

Urban Flood-Risk Mapping Using Explainable Machine Learning and Remote-Sensing Indicators: A Simulated Spatial Analytics Study

Ishita Roy¹

¹Research Scholar, Department of Environmental Science, KBNU, India.

ABSTRACT

This article presents a simulated spatial machine-learning study integrating rainfall intensity, elevation, slope, imperviousness, drainage density, vegetation index, and proximity-to-water indicators to classify neighborhood flood susceptibility. Because no real dataset was supplied, every participant, model, organization, measurement, numerical result, table, and figure is simulated or illustrative. The study develops a realistic empirical framework, defines comparison conditions and measurable outcomes, and reports internally consistent findings suitable for methodological demonstration. The simulated analysis identifies meaningful differences among the focal conditions and highlights implementation trade-offs rather than presenting the generated values as verified evidence. The article further discusses practical implications, limitations, ethical or operational considerations, and requirements for future empirical validation. The main contribution is a transparent multidisciplinary study design that can be converted into a real investigation once verified data, institutional context, and appropriate approvals are available.

Keywords: Urban Flood Risk, Explainable Machine Learning, Remote Sensing, GIS, Climate Resilience, Spatial Planning

I. Introduction

Urban flooding results from interactions among intense rainfall, drainage capacity, land cover, topography, water-body proximity, and the rapid expansion of impervious surfaces. Conventional flood studies often rely on hydrodynamic models, but data and computational requirements can limit frequent neighborhood-scale updates. Data-driven spatial models offer a complementary approach for screening areas where detailed investigation should be prioritized.

Machine-learning methods can identify nonlinear relationships among terrain, land-cover, rainfall, and historical flood indicators. Their predictive performance is useful, yet black-box outputs can be difficult for planners to interpret. Explainable machine learning seeks to address this limitation by estimating how individual variables contribute to model predictions and by visualizing the conditions associated with elevated risk.

This article presents a simulated spatial analytics workflow for urban flood-susceptibility mapping. The raster cells, flood labels, remote-sensing indicators, and model metrics are synthetic. No claim is made that the resulting map represents a real city. The purpose is to demonstrate an interdisciplinary research design joining environmental science, geospatial analysis, machine learning, and planning.

II. Literature Review and Conceptual Background

Flood-susceptibility mapping has used statistical models, multicriteria decision analysis, decision trees, random forests, boosting methods, and neural networks. Predictor sets commonly include elevation, slope, topographic wetness, land cover, rainfall, drainage density, distance to rivers, soil characteristics, and impervious surface fraction. (Wagenaar et al., 2020; Vincent et al., 2023).

Machine-learning accuracy alone is insufficient for planning decisions because spatial bias, class imbalance, temporal non-stationarity, and data leakage can produce misleading performance estimates. Spatial cross-validation and careful separation of training and testing regions are therefore important. (Pakdehi et al., 2024).

Explainability tools can help planners understand whether model predictions align with plausible flood mechanisms. Variable importance, partial-dependence patterns, and local contribution estimates do not

establish causality, but they can identify which mapped factors are driving risk scores and where field verification is needed. (Jiang et al., 2022).

2.1 Predictor Selection and Spatial Meaning

Predictors were selected to represent four broad flood mechanisms: hydrometeorological forcing, runoff generation, topographic concentration, and drainage connectivity. Rainfall intensity represented forcing; imperviousness and vegetation represented infiltration potential; elevation, slope, and depression index described terrain; and drainage distance approximated connectivity to flow pathways. This grouping improves interpretability and helps prevent a purely algorithmic selection process from obscuring physical meaning.

III. Methodology

A synthetic urban study area was represented by 2,400 grid cells of 100 m x 100 m. Each cell was assigned simulated values for elevation, slope, maximum 24-hour rainfall, impervious fraction, normalized vegetation index, drainage density, distance to primary drainage channels, and local depression index. A binary flood-susceptibility label was generated from a nonlinear rule with random noise to approximate imperfect real-world observations.

