



Omar Seleem | Georgy Ayzel | Arthur Costa Tomaz de Souza | Axel Bronstert |
Maik Heistermann

Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany

Suggested citation referring to the original publication:

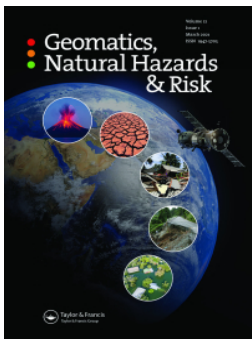
Geomatics, natural hazards and risk 13 (2022) 1, pp. 1640 - 1662
DOI <https://doi.org/10.1080/19475705.2022.2097131>
ISSN 1947-5705, 1947-5713

Journal article | Version of record

Secondary publication archived on the Publication Server of the University of Potsdam:
Zweitveröffentlichungen der Universität Potsdam :
Mathematisch-Naturwissenschaftliche Reihe 1297
ISSN: 1866-8372
<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-576806>
DOI: <https://doi.org/10.25932/publishup-57680>

Terms of use:

This work is licensed under a Creative Commons License. This does not apply to quoted content from other authors. To view a copy of this license visit <https://creativecommons.org/licenses/by/4.0/>.



Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany

Omar Seleem, Georgy Ayzel, Arthur Costa Tomaz de Souza, Axel Bronstert & Maik Heistermann

To cite this article: Omar Seleem, Georgy Ayzel, Arthur Costa Tomaz de Souza, Axel Bronstert & Maik Heistermann (2022) Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany, *Geomatics, Natural Hazards and Risk*, 13:1, 1640-1662, DOI: [10.1080/19475705.2022.2097131](https://doi.org/10.1080/19475705.2022.2097131)

To link to this article: <https://doi.org/10.1080/19475705.2022.2097131>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 12 Jul 2022.



[Submit your article to this journal](#)



Article views: 1673



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)



Towards urban flood susceptibility mapping using data-driven models in Berlin, Germany

Omar Seleem, Georgy Ayzel, Arthur Costa Tomaz de Souza, Axel Bronstert and Maik Heistermann

Institute of Environmental Science and Geography, Chair for Hydrology and Climatology, University of Potsdam, Potsdam, Germany

ABSTRACT

Identifying urban pluvial flood-prone areas is necessary but the application of two-dimensional hydrodynamic models is limited to small areas. Data-driven models have been showing their ability to map flood susceptibility but their application in urban pluvial flooding is still rare. A flood inventory (4333 flooded locations) and 11 factors which potentially indicate an increased hazard for pluvial flooding were used to implement convolutional neural network (CNN), artificial neural network (ANN), random forest (RF) and support vector machine (SVM) to: (1) Map flood susceptibility in Berlin at 30, 10, 5, and 2 m spatial resolutions. (2) Evaluate the trained models' transferability in space. (3) Estimate the most useful factors for flood susceptibility mapping. The models' performance was validated using the Kappa, and the area under the receiver operating characteristic curve (AUC). The results indicated that all models perform very well (minimum AUC = 0.87 for the testing dataset). The RF models outperformed all other models at all spatial resolutions and the RF model at 2 m spatial resolution was superior for the present flood inventory and predictor variables. The majority of the models had a moderate performance for predictions outside the training area based on Kappa evaluation (minimum AUC = 0.8). Aspect and altitude were the most influencing factors on the image-based and point-based models respectively. Data-driven models can be a reliable tool for urban pluvial flood susceptibility mapping wherever a reliable flood inventory is available.

ARTICLE HISTORY

Received 3 March 2022
Accepted 29 June 2022


KEYWORDS

Urban pluvial flood susceptibility; convolutional neural network; deep learning; random forest; support vector machine; spatial resolution; flood predictors

1. Introduction

Urbanization increases losses caused by floods (Karamouz et al. 2011; Cherqui et al. 2015; Zhou et al. 2019). Floods can be classified based on their generation mechanisms into different types such as river floods, urban pluvial floods, flash floods, and

CONTACT Omar Seleem  seleem@uni-potsdam.de

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/19475705.2022.2097131>.

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

coastal floods (Kundzewicz et al. 2014). Urban pluvial floods usually occur due to inundation caused by excess runoff before it enters stormwater drainage system (Falconer et al. 2009) or due to intense rainfall that leads to overwhelming the stormwater drainage system's capacity (Schmitt et al. 2004). They could occur anywhere subject to the occurrence of high intensity rainstorms and the existence of a critical area for runoff generation (Zhang and Pan 2014). The ubiquity of this hazard highlights the importance of accurate flood susceptibility mapping to support urban pluvial flood risk management.

Commonly, physical hydrodynamic models are used to simulate urban pluvial flooding. They can be divided into one-dimensional (1 D) hydrodynamic models such as SWMM model (Barco et al. 2008), two-dimensional (2 D) hydrodynamic models such as TELEMAC-2D model (Seleem et al. 2021), and 1 D-2D hydrodynamic models such as MIKE URBAN (Bisht et al. 2016). These models solve the shallow water equation numerically, and are considered the best representation of the involved processes; yet, the computational costs are high. Therefore, they can only be applied to small areas using a fine spatial resolution and cannot be scaled to produce flood hazard maps for large areas (Petroselli 2012).

To overcome this limitation, (Zhang and Pan 2014; Balstrøm and Crawford 2018; Samela et al. 2020) applied simplified methods based on digital elevation models (DEM), where depressions in the terrain were considered as the inundated areas. (Jalayer et al. 2014; Huang et al. 2019; Kelleher and McPhillips 2020) proposed using topographic indices as indicators of urban pluvial flooding locations. However, these methods have limitations; for example, they perform poorly with low precipitation depths, consider inundation only within the topographic depressions, and need to be calibrated for each precipitation depth of interest (Seleem et al. 2021).

Point-based data-driven models such as logistic regression (Al-Juaidi et al. 2018), the statistical index (Wi) method (Tehrany et al. 2019), random forests (RF) (Wang et al. 2015; Lee et al. 2017; Chen et al. 2020), support vector machines (SVM) (Tehrany et al. 2014; Tehrany et al. 2015), and artificial neural networks (ANN) (Bui et al. 2020) have been used as alternatives to map flood susceptibility for large areas. They can incrementally create high-level features from a raw dataset, and capture complex patterns in the dataset (Bui et al. 2020). They have demonstrated powerful performance in several areas worldwide (Tehrany et al. 2014; Tehrany et al. 2015; Zhao et al. 2018; Chen et al. 2020; Vafakhah et al. 2020; Costache et al. 2021). However, few studies have utilized such models for flood susceptibility mapping in urban areas because of the lack of inundation data or reliable flood inventories (Yang et al. 2016). Furthermore, the fundamental assumption of point-based models is that a relationship between local flood influencing factors and the local occurrence of flooding could be established (Zhao et al. 2020).

Recently, convolutional neural networks (CNN) have been used for flood susceptibility mapping (Wang et al. 2020; Zhao et al. 2020; Lei et al. 2021). Zhao et al. (2020) demonstrated that convolutional neural networks (CNN) could outperform point-based models. They could learn patterns in two dimensions and understand relevant patterns around a predicted location. However, (Zhao et al. 2020; Lei et al. 2021) used a dataset at a coarse spatial resolution (30 and 25 m respectively). The coarse

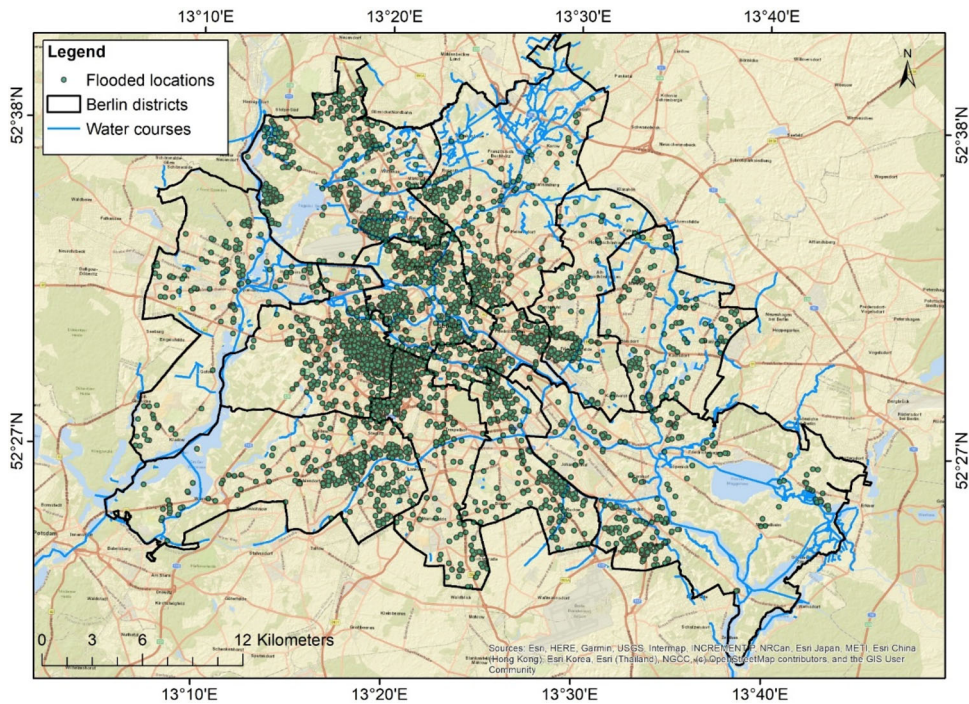


Figure 1. Spatial distribution of 4333 reported flooded locations between 2005 –2017 and administrative districts in Berlin, Germany.

spatial resolution could hide essential details of the urban surface that affect flow paths and inundation patterns (Komolafe et al. 2018; Arrighi and Campo 2019). So far, the impact of spatial resolution on the results of data-driven flood susceptibility mapping has scarcely been studied (Avand et al. 2022). Moreover, (Zhao et al. 2020) considered the CNN model as a black box and did not investigate the importance of the flood influencing factors.

