Journal of Electrical Engineering and Computer (JEECOM)
Vol. 8, No. 2 (2026), DOI: 10.33650/jeecom.v8i2.17328
p-ISSN: 2715-0410; e-ISSN: 2715-6427

Flood Susceptibility Mapping of Gorontalo Province Based on Climate and Spatial Data Using Machine Learning

Ade Irawaty Tolago[1*], Ulfatun Nadifa[2], Aristi Ayuningsi Ode Asri[3], Ikhsan Hidayat[4]
[1] Electrical Engineering, Universitas Negeri Gorontalo, Gorontalo, Indonesia
[2,4] Computer Engineering, Universitas Negeri Gorontalo, Gorontalo, Indonesia
[3] Civil Engineering, Universitas Negeri Gorontalo, Gorontalo, Indonesia

Article Info

Article history:

Received September 17, 2026
Revised September 22, 2026
Accepted October 15, 2026

Keywords:

Flood Susceptibility;
Machine Learning;
Random Forest;
Geographic nformation system;
Gorontalo;

ABSTRACT

Gorontalo Province regularly faces flood hazards driven by a combination of high rainfall, low-lying topography, and limited river capacity, while conventional hazard mapping approaches based on scoring and overlay remain limited in capturing the non-linear relationships among flood-causing variables. This study aims to produce a more accurate flood susceptibility map for Gorontalo Province by integrating climate and spatial data using machine learning. Seven physical variables (topography, slope, rainfall, land cover, soil type, DEMNAS elevation, and river distance) were paired with 186 historical flood points from 2010–2023 and non-flood control points, then modeled using five algorithms—Logistic Regression, Random Forest, Support Vector Machine, Multi-Layer Perceptron, and XGBoost—compared through 5-fold stratified cross-validation. The best-performing model was then used to predict the flood susceptibility index across the entire study grid and classify it into five hazard classes using the Natural Breaks (Jenks) method. The results show that Random Forest achieved the best performance, with a mean AUC of 0.9792 and a mean accuracy of 95.68%. Independent validation against historical flood points showed that 179 of 186 points (96.2%) fell within areas classified as hazardous, confirming the spatial validity of the resulting map. This study concludes that integrating climate and spatial data through machine learning produces more accurate flood susceptibility mapping than conventional approaches, and can serve as a decision-support tool for local government disaster mitigation policy.

Corresponding Author:
Ade Irawaty Tolago,
Electrical Engineering, Universitas Negeri Gorontalo, Gorontalo, Indonesia
Email: Adeirawty75@ung.ac.id ; Orcid ID : https://orcid.org/0009-0001-5144-1135

1. INTRODUCTION

Indonesia is a country vulnerable to hydrometeorological disasters—particularly flooding—due to its climatic characteristics and geographical conditions. Gorontalo Province is a region that frequently experiences flooding, resulting in losses for both the community and the environment [1]. National and regional disaster agency records show that flooding recurs almost every rainy season across several districts of the province; between 2015 and 2020 alone, at least 20 flood events were recorded in Gorontalo Regency, while Gorontalo City and Boalemo Regency have each reported repeated inundation affecting thousands of residents in recent years [32]. These conditions are driven by a combination of physical and hydrological factors, including rainfall, slope gradient, landform, drainage conditions, and river system capacity [2], [4], [6].

Hydrologically, Gorontalo Province is influenced by several major river systems, including the Limboto-Bolango-Bone, Paguyaman, and Randangan river basins. Topographic conditions dominated by lowlands, combined with river systems that discharge into Lake Limboto, influence the characteristics of inundation and flood susceptibility in the region [3]. Flood susceptibility levels are known to vary across the province: high susceptibility has been identified in Gorontalo Regency, while North Gorontalo exhibits a prevalence of moderate to high susceptibility classes [4], [5]. In Gorontalo City, rainfall, infiltration capacity, slope gradient, landform, and drainage conditions have been shown to jointly influence flood susceptibility [2], [6], and Sentinel-1 SAR-based remote sensing has additionally been used to map the spatial distribution of urban flood inundation [7]. Collectively, these studies confirm that flood susceptibility in Gorontalo is spatially heterogeneous and driven by multiple interacting physical factors, but most rely on descriptive scoring, overlay, or single-sensor remote sensing techniques applied to individual districts; none has yet integrated climate and multi-source spatial predictors within a single, quantitative, province-wide modeling framework, a gap that the present study addresses.

