hydrology

MDPI

Article

Deciphering Urban Flood Drivers: An Explainable Machine Learning Approach to Vulnerability Assessment in Indonesian Catchments

Ahyahudin Sodri [1,*], Geovanny Branchiny Imasuly [2], Nuraeni Nuraeni [1] and Annisa Layyina Ihsani [3]

[1] Department of Environmental Science, Graduate School of Sustainable Development, Universitas Indonesia, Central Jakarta 10430, Indonesia; nuraeni42@ui.ac.id
[2] Department of Petroleum Engineering, Faculty of Engineering, Universitas Pattimura, Ambon 97233, Indonesia; geovanny.imasuly@lecturer.unpatti.ac.id
[3] Nutrition Department, Faculty of Public Health, Universitas Indonesia, Depok 16424, Indonesia; annisa.layyina@ui.ac.id
* Correspondence: ahyaudin.sodri@ui.ac.id

Abstract

Flooding is one of the most frequent and damaging natural disasters, accounting for nearly half of global disasters and posing a major challenge in Indonesia, where floods represent approximately 77% of all nationally recorded disaster events. Rapid urbanisation, land-use change, and climate-induced extreme rainfall have intensified flood risks nationwide. However, existing vulnerability assessments remain fragmented and localised, limiting their relevance for national-scale adaptation planning. This study develops a measurable and explainable framework for assessing urban flood vulnerability across Indonesia using cloud-based geospatial data and interpretable machine learning. The approach integrates CEMS-GLOFAS (flood hazard), WorldPop (population exposure), SRTM (topography), and ESA WorldCover (land cover) datasets within Google Earth Engine (GEE). Flood vulnerability is quantified through a modified Flood Vulnerability Index (FVI) combining hazard, exposure, and physical vulnerability components. The Extreme Gradient Boosting (XGBoost) model predicts FVI values, while SHapley Additive exPlanations (SHAP) and Partial Dependence Plots (PDPs) enhance model transparency and identify the influence of key variables such as flood depth, population density, and elevation. The model achieved high predictive accuracy (R^2 = 0.89; RMSE = 0.04728 FVI units, dimensionless) and revealed substantial spatial heterogeneity across 514 districts, with the highest FVI (0.75–0.85) in Banda Aceh, Mojokerto, Pasuruan, Samarinda, and Merauke. The integration of GEE and explainable AI offers a transparent, scalable framework to support data-driven flood risk mitigation and urban climate resilience in Indonesia.

Keywords: climate resilience; explainable machine learning; flood risk mapping; hydrology; urban flood vulnerability

Check for updates

Academic Editors: Qijie Li, Reinhard Hinkelmann and Dongfang Liang

Received: 10 May 2026
Revised: 24 June 2026
Accepted: 6 July 2026
Published: 11 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Floods are among the most frequent and destructive natural disasters, occurring across all climates and levels of development. Between 2000 and 2019, they accounted for 44% of global disasters and affected over 1.6 billion people [1]. In recent years, severe flooding events in both urban and rural areas across the globe have caused loss of life, infrastructure damage, and disruption of socio-economic activities, resulting in total global economic

losses of around US$651 billion (22%) [1–4]. This phenomenon demonstrates that floods represent a global threat transcending regional and developmental boundaries, with an estimated 1.81 billion people (23% of the global population) living in high-risk flood areas across 188 countries [5–7]. Driven by population growth, land-use change, and rapid urbanisation, the frequency and extent of urban flooding have increased by approximately 34% globally between 2000 and 2018 [5]. The expansion of impervious surfaces, wetland conversion, and the 7% rise in extreme rainfall intensity per 1 °C of global warming have further exacerbated flood risk [8–10].

In tropical regions, flooding has emerged as an increasingly complex environmental and socio-economic challenge. The combination of monsoonal climate patterns and rapid urbanisation heightens inundation potential, reduces natural infiltration capacity, and disrupts drainage system functionality [11–13]. The impacts extend beyond infrastructure and agricultural loss, threatening livelihoods, domestic wastewater and community well-being [14,15]. Consequently, accurate, adaptive, and data-driven flood risk mapping has become a critical element of climate resilience strategies and sustainable development planning [16,17]. In Indonesia, floods remain the most frequent and destructive natural disaster, accounting for approximately 77.44% of all disaster events and affecting millions of people each year [18]. Rapid urbanisation, land-use change, and increasing extreme rainfall have significantly amplified flood exposure across many Indonesian cities, particularly in Java [19,20]. However, high-resolution, regularly updated national flood vulnerability maps remain limited, despite their critical role in supporting community preparedness and evidence-based mitigation policies [21]. These conditions make Indonesia a relevant natural laboratory on a national scale, given its diversity in topography, climate, land use, population density, and regional development. Consequently, it serves as an ideal case study for testing a scalable flood vulnerability assessment approach that is applicable to highly heterogeneous regions.

Various approaches have been developed to assess flood hazards, risk, and vulnerability at different spatial scales. Conventional hydrological and hydraulic models, such as HEC-HMS [22], HEC-RAS [22,23], SWAT, LISFLOOD, TELEMAC, MIKE FLOOD, FINEL, SWMM, JFLOW, SCSCN, TUFLOW, and Infloworks [12,24], have been widely applied in simulating runoff generation processes, river flow dynamics, and flood inundation [25,26]. These models can produce detailed, highly accurate simulations. However, they require extensive input data, a complex calibration process, and significant computational time, and entail licensing costs. These limitations make it difficult to consistently apply the models in heterogeneous regions, particularly in developing regions facing data and computational resource constraints, such as Indonesia [24].

The advancement of cloud-based geospatial technology, particularly Google Earth Engine (GEE), offers an efficient solution for large-scale spatial analysis using multi-source remote sensing data [27,28]. Several studies in Southeast Asia have demonstrated the capability of integrating GEE with Sentinel-1 Synthetic Aperture Radar (SAR) imagery for rapid flood mapping. Mai Sy et al. [29] developed a GEE that combines GIS-based socioeconomic exposure data to assess flood risk in Thái Nguyên (Vietnam), finding that floods affected 33 educational facilities, 4 health facilities, 116.46 ha of road networks, and more than 2100 ha of agricultural land and residential areas. Tew et al. [30] used Sentinel-1 data and GEE to map flood inundation in the Kelantan River Basin (Malaysia), achieving 57–60% accuracy and identifying an inundation area of 8.92 km^2 during a major flood. Meanwhile, Ruuhulhaq et al. [31] used Sentinel-1 data, a Lee filter, and the Otsu thresholding method in GEE to successfully analyse recurring flood patterns and spatio-temporal dynamics in Bandung, Indonesia. This approach has proven to be adaptive and reliable in tropical and archipelagic contexts with high environmental complexity [32].

However, most GEE-based studies still focus on flood detection and inundation mapping, while vulnerability aspects integrating physical, social, environmental, and exposure factors remain relatively limited. Although these studies have successfully identified flood-prone areas using remote sensing data, the information generated does not yet fully support comprehensive flood risk reduction planning and strategies to enhance flood resilience.

On a global scale, various vulnerability indices have been developed to integrate the dimensions of hazard, exposure, and social vulnerability, namely, the Vulnerability-Adjusted Risk Index (VARI-Flood) [33] and the Global Empirical Social Vulnerability Index (GlobE SoVI) [34]. Both approaches make a significant contribution to cross-country flood risk assessment through standardised indicators that are globally comparable. However, because they are designed for a global scale, these indices generally use relatively coarse spatial resolution, limiting their ability to represent local variations in detail [21,25]. This poses a challenge in Indonesia, which exhibits highly diverse biophysical characteristics, land use patterns, urbanisation levels, and socioeconomic conditions [19,20]. Therefore, this study proposes an Urban Flood Vulnerability (UFV) framework that utilises higher-resolution data, integrates hazard, physical, and environmental factors, and spatial data to produce a more appropriate representation of flood vulnerability for Indonesian conditions.

In addition, various studies have begun to utilise Machine Learning (ML) in mapping flood risk and vulnerability, as it improves accuracy and captures the nonlinear relationships between environmental factors and flood events. Various algorithms include logistic regression (LR), Support Vector Machine (SVM), Random Forest (RF), Artificial Neural Network, Gradient Boosting Machine (GBM), and Extreme Gradient Boosting (XGBoost) [35,36]. Among these algorithms, ensemble-based methods, particularly XGBoost, generally perform better than other approaches at effectively and efficiently capturing nonlinear interactions and relationships between variables [37]. This confirms that XGBoost is not only accurate but also robust to variations in data distribution [37–39]. However, most of these studies still focus on improving prediction accuracy and have not provided clear interpretations of each variable; thus, limitations in transparency and interpretability remain key challenges for supporting national-level flood vulnerability assessments in developing countries with limited data and high geographical diversity, such as Indonesia.

Based on a review of the existing literature, previous studies still have three main limitations, namely, (1) the limited scalability of conventional hydrological-hydraulic approaches at the national level, (2) a dominant focus on predictive performance at the expense of model interpretability, and (3) the lack of integration of hazard factors, physical exposure, and the environment into a single transparent and reproducible analysis. To address this gap, the present study develops a new framework for constructing an Urban FVI in Indonesia by integrating GEE and XGBoost, enhanced with explainable ML techniques, namely, Shapley Additive exPlanations (SHAP) and Partial Dependence Plots (PDPs) [40–44]. This integration advances previous studies by coupling cloud-based geospatial big-data processing with a high-performance predictive model and post hoc interpretability within a unified, scalable framework. Unlike earlier works that either relied on GEE for geospatial analysis without robust predictive learning or applied ML models without explicit explainability, the proposed framework ensures both predictive accuracy and transparency. By incorporating global datasets, such as CEMS GLOFAS (flood hazard), WorldPop (population exposure), SRTM (topography), and ESA WorldCover (land cover), to produce consistent, scalable, and transparent vulnerability maps that can be adapted to flood-prone regions with similar environmental and data constraints. The use of EML ensures that the model does not function as a "black box" [40,43,45,46], but as an interpretable analytical tool that can quantify the contribution of each variable to the model output [47,48]. This also contributes to achieving the Sustainable Development Goals

(SDGs), specifically Sustainable Cities and Communities (SDG 11), Industry, Innovation, and Infrastructure (SDG 9), and Climate Action (SDG 13) [49,50].

In line with these objectives, this study is designed to answer the following questions: (1) How can multi-source spatial data processed using GEE be integrated to develop a national-scale UFV framework in Indonesia? (2) To what extent can the XGBoost algorithm model the relationship between flood vulnerability and relevant physical, environmental, and exposure factors in Indonesia? (3) How can SHAP and PDP analyses improve the interpretability of flood vulnerability models and support evidence-based decision-making in flood resilience planning?

2. Materials and Methods

2.1. Study Area

Indonesia is the world's largest archipelago, consisting of more than 17,500 islands and located at geographical coordinates of approximately 95.0° E–141.0° E; 11.0° S–6.0° N (Figure 1) [51]. This region exhibits a high degree of diversity in climatic conditions, topography, land use, and socioeconomic characteristics, ranging from densely populated urban areas to relatively remote rural and mountainous regions [51]. Additionally, Indonesia has a tropical climate with year-round high rainfall, and climate variability is influenced by the El Niño and La Niña phenomena, which help control the occurrence of floods, droughts, and extreme weather [8,49]. In recent decades, rising temperatures and changes in rainfall patterns have increased the frequency and intensity of hydrometeorological events in various regions of Indonesia [49,51].

