A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina

Abstract

Rainfall-triggered landslides are commonly occurring across the Southern Appalachians, yet effective, localized "nowcasts" that fuse dynamic precipitation with terrain characteristics remain limited. We develop a near-real-time landslide risk model for Buncombe County, North Carolina, integrating daily PRISM precipitation windows with static hydrological data. The novel dataset containing 302 mapped landslides (1981–2024) was compiled by integrating USGS DEM data and SSURGO soil information. Across models, slope was the dominant feature, while short-duration rainfall (1–7 days and 3-day maxima) most strongly increased landslide probability. In particular, 30-day cumulative rainfall often exhibited a negative association with predicted risk, suggesting that prolonged rainfall without intense surges may reduce immediate triggering. The resulting model enables county-scale, data-driven susceptibility updates based on forecasted precipitation data, offering a practical foundation for near-nowcasting and a landslide alarm system during and after intense storms.

Keywords: Rainfall-triggered landslides, Nowcasts, Buncombe County, Random Forest, XGBoost, PRISM precipitation, Landslide susceptibility modeling

[Repository link redacted for review; will be included upon acceptance]

[Images that are under 300 dpi can be found in higher quality in this google docs]

Introduction

Landslides are devastating natural hazards that take place worldwide and cause wide-ranging damage to modern society, such as human casualties, economic losses, and infrastructure damage (Froude & Petley, 2018; Kirschbaum & Stanley, 2018; Petley, 2012). In the year 2024 alone, 766 instances of landslides took place around the world, killing 4,933 individuals (Petley, 2025). Landslides across the world incur an economic cost of US$20 billion annually, taking 17% of the yearly mean expenses caused by all natural disasters worldwide between 1980 and 2013 (Klose et al. 2016). Social infrastructures, especially transportation and communication, are highly susceptible to landslides (Sim, Lee, Remenyte-Prescott,

& Wong, 2022). Rural areas are likely to experience more impairments due to landslides, where these infrastructures are scarcer and scattered (Klose et al. 2015).

In search of effective landslide prediction methods, researchers have proposed various theories on the sources of these disasters, ranging from geological activity such as volcanic eruptions and earthquakes to human-induced sources such as modification of slopes and obstruction of hydrological flows (Sidle & Ochiai, 2006; Jaboyedoff et al., 2016). Among those factors, precipitation is regarded as the leading cause of landslides worldwide; extreme rainfall triggered 70% of all landslide events across the world between 2004 and 2010 (Petley, 2012). Fittingly, the role of precipitation in landslides will only continue to grow with climate change; Gariano and Guzzetti (2016) predict that landslides triggered by rainfall will be increasingly catastrophic and recurrent as climate change increases the frequency of short, intense storm episodes. The predicted increase in landslide prevalence forebodes the urgency of an effective rainfall-induced landslide prediction model.

The Western North Carolina region has historically been dominated by rainfall-induced shallow landslides due to geographical and meteorological factors (Wooten et al., 2008). The Blue Ridge Mountains, as a part of the Southern Appalachian mountain range, are characterized by steep slope gradients (Khashchevskaya et al., 2025). Its bedrock layer mostly comprises metamorphic and igneous rocks, which, when decomposed, create saprolite, which constitutes the top layer of the horizons (Hatcher, 2010; Watterson & Jones, 2006). The saprolite-heavy soil absorbs water easily, which increases pore pressure, resulting in a higher probability of shallow landslides (Aydin, 2006). The region's geomorphic susceptibility has been called for focus again after Hurricane Helene's extensive landslide damage in Western North Carolina in 2024 (Lin et al., 2024). Allstadt et al. (2025) state that the hurricane caused unprecedented geographic disruptions in the region, particularly in Buncombe, Henderson, Rutherford, and Yancey counties. Freshwater flooding contributed to 95 of the 176 direct deaths that Helene caused, most of which were landslide-related; around 2,000 independent cases of landslides were reported to be related to Helene, with the majority located over western North Carolina (National Hurricane Center, 2025). Helene significantly damaged the communications infrastructure in western North Carolina, considering WNC's rural characteristic, which exacerbated the situation (Office of State Budget and Management, 2024). Helene's unexpected landslide damage calls

for a real-time probabilistic forecasting model, localized for Western North Carolina.

At present, there is no near-nowcast model for Western North Carolina that captures dynamic changes in precipitation. In fact, nearly all of the current landslide hazard maps use minimal or no precipitation inputs. Specifically, the USGS Landslide Hazard Mapping Program's landslide susceptibility map identifies areas with elevated risk for landslips under extreme rainfall conditions (Fuemmeler et al., 2008); these, however, do not take account of real-time changes in precipitation since they were based on past correlations of landslides with geographical elements (Bauer et al., 2012). Academically, even though there have been numerous attempts at developing a physical/statistical model for the prediction of landslides (e.g., an ML-based model by Lin et al. (2024) that combines a variety of predictors such as the NC landslide database, soil surveys, digital elevation models, and water body locations), very few, if any, utilized dynamic rainfall as their primary predictor. To this end, an effective localized model of Western North Carolina for rainfall-induced landslides has proven necessary, as seen in the extensive damage Helene caused in the region.

Thus, to address the absence of a localized prediction model that incorporates dynamic rainfall variability, the proposed research will address the question: "How does incorporating static geographic data (elevation and soil depth) with the variability of rainfall data affect the predictability of shallow landslides in the Western North Carolina region?" Ultimately, this study hopes to create an efficient prediction model in near real-time for the area, developing an early warning system for residents of Western North Carolina, specifically, Buncombe County.

Global models based on rainfall-triggered landslides have been developed and utilized for several years, with generally positive results. They are considered ‘nowcasts,’ because they incorporate both static geological characteristics and dynamic observations of rainfall to provide nearly real-time forecasts (Kirschbaum et al., 2012). For example, NASA's LHASA v2 model is the first operational global system to utilize dynamic satellite rainfall data along with static geological and environmental characteristics (Stanley et al., 2021). It is based on a Logistic Regression algorithm and examines static predictor variables as well as IMERG precipitation (1 km resolution) to generate probabilistic hazard labels. The new

model is purported to be two-fold more accurate than the previous LHASA v1 system, which utilized a standard threshold algorithm (Kirschbaum & Stanley, 2018; NASA Earth Observatory, 2021).

Only recently the number of statistical models predicting landslides increased, from 31 in 2016 to 219 in 2024 (Ye et al., 2025). Supervised machine learning (ML) was a popular choice for its ability to capture nonlinear relationships between predictor types and landslide probability, resulting in a higher level of accuracy than conventional linear statistical models (Regorda et al., 2020). Mondini et al. (2023) developed a time-dependent landslide model based on Italy. Exploiting RNN's strength in sequential data, they used continuous rainfall data for a window of 30 days, excluding geological or environmental factors. Similarly, Chan et al. (2018) built a logistic regression-based model aimed at predicting landslides caused by a typhoon's heavy rain in the southern region of Taiwan. The model mainly used runoff flow depth, not pure precipitation data, taking into account soil saturation, which reached an accuracy of 80~85%. Kang et al. (2024) recently constructed a localized model for Yunnan Province, China, with its Random Forest model reaching an accuracy level of 0.906. Incorporating every aspect of geographic, environmental and hydrological factors into eight variables. All the above ML-based models center around rainfall data. Their high accuracy levels affirm the need for dynamic rainfall data in a localized landslide model.

Focusing on North America, Thomas et al. (2019) challenged if satellite rainfall data can replace in situ hydrological data to evaluate the soil saturation threshold for a slope failure, reflecting a recent trend of increasing remote geospatial data-based models (Akosah et al., 2024). The findings from this localized model of California demonstrated that rainfall (specifically satellite-measured rainfall data) cannot be relied upon to predict landslides and encourage a hydrogeological gauge that calibrates rainfall data with on-site hydrologic data to develop a soil wetness index. Other studies that support the aforementioned findings include a comprehensive model (Lee et al., 2023) blending precipitation duration/intensity and normalized soil moisture capacity, which dropped false alarms (FA) from ~26 to 3. Both studies emphasize the importance of contextualizing raw rainfall data with local hydrologic features in evaluating the soil saturation level.

The contribution of this study is twofold. We build a localized, near-real-time landslide prediction model. For the modeling, an equal number of non-events were randomly chosen based on the landslide inventory of Buncombe County (North Carolina Department of Environmental Quality, 2024). To avoid pattern biasing in specific constellations of events, stratification and a spatial block cross-validation approach were utilized. Three different ML algorithms were tested overall, with three models per algorithm (F0, F1, and F2) based on feature type. The assessment of performance and feature analysis was interpreted using evaluation metrics, SHAP plots, and permutation importance.

To fit and evaluate the model, we designed a hydrologic/environmental dataset that collates daily precipitation data (PRISM Climate Group, 2024), a topography map (USGS DEM), and soil depth (USDA SSURGO). The dataset is open and available for further research at the specified repository.

The next section describes the datasets and methodology of the research. Section 3 illustrates the results of the findings. Section 4 delivers the conclusion of the research.

Data and Methodology

Dataset

This research aimed to create a dynamic prediction model localized in Buncombe County, North Carolina, geographically located within the Blue Ridge Mountains (Figure 2a). The inventory of landslide events was extracted from the North Carolina Landslide Points dataset (via NC OneMap; North Carolina Department of Environmental Quality, 2024). The dataset included event geometry (points) and traceability fields: *IsLandslide* (event flag), *Event_Date*, *Sort_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, and *Source_Period*. The inventory was both a temporal anchor (*Event_Date*) for rainfall-window calculations referred to by dates and a spatial anchor (X, Y) for topography and soil data extraction. Additionally, an equal number of non-event samples were created for control, randomly selecting the same time (January 1981–December 2021) and same region of interest (Buncombe County) as those in the landslide dataset. Ultimately, equal numbers of events and non-events were included in the final dataset used for modeling.

Gridded precipitation (4 km resolution) was extracted from PRISM daily products (PRISM Climate Group, 2024) throughout the entire analysis period (spanning from January 1981 to December 2024). Using the reference dates of the events and non-events—the event date (*Event_Date*) or the control (*Random_Date*)—historical sums over windows of 1-day (*R1d*), 3-days (*R3d*), 7-days (*R7d*), and 30-days (*R30d*) were computed. Other temporal windows included maximum rainfall sums calculated for both 3-day (*Max_Rainfall_3day*) and 30-day (*Max_Rainfall_30day*) intervals. All precipitation measurements were in units of millimeters, and all the calculation windows' end dates coincided with the reference date of the observation (Table 1 contains the definition and sources of all features).

Topography was taken from a DEM (Digital Elevation Model) encompassing Buncombe County published by the U.S. Geological Survey (2022). The DEM already provided elevation (*Elevation_m*), while slope (*Slope_deg*) was computed from Horn's (1981) 3×3 finite difference gradients and converted into degrees (Figure 2; Equations 1, 2, and 3).

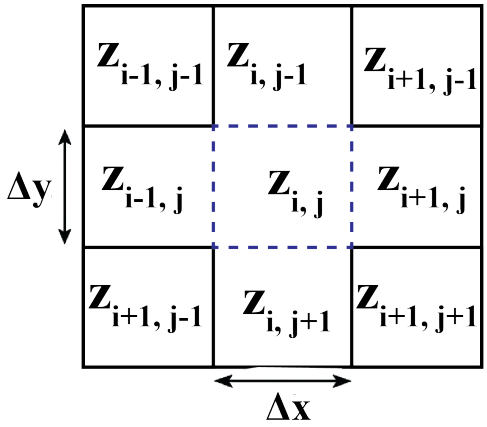

Figure 1. Visualization of Horn's (1981) 3×3 finite difference gradient calculator. $Z_{i,j}$ symbolizes the position of the elevation of a particular cell. Δx and Δy are the horizontal and vertical grid distances of the cells.

$$p = [(z_{i+1,j-1} + 2z_{i+1,j+1} - z_{i-1,j-1}) - (z_{i+1,j+1} + 2z_{i-1,j} + z_{i-1,j+1}\ )]/\ 8\Delta x \quad (1)$$

$$q = [(z_{i-1,j+1} + 2z_{i,j+1} - z_{i+1,j+1}) - (z_{i-1,j-1} + 2z_{i-1,j} + z_{i-1,j-1}\ )]\ /\ 8\Delta y \quad (2)$$

$$Slope = tan^{-1}(\sqrt{p^2 + q^2}) \quad (3)$$

The values p and q in Equations 1 to 3 represent the rate of elevation change in the east-west direction and the north-south direction, respectively. The DEM cell resolution for Buncombe County showed an equal width and height of 40 meters. To avoid unit discrepancy across different datasets, all datasets were under a consistent projected coordinate reference system (NAD83/North Carolina; EPSG:32119) and maintained a resolution of 40m.

Soil data originated from USDA NRCS SSURGO map units (U.S. Department of Agriculture, Natural Resources Conservation Service, 2024). A point-polygon comparison was used to match MUKEY and MUSYM to each sample point, linking depth information from the related tables. Soil depth was converted to a numeric variable (*Soil_Depth_cm*) by extracting data from raw strings (*Soil_Depth_cm_raw*) and converting the variable into a binary flag (*Soil_Depth_Deep200_Flag*), where entries larger than 200 were changed to 1, while any other entries were changed to 0. These soil properties were viewed as static covariates that reflect conditions of regolith associated with instability triggered by rainfall.

For map compilation and masking, administrative boundaries by county were employed to clip rasters and to define the sampling window for controls with a modest buffer to attenuate edge effects (U.S. Census Bureau, 2022). Land cover from NLCD was retained for descriptive mapping and voluntary sensitivity tests but was not included in the baseline prediction set (U.S. Geological Survey, 2021).

Gridded variables such as rainfall windows, elevation, and slope were sampled by bilinear interpolation at coordinates defined by points (X, Y) (Burrough & McDonnell, 1998). Soil data, specifically soil depth, came from the MUKEY and MUSYM area codes. Units of measurement have been standardized with precipitation in millimeters, elevation in meters, and slope measured in degrees. We verified the rainfall windows to check that the end dates coincide with the event reference date. Data instances with missing values were removed from the dataset.