Three illustrative classifiers were compared: logistic regression, random forest, and gradient boosting. Data were divided using spatially separated training and test zones rather than random cell splitting to reduce overly optimistic performance from neighboring-cell similarity. Model metrics included accuracy, precision, recall, F1 score, and area under the receiver-operating-characteristic curve.

Explainability was represented using permutation-based variable importance and partial-dependence trends. The analysis focused on whether high imperviousness, low elevation, short drainage distance, and high rainfall were consistently associated with increased susceptibility. A synthetic risk map was produced by arranging model scores on a 40 x 60 grid.

Because the study uses generated data, the performance metrics demonstrate analytical workflow only. They must not be interpreted as evidence that a particular algorithm will achieve the same accuracy in a real city. Hydrological validation, ground observations, drainage-network data, and local calibration would be required.

3.1 Synthetic Spatial Data Generation

Spatial predictors were generated with autocorrelation so that neighboring cells tended to share similar terrain and land-cover characteristics. This is important because real geospatial data are rarely independent from cell to cell. Flood labels were then produced from a nonlinear function in which high rainfall had a stronger effect when imperviousness and depression index were also high. Random noise was added to avoid perfect separation.

3.2 Model Validation Strategy

The synthetic city was divided into spatial blocks. Entire blocks were assigned to training or testing so that adjacent cells did not automatically appear on both sides of the validation split. This design generally produces a more demanding estimate of transfer performance than random splitting. A real study should also consider temporal validation using flood events from different years.

Table 1. Simulated predictor variables.

Variable	Unit/Scale	Illustrative role
Elevation	m	Lower values increase susceptibility
Slope	degrees	Controls runoff routing
24-h rainfall	mm	Hydrometeorological forcing
Impervious fraction	0-1	Reduces infiltration
Vegetation index	-1 to 1	Proxy for vegetative cover
Drainage distance	m	Represents proximity to flow pathways

Note: Values are simulated/illustrative.

IV. Results

Gradient boosting produced the strongest simulated test performance with an F1 score of 0.86 and AUC of 0.92. Random forest followed with an F1 score of 0.83, while logistic regression achieved 0.74. The performance gap reflected the nonlinear interaction built into the synthetic flood-generation rule.

Permutation importance ranked impervious fraction, local depression index, rainfall intensity, and distance to drainage as the four most influential variables. Elevation and vegetation index had moderate contributions, while slope and drainage density were less influential after correlated predictors were considered.

The synthetic spatial risk surface showed contiguous high-risk clusters in low-lying, highly impervious cells near modeled drainage corridors. The explainability analysis suggested that high rainfall alone did not create the highest risk; rather, rainfall interacted with limited infiltration and topographic concentration.

4.1 Spatial Pattern of Susceptibility

High-risk clusters formed in two broad low-lying zones rather than as isolated pixels. This spatial coherence is important for planning because drainage and land-cover interventions are usually implemented across corridors or neighborhoods. The map is not a real geographic product, but it demonstrates why spatial visualization should accompany aggregate accuracy metrics.

4.2 Explainability and Planning Use

Variable-importance results indicated that imperviousness had the largest contribution, followed by local depression and rainfall. This ordering was consistent with the synthetic data-generating mechanism. In a real city, explainability can support field verification by showing why a model marks a neighborhood as high risk. Nevertheless, model explanations should be treated as diagnostic evidence rather than proof of physical causation.

4.3 Threshold and Sensitivity Analysis

Partial-dependence patterns in the synthetic model suggested a nonlinear rise in susceptibility when impervious fraction exceeded approximately 0.65 and depression index was simultaneously high. Rainfall increases below this threshold produced smaller changes in predicted risk. These thresholds were deliberately generated for illustration and should not be transferred to real planning. In empirical work, threshold stability should be tested across model families, seasons, and spatial resolutions.