Given that flood susceptibility is influenced by a multitude of factors, mapping based on isolated parameters or a limited set of variables is often inadequate. Geographic Information Systems (GIS) have been widely used to integrate various spatial parameters—such as elevation, slope gradient, distance to rivers, land use, and soil type—for flood susceptibility mapping [4]–[10]. However, conventional approaches based on rating and overlay techniques tend to be descriptive and may have limitations in representing the complex relationships between environmental variables [11].

Advances in machine learning offer a means to model the complex, non-linear relationships between environmental characteristics and flood occurrence, and its application to flood susceptibility mapping is by now well established. Random Forest and gradient boosting-based methods are widely used for this purpose [12]–[16], and recent comparative studies have gone further by benchmarking a broad range of algorithms—including Extra Trees, Gradient Boosting, XGBoost, LightGBM, CatBoost, and stacking ensembles—showing that predictive performance depends more on data characteristics, predictor variables, and validation strategy than on the choice of algorithm family alone [17]–[20]. In Indonesia, Random Forest and Support Vector Machine algorithms have also been applied to the classification and prediction of flood events [23], [28]–[30]. What remains comparatively limited, however, is the systematic integration of climate and multiple spatial predictors within a single machine learning framework for a specific, data-scarce region such as Gorontalo Province, together with independent validation of the resulting susceptibility map against historical flood records; this is the gap the present study aims to address.

In the context of flood susceptibility mapping, integrating spatial and climatic data is crucial, as flood characteristics are influenced not only by topographic and hydrological conditions but also by rainfall variability. Combining these variables enables a more comprehensive representation of susceptibility than relying on a single parameter in isolation [20], [31].

Integrating climatic and spatial data is key to enhancing model precision. Spatial predictors—such as elevation, distance to rivers, and land use prove more representative of the soil saturation conditions that trigger floods when combined with climatic variables (e.g., 3- to 7-day accumulated rainfall) rather than relying solely on single-day data [16], [30]. For Gorontalo Province, integrating climatic data (rainfall, temperature, and humidity) with spatial data (elevation, slope, land use, and distance to rivers) has the potential to significantly improve the quality of flood susceptibility mapping. This approach not only yields data-driven susceptibility maps but also supports predictive analysis to anticipate future changes in environmental conditions. This research aims to generate quantifiable spatial information regarding flood susceptibility, serving as a foundation to support flood mitigation and risk management efforts in Gorontalo Province.

2. METHOD

2.1. Study Area

This research was conducted in Gorontalo Province, a region hydrologically influenced by the Bone, Bolango, Tamalate, and Alopohu river systems; of these, the Bone and Bolango rivers discharge directly into Tomini Bay through Gorontalo City, while the Alopohu and Tamalate rivers are among the tributaries feeding Lake Limboto [3]. The lowland areas surrounding Lake Limboto and along the coastal plain, which form the primary focus of the flood susceptibility analysis in this study, are characterized by elevations generally below 50 meters above sea level and slopes in the range of 0% to 8%, in contrast to the higher, more sloping terrain found elsewhere in the province; the Lake Limboto area additionally faces sedimentation issues that are progressively reducing its natural water storage capacity.

2.2. Data and Variables

The model was developed using two groups of predictors:

1. Spatial data: topography, DEMNAS elevation (DEM), slope gradient, land cover, and soil type, based on the boundaries of Gorontalo Province; distance to rivers is subsequently derived as an additional predictor, as described in Section 2.3.
2. Climate data: daily rainfall for the 2015–2025 period, utilizing CHIRPS satellite data.

It should be noted that the CHIRPS rainfall record (2015–2025) does not fully overlap with the historical flood inventory (2010–2023): flood events prior to 2015 could not be matched with contemporaneous daily rainfall and were represented using the nearest available CHIRPS composite, while flood events occurring in 2024–2025 had not yet been compiled into the disaster-agency inventory at the time this study's data were

ISSN: 2715-6427

assembled and are therefore not represented in the training sample. This partial temporal mismatch is a limitation of the current dataset and is revisited in the Conclusion.