The dataset created is a linked dataset, combining information from PRISM daily rainfall data, the USGS DEM, and the USDA SSURGO map specific to Buncombe county, and can be used for further analyses related to landslide risk. Linking information from the various datasets was not trivial, as it required

gathering chronological information and summarizing it for the selected data points before linking it to the main dataset (Longley et al., 2015; Burrough & McDonnell, 1998). Other challenges included incompatible reference systems that had to be reconciled. The complete dataset includes the outcome variable (*IsLandslide*) and full fields for past rainfall, topography, and soil depth with comprehensive metadata for provenance and audit purposes, which comprises *Event_Date/Random_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, *Source_Period*, MUKEY, MUSYM, and *Soil_Depth_cm_raw*.

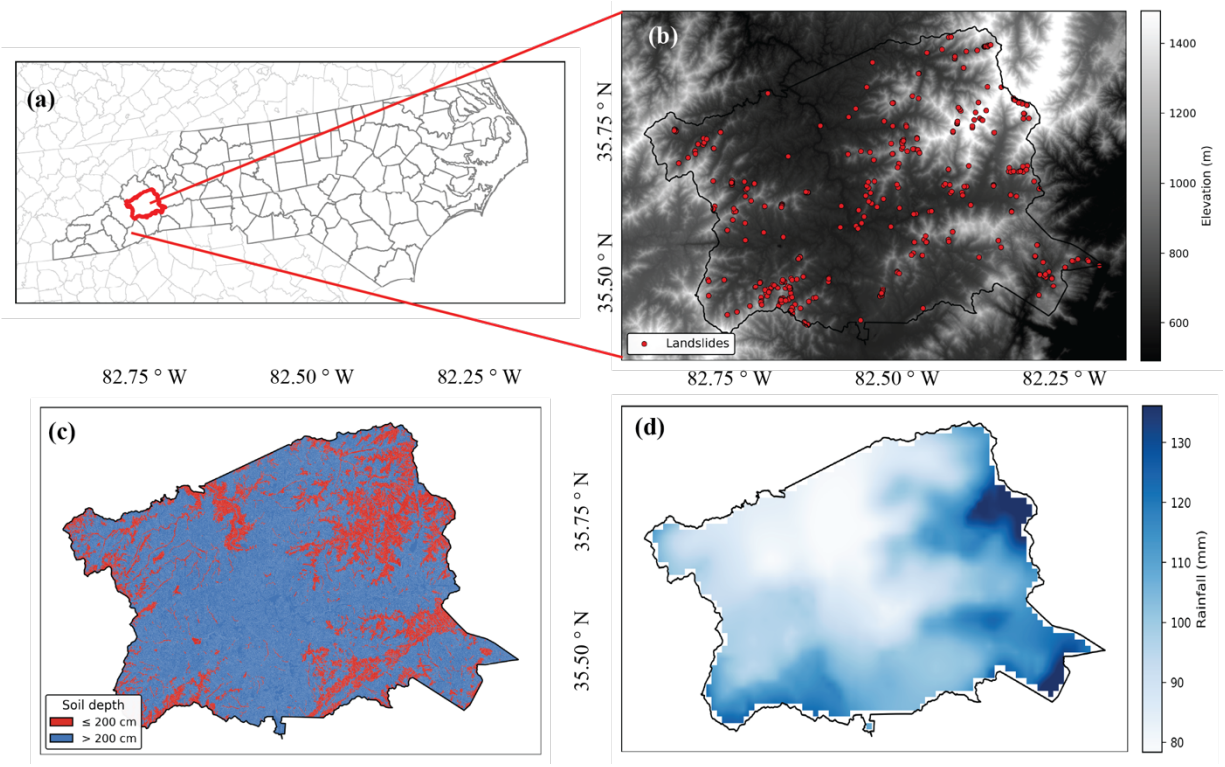

Figure 2: (a) Geographical location of the study area, (b) elevation, (c) soil depth, (d) 30 year average rainfall.

Symbol (unit)	How it was calculated	Source
R1d (mm)	Sum of daily PRISM rainfall for 1 day	PRISM (800m)
R3d (mm)	Sum of daily PRISM rainfall for 3 days	PRISM (800m)
R7d (mm)	Sum of daily PRISM rainfall for 7 days	PRISM (800m)

R30d (mm)	Sum of daily PRISM rainfall for 30 days	PRISM (800m)
Max_Rainfall_3day (mm)	Rolling maximum of 3-day totals in past 30 days	PRISM (800m)
Max_Rainfall_30day (mm)	Rolling maximum of 30-day totals in past 90 days	PRISM (800m)
Elevation_m (m)	Extracted directly from DEM	USGS DEM (40m)
Slope_deg (°)	Derived from DEM using slope algorithm	USGS DEM (40m)
Soil_Depth_Deep200_Flag (—)	Flag for soils deeper than 200 cm	SSURGO Soil

Table 1: Features definitions (Rainfall time-frames are from the reference point of the corresponding event)

All data collection, processing and model development steps were done in Python 3.9 in the PyCharm IDE. Libraries, including NumPy, Pandas, SciPy, Matplotlib, and Seaborn, were used for scientific computing; GeoPandas, Shapely, and Rasterio were used for data processing; and scikit-learn and XGBoost models combined with SHAP were chosen for model development and assessment. The models were run on a Mac Apple M2 processor.

Modeling

To assess which rainfall variables specifically influenced landslide probability, the temporal accumulation/maximum rainfall predictors were analyzed with all non-empty possible combinations. A primary concern was minimizing overfitting due to the large number of rainfall predictors (*R1d*, *R3d*, *R7d*, *R30d*, *Max_Rainfall_3day*, *Max_Rainfall_30day*). Thus, a training–test split was implemented via a stratified five-fold cross-validation approach using a constant seed and data point shuffling. In addition, the evaluation was repeated for every rainfall predictor subset. The model that showed the most accurate predictions on the test data was included in the final comparison across other types of models. The above processes were independently conducted for each machine learning (ML) algorithm model.

Three separate ML algorithms were employed for the actual study—Logistic Regression, Random Forest and XGBoost (Breiman, 2001; Chen & Guestrin, 2016) Due to logistic regression's simple and linear design, the model was effective in serving as a benchmark to compare with nonlinear models. The performance of the model was quantified through ROC curves, confusion tables, and summary measures: accuracy, precision, recall, ROC–AUC, and the Brier score (Brier, 1950). For logistic regression, the coefficient magnitudes summarized how each variable influenced the model prediction, while standard errors, p-values and confidence intervals measured the degree of statistical significance of that prediction.

A Random Forest algorithm was used as the baseline nonlinear model. Interpretation of the model involved a combination of intrinsic feature importance as well as SHAP analysis. Intrinsic feature importance was determined as the mean decrease in impurity, indicating the effectiveness of each predictor in splitting the data. Shapley values for each variable and data point were illustrated via SHAP bar plots, SHAP beeswarm plots and SHAP waterfall plots per sample (Lundberg & Lee, 2017). The bar plot visualized the average contribution of each feature to the overall prediction, while the beeswarm plots showed both the distribution as well as the direction of the predictor contributions. SHAP waterfall plots gave local explanations regarding the most likely landslide samples. A full dataset, including SHAP values as well as both the imputed and the original feature values, expected values, probabilities, and true labels, in addition to original metadata, was also provided so that both global and case-specific insights could be obtained. Model discrimination as well as calibration were also assessed as part of ROC curves and confusion matrices, as well as the same summary metrics that were used as part of Logistic Regression.

In answering the hypothesis, three separate models were created for each type of ML algorithm. Model F0 used just slope and elevation, aiming to assess the predictive potential based on the static predictor: terrain alone. Model F1 added rainfall accumulation intervals, alongside slope and elevation, as the basis to also evaluate the short- as well as longer-term rainfall triggers. Following on from F1, Model F2 also included soil depth so that the effect of the subsurface interacting with both rainfall and the terrain factors could be determined. Each algorithm was then trained across the three feature settings, ultimately leaving us with parallel models that allowed a direct comparison.

The F0, F1, and F2 configurations for each model (Logistic Regression, Random Forest, and XGBoost) were trained and tested on the same sample and therefore any discrepancies in results were not driven by sampling but rather variable differences.

Comparison analysis conducted on XGBoost involved classification accuracy, showing ROC curves and confusion matrices, as well as the complete set of metrics, including accuracy, precision, recall, F1 score, and ROC–AUC, as well as the Brier score. Interpretable techniques, such as SHAP dependence plots or permutation importance, are only presented for the Random Forest model.

Results and Discussion

The compiled landslide dataset contained 9,092 events recorded within North Carolina between 1940 and 2024, of which 398 occurred within the study area of Buncombe County (Figure 3). Because the PRISM Weather daily rainfall data before 1981 was unavailable for the public, the inventory was limited to the time frame between 1981 and 2024, limiting the data points to 302 landslide events. Corresponding to these landslide points, an equal number of control points were created by sampling sites and dates randomly within the same county and time frame. As such, the final dataset contained 604 total entries, split evenly between events and controls, thus reducing potential biases for certain variables and increasing its robustness toward new data.

All the records had a collection of hydrological covariates. Rainy features comprised both cumulative windows (*R1d*, *R3d*, *R7d*, *R30d*) as well as highest intensities (*Max_Rainfall_3day*, *Max_Rainfall_30day*). Geological features included elevation as well as slope at 40 m resolution. Soil depth was included as an added predictor expressed as a binary flag (>200 cm vs. ≤200 cm), as the vast majority of the sites had extremely deep soils.

The final dataset merged the rainfall, terrain, and soil characteristics for each observation. Its well-balanced design (302 events + 302 controls) ensured the risk of class bias was kept small, and the range of covariate variability enabled the models to incorporate long-term conditioning factors (e.g., slope, soil depth) as well as short-term triggers (e.g., extreme rains).

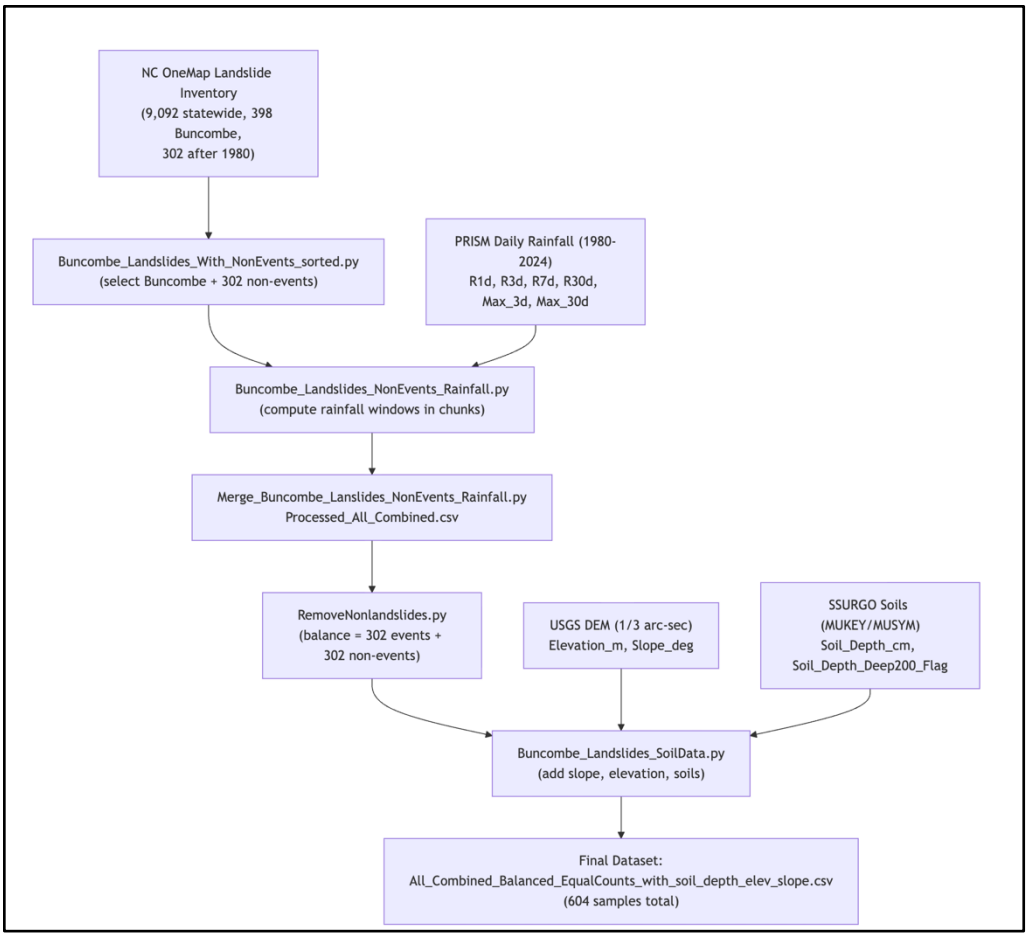

Figure 3: Flowchart of landslide events and non-events data processing

Feature selection:

Feature selection indicated consistency across the nonlinear and the linear models. In the Logistic Regression category, the optimum subset of rainfall data comprised *Max_Rainfall_30day*, *R3d*, *R7d*, and *R30d*, achieving an accuracy of 0.593 as well as an ROC–AUC of 0.614 (see Table 2). Short-period rainfall accumulation, including three-day as well as seven-day buildup, was frequently shown in the top-ranking combinations, showing how these predictors have a strong relationship with landslide risks. However, long-term predictors such as *Max_Rainfall_30day* and *R30d* also showed up occasionally, albeit not as frequently as short-term variables.

Logistic Regression (Top 5)	**Accuracy**	**ROC–AUC**
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.593	0.614

Max_Rainfall_30day, *Max_Rainfall_3day*, *R3d*, *R7d*	0.589	0.621
Max_Rainfall_30day, *R1d*, *R3d*, *R7d*, *R30d*	0.589	0.613
Max_Rainfall_30day, *R7d*, *R30d*	0.584	0.600
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.584	0.619

Table 2: Top-performing subsets of Rainfall Features for Logistic Regression

	Accuracy	**ROC–AUC**
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*	0.808	0.888
Max_Rainfall_30day, *R1d*, *R7d*, *R30d*	0.806	0.882
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.805	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*, *R7d*	0.801	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.801	0.883

Table 3: Top-performing subsets of Rainfall Features for Random Forest

In the Random Forest output, the best-performing subset included *Max_Rainfall_30day*, *Max_Rainfall_3day*, *R1d*, and *R3d*, yielding exceptional accuracy up to 0.808 and ROC-AUC reaching as high as 0.888 (see Table 3). Moreover, subsets including R3d and R7d also appeared among the very best, a point also supported by the results yielded by the Logistic Regression (see Table 2). However, the Random Forest model also brought to light that short-term rainfall predictors like *R1d* as well as the *Max_Rainfall_3day* significantly improved classification, bringing to the forefront nonlinear interactions. These findings indicated that although both nonlinear and linear prototypes invariably recognized the paramount significance of short-term rainfall, the Random Forest technique, in addition, captured the additional nuances of the lengths of the rainfall affecting the probabilities of the landslide occurrences (refer to Table 3).

