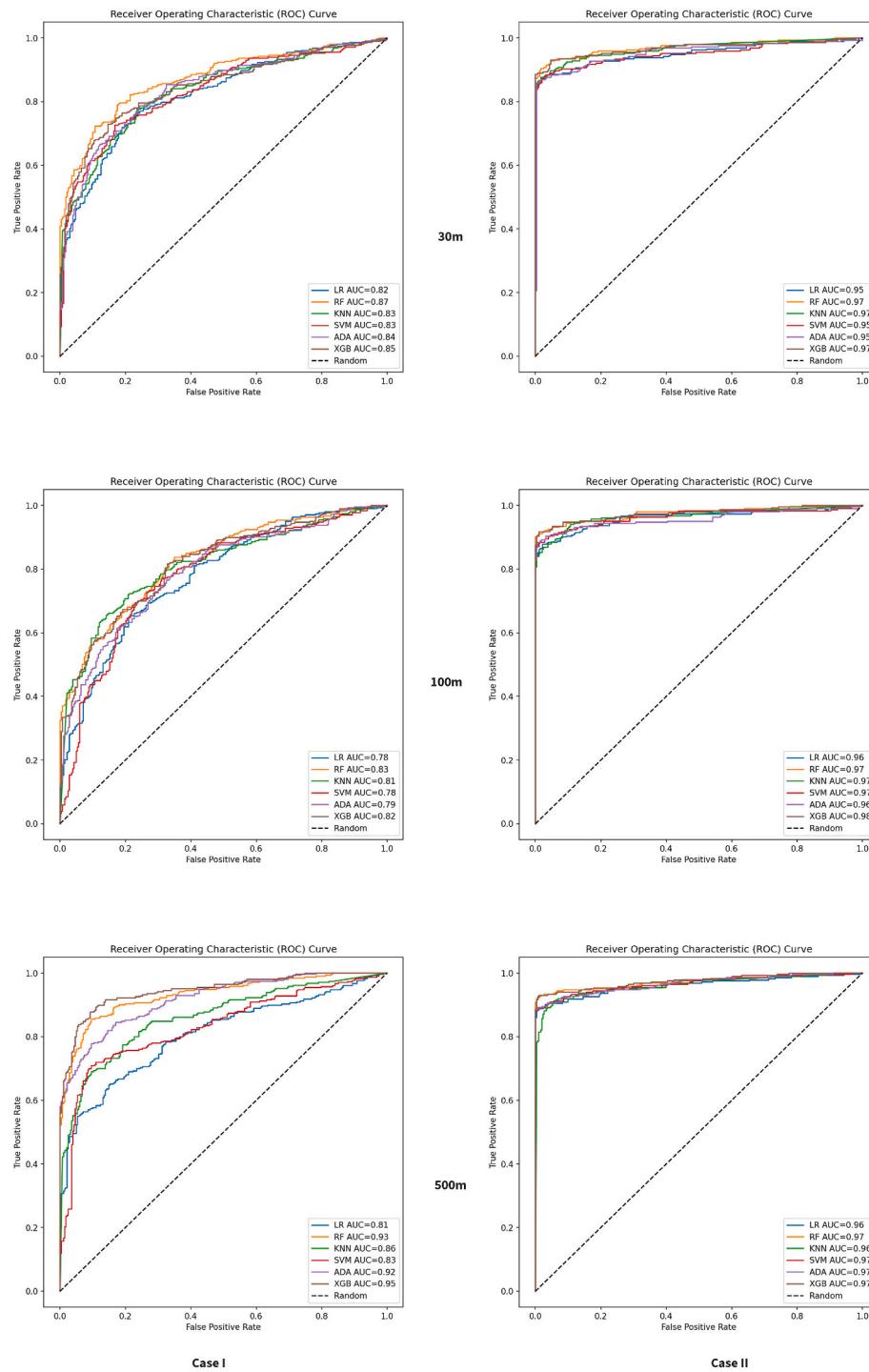


Fig. 7. ROC-AUC Curve of all Algorithms across multiple resolutions in two scenarios.

3.4. Accuracy assessment

A shift in the increase of AUC scores was observed when the training flood point inventories were created outside the stream buffer (Case II scenario). In Case I, at a resolution of 500 m, the AUC scores were higher than other resolutions. LR achieves a reasonable AUC score of 0.81, while RF outperforms LR with an AUC score of 0.93. KNN, SVM, and Ada boost algorithms obtained scores of 0.86, 0.83, and 0.92. The XGB model received the highest score of 0.95, indicating that the model performed well to classify the pixels as flood and non-flood. At a resolution of 100 m, RF maintains its dominance with an AUC score of 0.83, while LR and KNN achieve lower scores of 0.78 and 0.81, respectively. The XGB and Ada boost scores dropped to 0.82 and 0.79, respectively. In a finer resolution (30 m), RF and XGB obtained scores of 0.93 and 0.91, outperforming the remaining algorithms (Fig. 7).