Table 1. Input and Data Specification

TABLE I. INPUT AND DATA SPECIFICATION

Data Input	Data Specifications	Spatial Resolution	Data Source
Rainfall, Daily	Daily CHIRPS rainfall data for the period **2015–2025**	**0.05° × 0.05°** (±5 km × 5 km)	Climate Hazards Group InfraRed Precipitation with Station data (**CHIRPS**)
Elevation	National Digital Elevation Model (**DEMNAS**)	**0.27 arc-second** (±8 m)	Geospatial Information Agency (BIG)
Slope	Derived from DEMNAS data and classified into **five suitability classes**	**0.27 arc-second** (±8 m)	BIG/DEMNAS
Land Cover	Land cover classification data based on land use/land cover categories	**Adjusted according to the product used**	USGS
Soil Type	Soil characteristics and classification based on infiltration capacity and water storage capacity	**30 arc-seconds** (±1 km)	FAO/HWSD
Distance to River	**Euclidean distance** from the river network	30 × 30 m pixel size	River network data/DEMNAS
Administrative Boundary	Administrative boundaries of the **study area**	Vector	BIG/BPS

As response data, an inventory of historical flood event points in Gorontalo Province for the period 2010–2023 was used, compiled from records of the Regional Disaster Management Agency (BPBD), the National Disaster Management Agency (BNPB), the Sulawesi II Gorontalo River Basin Center (BWS), and verified news reports, complemented by non-flood points as a comparison class.

2.3. Preprocessing Dataset

Predictor and response data were integrated within a GIS environment, then modeled using five machine learning algorithms: Logistic Regression, Random Forest, Support Vector Machine, Multi-Layer Perceptron, and XGBoost. The data were split into training and testing data in a stratified manner, with 5-fold cross-validation to avoid overfitting and as a basis for selecting the best model. Model performance was evaluated using the Area Under the Curve (AUC-ROC) and accuracy metrics on cross-validation, complemented by RMSE, MAE, and R^2 on the test data to assess the closeness of predicted probabilities to the actual labels. The best-performing model was then used to predict the flood susceptibility index (in the form of a continuous probability of 0–1) across all grid cells in the study area, which was subsequently classified into five susceptibility classes using the Natural Breaks (Jenks) method: very low, low, moderate, high, and very high.

Before the modeling stage was carried out, the dataset underwent a series of preprocessing stages to ensure that the predictor and response data were in a uniform format, resolution, and coordinate system, and suitable for use as training data. This preprocessing stage consists of six parts as described below.

1. Spatial Data Integration and Reference Grid Construction

Six predictor data layers — topography, slope, rainfall, land cover, soil type, and DEMNAS elevation data — were integrated into a single reference grid with a resolution of 30 meters and coordinate system EPSG:32651 (UTM Zone 51N), using the topography raster as the reference for area coverage and coordinate system.

The integration process was carried out by reprojecting and resampling all raster layers to the reference grid. The choice of resampling method was adapted to the characteristics of each dataset: the nearest neighbor method was applied to categorical data (topography, land cover, and soil type) so that the integrity of class values remained preserved, while bilinear interpolation was applied to continuous data (slope, rainfall, and elevation) so that transitions between pixel values better represent actual surface conditions. Grid cells that had no value (NoData) in any raster layer were excluded from further analysis.

As an additional predictor variable, the distance of each cell to the nearest river network (distance_to_river) was calculated using the Euclidean Distance Transform (EDT) method, after the river network vector data had first been converted (rasterized) to raster format at the same resolution and coordinate system as the reference grid.

2. Extraction of Flood Event Sample Points and Non-Flood Control Points

Response data were built using a point-based sampling approach, unlike a full-grid labeling approach. Historical flood event points available in non-uniform coordinate formats (decimal, decimal degrees, and degrees-minutes-seconds) were processed through automatic parsing and coordinate format conversion, then transformed to the reference grid coordinate system (EPSG:32651) and matched to the nearest grid cell using nearest neighbor search based on a k-d tree. Each historical flood point was assigned a binary label of 1 (flood).

Non-flood control points were obtained through random sampling of grid cells located >= 300 meters from historical flood points, with the aim of ensuring that control points were in locations spatially sufficiently separated from actual flood events. The number of non-flood control points was set to twice the number of historical flood points, resulting in a class ratio of 1:2 (flood 183: non-flood 372) in the training dataset, with each control point assigned a binary label of 0 (no flood).