In previous studies, data-driven models for flood susceptibility mapping were typically evaluated in the same area used to train the models. Zhao et al. (2021) showed that using transfer-learning technique could improve the CNN model performance. The trained CNN performed poorly outside the training area, but its performance improved by adding more training data from the area outside the training area. Such techniques have not been yet investigated for RF and SVM which are considered as a benchmark for other models in the literature (Tehrany et al. 2015).

In summary, the use of data-driven models for urban pluvial flood susceptibility mapping still lacks a sufficient understanding of: (i) the effect of spatial resolution on different model types, (ii) the transferability of models in space, and (iii) the importance of specific predictor sets. On the basis of these gaps, this study aims to address the following research questions:

1. How do image-based models (CNN) and point-based models (RF, SVM and ANN) compare with regard to the spatial resolution (i.e., 30, 10, 5, and 2 m) of the input data?

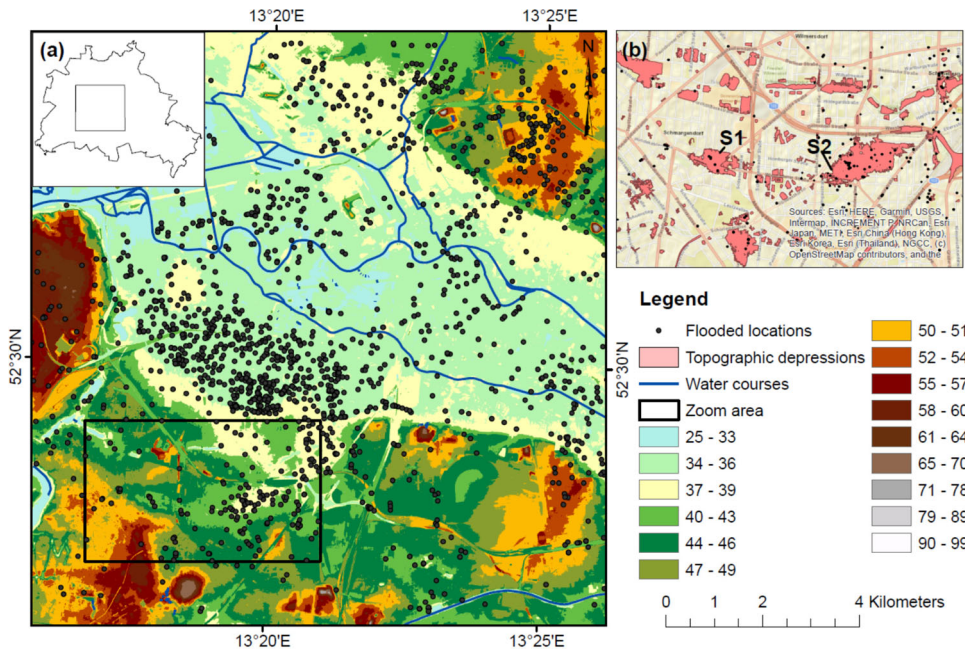


Figure 2. (a) Location of the models' training area and spatial distribution of the reported flooded locations within it. (b) Zoom area (the area used to compare flood susceptibility maps at different spatial resolutions) showing the topographic depressions and the flooded locations within it. The color bar represents the altitude (m.a.s.l.).

2. How transferable (in space) are the trained models?
3. Which factors are most useful for flood susceptibility mapping?

We will investigate these questions on the basis of a unique flood inventory that is available for the city of Berlin, Germany (Berghäuser et al. 2021).

2. Study area

Berlin is the capital and the largest city in Germany. It has 12 administrative districts as shown in Figure 1. The city's population was around 3.6 million in 2020. 55% of the city consists of built-up areas (Kottmeier et al. 2007). It is located in the northeast of Germany and has a relatively flat topography: 95% of the city has an altitude between 30 and 60 m above sea level, and 55% of the city has a slope angle less than 2° (as shown in Figure S1 in the supporting information). It has an oceanic climate (Köppen: Cfb) (Peel et al. 2007), with an average annual precipitation around 570 millimetres (Berghäuser et al. 2021), see also Figure S2 in the supporting information. Several pluvial urban flood events occurred in the past decades as a consequence of heavy summer precipitation. For example, the 24 hr precipitation depth for the 29th and 30th of the June 2017 event was around 170 mm and the 3 hr precipitation depth for the 2nd of August 2019 event was around 45 mm (Berghäuser et al. 2021).

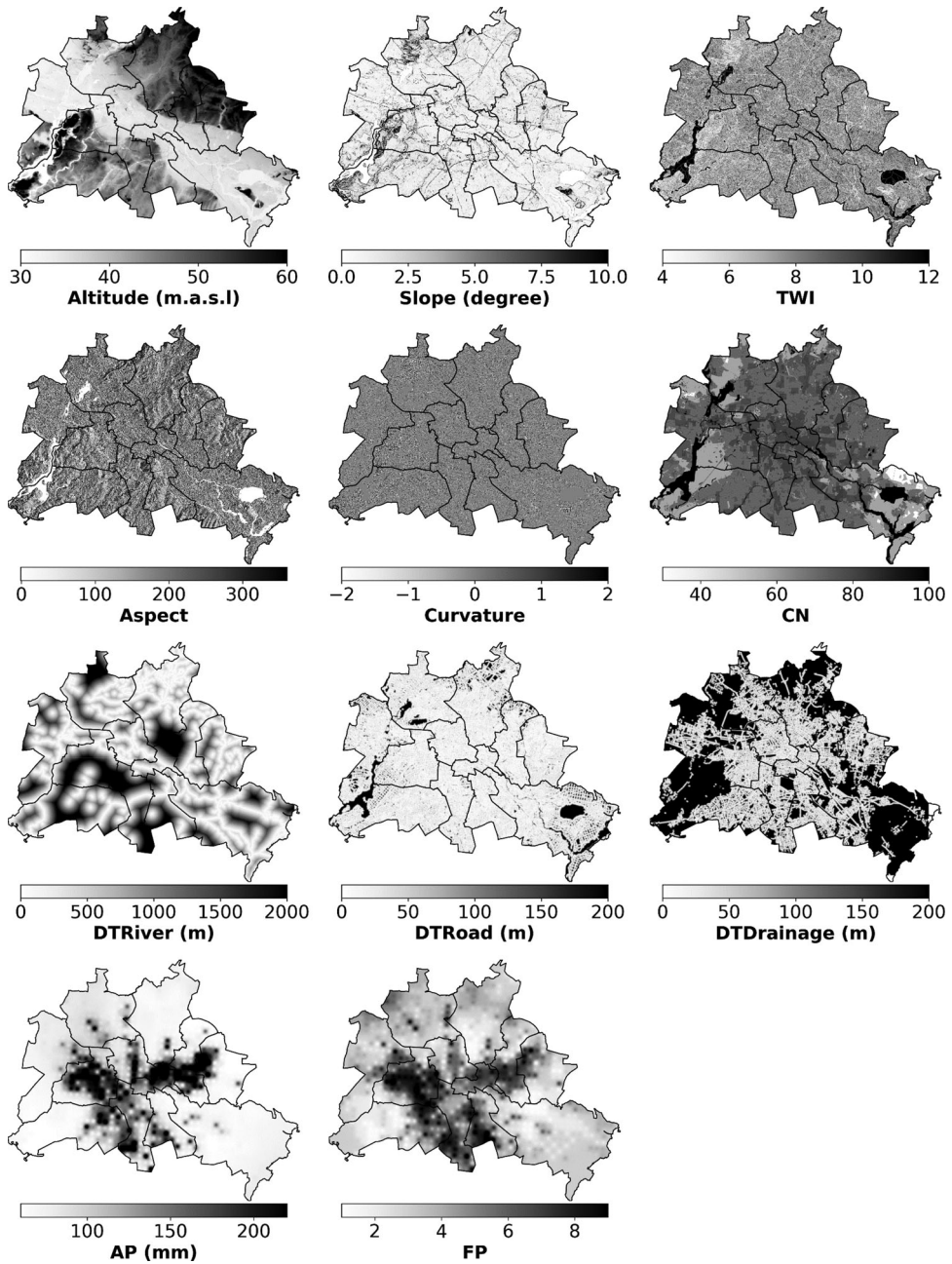


Figure 3. Spatial distribution of flood influencing factors used to develop the models.

3. Data and methods

The overall approach implemented in this study is as follow: Firstly, the flood inventory and eleven factors that potentially influence flood occurrence were used to prepare the training, validation, and testing datasets for models development. Then, the flood susceptibility maps were compared for a selected area within the training area

and the models' performance was compared based on selected performance indices. After that, the ability of the trained models to map flood susceptibility for the whole city was evaluated. Finally, the importance of the flood influencing factors was estimated for all the implemented models. [Figure S3](#) in the supporting information provides a graphical overview of the implemented methodology.

3.1. Flood inventory for Berlin

The flood inventory from Berlin Wasserbetriebe includes 4333 reported flood locations distributed all over the city as shown in [Figure 1](#). These were compiled based on reports of the fire brigade, from social media and from customer reports between the years 2005 to 2017.