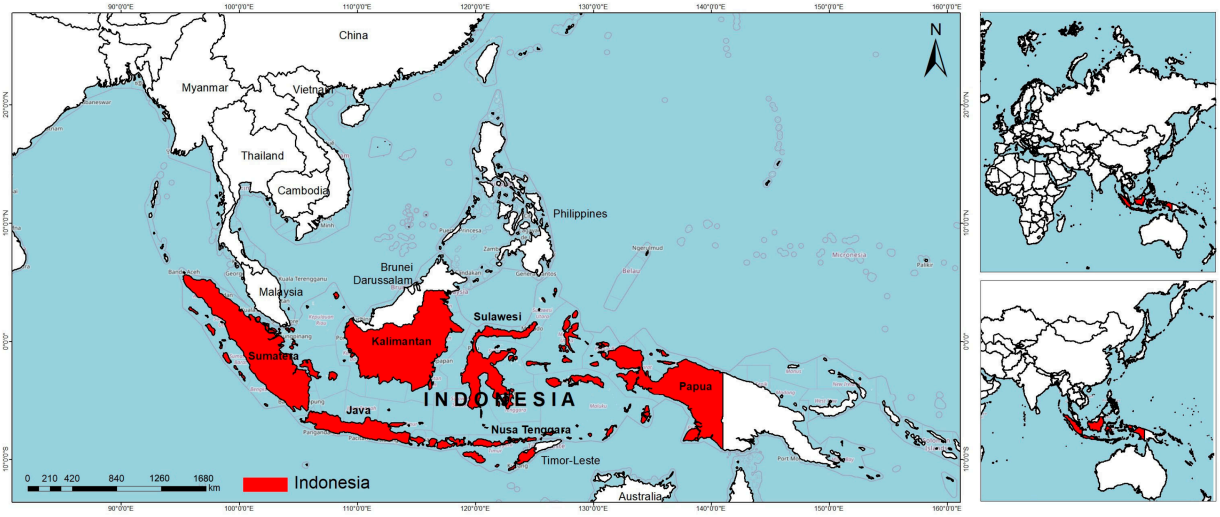


Figure 1. Location of the study area.

Rapid urbanisation, land-use changes, and population growth have heightened exposure and vulnerability to flood disasters [19,20]. Furthermore, topographic variations—including coastal lowlands, river basins, swamps, and mountainous regions—result in diverse hydrological characteristics across regions. Consequently, these conditions make flooding one of the most frequent natural disasters in Indonesia, causing significant impacts on communities, infrastructure, and economic activities [18]. Consequently, Indonesia serves as a representative region for national-scale flood research and mapping.

2.2. Data Sources

The datasets include flood hazard data from the Copernicus Emergency Management Service Global Flood Awareness System (CEMS GLOFAS), population data from WorldPop 2025, topographic data from the Shuttle Radar Topography Mission (SRTM), and land cover

data from ESA WorldCover 2020. The corresponding dataset references are provided in Table 1. The selection of these datasets was based on their broad spatial coverage, high resolution, and relevance in representing the main components of UFV in Indonesia. The overall research workflow is illustrated in Figure 2.

Table 1. Geospatial datasets used for flood vulnerability modelling.

Component	Dataset	Resolution	Year	Source
Flood hazard	CEMS-GLOFAS RP100 depth	~100 m	2022	Earth Engine Data
Population exposure	WorldPop unconstrained	100 m	2025	Earth Engine Data
Elev & derivatives	USGS SRTM GL1	30 m	2014	Earth Engine Data
Land cover	ESA WorldCover v100	10 m	2020	Earth Engine Data

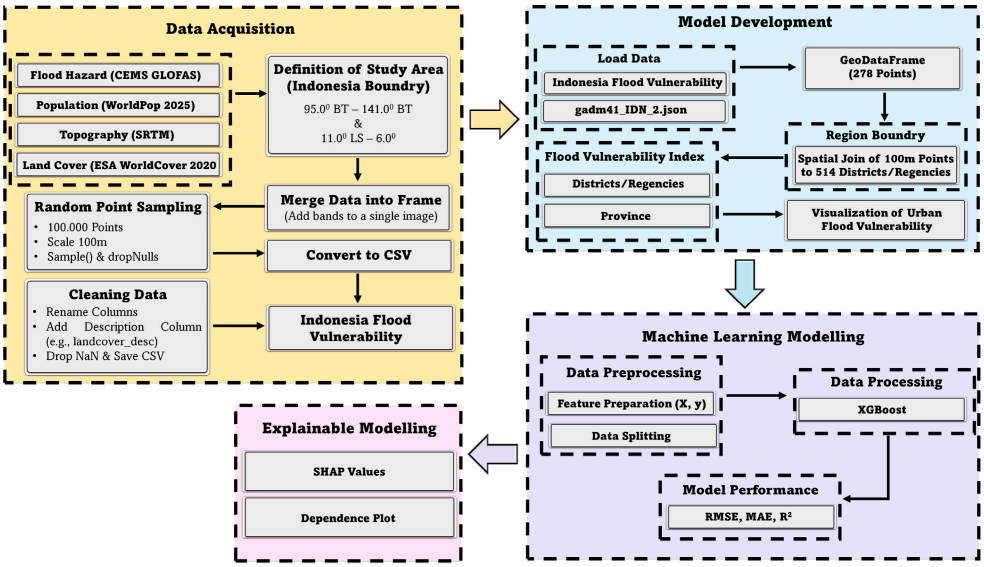


Figure 2. Research Flowchart.

Although the geospatial datasets used for this study were collected in different years—specifically 2014–2025—cross-temporal data integration is a common approach in flood vulnerability and risk studies because each variable represents a different dimension and has a non-uniform data update frequency [21,25]. Topographic variables derived from the USGS SRTM GL1 dataset (2014), such as elevation and slope, are considered relatively stable over time; geomorphological characteristics generally change very slowly. Meanwhile, flood hazard data were obtained from the Copernicus Emergency Management Service Global Flood Awareness System (CEMS GLOFAS RP100 Depth, 2022), land cover data from ESA WorldCover 2020, and population exposure data from WorldPop 2025. These datasets were selected based on their global coverage, high spatial consistency, and the availability of the most recent data products.

The use of datasets from different years is not intended to reconstruct flood conditions for a specific event or year, but rather to identify spatial patterns of flood vulnerability at the national scale. A similar approach has been widely applied in studies using global datasets for flood risk and vulnerability assessments, although with different data sources [21,25,33]. However, uncertainties may persist in areas experiencing rapid urbanisation and land-use change. Therefore, limitations arising from temporal differences between datasets have been taken into account in the interpretation of the research results.

Figure 2 shows the research workflow for UFV in Indonesia. The first stage involves acquiring and preprocessing four main geospatial datasets—Flood Hazard (CEMS-GLOFAS), Population Exposure (WorldPop), Topography (SRTM), and Land Cover (ESA

WorldPop)—using GEE to ensure consistency in extent, resolution, and projection. Next, approximately 100,000 randomly sampled points at a 100-m resolution were generated and integrated with all geospatial layers. Additionally, to ensure data consistency, only points with complete information across all variables were retained, while points with missing values in any layer were excluded. This process reduced the number of observations from the initial 100,000 points to 278 valid points due to differences in spatial coverage and data availability across datasets.

During the model development stage, the data were linked to administrative boundaries down to the regency/city level to calculate the Flood Vulnerability Index (FVI) and map the distribution of flood vulnerability across 514 regencies and cities in Indonesia. Additionally, in the final stage, XGBoost (version 2.1.4) and EML (SHAP version 0.49.1 and PDP version 0.3.0) were used. Model performance was evaluated using MAE, RMSE, and R^2, while the interpretation of factors influencing flood vulnerability was performed using SHAP and Dependence plots.

2.3. Data Acquisition

The data acquisition began by defining Indonesia's land area (95.0° E–141.0° E; 11.0° S–6.0° N) using the ee.Geometry.Rectangle function in GEE, which is the clipping extent for all datasets. All spatial layers were merged into a single image to ensure uniform extent, resolution, and projection. A total of 100,000 randomly sampled points at a 100 m spatial resolution were generated using the sample() function in GEE to represent the national spatial domain. Next, a multi-layer consistency check was performed to ensure complete data availability across all layers, while points with missing values in one or more variables were excluded from the analysis. This filtering process reduced the initial 100,000 points to 278 valid points, due to differences in spatial coverage and data availability across the datasets used. The extracted data were exported in CSV format, cleaned by renaming columns, adding descriptive variables (e.g., flood depth, population density, elevation, slope, aspect, land cover), and removing missing values (NaN). The resulting integrated dataset was prepared for modelling Indonesia's 2025 flood vulnerability using an EML–SHAP approach.

2.4. Model Development

2.4.1. Data Loading and Spatial Structuring

The process began with the retrieval of the acquired datasets, comprising the Indonesia Flood Vulnerability Dataset and the Level-2 administrative boundary data (district or city level) in GeoJSON format (gadm41_IDN_2.json). The administrative boundary dataset covers 514 districts or cities across 34 provinces, which were subsequently converted into a GeoDataFrame format to facilitate spatial integration and analysis.

Subsequently, data extraction was conducted using GEE through a random sampling method at a spatial resolution of 100 m, encompassing key variables such as flood depth, population density, elevation, slope, and land cover. At the initial stage, 100,000 random sample points were generated to represent the national spatial domain. However, due to strict multi-layer data integration requirements and differences in spatial coverage among geospatial datasets, only locations with complete observations across all variables were retained for analysis. This resulted in 278 sample points with fully consistent and valid multi-source geospatial information.

The final sample size of 278 points was not predetermined but rather resulted from a rigorous spatial quality filtering process that retained only areas with complete and consistent information across datasets. Each point therefore does not represent an individual administrative district but serves as a representative spatial unit measuring 100 m × 100 m,

reflecting the combined characteristics of flood hazard, exposure, and physical environmental conditions. This approach was adopted to ensure that the analysis focused solely on areas with reliable and consistent data while minimising potential bias caused by spatial data gaps or inconsistencies among layers. The representative points were subsequently aggregated at the district level to compute the national FVI, thereby preserving proportional spatial variation across Indonesia.

2.4.2. Flood Vulnerability Index (FVI) Computation

The FVI was calculated to assess the level of flood vulnerability in each administrative unit based on a combination of indicators representing the dimensions of hazard, exposure, and vulnerability. This approach adapts the conceptual model developed by [52–54], which emphasises the use of standardised indicators and weighting based on their relative contribution to flood risk [55].

Each indicator was normalised to a 0–1 scale using the rank percentile method, which is more robust to outliers compared with the traditional min–max normalisation. The FVI value was computed using Equation (1).

$$FVI = (0.40 \times H) + (0.35 \times E)[0.15 \times (1 - Elev)] + (0.10 \times BU) \quad (1)$$

The weighting of FVI components (0.40, 0.35, 0.15, and 0.10) was established based on the empirical contribution of each indicator to flood risk levels. This approach aligns with the principles of [54,55], who emphasised that weighting in flood vulnerability indices should reflect the relative influence of each factor on overall vulnerability.

The Hazard component (H, 0.40), represented by flood depth (flood_depth_m), received the highest weight as it directly determines the probability and intensity of flooding. The Exposure component (E, 0.35), measured by population density (pop_density_pers_per_pixel), ranked second due to its strong association with potential social and economic impacts, including the number of affected people and exposed assets. The Elevation ($Elev$, 0.15) and Built-up area (BU, 0.10) factors represent physical vulnerability elements influencing natural drainage capacity and surface runoff. These were assigned lower weights as their effects are moderating rather than primary when compared to hazard and exposure dimensions.