In contrast to the second scenario, all the algorithms obtained high AUC scores between 0.95 and 0.98, showcasing their strengths in capturing patterns at multiple resolutions. At 500m, RF and XGB stand out again with an exceptional AUC score of 0.97, showcasing their robustness in capturing patterns and predicting susceptibility accurately. SVM also performs well in this case, achieving a high AUC score of 0.97. At 100m, RF and ADA both achieve impressive AUC scores of 0.97, indicating their effectiveness in predicting susceptibility at a finer resolution. However, XGB showcases its superiority at 100m with an AUC score of 0.98. At 30m, RF, KNN, ADA, and XGB consistently deliver high AUC scores (AUC score 0.97), emphasizing their efficacy in predicting susceptibility at this fine resolution.

3.5. Stack generalization

The improved performance was observed in most cases while using the RF algorithm as the meta-learner (Fig. 8). However, when XGB was used as a meta-learner, it only outperformed the algorithms with low AUC scores like LR, SVM, and KNN. This concludes that the stack generalization positively improved the model response with lower AUC values, thereby making the final stacked model more robust as the stacking process incorporates the output from each algorithm used as base learners.

4. Discussion

4.1. Feature importance

Machine learning models leverage data features to make predictions,

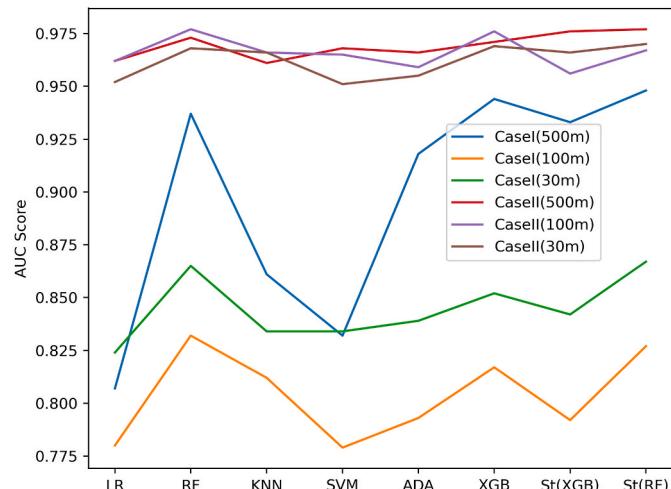


Fig. 8. XGB and RF-based Stack Generalization AUC-ROC scores plot for all algorithms over multiple resolutions and scenarios.

but not all features hold equal weight (Cox, 1958; Pradhan, 2010). Different models employ unique techniques to identify the most influential features, further influenced by the model's architecture and hyperparameter settings. For instance, Linear Regression (LR) assigns a score to each feature based on its absolute value (Cox, 1958; Pradhan, 2010). Higher scores indicate a stronger influence on the prediction. In our study, LR identified Hydrologic Soil Groups, Elevation, and Drainage Network System (DNS) as key factors. K-Nearest Neighbors (KNN) takes a different approach. Features with larger scales might have a bigger impact on distance calculations, potentially making them more important (Abu El-Magd et al., 2021). Here too, Hydrologic Soil Groups, Elevation, and DNS were prominent. Support Vector Machines (SVM) focus on "support vectors" to define boundaries between categories (Cortes and Vapnik, 1995). Features with non-zero weights in the model are considered support vectors and are more crucial for predictions. Using SVM, Hydrologic Soil Group, Elevation, Rainfall, and DNS emerged as important factors. Tree-based models like Random Forest (RF) and Extreme Gradient Boosting (XGB) directly assess feature importance (Breiman, 2001; Chen and Guestrin, 2016). They track a feature's contribution to separating data points within the model. Features leading to a larger decrease in randomness are considered more important. RF highlighted Hydrologic Soil Groups, DNS, Rainfall, Elevation, and Land Use Land Cover (LULC) as key factors, while XGB identified Hydrologic Soil Groups, Impervious Surface, Elevation, LULC, and DNS. AdaBoost, another model type, analyzes feature importance by monitoring their use and impact on classification errors (Freund and Schapire, 1997). Features used more frequently and leading to better classifications are considered more critical. Interestingly, AdaBoost prioritized features like Rainfall, Slope, and Flow Direction over Drainage Network System (DNS) compared to other models. Overall, analyzing feature importance allows us to understand the inner workings of machine learning models. By identifying the most influential features, we can refine models, focus on collecting relevant data, and ultimately achieve more accurate and interpretable predictions. Additionally, across the multiple resolutions and two scenarios examined in this study, we found DNS, Hydrologic soil group, Elevation, and Impervious Surface were the most influential variables in modeling flood hazard. These findings are consistent with recent studies highlighting the importance of these factors in flood occurrence (Pahm et al., 2021; Blum et al., 2020; Chu et al., 2020; Arabameri et al., 2020; Tehrany et al., 2015). Physiographic characteristics, such as soil and elevation, play a crucial role in shaping a region's hydrological regime and indirectly control its climate. These factors significantly influence various hydrological aspects, including flood volume, runoff coefficient, flood intensity, and discharge vulnerability (Arabameri et al., 2020; Abdel Hamid et al., 2020). Proximity to the river emerged as another significant factor influencing flood risk, with areas near streams showing high sensitivity to floods. Several researchers have conducted studies demonstrating the importance of the distance from the river as a critical factor affecting flood risk, with most areas near rivers exhibiting high sensitivity to floods (Pahm et al., 2021; Chu et al., 2020). Furthermore, several research have mentioned the impact of impervious surfaces on flood magnitude (Blum et al., 2020; Du et al., 2015). This consistency in ranking the most important features across resolutions suggests their robust influence on flooding prediction.