We selected an area within the city (170 km²) that has a high density of flooded locations to develop the models as shown in [Figure 2](#). There were 1967 reported flood locations within this area. Additionally, 1967 non-flooded locations were selected in area free of flooding. In comparison to previous studies, this is a very large and unique dataset (e.g., [Termeh et al. \(2018\)](#) with 53 flooded locations in an area of 5737 km²; [Choubin et al. \(2019\)](#): 51 locations in 126 km²; [Zhao et al. \(2020\)](#): 216 locations in 131 km²). Flooded and non-flooded locations within the model training area were randomly split into a training set (60%), a validation set (20%), and a testing set (20%) ([Yacoub et al. 2003](#); [Trost et al. 2014](#), [Raschka 2015](#)). The training dataset was used to fit the different models, the models' hyper-parameters were estimated based on the model performance on the validation dataset, then the testing dataset was used to evaluate the models performance. Evaluating the models based on a test dataset that the models have not seen before allowed us to get a less biased estimate of their ability to generalize to a new data. Flooded locations outside the training area and an equivalent number of non-flooded locations were used to evaluate the models transferability.

3.2. Flood influencing factors

According to the available data for Berlin and a literature review ([Arabameri et al. 2019](#); [Khosravi et al. 2019](#); [Zhao et al. 2020](#)), 11 factors were identified which potentially indicate an increased hazard for pluvial flooding. These factors represent the topographical, infrastructural, and hydrometeorological conditions: altitude, slope, topographic wetness index (TWI), curvature (Curve), distance to the river (DTRiver), distance to the road (DTRoad), distance to the stormwater drainage system (DTDrainage), curve number (CN), and the frequency (FP) and magnitude (AP) of extreme precipitation events.

- Altitude is one of the most important flooding triggering factors ([Tehrany et al. 2014](#)). In general, runoff tends to accumulate at lower elevation ([Zhang and Pan 2014](#); [Seleem et al. 2021](#)). A digital elevation model (DEM) with 1 × 1 m pixel size is openly available to download for the entire city of Berlin ([ATKIS 2020](#)).

- Slope affects runoff velocity and thus time available for infiltration, and also the speed for runoff concentration and hence accumulation (Rahmati et al. 2016).
- The TWI was originally proposed by (Kirkby 1975) for hydrological modelling in mountainous and hilly terrain. It is a physical property that indicates the level of geotechnical wetness (Chapi et al. 2017) and can be used to identify flood-prone areas (Jalayer et al. 2014; Seleem et al. 2021). It is calculated as follow:

$$TWI = \ln\left(\frac{a}{\tan \beta}\right)$$

The parameter a represents the upslope contributing area per grid length and β is the local slope angle.

- The DTRiver indicates the Euclidean distance between a point and the nearest river. The DTRiver is considered as one of the most important factors to map flood susceptibility (O'Neill et al. 2016). The river network was obtained from open street maps (Haklay and Weber 2008).
- Aspect indicates the direction of the maximum slope. It is directly related to the water flow direction and indicates flat areas (Regmi et al. 2014; Jaafari et al. 2015; Tehrani et al. 2019; Choubin et al. 2019).
- Curvature (Curve) represents the changes in slope inclination (Wilson and Gallant 2000). The curvature value indicates whether the surface is convex, concave or flat. Flood water tends to retain in concave surfaces, potentially increasing flooding susceptibility (Rejith et al. 2019).
- DTRoad indicates the Euclidean distance between a point and the nearest road. In pluvial flooding, the limited capacity of the stormwater drainage system can generate runoff that travels through the road network converting it to a preferential path for runoff (Yin et al. 2016; Singh et al. 2018). The road network was obtained from open street maps (Haklay and Weber 2008).
- DTDrainage indicates the Euclidean distance between a point and the nearest inlet to the stormwater drainage system. Pluvial flooding can occur when the capacity of the stormwater drainage system is exceeded. We downloaded the gullies' locations from (ATKIS 2020).
- CN is an empirical parameter that is used to calculate the direct runoff (Cronshey 1986). It represents the ability of the land surface to retain water, and is calculated using land-cover, soil type and soil texture. We used the CN map for Berlin generated by (Seleem et al. 2021).
- By definition, urban pluvial flooding is caused by heavy precipitation. Therefore, flood susceptibility mapping should consider both the spatial and temporal precipitation patterns (Wang et al. 2015; Zhao et al. 2018). To that end, we selected the annual maximum daily precipitation (AP) and the frequency of extreme precipitation storm (FP) (Zhao et al. 2020). AP indicates the maximum daily precipitation depth recorded between 2005 and 2017 and FP indicates the frequency of occurrence of daily precipitation depth greater than 50 mm between 2005 and 2017. Both AP and FP were calculated from radar precipitation estimates from the

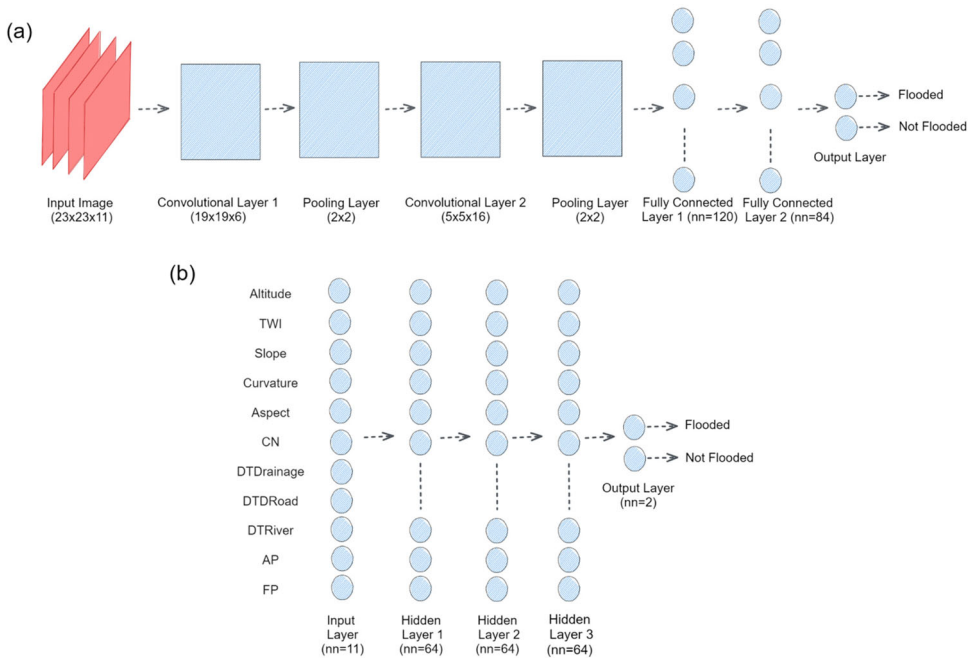


Figure 4. (a) The adapted LeNet-5 network architecture for the CNN models (Zhao et al. 2020). (b) The adapted ANN network architecture from (Bui et al. 2020).

German Weather Service (DWD) in the period between 2005 and 2017 (Winterrath et al. 2017; Kreklow et al. 2019).

Figure 3 shows the spatial distribution of flood influencing factors (predictors) in Berlin.

3.3. Models

The following sub-sections provide a brief summary of the different model types, designs and set-ups, including hyper-parameters which were applied and which can influence both model performance and feature importance (Probst et al. 2019), and need to be set by the user.

3.3.1. Convolutional neural networks (CNN)

The application of CNNs for flood susceptibility mapping is still rare in the literature (Zhao et al. 2020). In our study, we adopted the LeNet-5 architecture with an input image size of 23×23 pixels (Zhao et al. 2020). LeNet-5 has one input layer which is followed by two convolutional layers, each convolutional layer is followed by one pooling layer; then two fully connected neural network layers, and, a final output layer. The design is shown in Figure 4a. This study used the Rectified Linear Unit (ReLU) and the softmax functions as the activation and transfer functions respectively. The Adaptive moment estimation (Adam) (Kingma and Ba 2014) was used to update and optimize the weights of the CNN. A dropout strategy with a drop rate of 0.4 was implemented to the convolutional layers and the fully connected layer to

avoid overfitting. We considered the batch size and the learning rate as the hyper-parameters for the CNN models setup (Table S1 in the supporting information shows the best hyper-parameters combinations for the implemented CNN models).

3.3.2. Artificial neural network (ANN)

This study adopted the Artificial Neural Network (ANN) architecture from Bui et al. (2020) to generate urban pluvial flood susceptibility maps. ANN includes 3 hidden layers and 192 neurons as shown in Figure 4b. The ReLU and the sigmoid functions were used as the activation and transfer functions respectively, while the weights of the ANN were updated and optimized using Adam (Kingma and Ba 2014). Similar to the CNN, a dropout strategy with a drop rate of 0.4 was implemented to the hidden layers to avoid overfitting. We considered the batch size and the learning rate as the hyper-parameters for the ANN models setup (Table S2 in the supporting information shows the best hyper-parameters combinations for the implemented ANN models).

3.3.3. Random forest (RF)

RF was proposed by Breiman (2001). It has been widely used for flood susceptibility mapping (Lee et al. 2017; Chen et al. 2020; Zhao et al. 2020; Abu El-Magd 2022). It implements the bootstrap technique that divides the input data to several sub-samples and develops a tree model for each sub-sample. The final result is determined based on the majority result of all the tree models. This allows RF models to avoid problems such as outliers, noise, and overfitting. We considered the number of trees in the forest, the minimum number of samples necessary to split an internal node, the minimum number of samples required to be at a leaf node, and the maximum depth of the tree as the hyper-parameters for the RF models setup (Table S3 in the supporting information shows the best hyper-parameters combinations for the implemented RF models).