All calculations were performed using Python (version 3.9.9) (pandas version 2.1.4 and geopandas version 1.0.1 libraries) following the Weighted Linear Combination (WLC) approach [56–58], which maintains proportional balance among variables and ensures transparent interpretation within the global flood risk framework.

According to the study by [55], each indicator was normalised based on its statistical distribution and standardised into a common scale interval to prevent unit discrepancies from biasing the final results. The resulting FVI values were subsequently classified into six vulnerability levels, as presented in Table 2.

Table 2. Flood Vulnerability Index (FVI) Classification.

Vulnerability Classes	Distribution Interval
Very low	$0 \leq$ FVI < 0.2
Low	$0.2 \leq$ FVI < 0.4
Moderate	$0.4 \leq$ FVI < 0.6
High	$0.6 \leq$ FVI < 0.8
Very high	$0.8 \leq$ FVI < 0.95
Extreme	$0.95 \leq$ FVI ≤ 1.00

2.5. Machine Learning Modeling

2.5.1. Data Preprocessing

The data preprocessing stage involved a series of steps to prepare the dataset for model training. This included separating the features (X) as input variables and the target label (y) as the output variable, followed by the application of feature scaling, encoding, and data imputation techniques to enhance the quality and consistency of the dataset [59,60]. This stage is a fundamental step in the ML pipeline, as it determines the model's effectiveness and accuracy [61]. The feature set consisted of six main independent variables: flood_depth_m, pop_density_pers_per_pixel, elevation_m, slope_deg, aspect_deg, and landcover_encoded. The selection of these variables was based on their established relevance in influencing flood vulnerability, as supported by previous ML-based flood studies [62,63].

Subsequently, the dataset was divided into two subsets: a training set (80%) and a testing set (20%), using the hold-out validation method with controlled randomisation (random_state = 42) [64]. This method was employed to provide an objective estimate of model performance on unseen data, thereby enabling accurate assessment of generalisation capability [65,66]. Furthermore, the use of a fixed random_state ensured the reproducibility of the experimental results across repeated runs [66].

2.5.2. Data Processing

The data processing stage employed the eXtreme Gradient Boosting (XGBoost) algorithm, chosen for its proven ability to capture nonlinear relationships, computational efficiency, and high predictive accuracy in large-scale data-driven studies [37,38,67,68].

The selection of XGBoost was also motivated by the characteristics of Indonesia as a study area, which exhibits substantial spatial heterogeneity in topography, land cover, population distribution, and flood hazard conditions across more than 17,000 islands. These complex and nonlinear interactions among physical, environmental, and exposure factors require a modelling approach capable of handling heterogeneous geospatial data while maintaining robust predictive performance. Furthermore, the integration of SHAP and PDP was intended to enhance model transparency and support the interpretation of variable contributions, which is essential for evidence-based flood risk management and policy development at the national scale.

In this study, the XGBoost Regressor was employed, an ensemble decision tree method that applies gradient boosting optimisation to minimise the loss function and improve the model's generalisation capability [38,39,69]. The algorithm was selected for its scalability, robustness against overfitting, and effectiveness in hydrometeorological risk assessments [70].

Before training, both the training and testing datasets were converted into DMatrix, an internal data structure in XGBoost that enhances memory efficiency and computational speed [71]. The model was trained using the following main parameters: maximum tree depth (max_depth = 7), learning rate (learning_rate = 0.05), subsampling ratio (subsample = 0.8), and feature sampling ratio per tree (colsample_bytree = 0.8). The regression objective function (objective = "reg:squarederror") and histogram-based tree method (tree_method = "hist") were applied to improve model convergence.

The model training was performed for 2000 boosting iterations with an early stopping mechanism applied after 50 iterations to prevent overfitting and ensure convergence stability. The trained model was subsequently used to generate predictions of flood vulnerability values for both the training and testing datasets, enabling evaluation of the model's generalisation capability on unseen data [67]. In addition, feature importance analysis and SHAP interpretation were conducted to enhance model interpretability and to identify the most influential features contributing to the prediction outcomes [69].

2.5.3. Performance Model

The model performance was evaluated using three statistical metrics commonly applied in regression analysis: Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and the Coefficient of Determination (R^2). These metrics provide a comprehensive understanding of model accuracy and generalisation ability [72].

The XGBoost model was trained using the reg:squarederror objective function, which minimises the mean squared error (MSE) to produce continuous prediction values. The main parameters used included max_depth = 7, learning_rate = 0.05, subsample = 0.8, colsample_bytree = 0.8, and tree_method = "hist". The model was trained for a maximum of 2000 boosting iterations, with early stopping applied after 50 consecutive rounds without improvement in performance to prevent overfitting and to ensure convergence stability.

According to [73], the MAE represents the average of the absolute differences between the actual and predicted values, providing a proportional measure of model error regardless of its direction. The RMSE assigns a greater penalty to larger errors because it squares the deviations between observed and predicted values. As noted by [74], RMSE is particularly effective for evaluating models that are sensitive to outliers and for assessing prediction stability. The R^2, on the other hand, is used to evaluate how well the model explains the variance in the observed data. As described by [75], an R^2 value approaching 1 indicates a strong predictive performance, whereas a value close to 0 suggests weak predictive ability. The corresponding equations are presented as follows:

$$MAE \quad \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{2}$$

$$RMSE \quad \sqrt{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|^2} \tag{3}$$

$$R^2 \quad 1 - \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|^2}{\sum_{i=1}^{n}|y_i - \hat{y}_i|^2} \tag{4}$$

where n denotes the number of observations, y_i is the actual value at each data point, and $\hat{y}_i$ is the predicted value generated by the model. In Equations (2) and (3), the difference $|y_i - \hat{y}_i|$ refers to the prediction error for each prediction. The numerator in Equation (4), $\sum_{i=1}^{n}|y_i - \hat{y}_i|^2$, refers to the Sum of Squared Error (SSE), which is the sum of the squares of the differences between the observed values and the predicted values, and the denominator refers to the Sum of Squares Total (SST), which is the sum of the squares of the deviations of the observed values from their mean, used to measure the total variation in the observed data [76].

The combination of these three metrics is often employed simultaneously to provide a comprehensive evaluation of model performance. According to [77], the joint use of MAE and RMSE offers a more complete understanding of model errors, as RMSE is more sensitive to large deviations while MAE provides a more stable estimate of the average error magnitude.

In this study, the evaluation results of MAE, RMSE, and R^2 for both the training and testing datasets indicate the extent to which the XGBoost model is capable of generalising the relationships among variables that influence the level of UFV in Indonesia. Furthermore, these results strengthen the interpretability of the model through the application of the EML approach.

2.6. Explainable Modelling

Complex ML models such as XGBoost possess strong predictive capabilities but are often regarded as "black box" systems because their internal decision processes are difficult

to interpret directly [78]. To address this limitation, the EML approach was applied to provide transparent insights into how the model generates predictions and to quantify the contribution of each feature to the final outcome [79].

In this study, the XGBoost model was employed not only to predict the level of UFV but also to explain the relative influence of the contributing variables. Model interpretation was conducted using a post hoc explainability approach, which allows interpretation after the model has been trained without altering its internal structure [36,40].

The primary method adopted was SHAP [80], developed from game theory to measure the contribution of each feature to individual predictions. SHAP values indicate whether a given feature exerts a positive or negative effect on the model output, where positive values signify increased risk or vulnerability and negative values represent reduced risk contribution [81].

According to [40], SHAP provides advantages in terms of interpretative consistency and local accuracy, allowing scientific justification for the XGBoost model outputs. Consequently, each prediction can be traced back to its underlying variable combinations, enabling policymakers to identify and understand the most influential factors shaping flood vulnerability across different regions. In addition to SHAP, visualisation techniques such as the summary plot and dependence plot were employed to display the distribution of feature contributions at both global and individual levels [35,41]. These visual approaches facilitate the identification of dominant features such as extreme rainfall, elevation, population density, and land use, all of which play a significant role in determining the degree of flood vulnerability across regions. As a result, the outcomes are not only accurate but also interpretable, reproducible, and scientifically accountable within the broader context of disaster risk assessment and urban spatial planning in Indonesia.

3. Results

The increasing intensity of climate change and the rapid pace of urbanisation have significantly amplified flood risks across many Indonesian cities, particularly in areas with high population density and concentrated economic activities. Several studies have highlighted the rising frequency and intensity of flooding as a consequence of the combined effects of global climate change and uncontrolled expansion of built-up areas [5,82].

Therefore, spatial analysis of UFV is a critical step in supporting evidence-based policy formulation and adaptive disaster risk management, particularly in rapidly urbanising and flood-prone regions.

3.1. Visualisation of Urban Flood Vulnerability

Before presenting the FVI map, the spatial distributions of the predictor variables used in the modelling process are shown in Figure 3. These variables include flood hazard, population exposure, elevation, slope, aspect, and land cover. Visualising these input variables provides important context for understanding the spatial patterns of vulnerability shown in Figure 4.

Figure 3 shows significant spatial variation across Indonesia. Higher flood hazard values are generally concentrated in low-lying areas, river basins, and coastal zones. Meanwhile, population exposure is centred in major urban centres, particularly on the island of Java. Additionally, topographic variables—namely, elevation, slope, and aspect—reflect diverse terrain conditions. Meanwhile, land cover patterns distinguish built-up areas, agricultural land, and natural vegetation [83]. Thus, these variables form the basis for the FVI model and help explain the vulnerability patterns shown in Figure 4.

The FVI values calculated using Equation (1) were classified into six vulnerability categories as presented in Table 2. These values were then visualised spatially for all

514 districts and municipalities across Indonesia using a GeoDataFrame-based mapping approach. The resulting map illustrates the spatial distribution of flood vulnerability levels and identifies high-risk (hotspot) areas, as shown in Figure 4.