We observed varied algorithm responses influenced by different flood-controlling variables, underscoring the complexity of flood prediction and the sensitivity of algorithms to data processing (e.g., resampling and the making creation of non-flood inventory points). The variation within resolutions is expected due to data resampling. In our study, variable such as Aspect, Slope, Elevation, Flow Direction, Flow Accumulation, and TWI were derived from 30m DEM, Hydrologic Soil groups, Impervious layer, LULC, and DNS were resampled to coarser resolution. Variables like LST and AET were resampled to finer resolutions. Rainfall data was interpolated and generated at multiple resolutions. These resampled products may reflect different values, potentially

introducing variations due to information losses during resampling. Similarly, the variation observed in the second scenario is due to the differences within the non-flood inventory points. As the points are created outside the flood buffer areas, exhibited significantly different values, making it easier for algorithms to predict classifications based on DNS.

4.2. Algorithm's response and model stacking

The varied response of the algorithms across different resolutions highlight the importance of analyzing predictor variables at multiple scales. In the first scenario, coarser resolution dataset achieved high AUC score and F1 scores, although the differences were minimal. This is consistent with observation of improved scores in coarser resolution mentioned in (Tavakkoli Piralilou, 2022) for wildfire susceptibility prediction. In the second scenario, the model performance was consistent across all resolutions. This consistency likely stems from the creation of non-flood inventory points outside stream buffers, making it easier for algorithms to classify based on distance from streams. Nonetheless, high-resolution images for the analysis are recommended, the various studies mentioned the improvement of spatial resolution as a justified means to improve the accuracy of findings (Saha et al., 2021; Islam et al., 2023). Since most of our dataset was resampled from 30 m resolution, it did not capture fine-scale variations, leading to varied observations.

Stack generalization, which integrates multiple models to reduce bias and variance, proved effective in our study, particularly with an RF-based meta-model stacking approach. This technique has shown to enhance predictive accuracy in various applications, including flood susceptibility mapping, rainfall prediction, landslide assessment, and stream flow forecasting.

4.3. Flood susceptible maps and area

Variation in area were observed as weights were assigned according to percentage of feature importance of flood-controlling variables while developing the susceptibility maps. Since the importance of these factors varied with each algorithms, the flood-susceptibility maps generated from each model showed differences in area across the two scenarios and multiple resolutions. Such variations is expected, as ML algorithms respond differently to changes in input data. This is in accord with observations associated with the variation of the regions in different ML algorithms were observed in the studies conducted by Al-Azari et al., 2022; Saravanan and Abijith, 2022; El-Haddad et al., 2021; Saha et al., 2021; Towfiqul Islam et al., 2021; Khosravi et al., 2019.

The models chosen in this study proved viable and effective, considering their performance and ease of interpretation. However, a limitation of the study was the lack of critical hydrological data, such as flood depth, velocity, and discharge, which made developing a robust model challenging. Flood modeling is a complex task that involves uncertainties. Nevertheless, when reliable historical flood inventory maps are available, machine learning algorithms can effectively address these uncertainties (Janizadeh et al., 2019). To mitigate uncertainties, we created a flood inventory map of the study area using flood events from the NOAA storm database.