3.3.4. Support vector machine (SVM)

SVM is a machine learning technique proposed by (Cortes and Vapnik 1995). It is based on the risk minimization and statistical learning theory (Tien Bui et al. 2012). It has been widely implemented in flood susceptibility mapping (Tehrany et al. 2014; Tehrany et al. 2015; Wang et al. 2020; Zhao et al. 2020). It works on finding the optimal hyperplane that separates the non-flooded and flooded classes [0, 1] (Choubin et al. 2019). This study used the RBF kernel of the SVM as it outperformed other kernels in flood susceptibility mapping (linear, sigmoid, polynomial, see Tien Bui et al. 2012; Tehrany et al. 2015; Hong et al. 2018; Wang et al. 2020). We considered the penalty coefficient and the radial basis function bandwidth as the hyper-parameters for the SVM models setup (Table S4 in the supporting information shows the best hyper-parameters combinations for the implemented SVM models).

3.4. Feature importance

Previous studies have not investigated the importance of individual factors to predict pluvial flood susceptibility. In this study, the SHAP (SHapley Additive exPlanations)

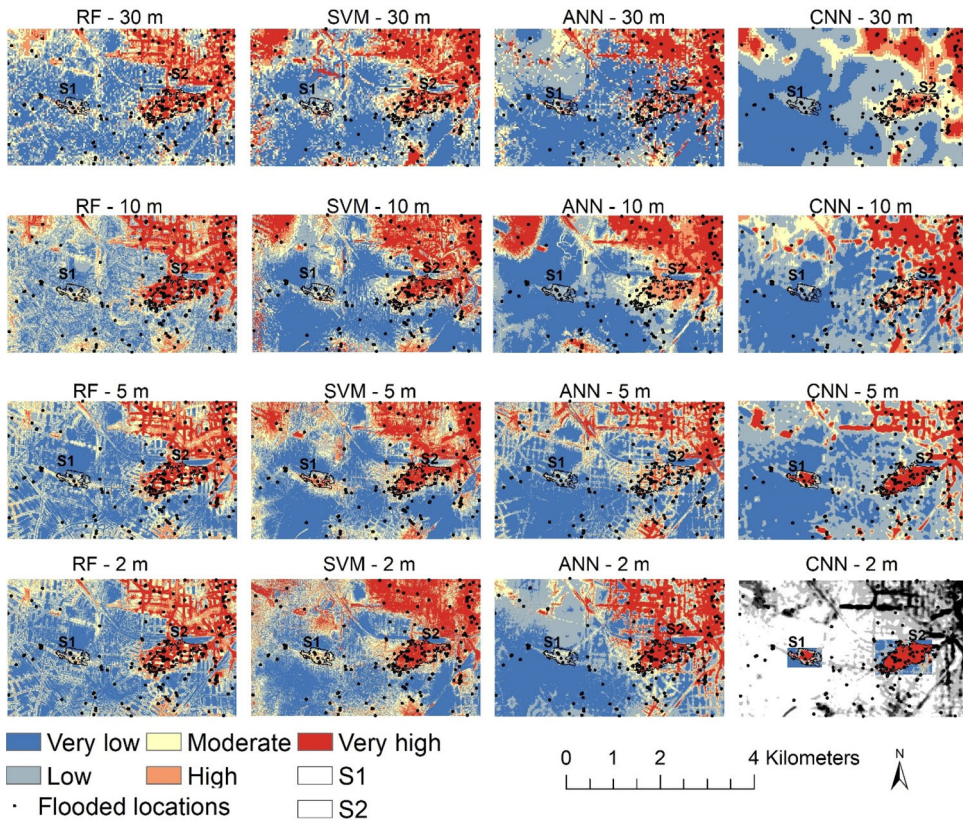


Figure 5. Flood susceptibility maps from all models at different spatial resolution for the zoom area in the training area.

Python package (Lundberg and Lee 2017) was used to determine the feature importance for all models. SHAP assigns an importance value for each feature and each prediction (Lundberg and Lee 2017). It can be used with a wide range of models, including tree-based models, linear models and neural networks. Compared to other techniques, SHAP does not only show feature importance, but also determines whether a feature has a positive or negative effect on the predicted values.

3.5. Evaluation of model performance

We used Kappa and the area under the receiver operating characteristic (ROC) curve (AUC) to evaluate model accuracy (additional indices are shown in Table S5 in the supporting information). Both indices have been widely used to evaluate flood susceptibility maps in literature (Tehrany et al. 2015; Bui et al. 2020; Zhao et al. 2020; Zhao et al. 2021). Kappa is calculated as follow:

$$Kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

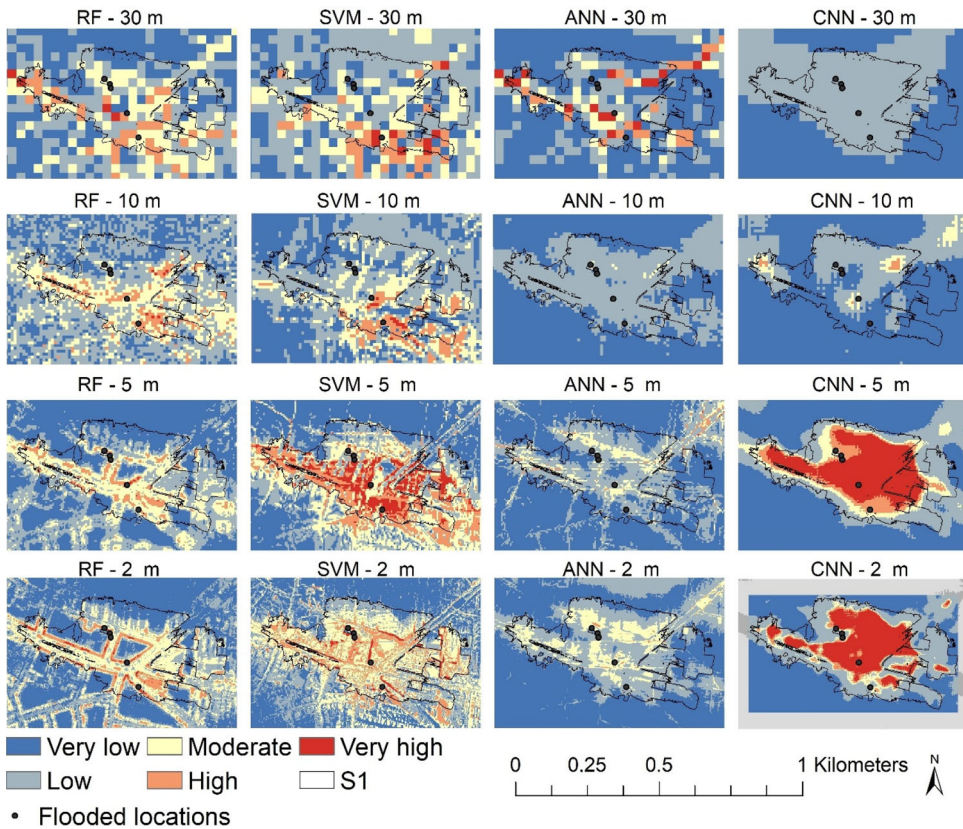


Figure 6. Flood susceptibility maps from all models at different spatial resolution for topographic depression S1.

Where observed agreement (P_o) and hypothetical probability of chance agreement (P_c) (Viera and Garrett 2005) can be calculated by comparing observations to model predictions. Kappa can range between -1 (a Kappa value less than 0 indicates chance agreement) and 1 (perfect agreement) while the AUC can range between 0 and 1, where 1 indicates a perfect model while a value of 0.5 marks the performance of a random prediction.

3.6. Computational details

The maps of the flood influencing factors were created using ArcGIS. The Keras Python package (Chollet et al. 2015) was used to implement both the CNN and ANN models while the RF and SVM models were implemented using the sklearn.ensemble.RandomForestClassifier (Pedregosa et al. 2011) and the sklearn.svm.SVC Python modules (Chang and Lin 2011), respectively. K-fold cross validation was applied to quantify the model performance.