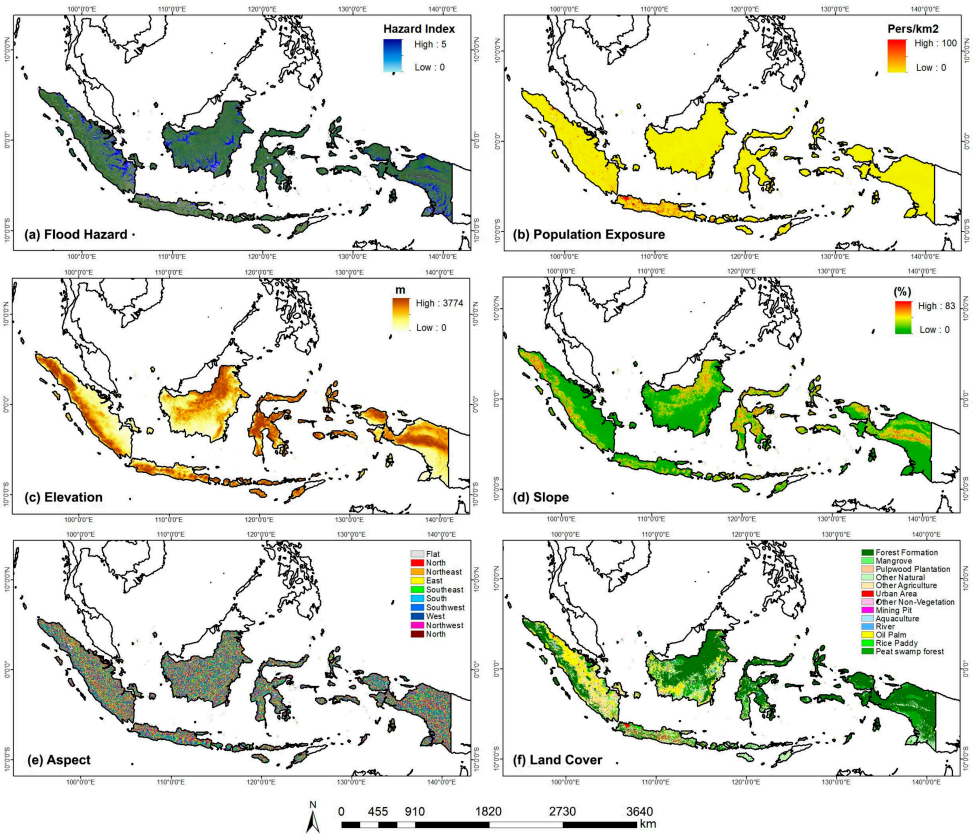


Figure 3. Spatial distribution of the predictor variables used for flood vulnerability assessment in Indonesia: (**a**) Flood Hazard, (**b**) Population Exposure, (**c**) Elevation, (**d**) Slope, (**e**) Aspect and (**f**) Land Cover. These variables were used as input features for flood vulnerability modelling.

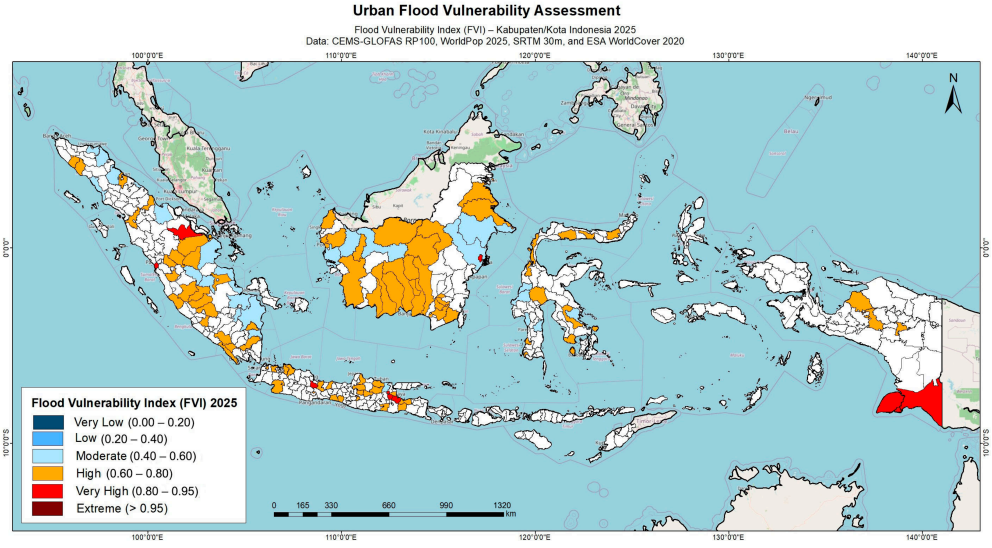


Figure 4. Urban Flood Vulnerability Assessment in Indonesia for 2025 based on the Flood Vulnerability Index (FVI), integrating datasets from CEMS-GLOFAS RP100, WorldPop 2025, SRTM 30m, and ESA WorldCover 2020.

3.2. Flood Vulnerability Index of Urban Areas in Indonesia

The results of the FVI analysis were subsequently aggregated at the administrative level of regencies and cities to identify spatial variations in flood vulnerability across

regions [52]. The ten regions with the highest FVI values in 2025 are presented in Table 3, representing urban areas with the greatest flood risk in Indonesia as determined using Equation (1).

Table 3. Flood vulnerability index (FVI) ranking of urban areas in Indonesia, 2025.

Rank	Province	City	FVI
1	Aceh	Banda Aceh	0.85
2	Jawa Timur	Mojokerto	0.85
3	Jawa Timur	Pasuruan	0.85
4	Kalimantan Timur	Samarinda	0.85
5	Jawa Barat	Kuningan	0.85
6	Riau	Siak	0.85
7	Papua	Merauke	0.85
8	Jawa Timur	Probolinggo	0.75
9	Jawa Timur	Kota Madiun	0.75
10	Jawa Timur	Lumanjang	0.75

Data Source: Calculated based on the Flood Vulnerability Index (FVI) model integrating data CEMS-GLOFAS RP100, WorldPop 2025, SRTM 30m, and ESA WorldCover 2020.

Based on the results in Table 3, which show the highest FVI values, Figure 5 presents a geographic map of the 10 urban areas with the highest flood vulnerability levels in Indonesia. Figure 5 provides spatial context for the FVI ranking results and clarifies flood vulnerability across Indonesia's major islands, thereby facilitating interpretation of national-level flood vulnerability.

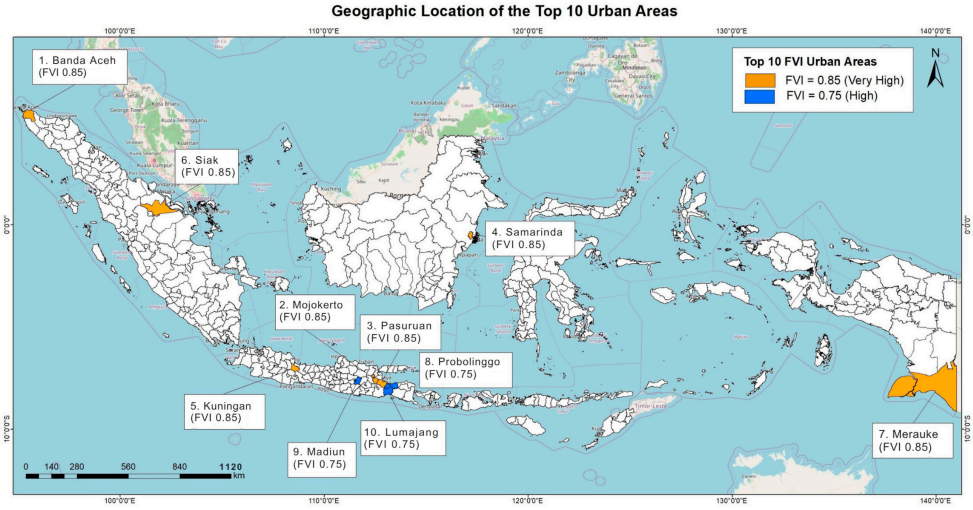


Figure 5. Geographic Distribution of the Top 10 Urban Areas with the Highest Flood Vulnerability Index (FVI) in Indonesia.

3.3. Model Performance and SHapley Additive exPlanations

The EML model developed in this study demonstrates a high level of predictive accuracy in estimating the FVI across urban areas in Indonesia. The model was constructed using an optimised XGBoost approach and enhanced with SHAP and PDP interpretability methods presented in Section 2.4. These methods were applied to identify the relative contribution of each variable to the FVI values.

3.3.1. Model Performance Evaluation

The model performance was evaluated using three primary metrics and Equations (2)–(4), applied to both the training and independent testing datasets. The testing dataset, com-

prising 20 percent of the samples, was used for hold-out validation to assess the model's generalisation capability.

As presented in Table 4, the XGBoost model demonstrated strong predictive performance on unseen data with an R^2 of 0.89, RMSE of 0.04728, and MAE of 0.03639, indicating a high level of agreement between predicted and observed FVI values. The training data achieved an R^2 of 0.99, showing the model's ability to effectively capture data patterns.

Table 4. Comparison of Model Performance Evaluation Metrics for XGBoost.

Metrics	Test	Train
RMSE	0.04728	0.00052
MAE	0.03639	0.00040
R^2	0.89	0.99

Note: RMSE and MAE are expressed in dimensionless FVI units.

Although a minor indication of overfitting was observed, which is common in complex models such as XGBoost, the high test performance (R^2 = 0.89) confirms that the model is robust and suitable for national-scale flood vulnerability mapping [84,85].

3.3.2. SHapley Additive Explanations and Feature Importance

The SHAP interpretability model was applied to evaluate the relative contribution of each variable to the model output, as visualised in Figures 6 and 7. The SHAP method quantifies how each input variable contributes to the model output by assigning an importance value to every feature in each observation. This allows model predictions to be interpreted in terms of environmental and socioeconomic drivers [40].

Global ranking refers to the dataset-level feature importance derived from the mean absolute SHAP values across all observations [40]. In simple terms, it indicates which variables are most influential in determining flood vulnerability at the overall dataset scale, regardless of whether their effect is positive or negative. Based on the SHAP analysis and the overall importance ranking derived from the mean absolute SHAP values, the five most influential variables affecting the FVI are flood depth (flood_depth_m), population density (pop_density_pers_per_pixel), elevation (elevation_m), slope (slope_deg), and aspect (aspect_deg).

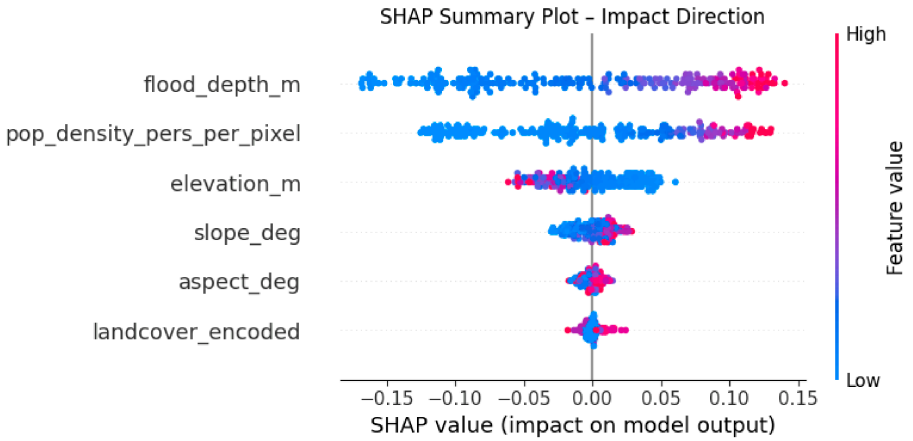


Figure 6. The SHAP summary plot illustrates the direction, magnitude, and distribution of individual predictor contributions to the prediction of the Flood Vulnerability Index (FVI) across urban catchments in Indonesia. Each point represents one observation; the horizontal position indicates the contribution of a variable to prediction (SHAP value), and the colour indicates whether the predictor value is relatively high (pink-purple) or low (blue).