3. Feature Standardization and Training-Test Data Split

The seven physical predictor variables (topography, slope, rainfall, land cover, soil type, DEMNAS, and distance_to_river) that had been matched to each sample point were standardized using StandardScaler so that all variables were on a comparable scale, considering that the original value ranges among variables differed considerably (meters, degrees, categorical classes). The combined dataset of flood points and non-flood control points was then split into training data (80%) and test data (20%) in a stratified manner to maintain the class proportion in both subsets.

4. Model Comparison

Five classification algorithms were trained and compared on the training data, namely Logistic Regression, Random Forest, Support Vector Machine (RBF kernel), Multi-Layer Perceptron, and XGBoost, with class imbalance handled through balanced class_weight parameters (for Logistic Regression, Random Forest, and SVM) and scale_pos_weight (for XGBoost). Comparison was performed through 5-fold stratified cross-validation, with mean Area Under the Curve (AUC) as the main metric for model selection, accompanied by the accuracy metric. The model with the highest mean AUC in cross-validation was designated as the final model.

5. Evaluation on Test Data

The final model was further evaluated on the previously separated test data (20%), using AUC and accuracy metrics to assess classification ability, as well as Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and coefficient of determination (R^2) to assess how close the predicted probability values were to the actual labels. ROC curves were also constructed to compare the performance of the five models visually.

6. Prediction of Susceptibility Index Across the Entire Grid

The final model was retrained on all sample data (flood points and non-flood control points) with standardized features, then applied to predict flood susceptibility probability for each grid cell in the study area. The resulting continuous probability values (0–1) are referred to as the flood susceptibility index (flood_susceptibility_index / indeks_kerawanan_banjir), which represents the vulnerability level of each cell in a graded manner, differing from the binary flood-prone/non-flood-prone labels used in the training stage.

3. RESULTS AND DISCUSSION

3.1. Flood Susceptibility Mapping

Flood susceptibility mapping is one of the main instruments in hydrometeorological disaster mitigation, aiming to identify areas with a level of inundation susceptibility based on the physical characteristics of the land. In this study, the flood susceptibility level of Gorontalo Province was mapped using a machine learning-based approach, with output in the form of a continuous susceptibility index classified into five susceptibility classes — very low, low, moderate, high, and very high — as commonly used in GIS-based disaster susceptibility mapping studies.

3.2. Analysis of Machine Learning Flood Susceptibility Mapping

The process of integrating six predictor raster layers into a 30-meter resolution reference grid produced 13,284,944 valid cells out of a maximum total of 23,928,562 cells in the reference raster coverage. Of the 186 available historical flood event points, 183 points (98.4%) were successfully processed through coordinate parsing, while 3 points failed to be parsed due to coordinate formats unreadable by the parser. Non-flood control points were taken from a distance of >= 300 meters from any of the 13,284,944 flood point data. The training dataset formed consists of 183 samples of the flood class and 372 samples of the non-flood class, with a class ratio of about 1:2, for the non-flood class (0:372) and flood events (1:183).

ISSN: 2715-6427

Five algorithms were compared through 5-fold stratified cross-validation on the training-validation data samples, with Area Under the Curve (AUC) as the main metric for model selection. Table II presents the average AUC and accuracy results from cross-validation for the five models.

Table II. 5-Fold Cross-Validation Results for Five Models

Model	TABLE II. **5-FOLD CROSS-VALIDATION RESULTS FOR FIVE MODELS**			
	AUC (Mean)	**AUC (SD)**	**Accuracy (Mean)**	**Accuracy (SD)**
Logistic Regression	0,9715	0,0153	91,71%	1,44%
Random Forest	0,9792	0,0148	95,68%	1,55%
SVM (RBF)	0,9670	0,0218	92,97%	2,01%
MLP (Neural Network)	0,9718	0,0173	95,32%	1,05%
XGBoost	0,9761	0,0209	95,50%	1,61%

Random Forest obtained the highest mean AUC (0.9792 ± 0.0148) and was designated as the final model, followed by XGBoost (0.9761), Multi-Layer Perceptron (0.9718), Logistic Regression (0.9715), and Support Vector Machine (0.9670). Random Forest also recorded the highest mean accuracy (95.68%) among the five models. Nevertheless, the AUC difference among models was relatively small (0.0122 between the highest and lowest AUC), indicating that the seven physical variables used had sufficiently strong discriminatory power regardless of the classification algorithm chosen. As an additional illustration on a single split scheme (80% train, 20% test), the five models were evaluated using classification metrics (AUC, accuracy) as well as metrics commonly used in continuous prediction evaluation (RMSE, MAE, R^2) against the resulting probability values. The evaluation results are presented in Table III