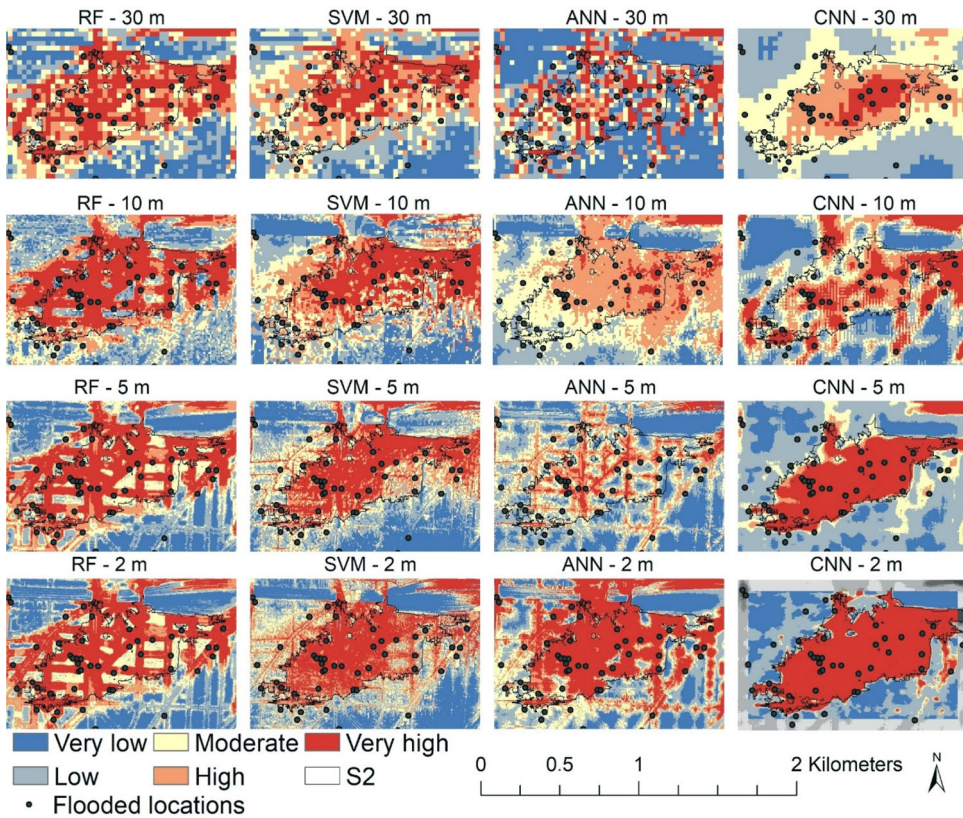


Figure 7. Flood susceptibility maps from all models at different spatial resolution for topographic depression S2.

4. Results and discussion

4.1. Comparison of predicted flood susceptibility

We used 11 factors which potentially affect flood susceptibility (altitude, slope, curvature, TWI, CN, DTRoad, DTRiver, DTDrainage, FP and AP), and used 60% of the flooded locations as training data to produce flood susceptibility maps for Berlin utilizing different data-driven models at different horizontal resolutions. Flood susceptibility values range from 0 (lowest) to 1 (highest). For visualisation, the flood susceptibility was categorized into five classes using the natural breaks (Jenks) method (Jenks 1967) which is widely utilized in flood susceptibility mapping (Chapi et al. 2017; Wang et al. 2020; Zhao et al. 2020): very low, low, moderate, high and very high. Figure 5 shows the flood susceptibility maps using the different models and the different horizontal resolutions for the zoom area inside the training area (Figure 2). Visually, the majority of the flooded locations coincide with high flood susceptibility values. Due to excessive computation costs, the CNN model at 2 m resolution was not applied for the entire area. Instead, we produced a flood susceptibility map only for topographic depressions S1 and S2 (Figure 2b).

We selected topographic depressions S1 and S2 to evaluate the models' ability to detect topographic depressions (where normally excess runoff would accumulate).

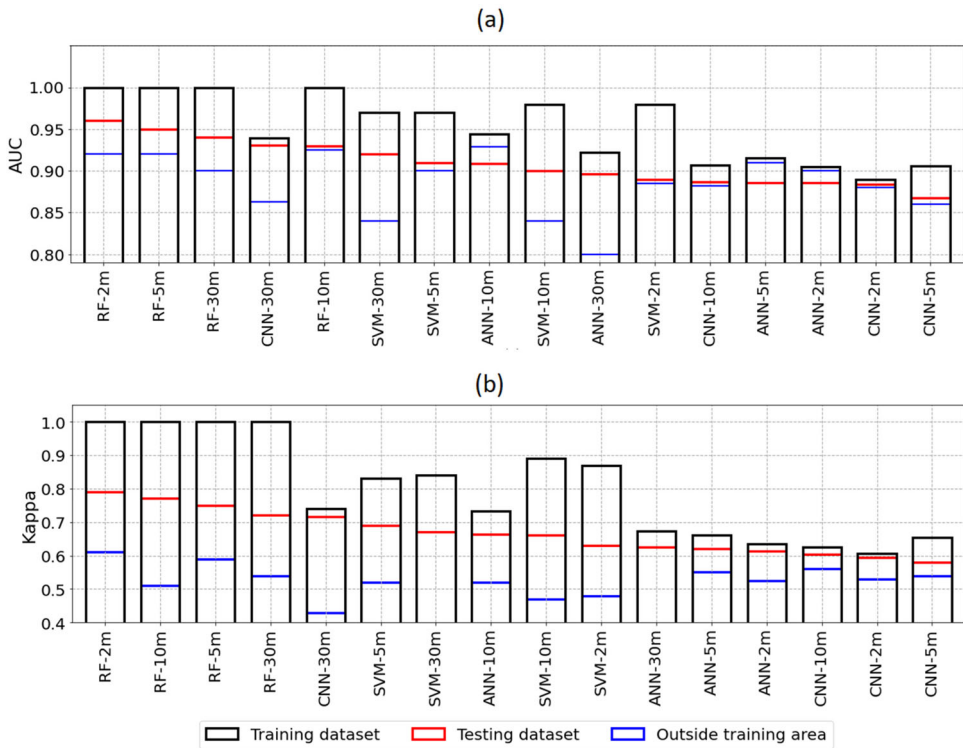


Figure 8. Calculated performance indices for the training dataset, testing dataset and the locations located outside training area, models are arranged horizontally in descending order according to their score for the testing dataset: (a) AUC, (b) Kappa.

Figure 6 shows the predicted flood susceptibility maps for S1. The figure highlights the impact of the spatial resolution. For the RF models, e.g., the impact of the road network on the identification of locations with high flood susceptibility becomes obvious only at resolutions of 5 and 2 m. Moreover, it is interesting to see that the CNN models (image-based) could only recognize S1 as a flood-susceptible area at the 5 and 2 m spatial resolutions. Figure 7 shows the produced flood susceptibility maps for S2. Again, it becomes obvious that a higher resolution allows the CNN model to identify the flood-prone area more clearly.

4.2. Model validation

The performance of the CNN, ANN, RF and SVM models at 30, 10, 5, and 2 m spatial resolutions was quantified based on the metrics AUC and Kappa (other metrics are shown in Table S5 in the supporting information). Figure 8 shows the results for the training dataset, the testing dataset (which was reserved as 20% of the flooded locations), and the points located outside the training area. Generally, all models achieved AUC values higher than 0.88, 0.87, and 0.8 for the three datasets respectively.

For the training dataset, the RF models outperformed the other models at all spatial resolutions. All the performance indices calculated for the RF models were equal to 1 which demonstrate the models' ability to perfectly distinguish between the flooded and

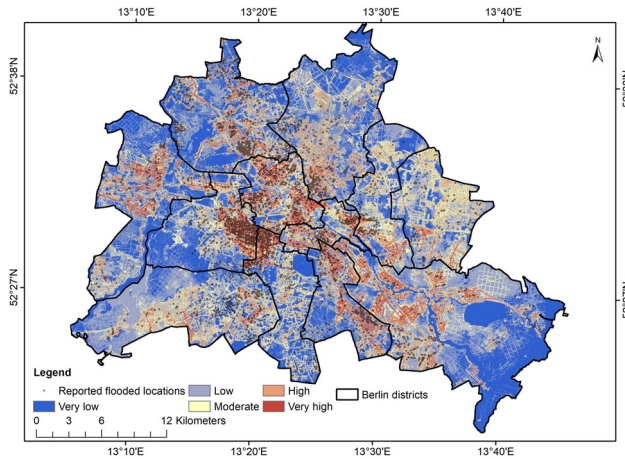


Figure 9. Flood susceptibility map for Berlin from RF - 2 m model (best performance model) and the reported flooded locations between 2005 and 2017.

non-flooded locations. The CNN – 2 m model had the lowest performance indices: $AUC = 0.88$ and $\kappa = 0.59$, which represent a moderate performance ($0.41 < \kappa < 0.60$) based on the kappa evaluation criteria (Viera and Garrett 2005).

For the testing dataset, the RF models outperformed the other models at all spatial resolutions, too. The RF model at 2 m resolution had the highest AUC (0.96) while the CNN-model at 5 m resolution had the lowest AUC (0.87). According to the kappa evaluation criteria, the predictions from the RF – 2 m model ($\kappa = 0.79$) demonstrated substantial agreement with the observations ($0.61 < \kappa < 0.80$) while the predictions from CNN – 5 m model ($\kappa = 0.58$) showed moderate agreement.

4.3. Flood susceptibility for Berlin

The trained models were then used to predict flood susceptibility for all of Berlin. The flooded and sampled non-flooded locations outside the training area were then used to assess the model performance outside the training area, and hence the model transferability in space. For illustration, Figure 9 shows flood susceptibility from the RF model at 2 m resolution (best model on training and testing data sets). Visually, areas with high flood susceptibility coincide with the flooded locations.

Figure 8 shows the performance indices for the points located outside the training area. The RF models were superior to the other models. While the RF – 2 m had the best performance ($AUC = 0.92$ & $\kappa = 0.61$), ANN – 30 m had the least performance ($AUC = 0.8$ & $\kappa = 0.02$). Figure 8, as well as Figure S4 and S5 in the supporting information show that despite the superiority of the RF models to other models, the CNN and ANN models had a smaller relative loss for the testing dataset and locations outside the training area.