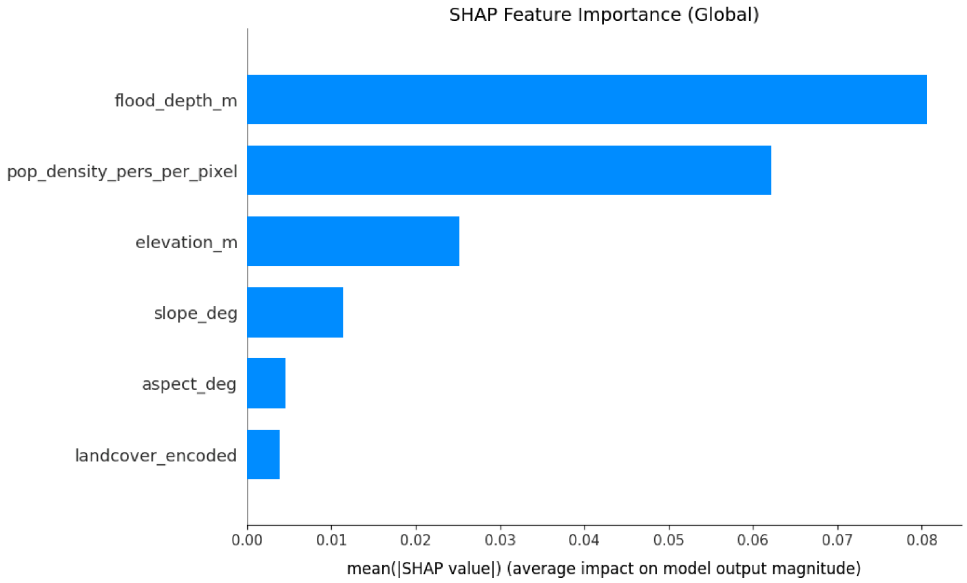


Figure 7. Feature importance ranking based on mean absolute SHAP values, highlighting the relative influence of hydrological, topographic, land-cover, and demographic variables on Flood Vulnerability Index (FVI) predictions across Indonesian urban catchments.

In simple terms, Figure 6 shows which environmental and socioeconomic conditions increase or reduce flood risk at each location. Features are ordered by overall importance, with those at the top having the strongest influence on model predictions. Among the predictors, flood depth (flood_depth_m) and population density (pop_density_pers_per_pixel) exert the strongest positive influence on FVI values. This finding suggests that areas with deeper inundation and higher population density tend to exhibit greater flood vulnerability. Conversely, elevation (elevation_m) shows a negative relationship with FVI, implying that low-lying areas are more prone to flooding. Slope and aspect (slope_deg and aspect_deg) exhibit minor protective effects, as steeper, well-oriented slopes can help reduce surface water accumulation. Meanwhile, land cover (landcover_encoded) has a relatively small but positive contribution to flood risk, particularly in densely built-up areas with impervious surfaces.

Figure 7 shows the overall ranking of variables by importance based on the average absolute SHAP values across the entire dataset. Figure 3, on the other hand, indicates whether a variable increases or decreases flood vulnerability but shows only its overall level of influence. Based on Figure 4, the flood depth variable (flood_depth_m) emerges as the most influential factor in determining flood vulnerability, followed by population density (pop_density_pers_per_pixel) and elevation (elevation_m). These three variables indicate that hydrological conditions and demographic pressures are the primary determinants of increased FVI values.

Meanwhile, elevation (elevation_m) makes a moderate contribution, reflecting topography's role in influencing surface water accumulation. The slope (slope_deg) and aspect (aspect_deg) variables show smaller yet relevant effects on surface water flow dynamics. The land cover variable (landcover_encoded) ranks last in relative importance but remains essential for identifying areas with impermeable surfaces that are more susceptible to flooding. These findings are consistent with the results presented in Table 5, which summarise the average SHAP value rankings of the five dominant features.

Table 5. The top five drivers of high vulnerability.

Feature	Importance
Flood_depth_m	2358
pop_density_pers_per_pixel	1896
elevation_m	1608
slope_deg	1357
aspect_deg	1056

3.3.3. Partial Dependence Plot and Interaction Insights

The relationships and interactions among the main variables influencing the flood risk prediction model are presented in Figure 8, which illustrates the dependence values of each feature based on their SHAP scores. The analysis shows that flood_depth_m exhibits the highest SHAP value of approximately 0.12, indicating that flood depth is the most dominant factor in increasing flood risk, with its effect stabilising at depths around five meters. Elevation_m demonstrates a negative SHAP value of about −0.06, suggesting that higher elevation significantly reduces flood risk.

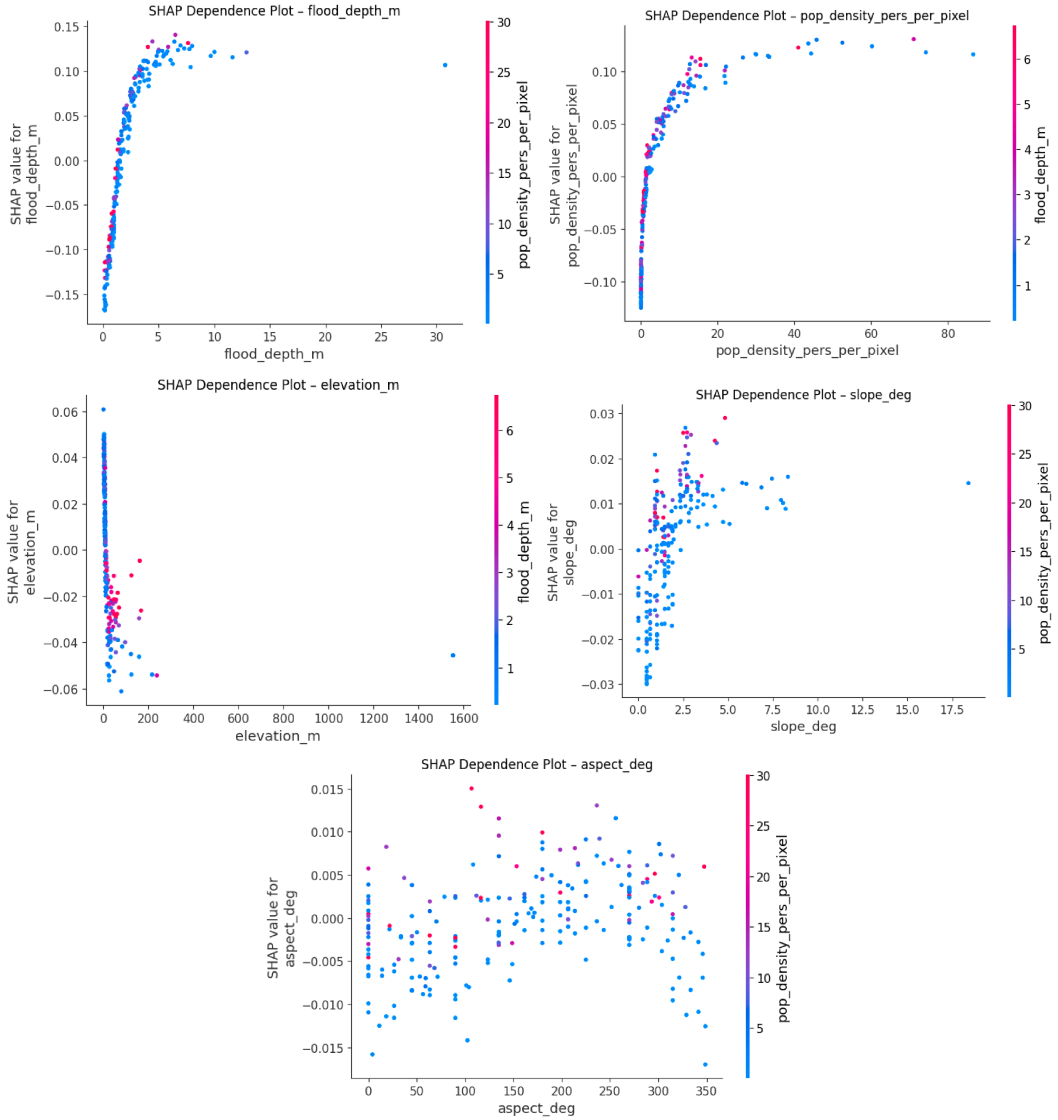


Figure 8. SHAP dependence plots illustrating the marginal effects and interaction patterns of flood depth, population density, elevation, slope, and aspect on Flood Vulnerability Index (FVI) predictions in Indonesia.

For the social variable, pop_density_pers_per_pixel shows a positive SHAP contribution of up to 0.11, indicating that high population density increases flood risk, particularly in areas experiencing deep inundation. Slope_deg contributes positively with a value of around 0.02 at gentle to moderate slopes (2–7 degrees), while aspect_deg exhibits a gradual increase up to 0.01, suggesting that slope orientation influences the spatial distribution of floodwater. Overall, the model effectively captures realistic interactions among flood depth, topography, and population density, jointly explaining the spatial variability of flood risk across the study area.

4. Discussion

4.1. Spatial Patterns of Urban Flood Vulnerability in Indonesia

Based on Figure 4, the spatial distribution of UFV in Indonesia in 2025 shows a heterogeneous pattern, influenced by topography, population density, and urbanisation. Areas categorised as having "High" to "Very High" vulnerability are predominantly concentrated along the northern coast of Java, parts of the western coast of Sumatra, and the southern region of Papua. These areas reflect a combination of low elevation, intense urban development, and high exposure to extreme rainfall runoff and potential sea-level rise [19,24,86].

Conversely, regions with "Low" to "Very Low" vulnerability are largely found in central Kalimantan, benefiting from extensive vegetation cover, limited urbanisation, and high soil infiltration capacity [20]. In contrast, most parts of northern Papua and several small islands in eastern Indonesia display either limited data availability or relatively low vulnerability. Areas classified as "Moderate" are generally distributed in transitional zones between highlands and coastal plains, indicating increased risk due to rapid settlement expansion unaccompanied by adequate drainage infrastructure [87].

This analysis underscores that the FVI is shaped by the integration of four key components. The spatial findings highlight the urgent need for place-based adaptation strategies and climate-resilient spatial planning, integrated across environmental, social, and infrastructural dimensions, to mitigate future urban flood risks in Indonesia effectively. The FVI values presented in Table 3 range from 0.75 to 0.85, corresponding to the "High" to "Very High" vulnerability categories as illustrated in the legend of Figure 4. Banda Aceh, Mojokerto, Pasuruan, Samarinda, Kuningan, Siak, and Merauke rank at the top with an FVI of 0.85, while Probolinggo, Madiun City, and Lumajang show slightly lower values but remain within the "Very High" category.

These results indicate that East Java Province dominates the list, with five of its cities included among the top ten. This reflects a combination of high population density, intensive land-use change, and limited drainage infrastructure [2], particularly in the Brantas River Basin. Meanwhile, coastal areas such as Banda Aceh, Siak, and Merauke face additional risks from tidal flooding and sea level rise [88], which exacerbate vulnerability in low-lying regions. This spatial pattern aligns with the zones shown in Figure 5, where the "High" to "Very High" categories are concentrated in coastal urban areas characterised by low elevation and high urbanisation. The interaction among geomorphological factors (low elevation), hydrometeorological processes (extreme runoff), and anthropogenic pressures (built-up areas and dense populations) explains why urban flood risk is unevenly distributed and tends to cluster in these regions [89].

From a disaster risk management perspective, these results highlight the urgency of strengthening adaptive infrastructure, such as integrated drainage systems, expanded green open spaces, and early warning systems. The implementation of risk-informed spatial planning and climate-resilient urban development is crucial to minimise losses in areas with high FVI values [90]. Overall, aggregating FVI at the district and city levels

not only complements the spatial insights shown in Figure 2 through quantitative data per region but also provides a scientific foundation for further analysis, including the application of the EML model (XGBoost SHAP) in the following subsection to identify the dominant factors influencing national flood vulnerability.