Table III. Model Evaluation Results on Test Data

Model	TABLE III. **MODEL EVALUATION RESULTS ON TEST DATA**				
	RMSE	**MAE**	**R^2**	**AUC (Split)**	**Accuracy**
Logistic Regression	0,2291	0,1011	0,7639	0,9737	94,59%
Random Forest	0,1937	0,0822	0,8311	0,9887	95,50%
SVM (RBF)	0,1895	0,0927	0,8383	0,9909	94,59%
MLP (Neural Network)	0,1788	0,0594	0,8562	0,9839	96,40%
XGBoost	0,2123	0,0597	0,7972	0,9876	94,59%

In this single split scheme, Random Forest recorded an AUC of 0.9887 and an accuracy of 95.50%, with an RMSE of 0.1937 and an R^2 of 0.8311. As a note, MLP showed the highest R^2 (0.8562) and the lowest RMSE (0.1788) in this single split, although its mean AUC in the 5-fold cross-validation remained lower than that of Random Forest. This difference in ranking between the cross-validation results and the single split is consistent with the nature of a single split, which is more susceptible to variations in data partitioning, so the average results of the 5-fold cross-validation are more representative as a basis for model selection than the results of a single split.

The final Random Forest model was applied to predict the flood susceptibility index (continuous probability 0–1) across all 13,284,944 grid cells. Given that the distribution of grid points is not entirely regular at that scale, the susceptibility index was interpolated to a raster grid of 400×400 cells for rasterization purposes. Cells located more than 997.7 meters from the nearest original data point were assigned as empty (transparent) values to prevent interpolation from bridging areas outside the study area coverage (for example, sea areas along indented coastlines), so that 65,700 of the 160,000 cells in the raster grid were excluded and 94,300 valid cells remained. The flood susceptibility mapping results are also presented in the form of an interactive map that can be accessed online via the following page: Flood Mapping Link.

Figure 1. Interactive Map

The susceptibility index for those 94,300 cells was classified into five susceptibility classes using the Natural Breaks (Jenks) method, with the resulting class breaks as follows: Very Low (0–0.057), Low (0.057–0.182), Moderate (0.182–0.363), High (0.363–0.674), and Very High (0.674–1.0). The spatial distribution of these five classes across the study area is shown in Figure 2. The distribution of the number of cells per class is 74,229 cells (78.7%) Very Low, 9,349 cells (9.9%) Low, 6,201 cells (6.6%) Moderate, 2,650 cells (2.8%) High, and 1,871 cells (2.0%) Very High.

Figure 2. Flood Susceptibility Classes

As an independent validation of the final susceptibility map (outside the model evaluation metrics in Table II and Table III), all 186 historical flood points were checked for their positions on the five-class susceptibility map resulting from the Jenks classification, to see whether locations that had experienced flood events were classified into appropriate susceptibility classes. This examination covered all 186 flood points as recorded in the field data, including 3 points that had previously failed to be processed automatically during the coordinate parsing stage for model training purposes.

The examination results showed that 179 of the 186 historical flood points (96.2%) were located in areas classified as low to very high susceptibility classes, while 7 points (3.8%) were located in areas classified as Very Low, the lowest susceptibility class in this classification scheme. These seven points (3.8%) are locations that had experienced flood events in the historical data but were predicted to have very low risk by the final susceptibility map, and can thus be regarded as detection failures in the context of map validation. The following 7 locations are presented in Table IV.

Table IV. Historical Flood Points Classified in the Low Class

TABLE IV. HISTORICAL FLOOD POINTS CLASSIFIED IN THE LOW CLASS

District	Regency/City	Latitude (N)	Longitude (E)	Latitude, Longitude	Class
Paguyaman	Boalemo	0°34'57.58"N	121°52'03.62"E	0.582661, 121.867672	Low
Patilanggio	Pohuwato	0°34'57.58"N	121°52'03.62"E	0.582661, 121.867672	Low

ISSN: 2715-6427

The seven locations in Table IV are spread across three regencies, namely North Gorontalo (2 locations), Boalemo (3 locations), and Pohuwato (2 locations). Possible causes of this discrepancy include the local characteristics of flooding that are not fully represented by the seven 30-meter-scale physical variables used (for example, micro-drainage conditions, channel blockage, or sea tides in coastal areas), as well as the possibility that flood events at these points were triggered by non-physical factors such as limited flood control infrastructure. Field verification at these seven locations is recommended as a follow-up to determine the specific causes of the discrepancy before the final susceptibility map is used as a basis for mitigation policy.