4.4. Feature importance

We used the SHAP algorithm to evaluate feature importance for all competing models. As mentioned before, the SHAP values do not only show the feature importance

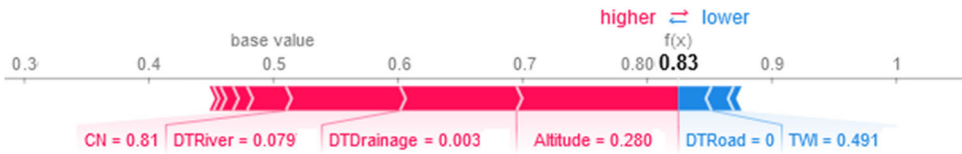


Figure 10. SHAP values of each feature and their impact on the model prediction for a certain location by the RF model at 2 m resolution. The values of the input features at this location (shown here as normalized values between 0 and 1) moved the model prediction from 0.5 to 0.83 (final prediction). Features that decreased the probability of classifying the location as flooded are colored in blue while features increased the probability of classifying the location as flooded are colored in red. The features visual size show the magnitude of their impact on the prediction. For example, altitude had the largest impact on the prediction at this location.

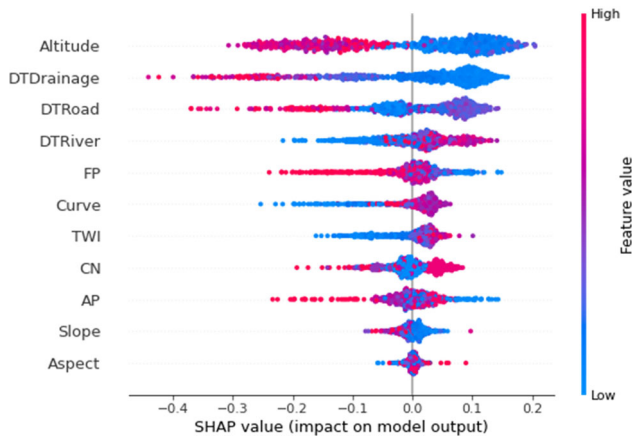


Figure 11. SHAP values for the testing dataset using RF - 2 m model. The features are arranged vertically based on their importance in descending order. The horizontal axis shows the SHAP values, a positive value means that it increased the probability of classifying the location as flooded and a negative value means that it increased the probability of classifying the location as non-flooded, the colour shows whether the feature value is low or high (The values are shown in Figure 3), the SHAP values for each feature at every location are represented by dots. The dots tend to pile up along each feature row to represent the dots density.

but also whether a feature affects the predicted values positively or negatively. Flood susceptibility values range from 0 to 1. A balanced training dataset (same number of flooded and non-flooded locations) was used to develop the models. Therefore, the expected model prediction would be 0.5 if we did not know the input feature values at the predicted location. Then, the prediction would change from 0.5 to the final prediction based on the values of the input feature as shown in Figure 10. The feature effects are represented by the SHAP values in Figure 11 for the RF model at 2 m resolution. Figure 11 shows that the predictor variables affect the model prediction (flooded or non-flooded) depending on the features at the predicted location and floods tend to occur in locations at low altitudes close to drainage system inlets and roads (RF – 2 m model).

For the image-based CNN model, the SHAP algorithm can detect which pixels increase the probability of a certain prediction (flooded or non-flooded) as shown in Figure 12. Figure 12 shows images for two locations, the top images represent a

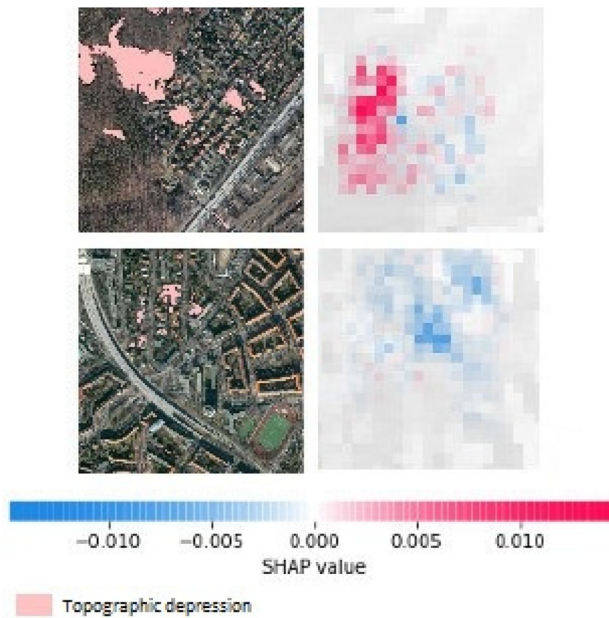


Figure 12. The left figures show images of two locations for the CNN model including the topographic depression in the images and the right figures show the calculated SHAP values for the two images based on the CNN - 30 m model. The intense red coloured pixels increase the probability that the image is predicted as flooded, whereas intense blue coloured pixels increase the probability that the image is predicted as non-flooded. The sum of the SHAP values of the pixels equals the difference between the expected model prediction (0.5 based on the training dataset) and the final prediction.

flooded location while the bottom images represent a non-flooded location. The model correctly predicted both locations. Figure 12 shows an advantage of using the image-based CNN model as it shows the importance of considering the area surrounding a location and how they influenced the prediction.

SHAP value can be positive or negative, depending on whether the feature increases or decreases the probability of a certain class (flooded or non-flooded). The higher the absolute SHAP value, the more impact the feature has on the model prediction. Therefore, the mean absolute SHAP values could indicate the importance of each feature on the models' prediction. Figure 13 and Table S6 in the supporting information show the calculated mean absolute SHAP values for each feature for all the developed models. For the point-based models (RF, SVM, and ANN), the feature importance depended on the model and the horizontal resolution of the used dataset. The mean absolute SHAP value for the altitude was always significantly higher than for other features. Altitude, DTDrainage, DTRiver, AP and FP had high values while TWI, slope, curve, aspect and CN had low values which demonstrated that they had a low impact on the models' prediction. Despite the CN had no impact on the models' prediction at 30 m spatial resolution, it had impact at finer spatial resolutions. Similarly, the importance of DTRoad on the RF models prediction increased with finer resolutions which demonstrates that the importance of the features could change with different spatial resolutions.

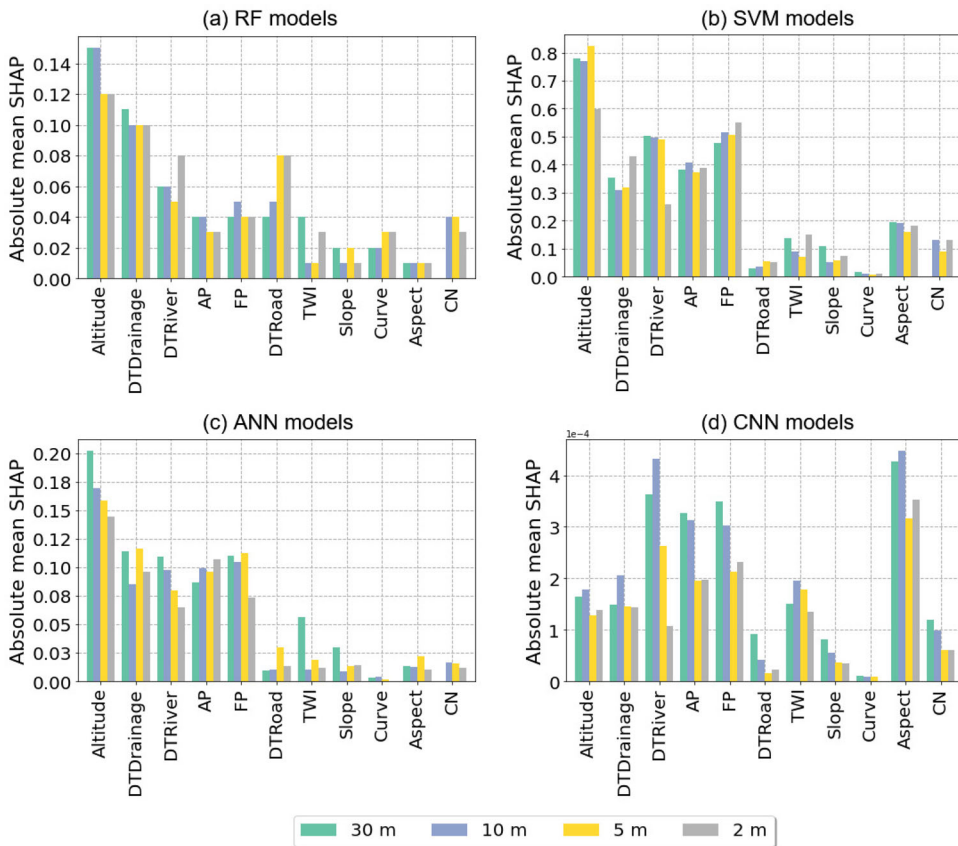


Figure 13. Feature importance for all models based on the calculated absolute mean SHAP values. The calculated values are shown in Table S6 in the supporting information.

The common most important features for the CNN models were aspect, DTRiver, FP, and AP while curvature, slope and DTRoad were the common least important features as shown in Figure 13d. Although altitude was the most important feature for the point-based models, it was only moderately important for the CNN models. The SHAP values are not explaining how the CNN models are working but only show which features and pixels influenced the model predictions. Please note that the magnitude of SHAP values from the CNN models was different from other models because SHAP uses different algorithms to explain the prediction of different models and the importance is distributed over the image pixels (Lundberg and Lee 2017).

5. Conclusions

CNN, ANN, RF and SVM models were used to map flood susceptibility for the city of Berlin using 11 predictor variables at 30, 10, 5, and 2 m spatial resolution. A detailed urban flood inventory served as reference data for training, testing, and validation. The key findings are summarized in the following sections:

5.1. Model performance

Based on the calculated performance indices, we found that all models performed well on the training and testing datasets, while the RF model outperformed the others at all spatial resolutions, and the RF model at 2 m spatial resolution performed best. We evaluated CNN and ANN architectures that had been used in the literature (Bui et al. 2020; Zhao et al. 2020; Zhao et al. 2021). Both architectures performed well. In contrast to (Zhao et al. 2020) findings, we found that the RF and SVM models outperformed the implemented CNN models.