4.2. Dominant Drivers of Flood Vulnerability

From a hydrological perspective, the results align with established flood processes. Higher flood depths increase inundation extent, infrastructure damage, and socioeconomic disruption, thereby raising overall vulnerability [3,5,91]. Population density further amplifies impacts by concentrating exposure in flood-prone areas. The inverse relationship between elevation and FVI is also consistent with hydrological modelling, as low-lying areas tend to accumulate runoff and experience prolonged inundation. Overall, the SHAP results indicate that the model captures meaningful statistical patterns and is generally consistent with established hydrological principles. However, SHAP-based interpretations describe the statistical associations captured by the model, rather than physical mechanisms explicitly derived from the underlying hydrological processes.

Based on the results presented in Table 5, flood depth and population density emerged as the most influential predictors of UFV within the modelling framework used, while elevation and slope act as key topographic controls regulating surface water accumulation. These findings suggest that the variations in flood depth and exposure levels have stronger effects on vulnerability than other variables. Accordingly, risk reduction strategies should prioritise reducing inundation severity and managing population exposure in high-risk zones [2,3]. More broadly, flood vulnerability in Indonesia is shaped by the interaction between hydrological variables (flood_depth_m, elevation_m, slope_deg, aspect_deg) and anthropogenic factors (pop_density_pers_per_pixel), highlighting the combined influence of physical and human systems on urban flood risk.

The SHAP dependence analysis further supports consistency between model interpretation and spatial FVI patterns (Figure 8 and Table 3). High-impact variables, particularly flood depth (flood_depth_m = +0.12) and population density (pop_density_pers_per_pixel = +0.11), are associated with high FVI cities such as Banda Aceh, Samarinda, Mojokerto, and Pasuruan (FVI = 0.85), which are characterised by low elevation, dense settlements, and frequent inundation events. In contrast, lower-FVI cities such as Probolinggo, Madiun, and Lumajang (FVI = 0.75) exhibited weaker SHAP contributions for elevation (elevation_m = 0.06) and slope (slope_deg = +0.02), reflecting more favourable topographic conditions that reduce flood susceptibility. This consistency indicates that the EML model can reproduce general spatial patterns of vulnerability, with hydrological, topographical, and social factors collectively associated with spatial variations in urban flood risk across Indonesia. However, the patterns identified do not represent causal mechanisms; rather, they indicate consistent spatial associations between predictor variables and flood vulnerability within the dataset.

The SHAP results indicate that the model captures patterns consistent with known hydrological and socio-spatial processes, but do not provide direct evidence of physical causality. The integration of SHAP and FVI mapping demonstrates the model's robustness from both a predictive and interpretative perspective. Nevertheless, this robustness cannot be regarded as a causal validation of the underlying physical processes of flooding. It is important to emphasise that SHAP scores quantify the statistical associations learned by the XGBoost model, rather than direct physical causal relationships. Therefore, feature importance scores should not be interpreted as evidence of causality. Instead, SHAP scores reflect the contribution of each predictor to the model's output within the constraints of the dataset and the modelling framework used.

Building on these findings, the following section compares Indonesia's flood vulnerability drivers with those reported in other tropical regions, highlighting similarities and regional specificities in flood generation mechanisms.

4.3. Comparison with Other Tropical Regions

The results of this study confirm that urban flood vulnerability in tropical regions is primarily influenced by interactions among hydrological, topographical, and socioeconomic factors. In line with recent EML studies, the use of XGBoost-SHAP has been shown to improve interpretability and spatial accuracy in flood risk modelling [62,73]. In this study, flood depth (flood_depth_m) and population density (pop_density_per_pixel) emerged as dominant factors shaping the FVI, supporting the finding that low-lying areas with high population density are vulnerable to flooding due to limited drainage capacity and high impervious surface coverage [5,6]. The consistency between SHAP results and the spatial patterns of FVI is also consistent with large-scale studies based on multi-source data and remote sensing [30,31]. In general, these findings align with the literature from various tropical regions, where rapid urbanisation, land-use changes, increased impervious surfaces, population growth, and inadequate drainage systems are the primary factors controlling flood risk [7,10–12,17]. These processes reduce infiltration capacity and increase surface runoff, thereby amplifying the frequency and intensity of flooding during extreme rainfall. Similar patterns are commonly observed in Southeast Asia and South Asia, where socioeconomic development often outpaces effective mitigation and spatial planning capacities [12,17].

A similar pattern is observed in the Mekong region, Thailand, and South Asia, where flood risk is strongly influenced by the interaction between hydrometeorological extremes and increased levels of exposure in flood-prone areas [12–14]. This is consistent with conditions in Indonesia, where densely populated lowland urban areas repeatedly exhibit high levels of vulnerability. On a global scale, various studies also indicate that flood risk is driven by the concentration of people and economic assets in disaster-prone areas [2,3,5,6], underscoring the dominant role of exposure in shaping flood vulnerability.

However, Indonesia has distinct characteristics that set it apart from many other tropical countries. The archipelagic geography and long coastline expose urban areas in Indonesia to various types of floods that interact with one another, including fluvial, pluvial, coastal, tidal, and subsidence-induced floods [24,86,88]. This multi-hazard configuration is particularly evident in coastal cities such as Banda Aceh, Siak, and Merauke, where coastal processes and sea-level rise exacerbate the impacts of flooding. Consequently, the spatial heterogeneity of vulnerability in Indonesia is higher compared to many continental tropical regions. Additionally, variations in topography, river basin characteristics, urban development patterns, and infrastructure capacity further reinforce the heterogeneity of flood control factors across Indonesia. This indicates that flood processes are highly context-dependent and scale-dependent. Thus, these findings are consistent with the international literature on tropical flood systems while highlighting the additional complexity arising from Indonesia's archipelagic conditions and multi-hazard environment.

4.4. Implications for Flood Risk Management

From both scientific and practical perspectives, this study's results indicate that EML provides a robust, interpretable framework for assessing urban flood vulnerability by linking complex AI-based models with policy-relevant information [40].

These findings confirm that areas with high FVI, such as Banda Aceh, Samarinda, Mojokerto, and Pasuruan, need to be prioritised in mitigation strategies. Specifically, this involves improving urban drainage systems, expanding green open spaces, and implement-

ing nature-based solutions. This approach aligns with efforts to reduce disaster risk and enhance urban resilience. Additionally, integrating SHAP results can support risk-based decision-making in more adaptive and evidence-based spatial planning. This also contributes to achieving the SDGs, particularly Industry, Innovation, and Infrastructure (SDG 9), Sustainable Cities and Communities (SDG 11), and Climate Action (SDG 13) [49,50].

4.5. Limitations and Future Research

This study has a major limitation: a reduction in sample size due to the multi-layer consistency filtering process. This procedure ensures that only observations with complete information across all spatial variables are retained, thereby maintaining consistency. However, it reduces the dataset from approximately 100,000 spatial points to 278 observations, which may affect the statistical representativeness, model stability, and the generalisability of the findings.

Although the final dataset represents a rigorously harmonised subset derived from multiple geospatial datasets, this reduction introduces potential spatial sampling bias. Areas with incomplete or inconsistent data coverage, particularly remote islands, mountainous regions, and sparsely populated areas, are underrepresented relative to urban regions with more complete data. Consequently, the results should be interpreted as relative indicators of flood vulnerability rather than exact or fully generalisable representations of local conditions. Despite this limitation, the retained observations remain spatially distributed across Indonesia and capture a wide range of environmental, topographic, and socioeconomic conditions. This enables the model to preserve meaningful large-scale spatial variability relevant to national-scale flood vulnerability assessment, although it does not fully capture fine-scale local heterogeneity. Therefore, additional validation using higher-resolution regional datasets is recommended to strengthen the robustness of the findings.

Further uncertainty arises from integrating multi-source geospatial datasets with heterogeneous spatial resolutions, temporal coverage (2014–2025), and varying data quality. The required harmonisation and resampling processes may reduce fine-scale spatial variability and limit sensitivity to local heterogeneity, while temporal inconsistencies may introduce bias in rapidly urbanising areas. These uncertainties propagate through both the FVI construction and the ML model. In particular, SHAP-based interpretations reflect statistical associations rather than causal relationships and should not be interpreted as physical causation. Overall, uncertainties associated with hazard, exposure, land cover, and topographic datasets influence model outputs. Therefore, the resulting vulnerability estimates should be interpreted as data-driven, relative indicators of vulnerability rather than precise or physically deterministic representations of flood risk. The framework is thus more suitable for identifying broad spatial patterns at the national scale than for detailed local-scale analysis.

Although this framework is designed to be scalable using global datasets and cloud-based geospatial processing, its transferability to other regions is limited by differences in environmental, hydrological, climatic and socioeconomic conditions. These differences may lead to distributional shifts affecting model performance. Therefore, external validation and recalibration are required before applying the framework outside the Indonesian context.

Future research should incorporate higher-resolution local datasets and conduct independent validation across diverse geographic settings to improve robustness. In addition, integrating CMIP6 climate projections and dynamic socioeconomic indicators would enable spatiotemporal forecasting under climate change and urbanisation scenarios, thereby enhancing the temporal robustness, predictive capability, and policy relevance of flood vulnerability assessments.

5. Conclusions

This study develops a scalable and interpretable Urban Flood Vulnerability (UFV) assessment framework by integrating Google Earth Engine (GEE) with Machine Learning and Explainable Machine Learning (EML), specifically XGBoost combined with SHAP-based interpretation. This proposed framework addresses a critical gap in flood risk research by linking high-resolution multi-source spatial data with robust predictive models, while maintaining transparency in the decision-making process. Unlike conventional approaches that typically focus only on simulation details or predictive performance separately, this study demonstrates that accuracy and interpretability can be achieved simultaneously within a single integrated analysis workflow.

Beyond its methodological contributions, this study's main contribution lies in its broad implications for data-driven disaster risk management, particularly in regions with limited data and highly diverse geographical characteristics. The integration of EML enables complex flood vulnerability patterns to be broken down into meaningful physical and social factors, thereby supporting more transparent and evidence-based planning and decision-making. This approach is highly relevant for tropical countries experiencing rapid urbanisation, where flood risks are shaped by interactions among hydrological processes, land-use changes, and high population concentrations, often amid limited monitoring capacity.

From an applied perspective, the developed framework provides a scalable framework for supporting national-scale flood resilience strategies and climate change adaptation planning by generating consistent vulnerability maps across various environmental conditions. Furthermore, this approach can support the prioritisation of risk reduction policies, infrastructure investments, and land-use interventions aligned with the sustainable development agenda, including climate adaptation and sustainable urban development. Overall, this research demonstrates that integrating geospatial big data analytics with explainable AI provides a practical pathway toward a flood risk assessment system that is more transparent, scalable, and aligned with policy needs. Nevertheless, the framework developed in this study should primarily be interpreted as a national-scale decision-support tool. Its application to regions with diverse geographical, hydrological, climatic, and socioeconomic characteristics, as well as to finer-scale local assessments, requires external validation and recalibration to ensure the reliability of the resulting outputs.

Author Contributions: Conceptualisation, A.S., G.B.I., N.N. and A.L.I.; methodology, A.S. and G.B.I.; software, G.B.I.; validation, A.S. and N.N.; formal analysis, A.S. and G.B.I.; investigation, A.S. and G.B.I.; resources, A.S. and G.B.I.; data curation, A.S. and G.B.I.; writing—original draft preparation, A.S., G.B.I., N.N. and A.L.I.; writing—review and editing, A.S., G.B.I., N.N. and A.L.I.; visualisation, G.B.I.; supervision, A.S.; project administration, A.S. and A.L.I.; funding acquisition, A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Research and Development (Risbang), Universitas Indonesia, with a research scheme on the PUTI Q1 2024 and with contract numbers, grant number [NKB-550/UN2.RST/HKP.05.00/2024].