The proposed models offer a valuable and innovative approach combined with multi-criteria analysis to manage flood threats in humid and semi-arid regions like Kansas. However, like any other study, the results of this research are susceptible to errors and uncertainties. Factors such as the subjective classification of flood-influencing factors and the selection of performance indicators subjected to temporal variations contribute to these uncertainties. Further research is required to investigate the impact of these uncertainties on the final flood susceptibility maps and subsequent decision-making processes. Additionally, analyzing the two scenarios in the research, we can emphasize the variation in response of models posed due to variations in non-flooding

control points. Future studies should consider incorporating other flood factors, such as daily or sub-daily rainfall, collaborative classification of flood factors with stakeholders, sensitivity analysis of observed dataset classification, and evaluation of alternative goodness-of-fit measures for assessing the efficacy of the six methods (Ahmadisharaf et al., 2017, 2019; Janizadeh et al., 2019).

5. Conclusion

This study investigated flood susceptibility using machine learning and two scenarios for flood control points (recorded flood events vs. randomly generated non-flooded points). Six algorithms were employed to classify flood locations and develop flood susceptibility maps within a spatial framework. RF and XGB achieved superior performance compared to other algorithms across resolutions. Stacking, which combines multiple models, further improved predictions. Flood susceptibility maps were generated using these findings. The resulting maps highlighted highly vulnerable areas and can inform flood risk management strategies. However, limitations include potential false alarms and the need for more comprehensive data (e.g., flood depth, velocity, discharge). Additionally, the selection of flood control points and classification of flood-influencing factors can introduce uncertainties. Future research can incorporate these factors, refine flood susceptibility modeling, and evaluate the impact of uncertainties on decision-making.

CRediT authorship contribution statement

Zelalem Demissie: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Prashant Rimal:** Writing – original draft, Visualization, Software, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Wondwosen M. Seyoum:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Project administration, Methodology, Investigation, Conceptualization. **Atri Dutta:** Visualization, Funding acquisition, Conceptualization. **Glen Rimmington:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.acags.2024.100183>.

References

- Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Sci. Data* 5, 170191.
- Abdel Hamid, H.T., Wenlong, W., Qiaomin, L., 2020. Environmental sensitivity of flash flood hazard using geospatial techniques. *Glob. J. Environ. Sci. Manag.* 6, 31–46. <https://doi.org/10.22034/GJESM.2020.01.03>.
- Abedi, R., Costache, R., Shafizadeh-Moghadam, H., Pham, Q.B., 2022. Flash-flood susceptibility mapping based on XGBoost, random forest and boosted regression trees. *Geocarto Int.* 37, 5479–5496. <https://doi.org/10.1080/10106049.2021.1920636>.
- Abu El-Magd, S.A., Ali, S.A., Pham, Q.B., 2021. Spatial modeling and susceptibility zonation of landslides using random forest, naive bayes and K-nearest neighbor in a