The models' ability to identify topographic depressions was evaluated using two topographic depressions (S1 and S2). Although all the models could predict depression S2 as flood-susceptible area at all spatial resolutions, the performance varied for depression S1. RF models at fine resolution (5 and 2 m) could recognize the streets in depression S1 as a flood-susceptible area, the CNN models recognized the topographic depression as a flood-susceptible area only at fine resolution (5 and 2 m). The maps showed that the models could better understand the complex urban environment using finer horizontal resolution. Moreover, the relative loss of the performance indices from inside to outside the training area is particularly high for the 30 m models. Therefore, the 30 m resolution might in fact hide generalizable patterns. The literature still lacks models which use fine resolution datasets (Zhao et al. 2020; Lei et al. 2021).

5.2. Model transferability in space

Model transferability in space could enable the prediction of flood susceptibility for areas outside the model's training area. It is still a new rising topic for flood mapping (Zhao et al. 2021). Our findings show that the majority of the models had a moderate performance for predictions outside the training area based on Kappa evaluation. Although the predicted maps had a relative error up to 20% (minimum AUC = 0.8), it is a quick method to generate flood susceptibility maps for urban areas. The RF at 2 m spatial resolution outperformed all the other models. It had substantial performance based on kappa evaluation (kappa = 0.61) and AUC = 0.92. The model performance outside the training area could be improved by adding more training data to the trained model or using transfer-learning techniques (Zhao et al. 2021). Future research requires testing transferability further in environments with different characteristics (particularly with cities in more mountainous environments).

5.3. Feature importance

So far, data-driven models in pluvial flood susceptibility mapping were considered as black boxes. In this study, we investigated the importance of individual factors to predict pluvial flood susceptibility and explained the models' predictions based on the input features values using SHAP.

The spatial resolution affected the importance of the feature for the model prediction, for example, the importance of the DTRoad increased for the RF models for the fine resolution (5 and 2 m). The point-based models agreed with the findings in the

literature that low-lying areas located closer to the stormwater drainage system, to riverbanks and roads are more flood-prone (Tehrany et al. 2014; Bui et al. 2020). The implemented CNN models found that aspect is the most important predictor for pluvial flood mapping which confirm (Löwe et al. 2021) findings. We used 11 flood-influencing factors that are widely used in the literature for flood susceptibility mapping. However, we found that many features were not important for the point-based model while feature importance was more evenly distributed across features for the CNN models. (Löwe et al. 2021) showed that using fewer predicting features could improve the CNN model performance. Therefore, we recommend further research to carry out a feature selection analysis to consider only features that would strongly affect the model prediction.

Flood susceptibility mapping using a data-driven model is considered as an alternative for the complicated hydrodynamic simulations for flood-prone area identification in large urban watershed. Hydrodynamic simulations still the best representation of the involved processes. Therefore, we recommend future research to use the output of hydrodynamic simulations as a reference for training data-driven models for urban flood management.

Overall, all the used models could map the urban flood susceptibility efficiently. The point-based model would be recommended for flood susceptibility mapping for large areas because the CNN models were both computationally and time-consuming in terms of input data preparation especially for fine resolution which is necessary to show the urban watershed characteristics.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Deutscher Akademischer Austauschdienst (DAAD). We acknowledge the support of the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of the University of Potsdam; German Academic Exchange Service.

Data availability statement

The DEM is openly available to download for the entire city of Berlin (ATKIS 2020). The used flood inventory that support the findings of this study is available from Berlin Wasserbetriebe. Restrictions apply to the availability of the flood inventory, which was used under license for this study.

References

- Abu El-Magd SA. 2022. Random forest and naïve Bayes approaches as tools for flash flood hazard susceptibility prediction, South Ras El-Zait, Gulf of Suez Coast, Egypt. *Arab J Geosci.* 15(3):1–12.

- Al-Juaidi AE, Nassar AM, Al-Juaidi OE. 2018. Evaluation of flood susceptibility mapping using logistic regression and GIS conditioning factors. *Arab J Geosci.* 11(24):1–10.
- Arabameri A, Rezaei K, Cerdà A, Conoscenti C, Kalantari Z. 2019. A comparison of statistical methods and multi-criteria decision making to map flood hazard susceptibility in Northern Iran. *Sci Total Environ.* 660:443–458.
- Arrighi C, Campo L. 2019. Effects of digital terrain model uncertainties on high-resolution urban flood damage assessment. *J. Flood Risk Manag.* 12(S2):e12530.
- ATKIS: Digitale geländemodelle—atkis dgm. 2020. [accessed 2022 Feb 22]. <https://www.stadtentwicklung.berlin.de/geoinformation/landesvermessung/atkis/de/dgm.shtml>.
- Avand M, Kuriqi A, Khazaei M, Ghorbanzadeh O. 2022. DEM resolution effects on machine learning performance for flood probability mapping. *J Hydro-Environ Res.* 40:1–16.
- Balstrøm T, Crawford D. 2018. Arc-malstrøm: a 1d hydrologic screening method for storm-water assessments based on geometric networks. *Comput Geosci.* 116:64–73.
- Barco J, Wong KM, Stenstrom MK. 2008. Automatic calibration of the U.S. EPA SWMM model for a large urban catchment. *J Hydraul Eng.* 134(4):466–474.
- Berghäuser L, Schoppa L, Ulrich J, Dillenardt L, Jurado OE, Passow C, Mohor GS, Seleem O, Petrow T, Thieken AH. 2021. Starkregen in Berlin: Meteorologische Ereignisrekonstruktion und Betroffenenbefragung [Heavy precipitation in Berlin: Meteorological event reconstruction and survey of those affected]. Potsdam: University of Potsdam.
- Bisht DS, Chatterjee C, Kalakoti S, Upadhyay P, Sahoo M, Panda A. 2016. Modeling urban floods and drainage using SWMM and MIKE URBAN: a case study. *Nat Hazards.* 84(2): 749–776.
- Breiman L. 2001. Random forests. *Mach Learn.* 45(1):5–32.
- Bui DT, Hoang ND, Martínez-Álvarez F, Ngo PTT, Hoa PV, Pham TD, Samui P, Costache R. 2020. A novel deep learning neural network approach for predicting flash flood susceptibility: a case study at a high frequency tropical storm area. *Sci Total Environ.* 701:134413.
- Bui DT, Pradhan B, Lofman O, Revhaug I. 2012. Landslide susceptibility assessment in Vietnam using support vector machines, decision tree, and naïve bayes models. *Math Probl Eng.* 2012:1–26.
- Chang CC, Lin CJ. 2011. Libsvm: a library for support vector machines. *ACM Trans Intell Syst Technol.* 2(3):1–27.
- Chapi K, Singh VP, Shirzadi A, Shahabi H, Bui DT, Pham BT, Khosravi K. 2017. A novel hybrid artificial intelligence approach for flood susceptibility assessment. *Environ Model Softw.* 95:229–245.
- Chen W, Li Y, Xue W, Shahabi H, Li S, Hong H, Wang X, Bian H, Zhang S, Pradhan B, et al. 2020. Modeling flood susceptibility using data-driven approaches of naïve bayes tree, alternating decision tree, and random forest methods. *Sci Total Environ.* 701:134979.
- Cherqui F, Belmeziti A, Granger D, Sourdril A, Le Gauffre P. 2015. Assessing urban potential flooding risk and identifying effective risk reduction measures. *Sci Total Environ.* 514: 418–425.
- Chollet F, et al. 2015. Keras. [accessed 2022 Feb 22]. <https://keras.io>.
- Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A. 2019. An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci Total Environ.* 651(Pt 2): 2087–2096.
- Cortes C, Vapnik V. 1995. Support vector machine. *Mach Learn.* 20(3):273–297.
- Costache R, Arabameri A, Elkhachy I, Ghorbanzadeh O, Pham QB. 2021. Detection of areas prone to flood risk using state-of-the-art machine learning models. *Geomatics, Nat. Hazards Risk.* 12(1):1488–1507.
- Cronshey R. 1986. Urban hydrology for small watersheds. Technical Report. US Dept. of Agriculture, Soil Conservation Service, Engineering Division.
- Falconer R, Cobby D, Smyth P, Astle G, Dent J, Golding B. 2009. Pluvial flooding: new approaches in flood warning, mapping and risk management. *J. Flood Risk Manag.* 2(3): 198–208.