Institutional Review Board Statement: Ethical review and approval were waived for this study due to the analysis of anonymous data collected through a survey method that does not reveal participant identities, focuses on non-sensitive issues, and poses no risk of physical or psychological harm to the individuals involved.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Acknowledgments: Thanks to Research and Development (Risbang) Universitas Indonesia for supporting this research and also to the Institute for Advanced Science, Social, and Sustainable Future (IASSSF), which helped improve grammar.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. United Nations Office for Disarmament Affairs. *Human Cost of Disasters*; United Nations Office for Disarmament Affairs: New York, NY, USA, 2020. [CrossRef]
2. Winsemius, H.C.; Aerts, J.C.J.H.; van Beek, L.P.H.; Bierkens, M.F.P.; Bouwman, A.; Jongman, B.; Kwadijk, J.C.J.; Ligtvoet, W.; Lucas, P.L.; van Vuuren, D.P.; et al. Global drivers of future river flood risk. *Nat. Clim. Change* **2016**, *6*, 381–385. [CrossRef]
3. McDermott, T.K.J. Global exposure to flood risk and poverty. *Nat. Commun.* **2022**, *13*, 3529. [CrossRef] [PubMed]
4. Yulaichah. Analyzing the impact of land use change on flood risk and social vulnerability using SCS-CN method and GIS. *ASEAN Nat. Disaster Mitig. Educ. J.* **2026**, *3*, 101–116. [CrossRef]
5. Tellman, B.; Sullivan, J.A.; Kuhn, C.; Kettner, A.J.; Doyle, C.S.; Brakenridge, G.R.; Erickson, T.A.; Slayback, D.A. Satellite imaging reveals increased proportion of population exposed to floods. *Nature* **2021**, *596*, 80–86. [CrossRef] [PubMed]
6. Rentschler, J.; Salhab, M.; Jafino, B.A. Flood exposure and poverty in 188 countries. *Nat. Commun.* **2022**, *13*, 3527. [CrossRef] [PubMed]
7. Idowu, D.; Zhou, W. Global Megacities and Frequent Floods: Correlation between Urban Expansion Patterns and Urban Flood Hazards. *Sustainability* **2023**, *15*, 2514. [CrossRef]
8. Trenberth, K.E. Changes in precipitation with climate change. *Clim. Res.* **2011**, *47*, 123–138. [CrossRef]
9. Shariat, R.; Roozbahani, A.; Ebrahimian, A. Risk analysis of urban stormwater infrastructure systems using fuzzy spatial multi-criteria decision making. *Sci. Total Environ.* **2019**, *647*, 1468–1477. [CrossRef] [PubMed]
10. Zhou, Q.; Leng, G.; Su, J.; Ren, Y. Comparison of urbanization and climate change impacts on urban flood volumes: Importance of urban planning and drainage adaptation. *Sci. Total Environ.* **2019**, *658*, 24–33. [CrossRef] [PubMed]
11. Feng, B.; Zhang, Y.; Bourke, R. Urbanization impacts on flood risks based on urban growth data and coupled flood models. *Nat. Hazards* **2021**, *106*, 613–627. [CrossRef]
12. Manandhar, B.; Cui, S.; Wang, L.; Shrestha, S. Urban Flood Hazard Assessment and Management Practices in South Asia: A Review. *Land* **2023**, *12*, 627. [CrossRef]
13. Chen, A.; Pokhrel, Y.; Chen, D.; Huang, H.; Dai, Z.; He, B.; Wang, J.; Li, J.; Wang, H.; Liu, J. Impact of tropical cyclones and socioeconomic exposure on flood risk distribution in the Mekong Basin. *Commun. Earth Environ.* **2024**, *5*, 704. [CrossRef]
14. Utari, A.W.; Herdiansyah, H. Using filtration as a technology to remove pollutants in domestic wastewater. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *725*, 012025. [CrossRef]
15. Yahya, A.S.; Syahputra, D.A. Community preparedness for flood disasters: The role of knowledge and attitudinal factors in disaster-prone areas. *Calam. A J. Disaster Technol. Eng.* **2025**, *3*, 58–67. [CrossRef]
16. Sava, E.; Cervone, G.; Kalyanapu, A. Multiscale Observation Product (MOP) for Temporal Flood Inundation Mapping of the 2015 Dallas Texas Flood. *Remote Sens.* **2023**, *15*, 1615. [CrossRef]
17. Dharmarathne, G.; Waduge, A.O.; Bogahawaththa, M.; Rathnayake, U.; Meddage, D.P.P. Adapting cities to the surge: A comprehensive review of climate-induced urban flooding. *Results Eng.* **2024**, *22*, 102123. [CrossRef]
18. BNPB. *Data Bencana Indonesia 2024*; BNPB: Jakarta, Indonesia, 2024.
19. Handayani, W.; Chigbu, U.E.; Rudiarto, I.; Putri, I.H.S. Urbanization and Increasing Flood Risk in the Northern Coast of Central Java—Indonesia: An Assessment towards Better Land Use Policy and Flood Management. *Land* **2020**, *9*, 343. [CrossRef]
20. Dewi, R.S.; Handayani, W.; Pratama, I.P.; De Vries, W.T.; Rudiarto, I.; Artiningsih, A. Assessing Flood Vulnerability from Rapid Urban Growth: A Case of Central Java—Indonesia. *Chin. J. Urban Environ. Stud.* **2023**, *11*, 2350020. [CrossRef]
21. Carr, A.B.; Trigg, M.A.; Haile, A.T.; Bernhofen, M.V.; Alemu, A.N.; Bekele, T.W.; Walsh, C.L. Using global datasets to estimate flood exposure at the city scale: An evaluation in Addis Ababa. *Front. Environ. Sci.* **2024**, *12*, 1330295. [CrossRef]
22. USACE. *Hydrologic Modeling System*; Hydrologic Engineering Center: Davis, CA, USA, 2000.
23. Shabani, A.; Woznicki, S.A.; Mehaffey, M.; Butcher, J.; Wool, T.A.; Whung, P.-Y. A coupled hydrodynamic (HEC-RAS 2D) and water quality model (WASP) for simulating flood-induced soil, sediment, and contaminant transport. *J. Flood Risk Manag.* **2021**, *14*, e12747. [CrossRef] [PubMed]
24. Pakati, S.S.; Shoko, C.; Dube, T. Integrated flood modelling and risk assessment in urban areas: A review on applications, strengths, limitations and future research directions. *J. Hydrol. Reg. Stud.* **2025**, *61*, 102583. [CrossRef]
25. de Moel, H.; Jongman, B.; Kreibich, H.; Merz, B.; Penning-Rowsell, E.; Ward, P.J. Flood risk assessments at different spatial scales. *Mitig. Adapt. Strateg. Glob. Change* **2015**, *20*, 865–890. [CrossRef] [PubMed]

26. Budiyono, Y.; Aerts, J.C.J.H.; Tollenaar, D.; Ward, P.J. River flood risk in Jakarta under scenarios of future change. *Nat. Hazards Earth Syst. Sci.* **2016**, *16*, 757–774. [CrossRef]
27. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
28. Tamiminia, H.; Salehi, B.; Mahdianpari, M.; Quackenbush, L.; Adeli, S.; Brisco, B. Google Earth Engine for geo-big data applications: A meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* **2020**, *164*, 152–170. [CrossRef]
29. Mai Sy, H.; Luu, C.; Bui, Q.D.; Ha, H.; Nguyen, D.Q. Urban flood risk assessment using Sentinel-1 on the google earth engine: A case study in Thai Nguyen city, Vietnam. *Remote Sens. Appl. Soc. Environ.* **2023**, *31*, 100987. [CrossRef]
30. Tew, Y.L.; Tan, M.L.; Juneng, L.; Osman, S.; Hassan, M.H.; Samat, N.; Chang, C.K.; Li, L. Flood inundation mapping in the Kelantan River Basin, Malaysia, using Sentinel-1 SAR and Google Earth Engine. *Malays. J. Soc. Space* **2024**, *20*, 13–25. [CrossRef]
31. Ruuhulhaq, M.S.; Rohman, A.; Shalih, O. Exploring Google Earth Engine for Flood Detection (A Case Study in Bandung City). *Int. J. Disaster Dev. Interface* **2025**, *5*, 223–245. [CrossRef]
32. Siddiq, A.; Sejati, A.W. Combining Landsat, VIIRS Night-time Light, and Sentinel-1 SAR for Spatial Flood Risk Assessment in Coastal Area: An Earth Engine Cloud Computing Process. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1264*, 012003. [CrossRef]
33. Fox, S.; Agyemang, F.; Hawker, L.; Neal, J. Integrating social vulnerability into high-resolution global flood risk mapping. *Nat. Commun.* **2024**, *15*, 3155. [CrossRef] [PubMed]
34. Reimann, L.; Koks, E.; de Moel, H.; Ton, M.J.; Aerts, J.C.J.H. An Empirical Social Vulnerability Map for Flood Risk Assessment at Global Scale ("GlobE-SoVI"). *Earth's Future* **2024**, *12*, e2023EF003895. [CrossRef]
35. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
36. Khamil, S.T.; Kosuke, M.; Supriatna, S. Resilient Greenport Conceptual Model: Integrating Masterplan with Flood and Land Subsidence Disaster Management. *J. Environ. Sci. Sustain. Dev.* **2025**, *8*, 101–126. [CrossRef]
37. Velarde, G.; Weichert, M.; Deshmunkh, A.; Deshmane, S.; Sudhir, A.; Sharma, K.; Joshi, V. Tree boosting methods for balanced and imbalanced classification and their robustness over time in risk assessment. *Intell. Syst. Appl.* **2024**, *22*, 200354. [CrossRef]
38. Hakkal, S.; Lahcen, A.A. XGBoost To Enhance Learner Performance Prediction. *Comput. Educ. Artif. Intell.* **2024**, *7*, 100254. [CrossRef]
39. A'syifa, C.R.G.A.; Pramudita, B.A.; Yudiansyah; Rohman, B.P.A.; Setiawan, D.P. Comparing Predictive Capabilities of Machine Learning Models in Weather Forecasting. In *Proceedings of the 2025 International Conference on Data Science and Its Applications (ICoDSA)*; IEEE: Piscataway, NJ, USA, 2025; pp. 973–978. [CrossRef]
40. Lundberg, S.M.; Lee, S. A Unified Approach to Interpreting Model Predictions. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–10. [CrossRef]
41. Greenwell, B.M. pdp: An R package for constructing partial dependence plots. *R J.* **2017**, *9*, 421–436. [CrossRef]
42. Veeramsetty, V. Shapley value cooperative game theory-based locational marginal price computation for loss and emission reduction. *Prot. Control Mod. Power Syst.* **2021**, *6*, 33. [CrossRef]
43. Chen, H.; Lundberg, S.M.; Lee, S.-I. Explaining a series of models by propagating Shapley values. *Nat. Commun.* **2022**, *13*, 4512. [CrossRef] [PubMed]
44. Atzmueller, M.; Fürnkranz, J.; Kliegr, T.; Schmid, U. Explainable and interpretable machine learning and data mining. *Data Min. Knowl. Discov.* **2024**, *38*, 2571–2595. [CrossRef]
45. Saranya, A.; Subhashini, R. A systematic review of Explainable Artificial Intelligence models and applications: Recent developments and future trends. *Decis. Anal. J.* **2023**, *7*, 100230. [CrossRef]
46. Nam, K.; Lee, Y.; Lee, S.; Kim, S.; Zhang, S. Explainable Artificial Intelligence (XAI) for Flood Susceptibility Assessment in Seoul: Leveraging Evolutionary and Bayesian AutoML Optimization. *Remote Sens.* **2025**, *17*, 2244. [CrossRef]
47. Derman, M.T.; Mete, M.O. Flood Risk Analysis with Explainable Geospatial Artificial Intelligence (GeoAI) Techniques. *Systems* **2025**, *13*, 1007. [CrossRef]
48. Tan, C.; Ke, E.; Shi, H. Leveraging Explainable Artificial Intelligence for Place-Based and Quantitative Strategies in Urban Pluvial Flooding Management. *ISPRS Int. J. Geo-Inf.* **2025**, *14*, 475. [CrossRef]
49. IPCC. *Climate Change 2022: Impacts, Adaptation and Vulnerability*; Cambridge University Press: Cambridge, UK, 2022.
50. Lakshmi, D.; Tiwari, R.S.; Dhanaraj, R.K.; Kadry, S. *Explainable AI (XAI) for Sustainable Development*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2024. [CrossRef]
51. The World Bank Group. *Climate Risk Country Profile: Indonesia*; World Bank: Washington, DC, USA; Asian Development Bank: Manila, Philippines, 2025. Available online: https://climateknowledgeportal.worldbank.org/sites/default/files/country-profiles/17254-WB_Indonesia-Country%20Profile-WEB.pdf (accessed on 24 June 2026).
52. Balica, S.F.; Wright, N.G.; van der Meulen, F. A flood vulnerability index for coastal cities and its use in assessing climate change impacts. *Nat. Hazards* **2012**, *64*, 73–105. [CrossRef]
53. Tajuddin, N.F.A.; Ahmad Tarmizi, Z.I.; Che Ros, F.; Rambat, S.; Along, N.Z.; Harun, A.N. Flood Vulnerability Index Based on Indicator Approach in Assessing Flood Risk in Selangor, Malaysia. *Int. J. Integr. Eng.* **2024**, *16*, 16–22. [CrossRef]