Actual Models:

Logistic Regression:

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.620	0.609	0.650	0.629	0.673	0.229
F1	0.702	0.688	0.733	0.710	0.771	0.193
F2	0.702	0.682	0.750	0.714	0.775	0.191

Table 4: Logistic Regression model performance across feature sets (F0, F1, F2).

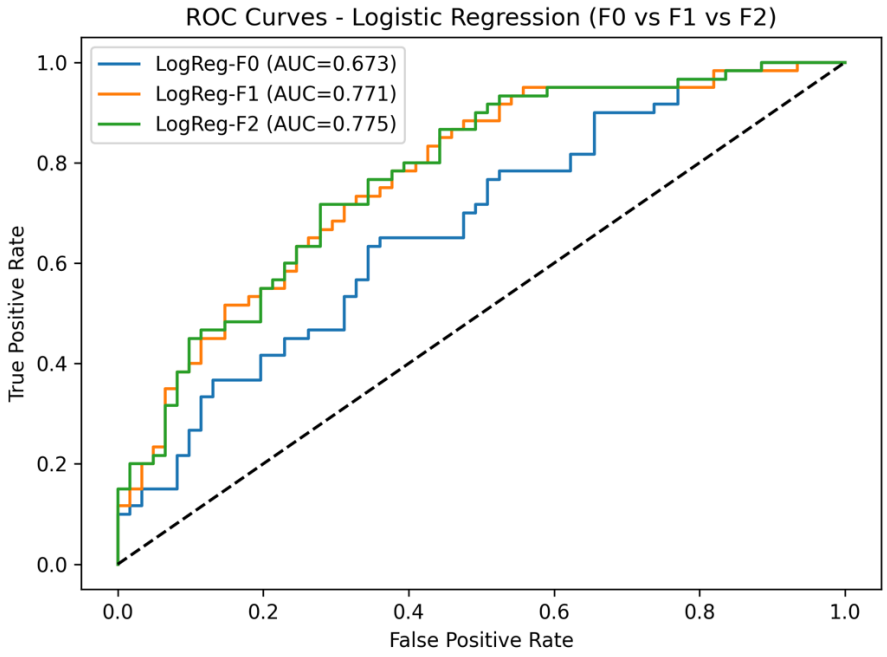

Figure 4: ROC Curves for the three Logistic Regression models F0, F1, and F2.

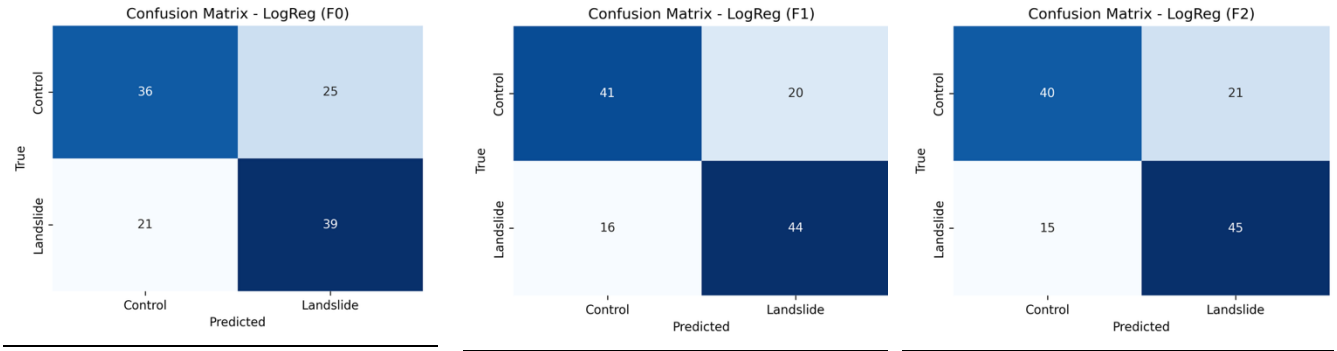

Figure 5-7: Confusion Matrices of F0, F1, F2 Logistic Regression Models

Feature	F0	F1	F2
Slope_deg	0.7150	1.3574	1.3813
Max_Rainfall_30day	-0.7287	0.5666	0.5871
R30d	0.2208	-0.5968	-0.6327
R7d	-0.1018	0.2192	0.2136
Elevation_m		-0.1886	-0.1099
R3d		-0.0571	-0.0620
Soil_Depth_Deep200_Flag			0.1943

Table 5: Regression coefficients for each feature of F0, F1, F2 Logistic Regression Models

Performance for Logistic Regression got progressively better as more predictors were added (Table 4). In the terrain-only version F0, accuracy was 0.62 with an AUC of 0.67. Inclusion of the rainfall predictors in F1 got the accuracy up to 0.70 and the AUC up to 0.77. Inclusion of the full set of predictors in the complete version F2 got the best accuracy at 0.70 as well as the best AUC at 0.78. All of these trends carry through on the ROC curves (Figure 4) as well as the confusion matrices (Figures 5-7): F0 got 36 controls and 39 landslide cases correct, but F2 got 40 correct controls as well as 45 correct landslides.

Model interpretation relied on the regression coefficients (Table 5). The slope was the strongest predictor variable, followed closely by *Max_Rainfall_30day*, both having a positive effect on the risk of landslides, except for the F0 model. Here, the cumulative 30-day rainfall (*R30d*) tended to have a negative coefficient, meaning long-term accumulation did not significantly increase the probability of slope failure after accounting for short-term accumulation or extreme events. These results strengthen the conclusion that the best explanation for landslides lies in the interaction between the region's steep slope and short-term, high-rate bursts of rain, shown in the Brier Scores of Table 4.

Random Forest:

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier

F0	0.661	0.679	0.623	0.650	0.741	0.216
F1	0.810	0.797	0.836	0.816	0.888	0.139
F2	0.785	0.778	0.803	0.790	0.880	0.144

Table 5: Random Forest model performance across feature sets (F0, F1, F2).

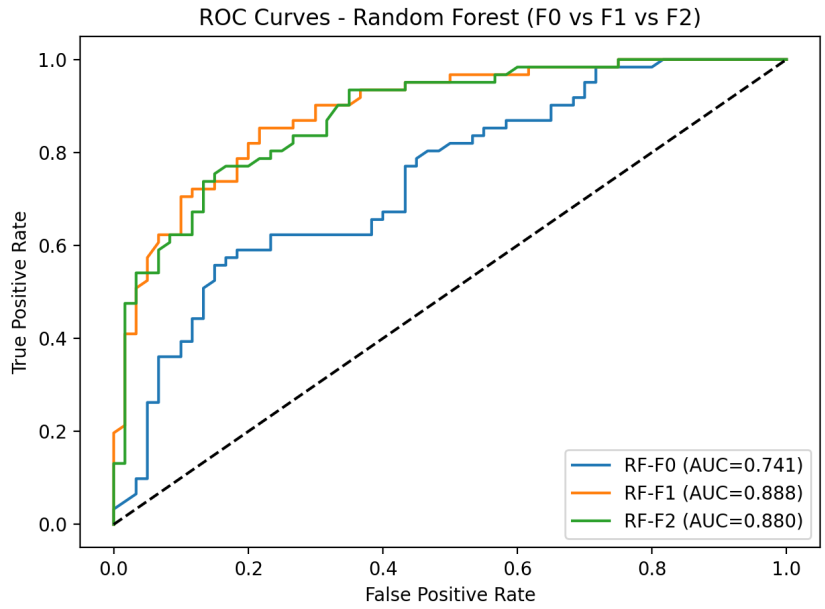

Figure 8: ROC Curves for the three models F0, F1, and F2 of Random Forest.

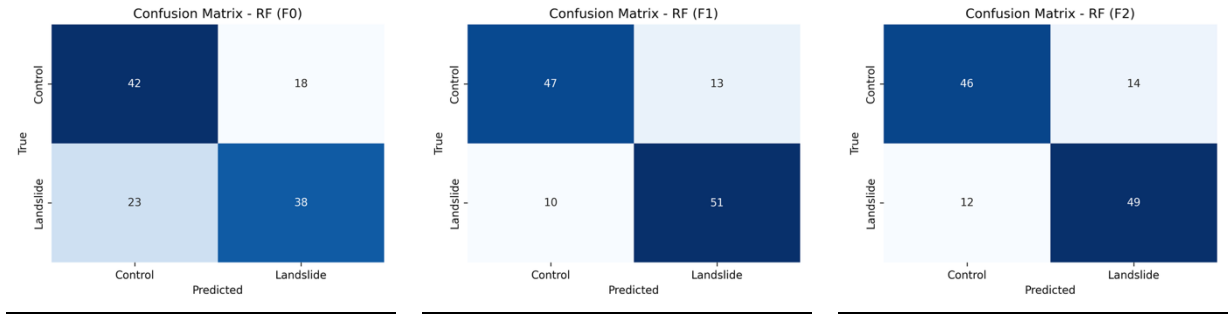

Figure 9-11: Confusion matrix of the three models F0, F1, and F2 of Random Forest.

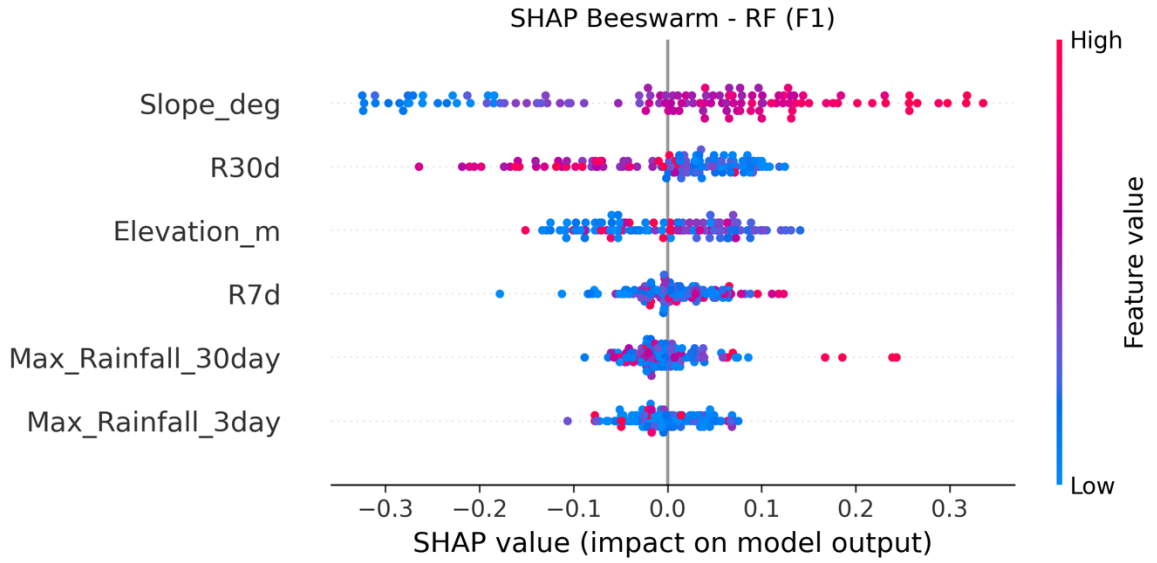

Figure 12: SHAP Beeswarm Chart for the F1 Random Forest Model.

The best-performing model, out of all three algorithms, was the Random Forest (RF) model. They were consistent across varying combinations of features (Table 5; Figure 8). Although the accuracy of the F0 model was 0.66 and the AUC 0.74—still notably higher than the accuracy of the Logistic Regression (LR) model—the F1 model setup, incorporating rainfall data, gave an accuracy of 0.81 and an AUC of 0.89, demonstrating the best performance of all. But the introduction of the soil feature in F2 detracted from the performance to an accuracy of 0.79 and an AUC of 0.88, departing from the trends of the LR. Still, the decrease in performance still holds that soil depth holds insignificant influence in prediction, especially for RF algorithms.

Considering F1's superior performance, feature analysis was only conducted on the F1 model, focusing on SHAP figures (Figure 12). The figure confirmed that slope is the primary predictor, exerting the most influence on the model. Subsequently, it was followed by *R30d*, *elevation*, and short-term rainfall features such as *R7d*, *Max_Rainfall_30day*, and *Max_Rainfall_3day*. According to the SHAP beeswarm plot (Figure 12), slope indicated a positive relationship with the risk of landslides, while elevation showed a bidirectional influence. Short-term bursts of rain, along with extreme maxima, exerted tremendously powerful effects in increasing the probability; conversely, cumulative rains over a 30-day period

revealed a negative outcome, which diminished the likelihood of landslide occurrence.