- complicated terrain. *Earth Sci. Inform.* 14, 1227–1243. <https://doi.org/10.1007/s12145-021-00653-y>.
- Adams, G.I., 1903. The Ogallala formation. *U. S. Geol. Surv. Water Supply Pap.* 104, 21–32.
- Ahmadvlou, M., Karimi, M., Alizadeh, S., Shirzadi, A., Parvinnejad, D., Shahabi, H., Panahi, M., 2019. Flood susceptibility assessment using integration of adaptive network-based fuzzy inference system (ANFIS) and biogeography-based optimization (BBO) and BAT algorithms (BA). *Geocarto Int.* 34, 1252–1272. <https://doi.org/10.1080/10106049.2018.1474276>.
- Ahmadisharaf, E., Camacho, R.A., Zhang, H.X., Hantush, M.M., Mohamoud, Y.M., 2019. Calibration and validation of watershed models and advances in uncertainty analysis in TMDL studies. *J. Hydrol. Eng.* 24, 03119001 [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001794](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001794).
- Ahmadisharaf, E., Kalyanapu, A., Chung, E.-S., 2017. Sustainability-based flood hazard mapping of the Swannanoa river watershed. *Sustainability* 9, 1735. <https://doi.org/10.3390/su9101735>.
- Al-Aizari, A.R., Al-Masny, Y.A., Ayddha, A., Zhang, J., Ullah, K., Islam, A.R.M.T., Habib, T., Kaku, D.U., Nizeyimana, J.C., Al-Shaibah, B., Khalil, Y.M., Al-Hameedi, W.M.M., Liu, X., 2022. Assessment analysis of flood susceptibility in tropical desert area: a case study of Yemen. *Remote Sens. (Basel)* 14, 4050. <https://doi.org/10.3390/rs14164050>.
- Arabameri, A., Saha, S., Chen, W., Roy, J., Pradhan, B., Bui, D.T., 2020. Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques. *J. Hydrol. (Amst.)* 587, 125007. <https://doi.org/10.1016/j.jhydrol.2020.125007>.
- Band, S.S., Janizadeh, S., Chandra Pal, S., Saha, A., Chakrabortty, R., Mellesse, A.M., Mosavi, A., 2020. Flash flood susceptibility modeling using new approaches of hybrid and ensemble tree-based machine learning algorithms. *Remote Sens. (Basel)* 12, 3568. <https://doi.org/10.3390/rs12213568>.
- Blum, A.G., Ferraro, P.J., Archfield, S.A., Ryberg, K.R., 2020. Causal effect of impervious cover on annual flood magnitude for the United States. *Geophys. Res. Lett.* 47 <https://doi.org/10.1029/2019gl086480>.
- Bowen, M.W., Juracek, K.E., 2011. Assessment of the geomorphic effects of large floods using streamgage data: the 1951 floods in eastern Kansas, USA. *Phys. Geogr.* 32, 52–77. <https://doi.org/10.2747/0272-3646.32.1.52>.
- Breiman, L., 2001. Random forests. *Mach. Learn.* 45 (1), 5–32. <https://doi.org/10.1023/a:1010933404324>.
- Bui, D.T., Panahi, M., Shahabi, H., Singh, V.P., Shirzadi, A., Chapi, K., Khosravi, K., Chen, W., Panahi, S., Li, S., Ahmad, B.B., 2021. Author correction: novel hybrid evolutionary algorithms for spatial prediction of floods. *Sci. Rep.* 11, 15152 <https://doi.org/10.1038/s41598-021-93957-4>.
- Bui, Q.-T., Nguyen, Q.-H., Nguyen, X.L., Pham, V.D., Nguyen, H.D., Pham, V.-M., 2020. Verification of novel integrations of swarm intelligence algorithms into deep learning neural network for flood susceptibility mapping. *J. Hydrol. (Amst.)* 581, 124379. <https://doi.org/10.1016/j.jhydrol.2019.124379>.
- Centre for Research on the Epidemiology of Disasters (CRED), 2022. Disaster Data: A Balanced Perspective. Centre for Research on the Epidemiology of Disasters. (2022). CRED, Brussels.
- Chapi, K., Singh, V.P., Shirzadi, A., Shahabi, H., Bui, D.T., Pham, B.T., Khosravi, K., 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ. Model. Software* 95, 229–245. <https://doi.org/10.1016/j.envsoft.2017.06.012>.
- Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, New York, NY, USA. <https://doi.org/10.1145/2939672.2939785>.
- Chu, H., Wu, W., Wang, Q.J., Nathan, R., Wei, J., 2020. An ANN-based emulation modelling framework for flood inundation modelling: application, challenges and future directions. *Environ. Model. Software* 124, 104587. <https://doi.org/10.1016/j.envsoft.2019.104587>.