54. Devi, K.; Reddy, C.C.; Rahul, K.; Khuntia, J.R.; Das, B.S. A holistic methodology for evaluating flood vulnerability, generating flood risk map and conducting detailed flood inundation assessment. *Sci. Rep.* **2025**, *15*, 28253. [CrossRef] [PubMed]
55. Buta, C.; Mihai, G.; Stănescu, M. Flood Risk Assessment Based on Flood Hazard and Vulnerability Indexes. *Ovidius Univ. Ann. Constanta Ser. Civ. Eng.* **2020**, *22*, 127–137. [CrossRef]
56. Malczewski, J. On the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. *Trans. GIS* **2000**, *4*, 5–22. [CrossRef]
57. Sanyal, J.; Lu, X.X. GIS-based flood hazard mapping at different administrative scales: A case study in Gangetic West Bengal, India. *Singap. J. Trop. Geogr.* **2006**, *27*, 207–220. [CrossRef]
58. Chen, Y.; Yu, J.; Khan, S. Spatial sensitivity analysis of multi-criteria weights in GIS-based land suitability evaluation. *Environ. Model. Softw.* **2010**, *25*, 1582–1591. [CrossRef]
59. Tripathi, R.N.; Kumar, D. Comprehensive Data Preprocessing and Feature Engineering for Optimized Machine Learning Models. In *Artificial Intelligence and Machine Learning*; RADemics Research Institute: Coimbatore, India, 2024; pp. 153–184. [CrossRef]
60. Kumaraswamy, R.S.B.; Agarwal, V.; Jain, A.G. A Comparative Study of Different Data Pre-processing Methods for Machine Learning. *Int. J. Multidiscip. Res.* **2025**, *7*, 1–8. [CrossRef]
61. Kiran, C.Y.P.; Manoj, P.B. The Novel Method for Data Preprocessing CLI. *Adv. Intell. Syst. Technol.* **2022**, *30*, 117–120. [CrossRef] [PubMed]
62. Liu, Y.; Liu, L.; Sun, H.; Chen, B.; Ma, X.; Ning, Y.; Qi, S. Flood Risk Assessment Combining Machine Learning with Multi-criteria Decision Analysis in Jiangxi Province, China. *Int. J. Disaster Risk Sci.* **2025**, *16*, 858–869. [CrossRef]
63. Muafiroh, S.; Pradono, K.A.; Sunandar, P.; Manurung, P. Application of machine learning and remote sensing in monitoring land use dynamics in tourism area. *Remote Sens. Technol. Def. Environ.* **2025**, *2*, 17–31. [CrossRef]
64. Xu, Y.; Goodacre, R. On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Sfile:///C:/Users/RHG-HORAS/Downloads/2205.03987v1.pdfupervised Learning. *J. Anal. Test.* **2018**, *2*, 249–262. [CrossRef] [PubMed]
65. Bennett, M.; Nekouei, M.; Mehta, A.P.R.; Kleczyk, E.; Hayes, K. Methodology to Create Analysis-Naive Holdout Records as well as Train and Test Records for Machine Learning Analyses in Healthcare. *arXiv* **2022**, arXiv:2205.03987.
66. Bami, Z.; Behnampour, A.; Doosti, H. A New Flexible Train-Test Split Algorithm, an approach for choosing among the Hold-out, K-fold cross-validation, and Hold-out iteration. *arXiv* **2025**, arXiv:2501.06492.
67. Sanches, R.G.; Miani, R.S.; dos Santos, B.C.; Moreira, R.M.; Neves, G.Z.d.F.; Bourscheidt, V.; Rios, P.A.T. Using Xgboost Models for Daily Rainfall Prediction. 2024. Available online: https://ssrn.com/abstract=4778138 (accessed on 11 December 2025).
68. Zhou, X.; Liu, G.; Wu, Y.; Wang, G.; Xu, Y.; Liu, T. XGBoost-SHAP based multi-scale evaluation and nonlinear response of river health in Guangdong Province. *Ecol. Indic.* **2025**, *178*, 114138. [CrossRef]
69. Xing, K.; Feng, M.; Bonev, I.A.; Champliaud, H.; Cheriet, M.; Liu, Z. XGBoost-Based Digital Twin Model for Predicting Trajectory Errors in a Hexapod Coordinated Machining System Using Positioning Accuracy and Vibration Data. *Sensors* **2025**, *25*, 7142. [CrossRef] [PubMed]
70. Törnquist, E.; Santos, W.C.; Pogue, T.; Wingle, N.; Caulk, R.A. Balancing Computational Efficiency and Forecast Error in Machine Learning-based Time-Series Forecasting: Insights from Live Experiments on Meteorological Nowcasting. *arXiv* **2023**, arXiv:2309.15207.
71. Abhinaya, P.; Reddy, C.K.K.; Ranjan, A.; Ozer, O. Explicit Monitoring and Prediction of Hailstorms with XGBoost Classifier for Sustainability. In *AI and IoT for Proactive Disaster Management*; IGI Global Scientific Publishing: Hershey, PA, USA, 2024; pp. 107–132. [CrossRef]
72. Dangeti, P. *Statistics for Machine Learning*; Packt Publishing Ltd.: Birmingham, UK, 2017.
73. Hyndman, R.J.; Koehler, A.B. Another look at measures of forecast accuracy. *Int. J. Forecast.* **2006**, *22*, 679–688. [CrossRef]
74. Chai, T.; Draxler, R.R. Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **2014**, *7*, 1247–1250. [CrossRef]
75. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning: With Applications in R*; Springer: New York, NY, USA, 2021; pp. 553–595. [CrossRef]
76. Truong, D. *Introduction to Data Mining and Data Science*; Chapman and Hall/CRC: Boca Raton, FL, USA, 2024. [CrossRef]
77. Willmott, C.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]
78. Pradhan, B.; Dikshit, A.; Lee, S.; Kim, H. An explainable AI (XAI) model for landslide susceptibility modeling. *Appl. Soft Comput.* **2023**, *142*, 110324. [CrossRef]
79. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*; Leanpub: Victoria, BC, Canada, 2022.
80. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef] [PubMed]

81. Ado, M.; Amitab, K.; Maji, A.K.; Jasińska, E.; Gono, R.; Leonowicz, Z.; Jasiński, M. Landslide Susceptibility Mapping Using Machine Learning: A Literature Survey. *Remote Sens.* **2022**, *14*, 3029. [CrossRef]
82. Hallegatte, S.; Rentschler, J.; Rozenberg, J. *Adaptation Principles: A Guide for Designing Strategies for Climate Change Adaptation and Resilience*; World Bank: Washington, DC, USA, 2020. [CrossRef]
83. Nugrahani, F.D. Land use transformation and sustainable spatial design lessons from a decade of regional development. *J. Placemak. Streetscape Des.* **2025**, *3*, 68–79. [CrossRef]
84. Peng, Z.; Huang, Q.; Han, Y. Model Research on Forecast of Second-Hand House Price in Chengdu Based on XGboost Algorithm. In *Proceedings of the 2019 IEEE 11th International Conference on Advanced Infocomm Technology (ICAIT)*; IEEE: Piscataway, NJ, USA, 2019; pp. 168–172. [CrossRef]
85. Prakash, A.; Thangaraj, J.; Roy, S.; Srivastav, S.; Mishra, J.K. Model-Aware XGBoost Method Towards Optimum Performance of Flexible Distributed Raman Amplifier. *IEEE Photonics J.* **2023**, *15*, 1–10. [CrossRef]
86. Yuwono, B.D.; Abidin, H.Z.; Poerbandono; Andreas, H.; Pratama, A.S.P.; Gradiyanto, F. Mapping of flood hazard induced by land subsidence in Semarang City, Indonesia, using hydraulic and spatial models. *Nat. Hazards* **2024**, *120*, 5333–5368. [CrossRef]
87. Rahayu, P. Mapping Vulnerability to Flood in Urbanizing Watershed: The Case of Surakarta Urban Region. *IOP Conf. Ser. Earth Environ. Sci.* **2023**, *1264*, 012013. [CrossRef]
88. Vousdoukas, M.I.; Mentaschi, L.; Voukouvalas, E.; Verlaan, M.; Jevrejeva, S.; Jackson, L.P.; Feyen, L. Global probabilistic projections of extreme sea levels show intensification of coastal flood hazard. *Nat. Commun.* **2018**, *9*, 2360. [CrossRef] [PubMed]
89. Fariz, T.R.; Rizky, T.B.; Holeng, V.A.; Heriyanti, A.P. Mapping household physical vulnerability to tidal floods through indicator-based scoring analysis. *Spat. Rev. Sustain. Dev.* **2025**, *2*, 111–128. [CrossRef]
90. Meerow, S.; Newell, J.P. Urban resilience for whom, what, when, where, and why? *Urban Geogr.* **2019**, *40*, 309–329. [CrossRef]
91. Zhang, K.; Shalehy, M.H.; Ezaz, G.T.; Chakraborty, A.; Mohib, K.M.; Liu, L. An integrated flood risk assessment approach based on coupled hydrological-hydraulic modeling and bottom-up hazard vulnerability analysis. *Environ. Model. Softw.* **2022**, *148*, 105279. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.