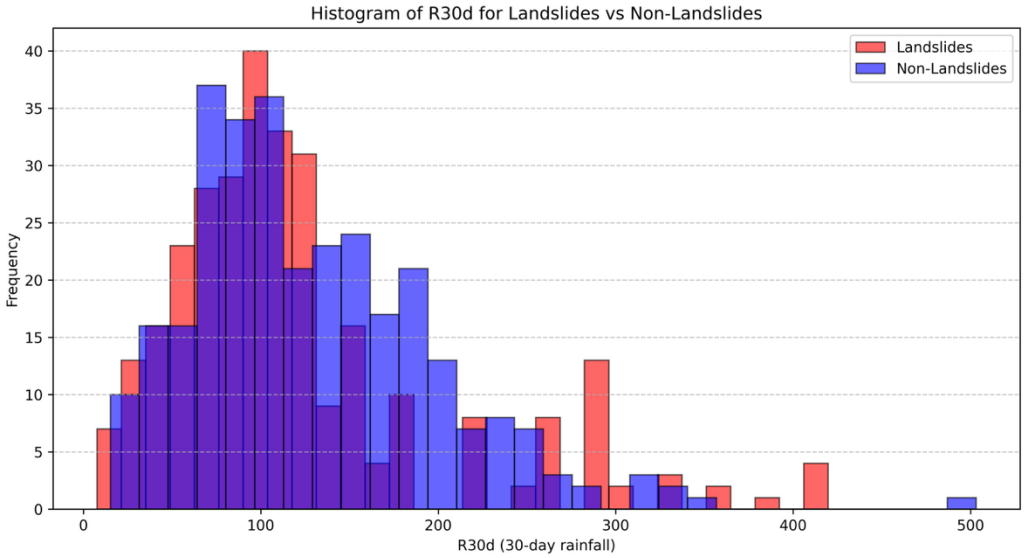

Figure 13: Distribution of 30-Day Cumulative Rainfall (R30d) for Landslide and Non-Landslide Events

As shown in the R30d histogram (Figure 13), both distributions of rainfall values were comparable, confirming that rainfall data were similarly distributed across both categories, thus strengthening the model's credibility. These findings underscore the capability of Random Forest to capture both anticipated and unexpected dynamics regarding landslide susceptibility. Intriguing is the discovery of the negative correlation between long-term cumulative rains and the model's predicted likelihood. This suggests that long-term rain accumulation, when not accompanied by brief extremes, may actually contribute to stabilizing the slope by facilitating gradual infiltration and percolation before pore pressurization persists long enough to initiate brief shallow failures. The results derived match past works citing landslide behavior on highly stepped landscapes as being mostly due to short-term high-magnitude events instead of being entirely the product of prolonged wet spells (e.g., Crozier, 2010; Tiranti & Rabuffetti, 2010; Bogaard & Greco, 2018). In addition, the decreased performance for F2 implies soil depth had little bearing on enhancing predictive performance at the particular resolution utilized in this study. In conclusion, the results from applying the Random Forest model indicate that the landslide hazard in Buncombe is primarily caused by a combination of a highly stepped landscape on slope discontinuities and short

periods of increased rainfall, rather than long periods of rainfall acting as an instantaneous trigger.

XGBoost

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.711	0.719	0.683	0.701	0.814	0.179
F1	0.802	0.757	0.883	0.815	0.879	0.141
F2	0.802	0.765	0.867	0.813	0.885	0.138

Table 6: XGBoost model performance across feature sets (F0, F1, F2).

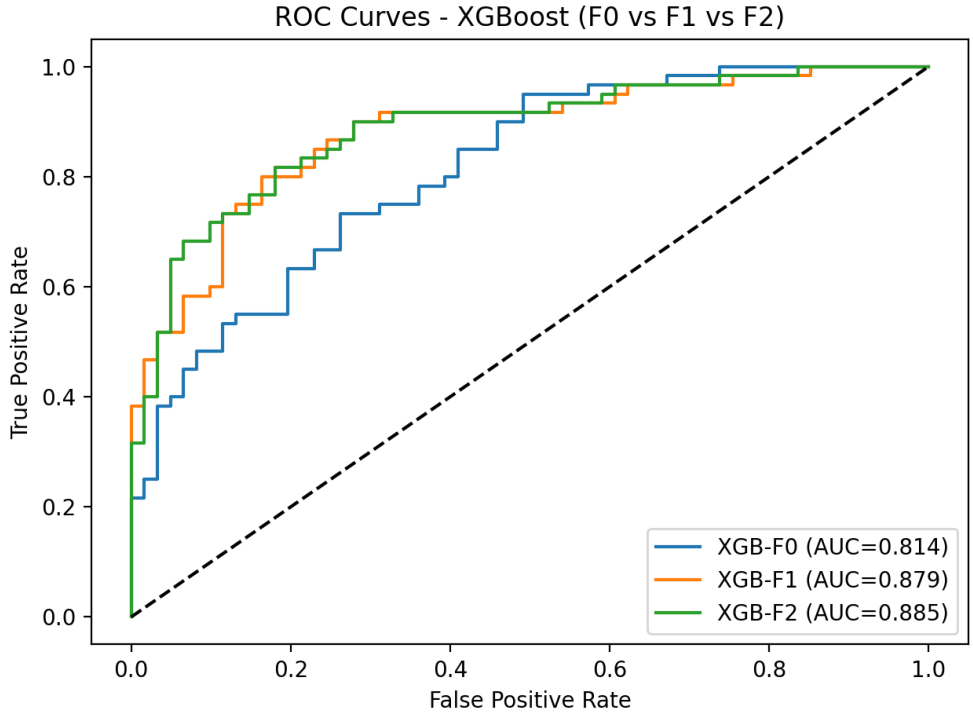

Figure 14: ROC curves of the three XGBoost models, F0, F1, and F2.

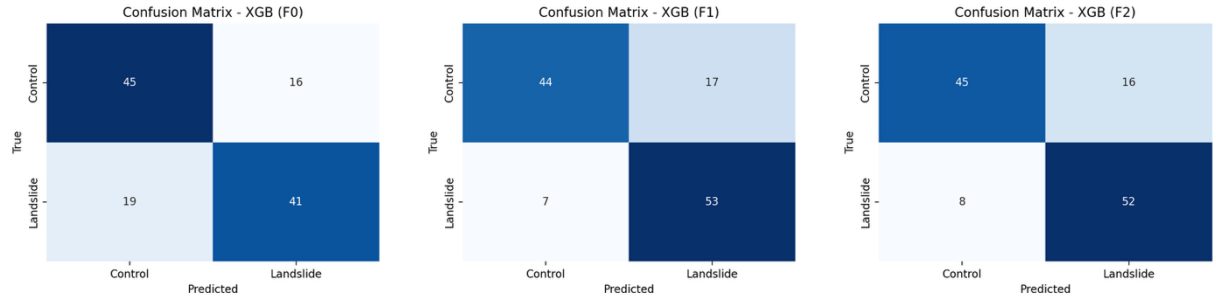

Figure 15-17: Confusion Matrix of the three models F0, F1, and F2 of XGBoost.

Surprisingly, XGBoost did not outperform the random forest model—instead, its F1 and F2 models showed similar prediction accuracy to the Random Forest Model (Table 6; Figure 14). The F0 XGBoost architecture excelled in model strength through an accuracy of 0.74 and an AUC of 0.81 in comparison to the other two F0 methods: Logistic Regression (LR) (accuracy 0.62, AUC 0.67) and Random Forest (RF) (accuracy 0.66, AUC 0.74). As expected, adding rainfall features improved the accuracy to 0.80 and AUC to 0.88, still better than the LR model, yet similar to the performance of the RF model (accuracy: 0.81, AUC: 0.89). Likewise, these trends appear to the full model (F2) equivalently. The full model (F2), which incorporated soil depth, produced the highest overall performance with an accuracy of 0.82 and an AUC of 0.89. The confusion matrix for F2 (Figure 17) highlights its balance, correctly classifying 45 controls and 52 landslides.

Compared to Logistic Regression and Random Forest, XGBoost delivered higher accuracy and consistently stronger AUC across feature sets, confirming its robustness in handling complex nonlinear relationships. The results reinforce that steep slopes and short-term rainfall extremes are primary triggers, while long-term rainfall can act as a dampening factor rather than a direct trigger (Crozier, 2010; Bogaard & Greco, 2018). The improved performance of the full F2 model also suggests that soil depth, while secondary to slope, provides additional predictive power in boosted frameworks.

Accuracy improved across all models when rainfall was added to terrain predictors. Logistic Regression rose modestly from 0.62 (F0) to 0.70 (F2), Random Forest increased more sharply from 0.66 (F0) to 0.81 (F1) but dropped slightly with soil depth (0.79, F2), while XGBoost achieved the highest accuracies at every stage, from 0.74 (F0) to 0.82 (F2).

Conclusion

This study developed a localized near-real-time landslide prediction framework for Buncombe County, North Carolina, using machine learning models and environmental predictors. By compiling a balanced dataset of 302 landslides and 302 non-events, rainfall windows from PRISM, topography from USGS DEMs, and soil depth from SSURGO, we trained Logistic Regression, Random Forest, and XGBoost classifiers. The pipeline incorporated stratified and spatial-block cross-validation, allowing us to rigorously compare models while preserving spatial

representativeness. SHAP analysis and intrinsic feature importance made it possible to understand how predictors shaped landslide susceptibility in this mountainous region.

Across models, Random Forest consistently achieved the highest predictive performance (accuracy ≈ 0.82, AUC ≈ 0.89), followed by XGBoost and Logistic Regression. In all frameworks, slope emerged as the dominant predictor, confirming prior work showing terrain steepness as the primary control on landslide occurrence (Zheng et al., 2025; Xiao et al., 2024). Short-term rainfall accumulations (1–7 days, maxima over 3 days) were strongly associated with failures, consistent with global findings that intense short-duration rainfall is the principal trigger for shallow landslides (Crosta, 2004; Pennington et al., 2014; Barthélemy et al., 2024). A notable theoretical contribution was the identification of long-term cumulative rainfall (R30d) as a *negative* predictor in several models, suggesting that extended wet periods may sometimes stabilize slopes by allowing gradual drainage and infiltration rather than immediate pore-pressure buildup (Crosta, 2004; Fan et al., 2020). Soil depth contributed marginally, adding value only in boosted frameworks, which echoes recent findings that subsurface factors are often secondary to slope–rainfall interactions (Lee et al., 2023).

These results underscore both the practical and scientific value of localized machine-learning–based landslide modeling. The framework demonstrates that integrating rainfall windows with terrain predictors can generate near-real-time susceptibility maps, offering actionable information for hazard managers after storms like Hurricane Helene. At the same time, the findings reinforce a broader theoretical consensus: steep slopes combined with short-term rainfall extremes are the key drivers of shallow failures, while long-term rainfall plays a preparatory but not directly triggering role (Pennington et al., 2014; Gariano & Guzzetti, 2016). Limitations of this study include a simplified binary soil depth representation, a relatively small sample size, and a lack of landslide-type differentiation. Future work should incorporate higher-resolution soil and land cover data, expand temporal coverage with satellite rainfall products, and test the transferability of this framework across Western North Carolina. Ultimately, coupling such models with rainfall forecasts could enable truly real-time landslide “nowcasting,” a direction increasingly emphasized in both local and global hazard research (Stanley et al., 2021; Khan et al., 2022).

References

Akosah, S., Gratchev, I., Kim, D.-H., & Ohn, S.-Y. (2024). Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review. *Remote Sensing*, *16*(16), 2947. https://doi.org/10.3390/rs16162947

Allstadt, K., McBride, S. K., Godt, J., Slaughter, S., Baxstrom, K., Sobieszczyk, S., & Stull, A. (2025). *Preliminary field report of landslide hazards following Hurricane Helene* (Report Nos. 2025–1028; Open-File Report). USGS Publications Warehouse. https://doi.org/10.3133/ofr20251028

Aydin, A. (2006). Stability of saprolitic slopes: Nature and role of field scale heterogeneities. *Natural Hazards and Earth System Sciences*, *6*(1), 89–96. https://doi.org/10.5194/nhess-6-89-2006

Barthélemy, S., Bernardie, S., & Grandjean, G. (2025). Assessing rainfall threshold for shallow landslides triggering: A case study in the Alpes Maritimes region, France. *Natural Hazards*, *121*(4), 4023–4049. https://doi.org/10.1007/s11069-024-06941-2

Bauer, J., Fuemmeler, S., Wooten, R., Witt, A., Gillon, K., Douglas, T., Eberhardt, E., Froese, C., Turner, K., & Lerouell, S. (2012). Landslide hazard mapping in North Carolina–Overview and improvements to the program.

Landslides and Engineered Slopes: Protecting Society through Improved Understanding. 11th International Symposium on Landslides and 2nd North American Symposium on Landslides, 257–263. https://www.researchgate.net/publication/260077615_Landslide_hazard_mapping_in_North_Carolina-Overview_and_improvements_to_the_program

Bogaard, T., & Greco, R. (2018). Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences*, 18(1), 31–39. https://doi.org/10.5194/nhess-18-31-2018

Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brier, G. W. (1950). VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. *Monthly Weather Review*, 78(1), 1–3. https://doi.org/10.1175/1520-0493(1950)078%253C0001:VOFEIT%253E2.0.CO;2

Burrough, P., & McDonnell, R. (1998). *Principle of Geographic Information Systems*.

https://www.researchgate.net/publication/37419765_Principle_of_Geographic_Information_Systems

Chan, H.-C., Chen, P.-A., & Lee, J.-T. (2018). Rainfall-Induced Landslide Susceptibility Using a Rainfall–Runoff Model and Logistic Regression. *Water*, 10(10), 1354. https://doi.org/10.3390/w10101354

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd* ACM SIGKDD *International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Crozier, M. J. (2010). Deciphering the effect of climate change on landslide activity: A review. *Geomorphology*, 124(3–4), 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009

Fan, L., Lehmann, P., Zheng, C., & Or, D. (2020). Rainfall Intensity Temporal Patterns Affect Shallow Landslide Triggering and Hazard Evolution. *Geophysical Research Letters*, 47(1), e2019GL085994. https://doi.org/10.1029/2019GL085994

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fuemmeler, S. J. (2008, April 10). *Landslide Hazard Mapping Methodology and Examples from North Carolina*. The Geological Society of America: Southeastern Section - 57th Annual Meeting. https://gsa.confex.com/gsa/2008SE/webprogram/Paper136714.html?utm_source=chatgpt.com

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Hatcher, R. D. (2010). The Appalachian orogen: A brief summary. In R. P. Tollo, M. J. Bartholomew, J. P. Hibbard, & P. M. Karabinos, *From Rodinia to Pangea: The Lithotectonic Record of the Appalachian Region*. Geological Society of America. https://doi.org/10.1130/2010.1206(01)

Horn, B. K. P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1), 14–47. https://doi.org/10.1109/PROC.1981.11918

Jaboyedoff, M., Michoud, C., Derron, M.-H., Voumard, jérémie, Leibundgut, G., Sudmeier-Rieux, K., Nadim, F., & Leroi, E. (2016). *Human-Induced Landslides: Toward the analysis of anthropogenic changes of the slope environment* (pp. 217–232). https://doi.org/10.1201/b21520-20

587 Kang, J., Wan, B., Gao, Z., Zhou, S., Chen, H., & Shen, H. (2024). Research on
588 machine learning forecasting and early warning model for rainfall-
589 induced landslides in Yunnan province. *Scientific Reports*, 14(1), 14049.
590 https://doi.org/10.1038/s41598-024-64679-0