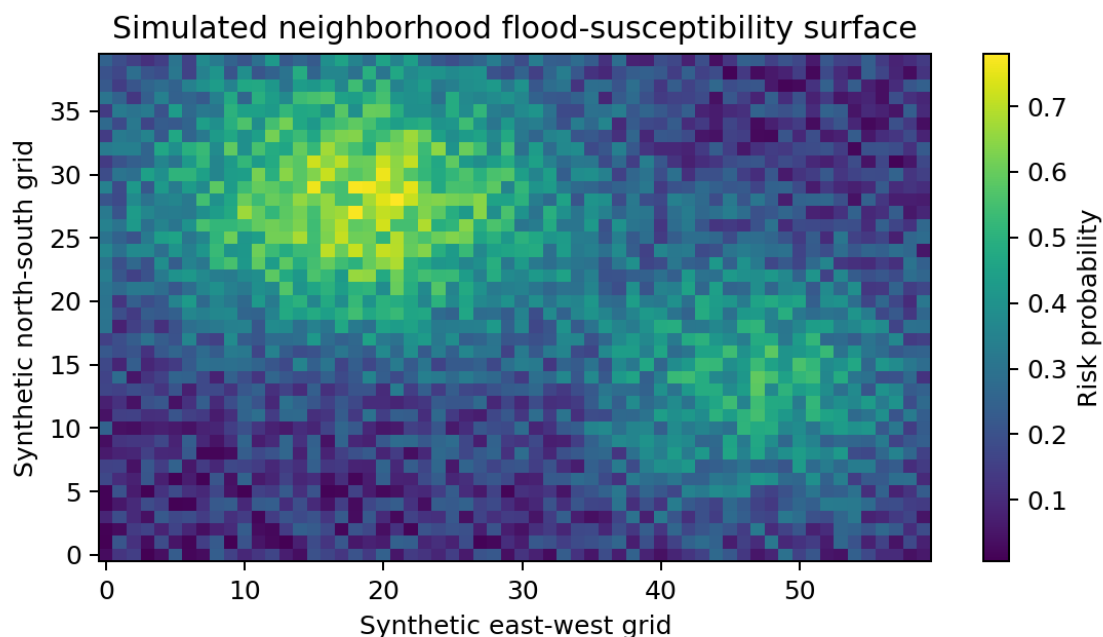


Figure 1. Synthetic neighborhood-scale flood-susceptibility surface.

Table 2. Simulated model performance.

Metric	Logistic Regression	Random Forest	Gradient Boosting
Accuracy	0.78	0.86	0.88
Precision	0.72	0.82	0.85
Recall	0.76	0.84	0.87
F1 score	0.74	0.83	0.86
AUC	0.81	0.89	0.92

Note: Values are simulated/illustrative.