4. CONCLUSION

This study set out to produce a flood susceptibility map for Gorontalo Province by integrating climate and spatial data through a machine learning approach, comparing five algorithms and validating the resulting map against independent historical flood records.

Among the five algorithms compared, Random Forest was the best-performing model, achieving the highest mean AUC in 5-fold cross-validation (0.9792 ± 0.0148) and a mean accuracy of 95.68%, slightly ahead of XGBoost, MLP, Logistic Regression, and SVM (RBF). On the held-out test data, the final Random Forest model obtained an AUC of 0.9887, an accuracy of 95.50%, an RMSE of 0.1937, and an R^2 of 0.8311, indicating good and consistent classification ability. The resulting susceptibility index, classified into five classes using the Natural Breaks (Jenks) method, showed that most of the study area falls into the Very Low class, with progressively smaller proportions in the High and Very High classes. Independent validation against 186 historical flood points showed that 96.2% of these points fell within areas classified as low to very high susceptibility, supporting the spatial validity of the resulting map.

These findings imply that integrating climate and multiple spatial predictors within a machine learning framework can produce a more accurate and spatially consistent flood susceptibility map than the descriptive, single-parameter scoring and overlay approaches commonly used in previous studies of Gorontalo Province. The resulting map, together with its interactive online version, can serve as a decision-support tool for local government agencies in prioritizing flood mitigation efforts and land-use planning.

This study nonetheless has several limitations. First, the historical flood inventory (2010–2023) and the CHIRPS rainfall record (2015–2025) do not fully overlap in time, and flood events from 2024–2025 were not yet available for inclusion, which may affect how well the model captures the most recent flood dynamics. Second, the seven physical predictors used are derived at a 30-meter resolution and do not capture local factors such as micro-drainage conditions, channel blockage, or flood control infrastructure; the 7 of 186 historical flood points (3.8%) that were misclassified as Very Low susceptibility are concentrated in North Gorontalo, Boalemo, and Pohuwato and are consistent with this limitation. Third, because nearby grid cells tend to share similar predictor values, the high cross-validation performance should be interpreted with some caution, as spatial autocorrelation can inflate apparent accuracy relative to a fully independent, spatially-blocked validation. Future work is therefore recommended to (1) update the flood inventory and rainfall record to a fully overlapping and more recent period, (2) evaluate model performance using spatial-based cross-validation to better estimate transferability to unseen areas, (3) incorporate local-scale variables such as micro-drainage and flood control infrastructure, and (4) conduct field verification at the seven misclassified locations before the map is adopted as a basis for mitigation policy.

Overall, this study demonstrates that integrating climate and spatial data through machine learning can produce flood susceptibility mapping that captures multivariate relationships among environmental factors more effectively than conventional descriptive approaches, while underscoring the need for further validation before the results are used as a sole basis for policy or final publication.

ACKNOWLEDGEMENTS

The authors thank Universitas Negeri Gorontalo (UNG), particularly LPPM, for funding and support through the Riset Akselerasi Publikasi Nasional (RAPN) 2026 scheme. This support enabled the research and preparation of this publication.

REFERENCES

[1]. Abdul, F. F., Maryati, S., & Koem, S. (2024). Analisis Persepsi dan Strategi Adaptasi Masyarakat terhadap Bencana Banjir di Kota Gorontalo. *Geography: Jurnal Kajian, Penelitian dan Pengembangan Pendidikan, 12*(1), 493–504.

[2]. Laya, A. (2008). *Kajian hidrogeomorfologi banjir di Kota Gorontalo* [Tesis]. Universitas Gadjah Mada.

[3]. Asri, A. A. O., Husnana, R., & Utamaa, K. A. (2025). Analisis Transpor Sedimen Dasar Sungai Alopohu yang Bermuara di Danau Limboto. *Rekonstruksi Tadulako Civil Engineering Journal on Research and Development, 6*(2), 121–128.