591 Khan, S., Kirschbaum, D. B., Stanley, T. A., Amatya, P. M., & Emberson, R. A.
592 (2022). Global Landslide Forecasting System for Hazard Assessment and
593 Situational Awareness. *Frontiers in Earth Science*, *Volume* 10-2022.
594 https://doi.org/10.3389/feart.2022.878996

595 Khashchevskaya, D., Owen, L. A., Wegmann, K., Scheip, C., & Figueiredo, P. M.
596 (2025). The characteristics and timing of multiphase major landslides
597 along the Blue Ridge Escarpment of the southern Appalachians revealed
598 by combined cosmogenic nuclide dating and Schmidt hammer rebound
599 measurements. *Geomorphology*, 485, 109857.
600 https://doi.org/10.1016/j.geomorph.2025.109857

601 Kirschbaum, D. B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., & Lerner-
602 Lam, A. (2012). Advances in landslide nowcasting: Evaluation of a global
603 and regional modeling approach. *Environmental Earth Sciences*, 66(6),
604 1683–1696. https://doi.org/10.1007/s12665-011-0990-3

605 Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-
606 Triggered Landslide Hazard for Situational Awareness. *Earth's Future*,
607 6(3), 505–523. https://doi.org/10.1002/2017EF000715

608 Klose, M., Damm, B., & Terhorst, B. (2015). Landslide cost modeling for
609 transportation infrastructures: A methodological approach. *Landslides*,
610 12(2), 321–334. https://doi.org/10.1007/s10346-014-0481-1

611 Klose, M., Maurischat, P., & Damm, B. (2016). Landslide impacts in Germany: A
612 historical and socioeconomic perspective. *Landslides*, 13(1), 183–199.
613 https://doi.org/10.1007/s10346-015-0643-9

614 Lee, S., Oh, S., Ray, Ram. L., Lee, Y., & Choi, M. (2023). Three-dimensional
615 hydrological thresholds to predict shallow landslides. *Terrestrial,*
616 *Atmospheric and Oceanic Sciences*, 34(1), 20.
617 https://doi.org/10.1007/s44195-023-00052-4

618 Lin, S., Chen, S., Rasanen, R. A., Zhao, Q., Chavan, V., Tang, W., Shanmugam, N.,
619 Allan, C., Braxtan, N., & Diemer, J. (2024). Landslide Prediction Validation
620 in Western North Carolina After Hurricane Helene. *Geotechnics*, 4(4),
621 1259–1281. https://doi.org/10.3390/geotechnics4040064

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic Information Science and Systems*. Wiley. https://books.google.com/books?id=C_EwBgAAQBAJ

Lundberg, S., & Lee, S.-I. (2017). A *Unified Approach to Interpreting Model Predictions* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Mondini, A. C., Guzzetti, F., & Melillo, M. (2023). Deep learning forecast of rainfall-induced shallow landslides. *Nature Communications*, 14(1), 2466. https://doi.org/10.1038/s41467-023-38135-y

NASA Earth Observatory. (2021, June 10). *Machine Learning Model Doubles Accuracy of Global Landslide 'Nowcasts'—NASA*. https://www.nasa.gov/missions/gpm/machine-learning-model-doubles-accuracy-of-global-landslide-nowcasts/

National Hurricane Center. (2025). *Tropical Cyclone Report: Hurricane Helene (AL092024)* (p. 17). https://www.nhc.noaa.gov/data/tcr/AL092024_Helene.pdf

North Carolina Department of Environmental Quality. (2024). *North Carolina Landslide Inventory Points* [Dataset]. NC OneMap.

https://www.nconemap.gov/datasets/01965a193482438cb70332e5e524e38b_0/about

Office of State Budget and Management. (2024). *Hurricane Helene Damage and Needs Assessment*. https://www.osbm.nc.gov/hurricane-helene-dna/open?utm_source=chatgpt.com

Pennington, C., Dijkstra, T., Lark, M., Dashwood, C., Harrison, A., & Freeborough, K. (2014). Antecedent Precipitation as a Potential Proxy for Landslide Incidence in South West United Kingdom. In K. Sassa, P. Canuti, & Y. Yin (Eds.), *Landslide Science for a Safer Geoenvironment* (pp. 253–259). Springer International Publishing. https://doi.org/10.1007/978-3-319-04999-1_34

Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, 40(10), 927–930. https://doi.org/10.1130/G33217.1

Petley, D. (2025, January 6). Global fatal landslides in 2024. *Eos*. https://eos.org/thelandslideblog/fatal-landslides-in-2024

PRISM Climate Group. (2024). *PRISM Daily Precipitation Data (1981–present)* [Dataset]. Oregon State University. https://prism.oregonstate.edu/

Regorda, A., Lardeaux, J.-M., Roda, M., Marotta, A. M., & Spalla, M. I. (2020). How many subductions in the Variscan orogeny? Insights from numerical

659 models. *Geoscience Frontiers*, *11*(3), 1025–1052.
660 https://doi.org/10.1016/j.gsf.2019.10.005

661 Sidle, R. C., & Ochiai, H. (2006). *Landslides: Processes, Prediction, and Land Use*
662 (Vol. 18). American Geophysical Union. https://doi.org/10.1029/WM018

663 Sim, K. B., Lee, M. L., Rasa Remenyte-Prescott, & Soon Yee Wong. (2022,
664 August). *An Overview of Causes of Landslides and Their Impact on*
665 *Transport Networks*. Advances in modelling to improve network
666 resilience, https://op.europa.eu/en/publication-detail/-
667 /publication/c81e8bc9-1469-11ed-8fa0-01aa75ed71a1/language-en.
668 https://www.researchgate.net/publication/369437132_An_Overview_
669 of_Causes_of_Landslides_and_Their_Impact_on_Transport_Network
670 s

671 Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M.,
672 Medwedeff, W., & Clark, M. K. (2021). Data-Driven Landslide Nowcasting
673 at the Global Scale. *Frontiers in Earth Science*, 9, 640043.
674 https://doi.org/10.3389/feart.2021.640043

675 Thomas, M. A., Collins, B. D., & Mirus, B. B. (2019). Assessing the Feasibility of
676 Satellite-Based Thresholds for Hydrologically Driven Landsliding. *Water*

Resources Research, 55(11), 9006–9023. https://doi.org/10.1029/2019WR025577

Tiranti, D., & Rabuffetti, D. (2010). Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation. *Landslides*, 7(4), 471–481. https://doi.org/10.1007/s10346-010-0198-8

U.S. Census Bureau. (2022). *Cartographic Boundary Shapefiles – Counties* [Dataset]. https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html

U.S. Department of Agriculture, Natural Resources Conservation Service. (2024). *Soil Survey Geographic* (SSURGO) *Database* [Dataset]. https://websoilsurvey.nrcs.usda.gov/app/

U.S. Geological Survey. (2021). *National Land Cover Database* (NLCD) 2021 [Dataset]. https://www.usgs.gov/centers/eros/science/national-land-cover-database

U.S. Geological Survey. (2022). *3D Elevation Program* (3DEP), *1/3 arc-second DEM seamless products* [Dataset]. https://www.sciencebase.gov/catalog/item/627f3798d34e3bef0c9a3198

695 Watterson, N. A., & Jones, J. A. (2006). Flood and debris flow interactions with
696 roads promote the invasion of exotic plants along steep mountain
697 streams, western Oregon. *Geomorphology*, 78(1–2), 107–123.
698 https://doi.org/10.1016/j.geomorph.2006.01.019

699 Wooten, R. M., Gillon, K. A., Witt, A. C., Latham, R. S., Douglas, T. J., Bauer, J. B.,
700 Fuemmeler, S. J., & Lee, L. G. (2008). Geologic, geomorphic, and
701 meteorological aspects of debris flows triggered by Hurricanes Frances
702 and Ivan during September 2004 in the Southern Appalachian Mountains
703 of Macon County, North Carolina (southeastern USA). *Landslides*, 5(1), 31–
704 44. https://doi.org/10.1007/s10346-007-0109-9

705 Xiao, X., Zou, Y., Huang, J., Luo, X., Yang, L., Li, M., Yang, P., Ji, X., & Li, Y. (2024).
706 An interpretable model for landslide susceptibility assessment based on
707 Optuna hyperparameter optimization and Random Forest. *Geomatics,*
708 *Natural Hazards and Risk*, 15(1), 2347421.
709 https://doi.org/10.1080/19475705.2024.2347421

710 Ye, C., Wu, H., Oguchi, T., Tang, Y., Pei, X., & Wu, Y. (2025). Physically Based and
711 Data-Driven Models for Landslide Susceptibility Assessment: Principles,
712 Applications, and Challenges. *Remote Sensing*, 17(13), 2280.
713 https://doi.org/10.3390/rs17132280

Zheng, W., Fan, W., Cao, Y., Nan, Y., & Jing, P. (2025). Landslide Hazard Assessment Under Record-Breaking Extreme Rainfall: Integration of SBAS-InSAR and Machine Learning Models. *Remote Sensing*, *17*(13), 2265. https://doi.org/10.3390/rs17132265

Mentor Disclosure

The author conducted this research independently.

AI Disclosure

AI tools (ChatGPT) were used solely for grammar and formatting edits; all analysis and interpretation are original to the author.

Review- A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina

The authors develop a model to predict landslides with better information, accuracy, and timeliness. They do an excellent job articulating the importance and value of this topic, and their review of existing literature is extensive and very well written. Their methodology and approach to developing their models and evaluating them to identify the unique contributions of rainfall above and beyond other more static factors.

I just have two minor critiques:

1. If dynamic rainfall data is the primary contribution of this paper, why not have F0 and F1 models be the static variables and the F2 model be the addition of dynamic variables? It would be more coherent and clearly show the contribution of rainfall above and beyond the static factors.
2. The sample size is pretty small, especially for a powerful model like XGBoost. I do wonder if it may outperform Random Forest with a larger, more complex dataset, but don't think this necessarily detracts from the overall value and impact of this paper

Overall, this is an excellent paper and I am happy to recommend it for publication!

Recommendation: Accept as is

Review of Paper: “A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina”

This study addresses an important natural hazard – landslides – which are often deemed geologic risks, but as this study addresses, they may interact, particularly non-linearly, with various hydrologic factors, including rain accumulation and rain rates on various time scales, the latter of which are often triggers beyond particular USGS (United States Geological Survey) thresholds depending on context and other factors. This study introduces a landslide-risk model with PRISM gridded precipitation data, as well as static USGS Digital Elevation Model data and SSURGE (acronym not defined) soil depth data using various Machine-Learning (ML) methods. While not an inherent forecast model per se, the implications of this study, as is stated, “lay the foundation for near-nowcasting,” critical for gauging near-term risk and therefore enabling the potential to create alarms during and after hydrological triggers of concern. The primary conclusions of this study are that while the primary static feature – namely the slope, is the dominant correlate with landslides, shorter-term rain rates add significant skill in predicting landslides, whereas the accumulation of rainfall over a month or longer, particularly precipitation that is not intense but is able to percolate into soils, makes areas less susceptible for landslides – these non-linear precipitation interactions seem to be at the heart of the novelty of this study, and may be consistent or even corroborative with past work, though the framework here potentially allows new separation of variables in a real-time way given fairly high-resolution observational data and sufficiently dense static data from the USGS.

Overall, the content and the presentation seem quite compelling, with some degree of modification needed here and there (further elucidated in the specific comments below), though there are some assumptions that the writer may have in terms of *a priori* knowledge of the reader that may not be deemed to be common, general knowledge. While this is a modest hindrance of accessibility of the study overall, it could fairly straightforwardly be ameliorated with some additional description and definition of some of the statistical/machine learning terms including, but certainly not an exhaustive list - Logistic Regression, Random Forest, and XGBoost. Furthermore, several acronyms are used without being spelled out or necessarily defined. Eventually, a reader may be able to deduce how to interpret these, but as this is not a highly specialized journal, spelling out what key acronyms mean is critical so that nearly any reader can walk away with key takeaways of the study without extensive background reading first. In general, unless explicitly stated by a journal, most, if not all acronyms should be defined once before using them throughout.

Related to the lack of some of the key definitions utilized throughout the manuscript, a number of the figures, which though decipherable for someone in the field of certain areas of specialized statistics or ML, may not be as easy to interpret for a generalist in hydrology, geology, mathematics, atmospheric sciences, or even traditional statistics. These include the ROC curves (Figures 4 and 8), the SHAP beeswarm plot (Figure 12), and perhaps to a lesser extent the Confusion Matrices (which might be mislabeled in terms in terms of the legends and figure numbers – more on that below). A brief description and interpretation may be needed for each of these, rather than merely explaining the results. One reason I am bringing this up is that many readers may be most familiar with scatter plots relating two or more variables, with degree of goodness of fit, often presented in the form of a correlation coefficient, is measured in terms of closeness to the diagonal (at least for linear relationships), but the best performance for an ROC (Receiver Operating Characteristic) curve is towards the vertical (top left corner), and on an ROC curve the diagonal simply suggests a random classifier. As

someone not as well-versed in ROC curves, it took me some additional time to understand quantitively why model sets F1 and F2 are generally superior to F0 based purely on the figures alone. The text generally is satisfactory at explaining why rainfall metrics improve the models, but the characterization of some of the figures needs additional information in the manuscript.

Finally, there is some consolidation that can likely be done which might have the dual objective of saving some space, reducing the number of figures, and better allowing a more direct comparison among the different ML methods – and that would be including all of the ROC curves (from Logistic Regression, Random Forest, and XGBoost) onto one panel with perhaps different symbols on some of the curves to properly distinguish them. Alternatively, consider a multi-panel figure with all three models that save the reader from flipping to different pages to compare on different pages.

Overall though, while some improvement is possible with regards to accessibility, presentation, and at some points even some ambiguity of the interpretation of some of the rainfall variables, this paper should be accepted pending mostly satisfactory minor (and maybe some moderate) revision. With regards to grammar, the paper is generally well written, though there are a few instances in which tenses switch to past tense whereas present tense would be more appropriate. A thorough (though non-exhaustive) list of these technical and grammatical points are provided below.