- Clement, R.W., 1987. Floods in Kansas and Techniques for Estimating Their Magnitude and Frequency on Unregulated Streams. US Geological Survey. <https://doi.org/10.3133/wri74008>.
- Clement, R.W., Johnson, D.G., 1982. Flood of june 15, 1981. In: Great Bend and Vicinity, Central Kansas. U.S. Geological Survey. <https://doi.org/10.3133/wri824123>.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297. <https://doi.org/10.1007/bf00994018>.
- Costache, R., Zaharia, L., 2017. Flash-flood potential assessment and mapping by integrating the weights-of-evidence and frequency ratio statistical methods in GIS environment – case study: Bâsca Chiojdului River catchment (Romania). *J. Earth Syst. Sci.* 126 <https://doi.org/10.1007/s12040-017-0828-9>.
- Cox, D.R., 1958. The regression analysis of binary sequences. *J. Roy. Stat. Soc. Ser. B Stat. Methodol.* 20, 215–232. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>.
- Daly, C., Taylor, C.H., Gibson, W.P., 1997. The PRISM approach to mapping precipitation and temperature. In: Reprints of 10th Conference on Applied Climatology, pp. 20–23.
- National Centers for Environmental Information (NCEI), 2022. Billion-Dollar Weather and Climate Disasters [WWW Document]. Noaa.gov. <https://www.ncei.noaa.gov/access/billions/summary-stats>.
- [Dataset] National Oceanic and Atmospheric Administration (NOAA) and National Centers for Environmental Information (NCEI), 2022. Storm Events Database [WWW Document]. Noaa.gov. <https://www.ncdc.noaa.gov/stormevents/>. (Accessed 6 September 2022).
- Demissie, Z.S., Bedassa, G., Rattani, A., Nigussie, W., Kebede, H., Muhabaw, Y., Haridasan, S., 2023. The significance of volcanic segments and rifts in faults characterization within the Amagmatic graben of the Afar Depression, Ethiopia. *Journal of Structural Geology* 174, 104914. <https://doi.org/10.1016/j.jsg.2023.104914>.
- Demissie, Z.S., Rimmington, G., 2022. Surface displacements mechanism of the Dobi Graben from ASAR time-series analysis of InSAR: Implications for the tectonic setting in the central afar depression, Ethiopia. *Remote Sensing* 14 (8), 1845. <https://doi.org/10.3390/rs14081845>.
- Dewitz, J., 2021. National land cover database (NLCD) 2019 products. U.S. Geological Survey data release. Available online at: <https://www.mrlc.gov/data/nlcd-2019-land-cover-conus>.
- Du, S., Shi, P., Van Rompaey, A., Wen, J., 2015. Quantifying the impact of impervious surface location on flood peak discharge in urban areas. *Nat. Hazards* 76, 1457–1471. <https://doi.org/10.1007/s11069-014-1463-2>.
- Downton, M.W., Miller, J.Z.B., Pielke Jr., R.A., 2005. Reanalysis of U.S. national weather service flood loss database. *Nat. Hazards Rev.* 6, 13–22. [https://doi.org/10.1061/\(ASCE\)1527-6988\(2005\)6:1\(13\)](https://doi.org/10.1061/(ASCE)1527-6988(2005)6:1(13).
- El-Haddad, B.A., Youssef, A.M., Pourghasemi, H.R., Pradhan, B., El-Shater, A.-H., El-Khashab, M.H., 2021. Flood susceptibility prediction using four machine learning techniques and comparison of their performance at Wadi Qena Basin, Egypt. *Nat. Hazards* 105, 83–114. <https://doi.org/10.1007/s11069-020-04296-y>.
- El-Magd, S.A.A., Pradhan, B., Alamri, A., 2021. Machine learning algorithm for flash flood prediction mapping in Wadi El-Laqaite and surroundings, Central Eastern Desert, Egypt. *Arabian J. Geosci.* 14 <https://doi.org/10.1007/s12517-021-06466-z>.
- Farhadi, H., Najafzadeh, M., 2021. Flood risk mapping by remote sensing data and random forest technique. *Water (Basel)* 13, 3115. <https://doi.org/10.3390/w13213115>.
- Freund, Y., Schapire, R.E., 1997. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.* 55, 119–139. <https://doi.org/10.1006/jcss.1997.1504>.
- Follansbee, R., Spiegel, J.B., 1935. Flood on Republican and Kansas Rivers May and June 1935.
- Goodin, D.G., Mitchell, J.E., Knapp, M.C., Bivens, R.E., 2004. Climate variability in Kansas. *Trans. Kans. Acad. Sci.* 107 (1/2), 17–32.