[4]. Hulantu, M. N., Ayuba, S. R., & Zees, D. S. (2023). Analisis Daerah Rawan Banjir melalui Pemanfaatan Sistem Informasi Geografis (SIG) di Wilayah Kabupaten Gorontalo Provinsi Gorontalo. *Jurnal Ilmu Komputer (JUIK), 3*(2), 30–44.
[5]. Saleh, C., Rusiyah, & Pambudi, M. R. (2026). Pemetaan Daerah Rawan Bencana Banjir Di Kecamatan Kwandang Kabupaten Gorontalo Utara Provinsi Gorontalo. *Jurnal Riset dan Pengabdian Interdisipliner, 3*(1), 190–201. https://doi.org/10.37905/jrpi.v3i1.36782
[6]. Arifin, Y. I., & Kasim, M. (2012). Penentuan Zonasi Daerah Tingkat Kerawanan Banjir di Kota Gorontalo Propinsi Gorontalo untuk Mitigasi Bencana. *Jurnal Sainstek, 6*(6).
[7]. Butarbutar, S. R., Lihawa, F., & Pratama, M. I. L. (t.t.). Analisis Spasial Genangan Banjir Perkotaan Menggunakan Citra Sentinel-1 SAR Berbasis Google Earth Engine di Kota Gorontalo. *[nama jurnal, volume, dan tahun perlu dilengkapi]*.
[8]. Doda, N. (2013). Analisis Daerah Rawan Banjir Kota Gorontalo Berbasis Sistem Informasi Geografis (SIG). *Radial,1*(2),112–125. https://doi.org/10.37971/radial.v1i2.33
[9]. Madani, I., Bachri, S., & Aldiansyah, S. (2022). Pemetaan Kerawanan Banjir di Daerah Aliran Sungai (DAS) Bendo Kabupaten Banyuwangi Berbasis Sistem Informasi Geografis. *Jurnal Geosaintek, 8*(2), 192–199.
[10]. Lodjo, L., & Sulfiani. (2026). Analisis Kerawanan Banjir Berbasis Sistem Informasi Geografis (SIG) sebagai Upaya Mitigasi Bencana dan Pengelolaan Lingkungan di Kelurahan Buliide, Kota Gorontalo. *Jurnal Ilmu Kesejahteraan Sosial HUMANITAS, 8*(1). https://doi.org/10.23969/humanitas.v8i1.60279
[11]. Arabameri, A., et al. (2020). Flood susceptibility mapping using machine learning methods.
[12]. Vojtek, M., Držík, D., Kapusta, J., & Vojteková, J. (2026). Fluvial flood extent modeling using machine learning algorithms trained on benchmark flood maps: new insights for computationally efficient flood mapping. *Water Cycle, 7*, 617–631.
[13]. Wahba, M., Sharaan, M., Elsadek, W. M., Kanae, S., & Hassan, H. S. (2024). Examination of the efficacy of machine learning approaches in the generation of flood susceptibility maps (Ibaraki Prefecture, Japan). *Environmental Earth Sciences, 83*, 429.
[14]. Dey, H., Shao, W., Peter, B. G., & VanDyke, M. (2024). Simulating flood risk in Tampa Bay using a machine learning driven approach. *Nature: Scientific Reports*.
[15]. Asrade, T. M., Abebe, S. A., Tadesse, K. B., Kerebih, M. S., & Meshesha, T. M. (2026). Flood susceptibility assessment using three machine learning techniques and comparison of their performance. *Scientific Reports*.
[16]. Nuritha, I., Widartha, V. P., & Wulandari, D. A. R. (2026). Perbandingan Algoritma Machine Learning untuk Pengembangan Sistem Prediksi Risiko Banjir Kabupaten Jember Berbasis Data BMKG Multi-Stasiun. *Informatics Journal (INFORMAL), 11*(1), 33–44.
[17]. Alwathaf, Y., Al-Areeq, A. M., Al-Masnay, Y. A., et al. (2025). Enhancing flood susceptibility mapping in Sana'a, Yemen with Random Forest and eXtreme gradient boosting algorithms. *Geocarto International, 40*(1), 2482707. https://doi.org/10.1080/10106049.2025.2482707
[18]. Liu, Y., Liu, L., Sun, H., Chen, B., Ma, X., Ning, Y., & Qi, S. (2025). Flood Risk Assessment Combining Machine Learning with Multi-criteria Decision Analysis in Jiangxi Province, China. *International Journal of Disaster Risk Science, 16*(5), 858–869. https://doi.org/10.1007/s13753-025-00669-8
[19]. Shrestha, S., Dahal, D., Bhattarai, N., Regmi, S., Sewa, R., & Kalra, A. (2025). Machine Learning-Based Flood Risk Assessment in Urban Watershed: Mapping Flood Susceptibility in Charlotte, North Carolina. *Geographies, 5*(3), 43. https://doi.org/10.3390/geographies5030043
[20]. Al-Kindi, K. M., & Alabri, Z. (2024). Investigating the role of the key conditioning factors in flood susceptibility mapping through machine learning approaches. *Earth Systems and Environment, 8*(1), 63–81.
[21]. Asri, A. A. O., Rohmat, F. I. W., & Kardhana, H. (2025). The effect of rainfall centroid position on lead time in the upstream Citarum River. *BIO Web of Conferences, 155*, 03001.
[22]. Asri, A. A. O., Rohmat, F. I. W., Kardhana, H., Kuntoro, A. A., & Farid, M. (2023). Analyzing lead time for flood early warning system in the upstream Citarum River. *E3S Web of Conferences, 467*, 02004.
[23]. Cahyaningtyas, C., Yuliana, Y., Noviyanti, N., & Saputro, T. V. D. (2026). Analisis Titik Genangan Air Akibat Banjir Menggunakan Metode Random Forest di Kecamatan Ledo. *Jurnal TIMES, 15*(1).
[24]. Demissie, Z., Rimal, P., Seyoum, W. M., Dutta, A., & Rimmington, G. (2024). Flood susceptibility mapping: Integrating machine learning and GIS for enhanced risk assessment. *Applied Computing and Geosciences, 23*, 100183. https://doi.org/10.1016/j.acags.2024.100183
[25]. Feizbahr, M., Brake, N., Arbabkhah, H., Asli, H. H., & Woods, K. (2025). Flood Susceptibility Mapping Using Machine Learning and Geospatial-Sentinel-1 SAR Integration for Enhanced Early Warning Systems. *Remote Sensing, 17*(20), 3471. https://doi.org/10.3390/rs17203471