Specific Points:

1) Lines 10-11: USGS DEM appears with being spelled out – in an abstract it should be written as "USGS Digital Elevation Model (DEM)". The same issue exists with SSURGO. It is not until line 183 that DEM is written out. If SSURGO is not written out anywhere else (I couldn't find it anywhere), it should be in the table of definitions (Table 1), and at the very least on line 11 add "depth" after "SSURGO soil".
2) Lines 90-92: Why is the question in quotes? Is this to bring additional emphasis, and are these words original to the author?
3) Lines 93-94 and more generally: What lead time is being sought preferentially with regards to the proposed "early warning system?" How much lead time is necessary to potentially reduce mortality related to landslides under more extreme conditions?
4) Lines 108-109: These numbers referred to in the citation (Ye et al., 2025) refer to the number of publications of landslide susceptibility studies, rather than the number of statistical models, it appears. Please clarify.
5) Line 114: What does RNN stand for? It is not defined anywhere and only seems to appear once here. My best guess (after searching a bit) is Recurrent Neural Network, but please spell out if this is this case.
6) Lines 250-251, Table 1, and more generally: I notice that there are not maximum rain rates at shorter time scales (subdaily, even hourly), which might be important as well. Speaking of which, the USGS often has millimeter (mm)/hr or inches (in)/hr thresholds, which if breached, can be triggers of both landslides and mudslides, so wonder about the appropriateness of PRISM only data for very intense, short events? I am thinking here also more about the generalizability of some of the results of this study – including landslides/mudslides following burn areas on hillsides/mountains with steep slopes and denuded vegetation; 5-minute to hourly rain rates can be critical at determining the potential susceptibility of landslide hazards. Or, would high temporal resolution rain rates be beyond the scope of the research/model presented here? The

latter could possibly come from radar data (https://mrms.nssl.noaa.gov/qvs/product_viewer/) or IMERG precipitation data, which is merged satellite-gauge data (https://gpm.nasa.gov/data/imerg).

7) Line 262 and more generally: The first instance of ROC appears to occur here on line 262, but it is not defined, even though it is a heavily used metric/concept throughout the study, and also in multiple figures/tables. I believe that it refers to Receiver Operating Curve, but please confirm. ROC may not be a common acronym, so for the benefit of readers, I would suggest spelling it out during its first instance, and also perhaps defining a bit more, in particular additional characterization of the plots and how a non-specialist reader should interpret the ROC curves. This also applies to the acronym AOC (Area Under the Curve).
8) Lines 288-289 and Lines 400-402: As the authors rightly note, including soil depth does not necessarily improve results beyond including rain rates/rain accumulation in explaining landslides – any improvement at times does not seem statistically significant. Why do you think this might be?
9) Lines 329 and beyond – More on the tenses switching to past when present would be more appropriate: It's best to be consistent and maintain present as much as possible. The exception to this is if an action was done at a particular point in time – e.g. "Data were collected on XX February XX", in which case past tense is correct.
10) Lines 329-337 and Results in Tables 2 & 3: The author emphasizes that "short-period rainfall accumulation" … "have a strong relationship with landslide risks", but "Max_Rainfall_30day" is such a prominent variable in all five of the top five logistic regression groups in Table 2. Perhaps a bit surprising to me, "Max_Rainfall_3day" only appears in one subset. The predictors for the Random Forest Table (Table 3) are somewhat different, with Max_Rainfall_3day making an appearance twice, although the 30-day equivalent is present in all five of the subsets. Indeed, when I look at all five subsets of Table 3, it seems that rainfall at the short-term (often known as synoptic scales of a few days up to two weeks) to monthly are broadly and nearly equally important.
11) Figures 4 (and 8 and 14): Particularly for those with varying degrees of color blindness, I would highly suggest altering the styles of the lines to better distinguish them, or perhaps even adding symbols to one or more to help differentiate. Even for someone without color blindness, the blue and green (F0 and F2 curves, respectively) are a bit tricky to distinguish (a good way of checking how figures appear to someone with varying degrees of color blindness can be found here: https://www.color-blindness.com/coblis-color-blindness-simulator/)

    Somewhat relatedly, for more general readers, I would be sure to specify that the thick dashed line is the "No Skill" line – not everyone knows how to interpret these ROC curves without some additional guidance. This also needs to be included in the figure legends.

12) Table 5: I am a little confused – why are rainfall parameters included in the F0 Logistic Regression Model – I thought only static slope/elevation parameters were used? What is the meaning of these coefficients in the context of F0?
13) Lines 370-371: This accuracy (of the F2 model) is the same (0.70) as when F1 was included as well – I'm not quite sure I follow?

14) Lines 373-374: While this statement sounds promising in terms of the improvement in correct identification of landslide cases correct from F0 versus F2, is this improvement statistically significant? Is there a way to formally test for this?
15) Table 5: The behavior of Max_Rainfall_30day for F1 and F2 is positive in terms of the sign of the regression coefficient, but strongly negative for F0 (again, with the caveat that I still do not know how to interpret the rainfall parameters for F0), but R30d is negative for F1/F2, which is a little difficult to digest. Max_Rainfall_30day is simply the rolling maximum rainfall of 30-days, is it not? Or, is it the maximum rainfall (of a day) within a 30-day period? Why is the behavior of R30d and Max_Rainfall_30day so different? Please help me understand what I am missing.
16) Lines 398-399: Yes, I see this in the ROC curves as well – why might the inclusion of soil depth slightly degrade the performance of F2 in terms of the Random Forest model?
17) Figure 13: I am confused about the two shades of blue/purple in this Figure 13 – what do the darker purple bars refer to? They are not defined in the legend or the caption. Also, no units are provided in Figure 13 – are they mm?
18) Please be certain to check all the DOIs in the Works Cited/References section are correct – I caught at least one in which the URL did not work for the reference in lines 544-547. It should be: https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2
19) Figure 12 and corresponding explanation in text: Please provide additional guidance for readers in how a SHAP Beeswarm chart is constructed and how to interpret it. Many non-specialists are not familiar with this type of plot.
20) Figures 5-7, 9-11, and 15-17: The three confusion matrices for each of the models do not require a separate figure number; they can simply be deemed a single figure, with multiple panels. I like these figures, but these so-called 9 figures really only need to be three. Furthermore, I might provide a little more narrative in the text about how to understand a Confusion Matrix, especially as many readers may not be that familiar with them.

Grammatical Suggestions:

1) First line of abstract (line 5): Change “Rainfall-triggered landslides are commonly occurring across the Southern” to “Rainfall-triggered landslides commonly occur across the Southern”
2) Lines 121-122: “Incorporating every aspect of geographic, environmental and hydrological factors into eight variables” is an incomplete sentence. I am not sure if the author intended to convey an additional thought here, or if this sentence got truncated?
3) Line 259: Need a period after the parentheses.
4) Line 273: Change “visualized” to “visualizes”
5) Line 274: Change “showed” to “show”
6) Line 275: An example of where the tense should be present (not past): Change “SHAP waterfall plots gave” to “SHAP waterfall plots give” (there are also multiple other instances in this paragraph in which the tense of statements should be present).
7) Line 400: Awkward Transition: “Thus” would be a better way to start the new sentence than “Still”

A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina

[name redacted by Managing Editor]
[school redacted by Managing Editor]
October 2025

Abstract

Rainfall-triggered landslides commonly occur across the Southern Appalachians, yet effective, localized "nowcasts" that fuse dynamic precipitation with terrain characteristics remain limited. We develop a near-real-time landslide risk model for Buncombe County, North Carolina, integrating daily PRISM precipitation windows with static hydrological data. The novel dataset containing 302 mapped landslides (1981–2024) was compiled by integrating USGS Digital Elevation Model (DEM) and Soil Survey Geographic Database (SSURGO) soil data set. Across models, slope was the dominant feature, while short-duration rainfall (1–7 days and 3-day maxima) most strongly increased landslide probability. In particular, 30-day cumulative rainfall often exhibited a negative association with predicted risk, suggesting that prolonged rainfall without intense surges may reduce immediate triggering. The resulting model enables county-scale, data-driven susceptibility updates based on forecasted precipitation data, offering a practical foundation for near-nowcasting and a landslide alarm system during and after intense storms.

Keywords: Rainfall-triggered landslides, Nowcasts, Buncombe County, Random Forest, XGBoost, PRISM precipitation, Landslide susceptibility modeling
[Repository link redacted for review; will be included upon acceptance]
[Images that are under 300 dpi can be found in higher quality in this google docs]

Introduction

Landslides are devastating natural hazards that take place worldwide and cause wide-ranging damage to modern society, such as human casualties, economic losses, and infrastructure damage (Froude & Petley, 2018; Kirschbaum & Stanley, 2018; Petley, 2012). In the year 2024 alone, 766 instances of landslides took place around the world, killing 4,933 individuals (Petley, 2025). Landslides across the world incur an economic cost of US$20 billion annually, taking 17% of the yearly

mean expenses caused by all natural disasters worldwide between 1980 and 2013 (Klose et al. 2016). Social infrastructures, especially transportation and communication, are highly susceptible to landslides (Sim, Lee, Remenyte-Prescott, & Wong, 2022). Rural areas are likely to experience more impairments due to landslides, where these infrastructures are scarcer and scattered (Klose et al. 2015).

In search of effective landslide prediction methods, researchers have proposed various theories on the sources of these disasters, ranging from geological activity such as volcanic eruptions and earthquakes to human-induced sources such as modification of slopes and obstruction of hydrological flows (Sidle & Ochiai, 2006; Jaboyedoff et al., 2016). Among those factors, precipitation is regarded as the leading cause of landslides worldwide; extreme rainfall triggered 70% of all landslide events across the world between 2004 and 2010 (Petley, 2012). Fittingly, the role of precipitation in landslides will only continue to grow with climate change; Gariano and Guzzetti (2016) predict that landslides triggered by rainfall will be increasingly catastrophic and recurrent as climate change increases the frequency of short, intense storm episodes. The predicted increase in landslide prevalence forebodes the urgency of an effective rainfall-induced landslide prediction model.

The Western North Carolina region has historically been dominated by rainfall-induced shallow landslides due to geographical and meteorological factors (Wooten et al., 2008). The Blue Ridge Mountains, as a part of the Southern Appalachian mountain range, are characterized by steep slope gradients (Khashchevskaya et al., 2025). Its bedrock layer mostly comprises metamorphic and igneous rocks, which, when decomposed, create saprolite, which constitutes the top layer of the horizons (Hatcher, 2010; Watterson & Jones, 2006). The saprolite-heavy soil absorbs water easily, which increases pore pressure, resulting in a higher probability of shallow landslides (Aydin, 2006). The region's geomorphic susceptibility has been called for focus again after Hurricane Helene's extensive landslide damage in Western North Carolina in 2024 (Lin et al., 2024). Allstadt et al. (2025) state that the hurricane caused unprecedented geographic disruptions in the region, particularly in Buncombe, Henderson, Rutherford, and Yancey counties. Freshwater flooding contributed to 95 of the 176 direct deaths that Helene caused, most of which were landslide-related; around 2,000 independent cases of landslides were reported to be related to Helene, with the majority located over western North Carolina (National Hurricane Center, 2025). Helene significantly

damaged the communications infrastructure in western North Carolina, considering WNC's rural characteristic, which exacerbated the situation (Office of State Budget and Management, 2024). Helene's unexpected landslide damage calls for a real-time probabilistic forecasting model, localized for Western North Carolina.

At present, there is no near-nowcast model for Western North Carolina that captures dynamic changes in precipitation. In fact, nearly all of the current landslide hazard maps use minimal or no precipitation inputs. Specifically, the USGS Landslide Hazard Mapping Program's landslide susceptibility map identifies areas with elevated risk for landslips under extreme rainfall conditions (Fuemmeler et al., 2008); these, however, do not take account of real-time changes in precipitation since they were based on past correlations of landslides with geographical elements (Bauer et al., 2012). Academically, even though there have been numerous attempts at developing a physical/statistical model for the prediction of landslides (e.g., an ML-based model by Lin et al. (2024) that combines a variety of predictors such as the NC landslide database, soil surveys, digital elevation models, and water body locations), very few, if any, utilized dynamic rainfall as their primary predictor. To this end, an effective localized model of Western North Carolina for rainfall-induced landslides has proven necessary, as seen in the extensive damage Helene caused in the region.

Thus, to address the absence of a localized prediction model that incorporates dynamic rainfall variability, the proposed research will address the question: How does incorporating static geographic data (elevation and soil depth) with the variability of rainfall data affect the predictability of shallow landslides in the Western North Carolina region? Ultimately, this study hopes to create an efficient prediction model in near real-time (rainfall data are updated daily) for the area, developing an early warning system for residents of Western North Carolina, specifically, Buncombe County. Because the model relies on rainfall windows that end on the event day, it produces a near-nowcast rather than a true forecast with a defined lead time. Generating operational lead-time predictions would require integrating short-term rainfall forecasts, which is beyond the scope of the present study but represents a clear direction for future work.

Global models based on rainfall-triggered landslides have been developed and utilized for several years, with generally positive results. They are considered

‘nowcasts,’ because they incorporate both static geological characteristics and dynamic observations of rainfall to provide nearly real-time forecasts (Kirschbaum et al., 2012). For example, NASA's LHASA v2 model is the first operational global system to utilize dynamic satellite rainfall data along with static geological and environmental characteristics (Stanley et al., 2021). It is based on a Logistic Regression algorithm and examines static predictor variables as well as IMERG precipitation (1 km resolution) to generate probabilistic hazard labels. The new model is purported to be two-fold more accurate than the previous LHASA v1 system, which utilized a standard threshold algorithm (Kirschbaum & Stanley, 2018; NASA Earth Observatory, 2021).

Only recently has the volume of published landslide susceptibility studies grown substantially, from 31 in 2016 to 219 in 2024 (Ye et al., 2025). Supervised machine learning (ML) was a popular choice for its ability to capture nonlinear relationships between predictor types and landslide probability, resulting in a higher level of accuracy than conventional linear statistical models (Regorda et al., 2020). Mondini et al. (2023) developed a time-dependent landslide model based on Italy. Exploiting the strength of Recurrent Neural Networks (RNNs) in sequential data, they used continuous rainfall data for a window of 30 days, excluding geological or environmental factors. Similarly, Chan et al. (2018) built a logistic regression-based model aimed at predicting landslides caused by a typhoon’s heavy rain in the southern region of Taiwan. The model mainly used runoff flow depth, not pure precipitation data, taking into account soil saturation, which reached an accuracy of 80~85%. Kang et al. (2024) recently constructed a localized model for Yunnan Province, China, with its Random Forest model reaching an accuracy level of 0.906. All the above ML-based models center around rainfall data. Their high accuracy levels affirm the need for dynamic rainfall data in a localized landslide model.