- Hauth, L.D., Carswell, W.J., 1978a. Floods in Kansas City. Missouri and Kansas September 12–13, 1977.
- Hauth, L.D., Carswell, L.D., 1978b. The 1977 Kansas Floods. Hauth, L.D., Carswell, L.D. (1978). U.S. Geological Survey Professional Paper 1089.
- Heimann, D.C., Weilert, T.E., Kelly, B.P., Studley, S.E., 2014. Flood-inundation maps and wetland restoration suitability index for the blue river and selected tributaries. In: *Geological Survey Scientific Investigations*. Kansas City, Missouri; U.S.
- Ho, T., 1995. Random decision forests (PDF). In: *Proceedings of the 3rd International Conference on Document Analysis and Recognition*. Montreal, QC, pp. 14–16.
- Islam, A.R.M.T., Bappi, M.M.R., Alqadhi, S., Bindajam, A.A., Mallick, J., Talukdar, S., 2023. Improvement of flood susceptibility mapping by introducing hybrid ensemble learning algorithms and high-resolution satellite imageries. *Nat. Hazards* 119, 1–37. <https://doi.org/10.1007/s11069-023-06106-7>.
- Janizadeh, S., Avand, M., Jaaafari, A., Van Phong, T., Bayat, M., Ahmadisharaf, E., Prakash, I., Pham, B.T., Lee, S., 2019. Prediction success of machine learning methods for flash flood susceptibility mapping in the Tafresh watershed, Iran. *Sustainability* 11, 5426. <https://doi.org/10.3390/su11195426>.
- Kansas Biological Survey, University of Kansas, 2022. National Hydrography Dataset (NHD). Available online at: <https://biosurvey.ku.edu/>.
- Kansas Department of Agriculture, 2022. Dam history in Kansas [WWW document]. Agriculture.ks.gov. URL https://agriculture.ks.gov/docs/default-source/dwr-ws-fact-sheets/history-of-dams.pdf?sfvrsn=ffd7aac1_6. (Accessed 25 September 2022).
- Kansas, 2022. USA - Yearly & Monthly Weather Forecast. Available online at: <https://www.weather.usa.com/en/kansas-usa-climate>.
- Khosravi, K., Shahabi, H., Pham, B.T., Adamowski, J., Shirzadi, A., Pradhan, B., Dou, J., Ly, H.-B., Gróf, G., Ho, H.L., Hong, H., Chapi, K., Prakash, I., 2019. A comparative assessment of flood susceptibility modeling using Multi-Criteria Decision-Making Analysis and Machine Learning Methods. *J. Hydrol. (Amst.)* 573, 311–323. <https://doi.org/10.1016/j.jhydrol.2019.03.073>.
- Liu, T., Shi, P., Fang, J., 2022. Spatiotemporal variation in global floods with different affected areas and the contribution of influencing factors to flood-induced mortality (1985–2019). *Nat. Hazards* 111, 2601–2625. <https://doi.org/10.1007/s11069-021-05150-5>.
- Loun, J.M., 2017. Summary of Hydrologic Conditions in Kansas, Water Year 2016. Fact Sheet. <https://doi.org/10.3133/f20173020>.
- Madhuri, R., Sistla, S., Srivinasa Raju, K., 2021. Application of machine learning algorithms for flood susceptibility assessment and risk management. *J. Water Clim. Chang.* 12, 2608–2623. <https://doi.org/10.2166/wcc.2021.051>.
- Mallakpour, I., Villarini, G., 2015. The changing nature of flooding across the central United States. *Nat. Clim. Change* 5, 250–254. <https://doi.org/10.1038/nclimate2516>.
- Meliho, M., Khattabi, A., Asinyo, J., 2021. Spatial modeling of flood susceptibility using machine learning algorithms. *Arabian J. Geosci.* 14 <https://doi.org/10.1007/s12517-021-08610-1>.
- National Centers for Environmental Information (NCEI), 2022. National Centers for Environmental Information. (2022). Billion-Dollar Weather and Climate Disasters. National Oceanic and Atmospheric Administration.
- National Weather Service, 2022. NWS Preliminary US Flood Fatality Statistics, "2022". Weather.gov. [WWW Document] <https://www.weather.gov/arkx/usflood>. (Accessed 8 June 2022)
- National Oceanic and Atmospheric Administration (NOAA) and National Centers for Environmental Information (NCEI), 2022. Storm Events Database [WWW Document], Noaa.gov.. <https://www.ncdc.noaa.gov/stormevents/>. (Accessed 6 September 2022).
- Pham, B.T., Prakash, I., 2019. Evaluation and comparison of LogitBoost Ensemble, Fisher's Linear Discriminant Analysis, logistic regression, and support vector