[26]. Harshasimha, A. C., & Bhatt, C. M. (2023). Flood Vulnerability Mapping Using MaxEnt Machine Learning and Analytical Hierarchy Process (AHP) of Kamrup Metropolitan District, Assam. *Environ. Sci. Proc., 25*, 73.

[27]. Mz, M. A., Nurhayati, O. D., & Suseno, J. E. (2026). Performance Comparison of Random Forest, XGBoost, and SVM for Flood Risk Prediction Using BNPB GIS Data. *Journal of Information Systems and Informatics, 8*(1),992–1010. https://doi.org/10.63158/journalisi.v8i1.1461

[28]. Purwati, S. E., & Pristyanto, Y. (2024). Model Random Forest and Support Vector Machine for Flood Classification in Indonesia. *Sinkron: Jurnal dan Penelitian Teknik Informatika, 8*(4), 2261–2268.

[29]. Qamarani, L. S., & Riasetiawan, M. (2024). Klasifikasi Level Banjir Menggunakan Random Forest dan Support Vector Machine. *IJEIS (Indonesian Journal of Electronics and Instrumentation Systems), 14*(2), 199–208. https://doi.org/10.22146/ijeis.97043

[30]. Rahimi, M., Malekmohammadi, B., Firozjaei, M. K., Kerachian, R., Arsanjani, J. J., Tan, M. L., ... & AghaKouchak, A. (2026). Integrating geospatial intelligence and machine learning for flood susceptibility mapping. *Scientific Reports*.

[31]. U. Nadifa and I. Hidayat, “Automatically Retrained Machine Learning System for Rice Yield Prediction Using Open-Meteo and BPS Data,” vol. 18, no. 1, pp. 34–39, 2026.

[32]. Badan Nasional Penanggulangan Bencana (BNPB). (2021). Rekapitulasi kejadian bencana banjir Kabupaten Gorontalo periode 2015–2020, sebagaimana dikutip dalam "Banjir di Gorontalo Berangsur Surut," Republika, 2 Oktober 2021.