Focusing on North America, Thomas et al. (2019) challenged if satellite rainfall data can replace in situ hydrological data to evaluate the soil saturation threshold for a slope failure, reflecting a recent trend of increasing remote geospatial data-based models (Akosah et al., 2024). The findings from this localized model of California demonstrated that rainfall (specifically satellite-measured rainfall data) cannot be relied upon to predict landslides and encourage a hydrogeological gauge that calibrates rainfall data with on-site hydrologic data to develop a soil wetness index. Other studies that support the aforementioned findings include a

comprehensive model (Lee et al., 2023) blending precipitation duration/intensity and normalized soil moisture capacity, which dropped false alarms (FA) from ~26 to 3. Both studies emphasize the importance of contextualizing raw rainfall data with local hydrologic features in evaluating the soil saturation level.

The contribution of this study is twofold. We build a localized, near-real-time landslide prediction model. For the modeling, an equal number of non-events were randomly chosen based on the landslide inventory of Buncombe County (North Carolina Department of Environmental Quality, 2024). To avoid pattern biasing in specific constellations of events, stratification and a spatial block cross-validation approach were utilized. Three different ML algorithms were tested overall, with three models per algorithm (F0, F1, and F2) based on feature type. The assessment of performance and feature analysis was interpreted using evaluation metrics, SHAP plots, and permutation importance.

To fit and evaluate the model, we designed a hydrologic/environmental dataset that collates daily precipitation data (PRISM Climate Group, 2024), a topography map (USGS DEM), and soil depth (USDA SSURGO). The dataset is open and available for further research at the specified repository.

The next section describes the datasets and methodology of the research. Section 3 illustrates the results of the findings. Section 4 delivers the conclusion of the research.

Data and Methodology

Dataset

This research aimed to create a dynamic prediction model localized in Buncombe County, North Carolina, geographically located within the Blue Ridge Mountains (Figure 2a). The inventory of landslide events was extracted from the North Carolina Landslide Points dataset (via NC OneMap; North Carolina Department of Environmental Quality, 2024). The dataset included event geometry (points) and traceability fields: *IsLandslide* (event flag), *Event_Date*, *Sort_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, and *Source_Period*. The inventory was both a temporal anchor (*Event_Date*) for rainfall-window calculations referred to by dates and a spatial anchor (X, Y) for topography and soil data extraction.

Additionally, an equal number of non-event samples were created for control, randomly selecting the same time (January 1981–December 2021) and same region of interest (Buncombe County) as those in the landslide dataset. Ultimately, equal numbers of events and non-events were included in the final dataset used for modeling.

Gridded precipitation (4 km resolution) was extracted from PRISM daily products (PRISM Climate Group, 2024) throughout the entire analysis period (spanning from January 1981 to December 2024). Using the reference dates of the events and non-events—the event date (*Event_Date*) or the control (*Random_Date*)—historical sums over windows of 1-day (*R1d*), 3-days (*R3d*), 7-days (*R7d*), and 30-days (*R30d*) were computed. Other temporal windows included maximum rainfall sums calculated for both 3-day (*Max_Rainfall_3day*) and 30-day (*Max_Rainfall_30day*) intervals. All precipitation measurements were in units of millimeters, and all the calculation windows' end dates coincided with the reference date of the observation (Table 1 contains the definition and sources of all features).

Topography was taken from a DEM (Digital Elevation Model) encompassing Buncombe County published by the U.S. Geological Survey (2022). The DEM already provided elevation (*Elevation_m*), while slope (*Slope_deg*) was computed from Horn's (1981) 3×3 finite difference gradients and converted into degrees (Figure 2; Equations 1, 2, and 3).

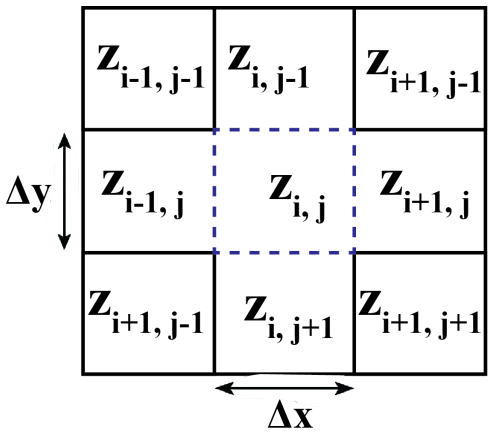

Figure 1. Visualization of Horn's (1981) 3×3 finite difference gradient calculator. $Z_{i,j}$ symbolizes the position of the elevation of a particular cell. Δx and Δy are the horizontal and vertical grid distances of the cells.

$$p = [(z_{i+1,j-1} + 2z_{i+1,j+1} - z_{i-1,j-1}) - (z_{i+1,j+1} + 2z_{i-1,j} + z_{i-1,j+1})] / 8\Delta x \quad (1)$$

$$q = [(z_{i-1,j+1} + 2z_{i,j+1} - z_{i+1,j+1}) - (z_{i-1,j-1} + 2z_{i-1,j} + z_{i-1,j-1})] / 8\Delta y \quad (2)$$

$$Slope = tan^{-1}(\sqrt{p^2 + q^2}) \quad (3)$$

The values p and q in Equations 1 to 3 represent the rate of elevation change in the east-west direction and the north-south direction, respectively. The DEM cell resolution for Buncombe County showed an equal width and height of 40 meters. To avoid unit discrepancy across different datasets, all datasets were under a consistent projected coordinate reference system (NAD83/North Carolina; EPSG:32119) and maintained a resolution of 40m.

Soil data originated from USDA NRCS Soil Survey Geographic Database (SSURGO) map units (U.S. Department of Agriculture, Natural Resources Conservation Service, 2024). A point-polygon comparison was used to match MUKEY and MUSYM to each sample point, linking depth information from the related tables. Soil depth was converted to a numeric variable (*Soil_Depth_cm*) by extracting data from raw strings (*Soil_Depth_cm_raw*) and converting the variable into a binary flag (*Soil_Depth_Deep200_Flag*), where entries larger than 200 were changed to 1, while any other entries were changed to 0. These soil properties were viewed as static covariates that reflect conditions of regolith associated with instability triggered by rainfall.

For map compilation and masking, administrative boundaries by county were employed to clip rasters and to define the sampling window for controls with a modest buffer to attenuate edge effects (U.S. Census Bureau, 2022). Land cover from NLCD was retained for descriptive mapping and voluntary sensitivity tests but was not included in the baseline prediction set (U.S. Geological Survey, 2021).

Gridded variables such as rainfall windows, elevation, and slope were sampled by bilinear interpolation at coordinates defined by points (X, Y) (Burrough & McDonnell, 1998). Soil data, specifically soil depth, came from the MUKEY and MUSYM area codes. Units of measurement have been standardized with precipitation in millimeters, elevation in meters, and slope measured in degrees. We

verified the rainfall windows to check that the end dates coincide with the event reference date. Data instances with missing values were removed from the dataset.

The dataset created is a linked dataset, combining information from PRISM daily rainfall data, the USGS DEM, and the USDA SSURGO map specific to Buncombe county, and can be used for further analyses related to landslide risk. Linking information from the various datasets was not trivial, as it required gathering chronological information and summarizing it for the selected data points before linking it to the main dataset (Longley et al., 2015; Burrough & McDonnell, 1998). Other challenges included incompatible reference systems that had to be reconciled. The complete dataset includes the outcome variable (*IsLandslide*) and full fields for past rainfall, topography, and soil depth with comprehensive metadata for provenance and audit purposes, which comprises *Event_Date/Random_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, *Source_Period*, *MUKEY*, *MUSYM*, and *Soil_Depth_cm_raw*.

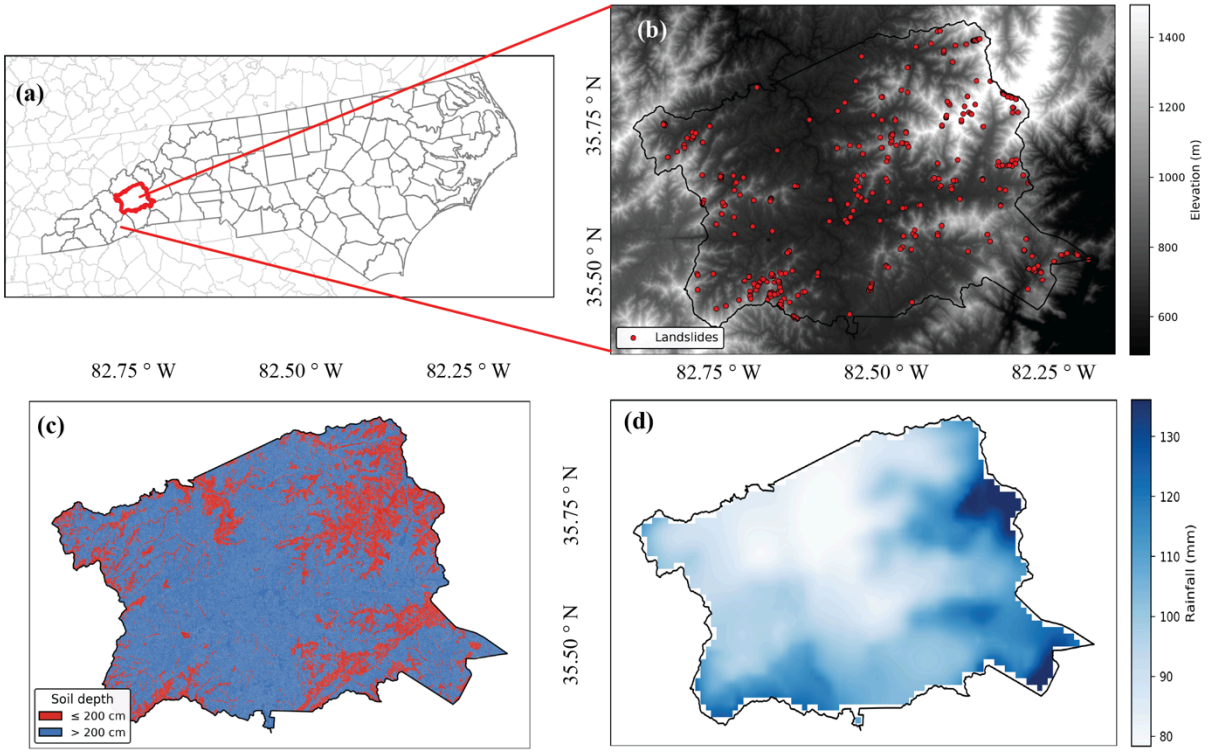

Figure 2: (a) Geographical location of the study area, (b) elevation, (c) soil depth, (d) 30 year average rainfall.

Symbol (unit)	How it was calculated	Source
R1d (mm)	Sum of daily PRISM rainfall for 1 day	PRISM (800m)
R3d (mm)	Sum of daily PRISM rainfall for 3 days	PRISM (800m)
R7d (mm)	Sum of daily PRISM rainfall for 7 days	PRISM (800m)
R30d (mm)	Sum of daily PRISM rainfall for 30 days	PRISM (800m)
Max_Rainfall_3day (mm)	Rolling maximum of 3-day totals in past 30 days	PRISM (800m)
Max_Rainfall_30day (mm)	Rolling maximum of 30-day totals in past 90 days	PRISM (800m)
Elevation_m (m)	Extracted directly from DEM	USGS DEM (40m)
Slope_deg (°)	Derived from DEM using slope algorithm	USGS DEM (40m)
Soil_Depth_Deep200_Flag (—)	Flag for soils deeper than 200 cm	SSURGO Soil

Table 1: Features definitions (Rainfall time-frames are from the reference point of the corresponding event)

All data collection, processing and model development steps were done in Python 3.9 in the PyCharm IDE. Libraries, including NumPy, Pandas, SciPy, Matplotlib, and Seaborn, were used for scientific computing; GeoPandas, Shapely, and Rasterio were used for data processing; and scikit-learn and XGBoost models combined with SHAP were chosen for model development and assessment. The models were run on a Mac Apple M2 processor.

Modeling

To assess which rainfall variables specifically influenced landslide probability, the temporal accumulation/maximum rainfall predictors were analyzed with all non-empty possible combinations. A primary concern was minimizing overfitting due to the large number of rainfall predictors (*R1d*, *R3d*, *R7d*, *R30d*, *Max_Rainfall_3day*, *Max_Rainfall_30day*). Thus, a training–test split was implemented via a stratified five-fold cross-validation approach using a constant

seed and data point shuffling. In addition, the evaluation was repeated for every rainfall predictor subset. The model that showed the most accurate predictions on the test data was included in the final comparison across other types of models. The above processes were independently conducted for each machine learning (ML) algorithm model.

Three separate ML algorithms were employed for the actual study—Logistic Regression, Random Forest and XGBoost (Breiman, 2001; Chen & Guestrin, 2016). Due to logistic regression's simple and linear design, the model was effective in serving as a benchmark to compare with nonlinear models. Receiver Operating Characteristic (ROC) curves plot the true positive rate against the false positive rate across probability thresholds. The Area Under the Curve (AUC) summarizes model discrimination. The diagonal dashed line represents the no-skill baseline, where the model performs no better than random chance. The performance of the model was quantified through ROC curves, confusion tables, and summary measures: accuracy, precision, recall, ROC–AUC, and the Brier score (Brier, 1950). For logistic regression, the coefficient magnitudes summarized how each variable influenced the model prediction, while standard errors, p-values and confidence intervals measured the degree of statistical significance of that prediction.

A Random Forest algorithm was used as the baseline nonlinear model. Interpretation of the model involved a combination of intrinsic feature importance as well as SHAP analysis. Intrinsic feature importance was determined as the mean decrease in impurity, indicating the effectiveness of each predictor in splitting the data. Shapley values for each variable and data point were illustrated via SHAP bar plots, SHAP beeswarm plots and SHAP waterfall plots per sample (Lundberg & Lee, 2017). The bar plot visualizes the average contribution of each feature to the overall prediction, while the beeswarm plots show both the distribution as well as the direction of the predictor contributions. SHAP waterfall plots give local explanations regarding the most likely landslide samples. A full dataset, including SHAP values as well as both the imputed and the original feature values, expected values, probabilities, and true labels, in addition to original metadata, was also provided so that both global and case-specific insights could be obtained. Model discrimination as well as calibration were also assessed as part of ROC curves and confusion matrices, as well as the same summary metrics that were used as part of Logistic Regression.