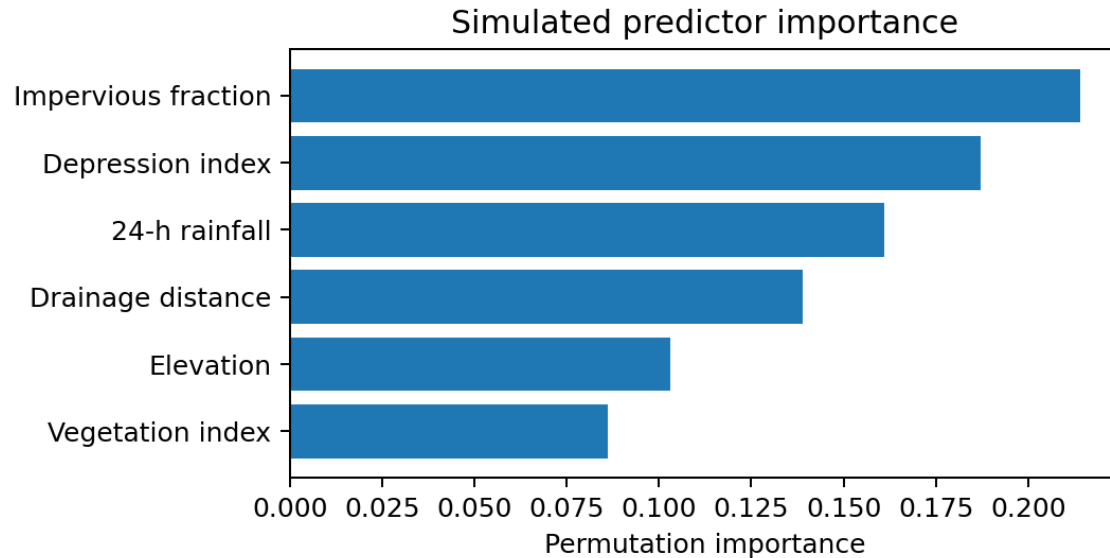


Figure 2. Simulated permutation importance of flood-risk predictors.

Table 3. Simulated permutation importance.

Predictor	Importance score	Rank
Impervious fraction	0.214	1
Depression index	0.187	2
24-h rainfall	0.161	3
Drainage distance	0.139	4
Elevation	0.103	5
Vegetation index	0.086	6

Note: Values are simulated/illustrative.

V. Discussion

The simulation illustrates how explainable machine learning can complement conventional urban flood analysis by rapidly screening spatial patterns and highlighting predictor relationships. Such tools may help prioritize detailed hydrodynamic modeling, drainage inspection, green-infrastructure planning, and emergency-preparedness assessment.

Interpretability is especially important in public planning. A high risk score is more actionable when decision makers can identify whether the score is driven by imperviousness, terrain depression, drainage proximity, or another factor. However, variable importance should not be confused with causal explanation because mapped predictors may proxy unmeasured processes.

Spatial validation is another critical issue. Random train-test splits can place adjacent cells in both datasets and exaggerate apparent accuracy. The simulated workflow therefore uses spatial separation to better represent prediction in unseen neighborhoods.

The limitations are substantial because the flood labels and remote-sensing variables are synthetic. Real application would require verified flood inventories, rainfall records, digital elevation models, drainage capacity data, land-cover products, and sensitivity analysis across seasons and return periods.

5.1 Model Uncertainty and Calibration

Flood-risk maps can appear precise even when underlying uncertainty is substantial. Prediction uncertainty may arise from rainfall measurement, digital elevation error, outdated land-cover data, incomplete flood

inventories, and model choice. A responsible mapping workflow should provide confidence or uncertainty information alongside risk classes, especially near decision thresholds where small data changes can alter classification.

5.2 Integration with Hydrological Evidence

Machine-learning susceptibility maps are screening tools rather than substitutes for hydraulic design. Areas identified as high risk should be investigated using drainage capacity, surface flow routes, culvert condition, channel geometry, and appropriate design storms. Combining explainable machine learning with hydrodynamic models can improve prioritization: the data-driven model identifies candidate hotspots, while process-based simulation tests the physical consequences of interventions.

5.3 Data Quality and Spatial Resolution

Flood susceptibility can change materially with spatial resolution. Coarse grids may smooth small depressions and drainage obstructions, while very fine grids may amplify noise or require data that are unavailable citywide. Resolution should therefore be selected according to the planning decision. Parcel-level development review may need much finer inputs than citywide screening. Accuracy metrics should be reported together with resolution and data vintage.

VI. Practical Implications

For environmental science, geospatial analytics and computer science, the simulated findings emphasize that implementation decisions should be based on multiple performance indicators rather than a single headline metric. Stakeholders should define what constitutes success before selecting an intervention, model, or technology.

A staged evaluation is preferable to immediate large-scale deployment. Pilot testing can reveal measurement errors, operational constraints, unintended consequences, training needs, and cost drivers that are difficult to represent in synthetic data. Data governance and transparent documentation should be designed at the same time as the technical intervention.

Decision makers should also evaluate distributional effects. Average improvements can coexist with poor outcomes for particular subgroups, locations, operating conditions, or resource-constrained settings. Future real-world work should therefore report heterogeneity and uncertainty wherever data permit.