- machines methods for landslide susceptibility mapping. *Geocarto Int.* 34, 316–333. <https://doi.org/10.1080/10106049.2017.1404141>.
- Pham, Q.B., Pal, S.C., Chakrabortty, R., Norouzi, A., Golshan, M., Ogunrinde, A.T., Janizadeh, S., Khedher, K.M., Anh, D.T., 2021. Evaluation of various boosting ensemble algorithms for predicting flood hazard susceptibility areas. *Geomatics, Nat. Hazards Risk* 12, 2607–2628. <https://doi.org/10.1080/19475705.2021.1968510>.
- Plake, S. (Ed.), 2019. Two Kansas City Area Reservoirs at Full Capacity; More Levee Breaches Expected. Scripps Local Media. Scripps Local Media, 27 May 2019.
- Piralilou, Sepideh Tavakkoli, 2022. Data Fusion Scenarios to Trustworthy Spatial Modelling of Natural Hazards. PhD diss.. Paris-Lodron Universität Salzburg.
- Pradhan, B., 2010. Flood susceptible mapping and risk area delineation using logistic regression, GIS and remote sensing. *J. Spatial Hydrol.* 9, 1–18.
- Qaiser, K., Ahmad, S., Johnson, W., 2012. Urbanization impacts on flood risks in the Kansas River basin. Qaiser, K., Ahmad, S., Johnson, W. (2012). *J. Hydrol. Eng.* 17 (4), 604–614.
- Saha, T.K., Pal, S., Talukdar, S., Debanshi, S., Khatun, R., Singha, P., Mandal, I., 2021. How far spatial resolution affects the ensemble machine learning based flood susceptibility prediction in data sparse region. *J. Environ. Manag.* 297, 113344 <https://doi.org/10.1016/j.jenvman.2021.113344>.
- Saravanan, S., Abijith, D., 2022. Flood susceptibility mapping of Northeast coastal districts of Tamil Nadu India using Multi-source Geospatial data and Machine Learning techniques. *Geocarto Int.* 37, 15252–15281. <https://doi.org/10.1080/10106049.2022.2096702>.
- Sesmero, M.P., Ledezma, A.I., Sanchis, A., 2015. Generating ensembles of heterogeneous classifiers using Stacked Generalization: generating ensembles of heterogeneous classifiers. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* 5, 21–34. <https://doi.org/10.1002/widm.1143>.
- Shafizadeh-Moghadam, H., Valavi, R., Shahabi, H., Chapi, K., Shirzadi, A., 2018. Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J. Environ. Manag.* 217, 1–11. <https://doi.org/10.1016/j.jenvman.2018.03.089>.
- Solomatine, D., See, L.M., Abrahart, R.J., 2008. *Data-driven modelling: concepts, approaches and experiences*. In: *Practical Hydroinformatics*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 17–30.
- Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, 2019. Soil Survey geographic database (SSURGO). Soil Survey Geographic (SSURGO) Database for Kansas State. <https://doi.org/10.15482/USDA.ADC/1242479>.
- Sophocleous, M., Stern, A.J., Perkins, S.P., 1996. Hydrologic impact of great flood of 1993 in south-central Kansas. *J. Irrig. Drain. Eng.* 122, 203–210. [https://doi.org/10.1061/\(asce\)0733-9437\(1996\)122:4\(203\)](https://doi.org/10.1061/(asce)0733-9437(1996)122:4(203)).
- Tang, X., Li, J., Liu, M., Liu, W., Hong, H., 2020. Flood susceptibility assessment based on a novel random Naïve Bayes method: a comparison between different factor discretization methods. *Catena* 190, 104536. <https://doi.org/10.1016/j.catena.2020.104536>.
- Tehrany, M.S., Pradhan, B., Jebur, M.N., 2015. Flood susceptibility analysis and its verification using a novel ensemble support vector machine and frequency ratio method. *Stoch. Environ. Res. Risk Assess.* 29, 1149–1165. <https://doi.org/10.1007/s00477-015-1021-9>.
- Towfiqul Islam, A.R.M., Talukdar, S., Mahato, S., Kundu, S., Eibek, K.U., Pham, Q.B., Kuriqi, A., Linh, N.T.T., 2021. Flood susceptibility modelling using advanced ensemble machine learning models. *Geosci. Front.* 12, 101075 <https://doi.org/10.1016/j.gsf.2020.09.006>.
- Veatch, N.T., 1952. The Kansas flood of 1951. *Journal (American Water Works Association)* 44, 765–774.
- Wang, Z., Lai, C., Chen, X., Yang, B., Zhao, S., Bai, X., 2015. Flood hazard risk assessment model based on random forest. *J. Hydrol. (Amst.)* 527, 1130–1141. <https://doi.org/10.1016/j.jhydrol.2015.06.008>.
- Wan, Z., Hook, S., Hulley, G., 2021. MODIS/Terra Land Surface Temperature/Emissivity 8-Day L3 Global 1km SIN Grid V061..
- World Meteorological Organization, 2021. *Wmo atlas of mortality and economic losses from weather CLIMATE AND WATER EXTREMES 1970–2019*.
- Wolpert, D.H., 1992. Stacked generalization. *Neural Network* 5, 241–259. [https://doi.org/10.1016/s0893-6080\(05\)80023-1](https://doi.org/10.1016/s0893-6080(05)80023-1).
- Xiao, Y., Wan, J., Hewings, G.J.D., 2013. Flooding and the Midwest economy: assessing the Midwest floods of 1993 and 2008. *Geojournal* 78, 245–258. <https://doi.org/10.1007/s10708-011-9415-9>.
- [Dataset] Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, 2019. Soil Survey Geographic Database (SSURGO). Soil Survey Geographic (SSURGO) Database for Kansas State. <https://doi.org/10.15482/USDA.ADC/1242479>.