- Haklay M, Weber P. 2008. Openstreetmap: user-generated street maps. *IEEE Pervasive Comput.* 7(4):12–18.
- Hong H, Tsangaratos P, Ilia I, Liu J, Zhu AX, Chen W. 2018. Application of fuzzy weight of evidence and data mining techniques in construction of flood susceptibility map of Poyang County. *Sci Total Environ.* 625:575–588.
- Huang H, Chen X, Wang X, Wang X, Liu L. 2019. A depression-based index to represent topographic control in urban pluvial flooding. *Water.* 11(10):2115.
- Jaafari A, Najafi A, Rezaeian J, Sattarian A. 2015. Modeling erosion and sediment delivery from unpaved roads in the north mountainous forest of Iran. *Int J Geomath.* 6(2):343–356.
- Jalayer F, De Risi R, De Paola F, Giugni M, Manfredi G, Gasparini P, Topa ME, Yonas N, Yeshitela K, Nebebe A, et al. 2014. Probabilistic GIS-based method for delineation of urban flooding risk hotspots. *Nat Hazards.* 73(2):975–1001.
- Jenks GF. 1967. The data model concept in statistical mapping. *Cartogr J.* 7:186–190.
- Karamouz M, Hosseinpour A, Nazif S. 2011. Improvement of urban drainage system performance under climate change impact: case study. *J Hydrol Eng.* 16(5):395–412.
- Kelleher C, McPhillips L. 2020. Exploring the application of topographic indices in urban areas as indicators of pluvial flooding locations. *Hydrol Process.* 34(3):780–794.
- Khosravi K, Shahabi H, Pham BT, Adamowski J, Shirzadi A, Pradhan B, Dou J, Ly HB, Gróf G, Ho HL, et al. 2019. A comparative assessment of flood susceptibility modeling using multi-criteria decision-making analysis and machine learning methods. *J Hydrol.* 573: 311–323.
- Kingma DP, Ba J. 2015. *A method for stochastic optimization*. Paper presentation at: 3rd International Conference on Learning Representations, ICLR 2015; May 7–9; San Diego, CA, US.
- Kirkby M. 1975. Hydrograph modelling strategies. Peel R, Chisholm M, Haggert P, Processes in Physical and Human Geography. London: Heineman, 69–90.
- Komolafe A, Herath S, Avtar R. 2018. Sensitivity of flood damage estimation to spatial resolution. *J Flood Risk Manage.* 11:S370–S381.
- Kottmeier C, Biegert C, Corsmeier U. 2007. Effects of urban land use on surface temperature in Berlin: case study. *J Urban Plann Dev.* 133(2):128–137.
- Kreklow J, Tetzlaff B, Kuhnt G, Burkhard B. 2019. A rainfall data intercomparison dataset of RADKLIM, RADOLAN, and rain gauge data for Germany. *Data.* 4(3):118.
- Kundzewicz ZW, Kanae S, Seneviratne SI, Handmer J, Nicholls N, Peduzzi P, Mechler R, Bouwer LM, Arnell N, Mach K, et al. 2014. Flood risk and climate change: global and regional perspectives. *Hydrol Sci J.* 59(1):1–28.
- Lee S, Kim JC, Jung HS, Lee MJ, Lee S. 2017. Spatial prediction of flood susceptibility using random-forest and boosted-tree models in Seoul metropolitan city, Korea. *Geomatics Nat Hazards Risk.* 8(2):1185–1203.
- Lei X, Chen W, Panahi M, Falah F, Rahmati O, Uuemaa E, Kalantari Z, Ferreira CSS, Rezaie F, Tiefenbacher JP, et al. 2021. Urban flood modeling using deep-learning approaches in Seoul, South Korea. *J Hydrol.* 601:126684.
- Löwe R, Böhm J, Jensen DG, Leandro J, Rasmussen SH. 2021. U-flood–topographic deep learning for predicting urban pluvial flood water depth. *J Hydrol.* 603:126898.
- Lundberg SM, Lee SI. 2017. A unified approach to interpreting model predictions. In: Proceedings of the 31st international conference on neural information processing systems, p. 4768–4777.
- O’Neill E, Brereton F, Shahumyan H, Clinch JP. 2016. The impact of perceived flood exposure on flood-risk perception: the role of distance. *Risk Anal.* 36(11):2158–2186.
- Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. 2011. Scikit-learn: machine learning in Python. *J Mach Learn Res.* 12:2825–2830.
- Peel MC, Finlayson BL, McMahon TA. 2007. Updated world map of the Köppen-Geiger climate classification. *Hydrol Earth Syst Sci.* 11(5):1633–1644.

- Petroselli A. 2012. Lidar data and hydrological applications at the basin scale. *GIsci Remote Sens.* 49(1):139–162.
- Probst P, Wright MN, Boulesteix AL. 2019. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews Data Min Knowl Discov.* 9(3):e1301.
- Rahmati O, Pourghasemi HR, Zeinivand H. 2016. Flood susceptibility mapping using frequency ratio and weights-of-evidence models in the Golastan Province, Iran. *Geocarto Int.* 31(1):42–70.
- Raschka S. 2015. *Python machine learning*. Birmingham: Packt Publishing Ltd.
- Regmi AD, Devkota KC, Yoshida K, Pradhan B, Pourghasemi HR, Kumamoto T, Akgun A. 2014. Application of frequency ratio, statistical index, and weights-of-evidence models and their comparison in landslide susceptibility mapping in Central Nepal Himalaya. *Arab J Geosci.* 7(2):725–742.
- Rejith R, Anirudhan S, Sundararajan M. 2019. Delineation of ground water potential zones in hard rock terrain using integrated remote sensing, GIS and MCDM techniques: a case study from Vamanapuram River basin, Kerala, India. [place unknown]:[publisher unknown].
- Samela C, Persiano S, Bagli S, Luzzi V, Mazzoli P, Humer G, Reithofer A, Essenfelder A, Amadio M, Mysiak J, et al. 2020. Safer_RAIN: a DEM-based hierarchical filling-&-spilling algorithm for pluvial flood hazard assessment and mapping across large urban areas. *Water.* 12(6):1514.,.
- Schmitt TG, Thomas M, Etrich N. 2004. Analysis and modeling of flooding in urban drainage systems. *J Hydrol.* 299(3-4):300–311.
- Seleem O, Heistermann M, Bronstert A. 2021. Efficient hazard assessment for pluvial floods in urban environments: a benchmarking case study for the city of Berlin, Germany. *Water.* 13(18):2476.
- Singh P, Sinha VSP, Vijhani A, Pahuja N. 2018. Vulnerability assessment of urban road network from urban flood. *Int J Disaster Risk Reduct.* 28:237–250.
- Tehrany MS, Kumar L, Jebur MN, Shabani F. 2019. Evaluating the application of the statistical index method in flood susceptibility mapping and its comparison with frequency ratio and logistic regression methods. *Geomatics Nat Hazards Risk.* 10(1):79–101.
- Tehrany MS, Pradhan B, Jebur MN. 2014. Flood susceptibility mapping using a novel ensemble weights-of-evidence and support vector machine models in GIS. *J Hydrol.* 512:332–343.
- Tehrany MS, Pradhan B, Mansor S, Ahmad N. 2015. Flood susceptibility assessment using GIS-based support vector machine model with different kernel types. *Catena (Amst).* 125: 91–101.
- Termeh SVR, Kornejady A, Pourghasemi HR, Keesstra S. 2018. Flood susceptibility mapping using novel ensembles of adaptive neuro fuzzy inference system and metaheuristic algorithms. *Sci Total Environ.* 615:438–451.
- Trost SG, Zheng Y, Wong WK. 2014. Machine learning for activity recognition: hip versus wrist data. *Physiol Meas.* 35(11):2183–2189.
- Vafakhah M, Loor SMH, Pourghasemi H, Katebikord A. 2020. Comparing performance of random forest and adaptive neuro fuzzy inference system data mining models for flood susceptibility map ping. *Arab J Geosci.* 13(11):1–16.
- Viera AJ, Garrett JM. 2005. Understanding interobserver agreement: the kappa statistic. *Fam Med.* 37(5):360–363.
- Wang Y, Fang Z, Hong H, Peng L. 2020. Flood susceptibility mapping using convolutional neural network frameworks. *J Hydrol.* 582:124482.
- Wang Z, Lai C, Chen X, Yang B, Zhao S, Bai X. 2015. Flood hazard risk assessment model based on random forest. *J Hydrol.* 527:1130–1141.
- Wilson JP, Gallant JC. 2000. *Terrain analysis: Principles and applications*. New York (NY): Wiley. Chapter 4, Secondary Topographic Attributes. p. 87–131.
- Winterrath T, Brendel C, Hafer M, Junghänel T, Klameth A, Walawender E, Weigl E, Becker A. 2017. Erstellung einer radargestützten Niederschlagsklimatologie [Creation of a radar-based precipitation climatology]. Offenbach am Main: Deutschen Wetterdienstes [German Meteorological agency].

- Yacoub S, Simske S, Lin X, Burns J. 2003. Recognition of emotions in interactive voice response systems. In *Interspeech*; Sep 1–4; Geneva; Switzerland
- Yang TH, Hwang GD, Tsai CC, Ho JY. 2016. Using rainfall thresholds and ensemble precipitation forecasts to issue and improve urban inundation alerts. *Hydrol Earth Syst Sci.* 20(12): 4731–4745.
- Yin J, Yu D, Yin Z, Liu M, He Q. 2016. Evaluating the impact and risk of pluvial flash flood on intra-urban road network: a case study in the city center of Shanghai, China. *J Hydrol.* 537:138–145.
- Zhang S, Pan B. 2014. An urban storm-inundation simulation method based on GIS. *J Hydrol.* 517:260–268.
- Zhao G, Pang B, Xu Z, Cui L, Wang J, Zuo D, Peng D. 2021. Improving urban flood susceptibility mapping using transfer learning. *J Hydrol.* 602:126777.
- Zhao G, Pang B, Xu Z, Peng D, Zuo D. 2020. Urban flood susceptibility assessment based on convolutional neural networks. *J Hydrol.* 590:125235.
- Zhao G, Pang B, Xu Z, Yue J, Tu T. 2018. Mapping flood susceptibility in mountainous areas on a national scale in China. *Sci Total Environ.* 615:1133–1142.
- Zhou Q, Leng G, Su J, Ren Y. 2019. Comparison of urbanization and climate change impacts on urban flood volumes: importance of urban planning and drainage adaptation. *Sci Total Environ.* 658:24–33.