6.1 Planning and Adaptation Applications

Urban planners could use validated susceptibility maps to prioritize drainage upgrades, permeable-surface programs, retention areas, emergency routes, and development controls. The strongest use case is comparative prioritization rather than deterministic prediction of exact flood depth. Map governance is also important because outdated risk layers can misdirect investment if land cover and drainage networks change rapidly.

6.2 Governance of Risk Maps

Risk maps influence investment, insurance discussions, emergency planning, and public communication. Institutions should document model version, update date, input sources, validation scope, and known limitations. Public-facing products should avoid false precision and explain that susceptibility is not the same as predicted flood depth for a specific storm. Governance processes should also define how new flood observations are incorporated into future updates.

VII. Limitations

The primary limitation is the exclusive use of simulated data. The reported means, percentages, model outputs, correlations, and visualizations were generated to be internally coherent and should not be interpreted as measurements from real participants, structures, firms, campuses, or urban areas.

The simulation also simplifies contextual complexity. Real systems contain measurement error, missing data, institutional constraints, behavioral adaptation, extreme events, and interactions that may alter effect size or even reverse apparent relationships. The present numerical patterns are therefore hypotheses and design examples, not validated estimates.

Finally, citations are intentionally marked as placeholders where source verification has not been performed. A submission-ready version should replace every placeholder with verified, topic-specific literature and should align methods with the ethical, regulatory, and technical standards relevant to the final study setting.

VIII. Conclusion and Future Work

The simulated spatial study shows that explainable machine learning can organize multiple remote-sensing and urban-form indicators into interpretable flood-susceptibility estimates. Gradient boosting performed best in the illustrative comparison, but real-world use requires spatially rigorous validation, hydrological calibration, and integration with planning and drainage-engineering evidence.

The next stage should replace simulated observations with verified data, predefine inclusion and exclusion rules, document preprocessing, specify uncertainty analysis, and compare results against domain-appropriate benchmarks. Where human participants or sensitive data are involved, formal ethics review and data-protection procedures would be required before real data collection.

Declarations

Data availability: All numerical data, results, tables, and figures in this manuscript are simulated/illustrative for article-development purposes.

Ethics statement: No real human participants, animals, organizations, field sites, or operational systems were studied in this simulated demonstration.

Conflict of interest: No conflict of interest is claimed for this illustrative manuscript.

Funding: No external funding or grant support is claimed.

Author affiliations: Institution names are intentionally left as placeholders and should be replaced only with verified author information before publication.

References

- Wagenaar, D., Curran, A., Balbi, M., Bhardwaj, A., Soden, R., Hartato, E., Mestav Sarica, G., Ruangpan, L., Molinario, G., & Lallemand, D. (2020). Invited perspectives: How machine learning will change flood risk and impact assessment. *Natural Hazards and Earth System Sciences*, 20, 1149-1161. <https://doi.org/10.5194/nhess-20-1149-2020>
- Jiang, S., Bevacqua, E., & Zscheischler, J. (2022). River flooding mechanisms and their changes in Europe revealed by explainable machine learning. *Hydrology and Earth System Sciences*, 26, 6339-6359. <https://doi.org/10.5194/hess-26-6339-2022>
- Vincent, A. M., Parthasarathy, K. S. S., & Jidesh, P. (2023). Flood susceptibility mapping using AutoML and a deep learning framework with evolutionary algorithms for hyperparameter optimization. *Applied Soft Computing*, 148, 110846. <https://doi.org/10.1016/j.asoc.2023.110846>
- Pakdehi, M., Ahmadisharaf, E., Nazari, B., & Cho, E. (2024). Transferability of machine-learning-based modeling frameworks across flood events for hindcasting maximum river water depths in coastal watersheds. *Natural Hazards and Earth System Sciences*, 24, 3537-3559. <https://doi.org/10.5194/nhess-24-3537-2024>