In answering the hypothesis, three separate models were created for each type of ML algorithm. Model F0 used just slope and elevation, aiming to assess the predictive potential based on the static predictor: terrain alone. Model F1 added rainfall accumulation intervals, alongside slope and elevation, as the basis to also evaluate the short- as well as longer-term rainfall triggers. Following on from F1, Model F2 also included soil depth so that the effect of the subsurface interacting with both rainfall and the terrain factors could be determined. Each algorithm was then trained across the three feature settings, ultimately leaving us with parallel models that allowed a direct comparison.

The F0, F1, and F2 configurations for each model (Logistic Regression, Random Forest, and XGBoost) were trained and tested on the same sample and therefore any discrepancies in results were not driven by sampling but rather variable differences.

Comparison analysis conducted on XGBoost involved classification accuracy, showing ROC curves and confusion matrices, as well as the complete set of metrics, including accuracy, precision, recall, F1 score, and ROC–AUC, as well as the Brier score. Interpretable techniques, such as SHAP dependence plots or permutation importance, are only presented for the Random Forest model.

Results and Discussion

The compiled landslide dataset contained 9,092 events recorded within North Carolina between 1940 and 2024, of which 398 occurred within the study area of Buncombe County (Figure 3). Because the PRISM Weather daily rainfall data before 1981 was unavailable for the public, the inventory was limited to the time frame between 1981 and 2024, limiting the data points to 302 landslide events. Corresponding to these landslide points, an equal number of control points were created by sampling sites and dates randomly within the same county and time frame. As such, the final dataset contained 604 total entries, split evenly between events and controls, thus reducing potential biases for certain variables and increasing its robustness toward new data.

All the records had a collection of hydrological covariates. Rainy features comprised both cumulative windows ($R1d$, $R3d$, $R7d$, $R30d$) as well as highest

intensities (*Max_Rainfall_3day*, *Max_Rainfall_30day*). Geological features included elevation as well as slope at 40 m resolution. Soil depth was included as an added predictor expressed as a binary flag (>200 cm vs. ≤200 cm), as the vast majority of the sites had extremely deep soils.

The final dataset merged the rainfall, terrain, and soil characteristics for each observation. Its well-balanced design (302 events + 302 controls) ensured the risk of class bias was kept small, and the range of covariate variability enabled the models to incorporate long-term conditioning factors (e.g., slope, soil depth) as well as short-term triggers (e.g., extreme rains).

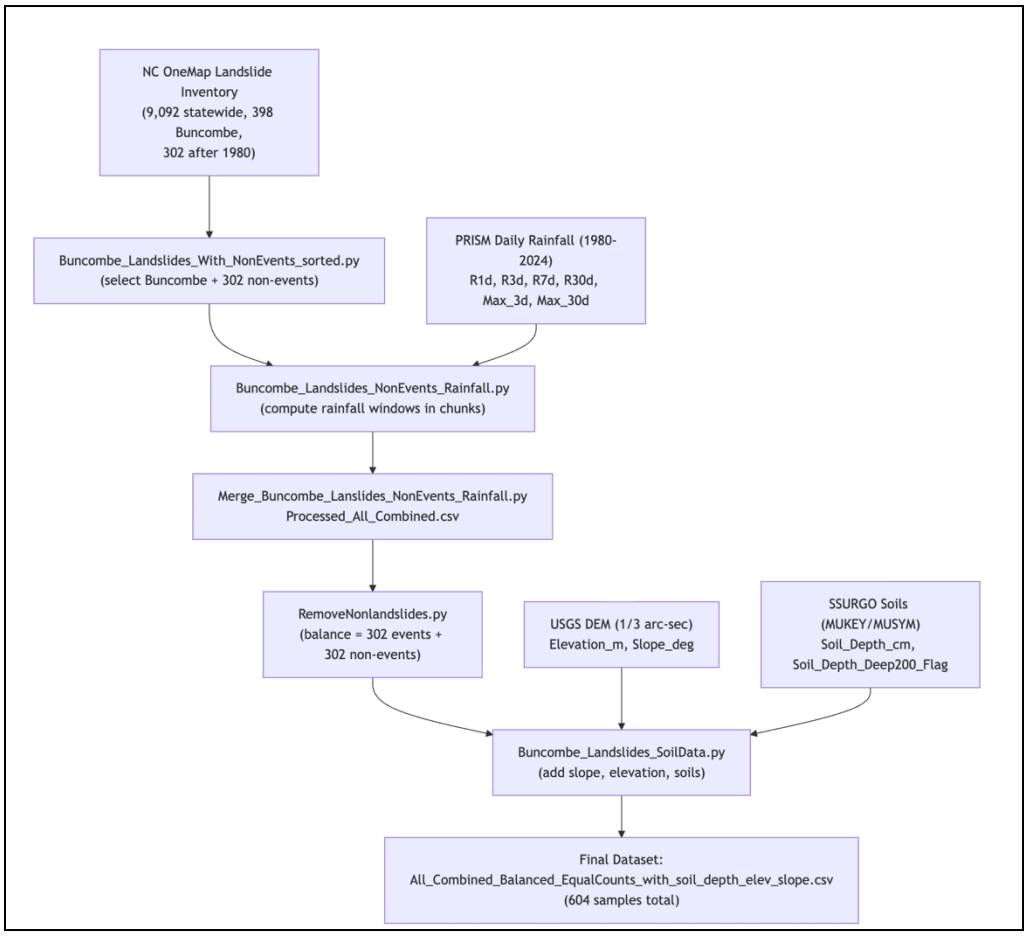

Figure 3: Flowchart of landslide events and non-events data processing

Feature selection:

Feature selection indicated consistency across the nonlinear and the linear models. In the Logistic Regression category, the optimum subset of rainfall data comprised *Max_Rainfall_30day*, *R3d*, *R7d*, and *R30d*, achieving an accuracy of 0.593

as well as an ROC–AUC of 0.614 (see Table 2). Short-period rainfall accumulation, including three-day as well as seven-day buildup, was frequently shown in the top-ranking combinations, showing how these predictors have a strong relationship with landslide risks. However, long-term predictors such as *Max_Rainfall_30day* and *R30d* also showed up occasionally, albeit not as frequently as short-term variables.

Logistic Regression (Top 5)	Accuracy	ROC–AUC
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.593	0.614
Max_Rainfall_30day, *Max_Rainfall_3day*, *R3d*, *R7d*	0.589	0.621
Max_Rainfall_30day, *R1d*, *R3d*, *R7d*, *R30d*	0.589	0.613
Max_Rainfall_30day, *R7d*, *R30d*	0.584	0.600
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.584	0.619

Table 2: Top-performing subsets of Rainfall Features for Logistic Regression

	Accuracy	ROC–AUC
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*	0.808	0.888
Max_Rainfall_30day, *R1d*, *R7d*, *R30d*	0.806	0.882
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.805	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*, *R7d*	0.801	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.801	0.883

Table 3: Top-performing subsets of Rainfall Features for Random Forest

In the Random Forest output, the best-performing subset included *Max_Rainfall_30day*, *Max_Rainfall_3day*, *R1d*, and *R3d*, yielding exceptional accuracy up to 0.808 and ROC-AUC reaching as high as 0.888 (see Table 3). Moreover, subsets including R3d and R7d also appeared among the very best, a point also supported by the results yielded by the Logistic Regression (see Table 2).

However, the Random Forest model also brought to light that short-term rainfall predictors like *R1d* as well as the *Max_Rainfall_3day* significantly improved classification, bringing to the forefront nonlinear interactions. These findings indicated that although both nonlinear and linear prototypes invariably recognized the paramount significance of short-term rainfall, the Random Forest technique, in addition, captured the additional nuances of the lengths of the rainfall affecting the probabilities of the landslide occurrences (refer to Table 3).

Actual Models:

Logistic Regression:

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.620	0.609	0.650	0.629	0.673	0.229
F1	0.702	0.688	0.733	0.710	0.771	0.193
F2	0.702	0.682	0.750	0.714	0.775	0.191

Table 4: Logistic Regression model performance across feature sets (F0, F1, F2).

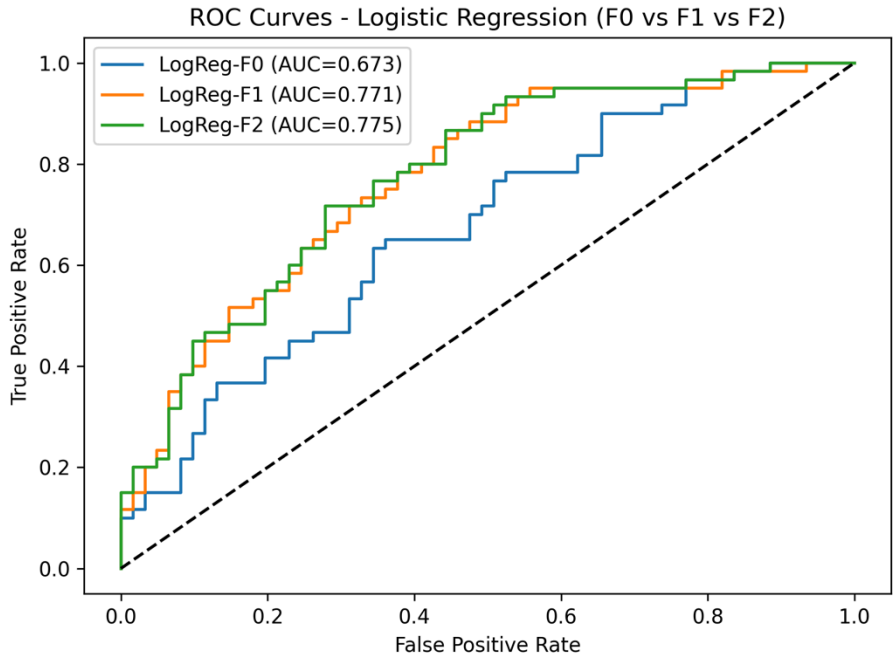

Figure 4: ROC Curves for the three Logistic Regression models F0, F1, and F2. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

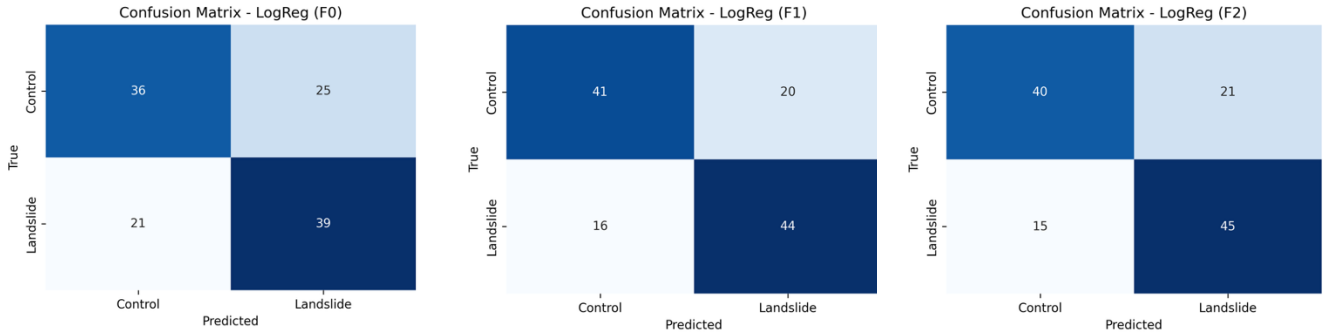

Figure 5-7: Confusion Matrices of F0, F1, F2 Logistic Regression Models

Feature	**F0**	**F1**	**F2**
Slope_deg	0.7150	1.3574	1.3813
Max_Rainfall_30day	-	0.5666	0.5871
R30d	-	-0.5968	-0.6327
R7d	-	0.2192	0.2136
Elevation_m	-0.0287	-0.1886	-0.1099
R3d		-0.0571	-0.0620
Soil_Depth_Deep200_Flag			0.1943

Table 5: Regression coefficients for each feature of F0, F1, F2 Logistic Regression Models. Note that rainfall-related coefficients are intentionally blank for the F0 model, as F0 includes only static terrain predictors (slope and elevation) and therefore does not estimate coefficients for any rainfall variables.

Performance for Logistic Regression got progressively better as more predictors were added (Table 4). In the terrain-only version F0, accuracy was 0.62 with an AUC (Area Under Curve evaluates how well the model distinguishes between cases and controls) of 0.67. Inclusion of the rainfall predictors in F1 got the accuracy up to 0.70 and the AUC up to 0.77 (Figure 4). Inclusion of the full set of predictors in the complete version F2 got the best accuracy at 0.70 as well as the best AUC at 0.78. The confusion matrices further corroborated this ranking of models (Figures 5-7): F2 got 40 correct controls as well as 45 correct landslides,

compared to 36 and 39 cases for the F0 model and 41 and 44 cases for the F1 model respectively.

No formal statistical test (e.g., McNemar's test or bootstrap confidence intervals) was conducted to evaluate whether these accuracy differences are statistically significant, and the improvements should therefore be interpreted cautiously.

Model interpretation relied on the regression coefficients (Table 5). The slope was the strongest predictor variable, followed closely by *Max_Rainfall_30day*, both having a positive effect on the risk of landslides, except for the F0 model. R30d represents the total cumulative rainfall in the 30 days preceding the event, whereas Max_Rainfall_30day represents the single wettest 30-day period within the past 90 days. These variables capture different hydrologic processes: R30d reflects gradual wetness buildup, while Max_Rainfall_30day captures past extreme episodes. Their differing definitions explain why one may be positive and the other negative in regression coefficients. Here, the cumulative 30-day rainfall (*R30d*) tended to have a negative coefficient, meaning long-term accumulation did not significantly increase the probability of slope failure after accounting for short-term accumulation or extreme events. These results strengthen the conclusion that the best explanation for landslides lies in the interaction between the region's steep slope and short-term, high-rate bursts of rain, shown in the Brier Scores of Table 4.

Random Forest:

FeatureSet	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**ROC_AUC**	**Brier**
F0	0.661	0.679	0.623	0.650	0.741	0.216
F1	0.810	0.797	0.836	0.816	0.888	0.139
F2	0.785	0.778	0.803	0.790	0.880	0.144

Table 5: Random Forest model performance across feature sets (F0, F1, F2).

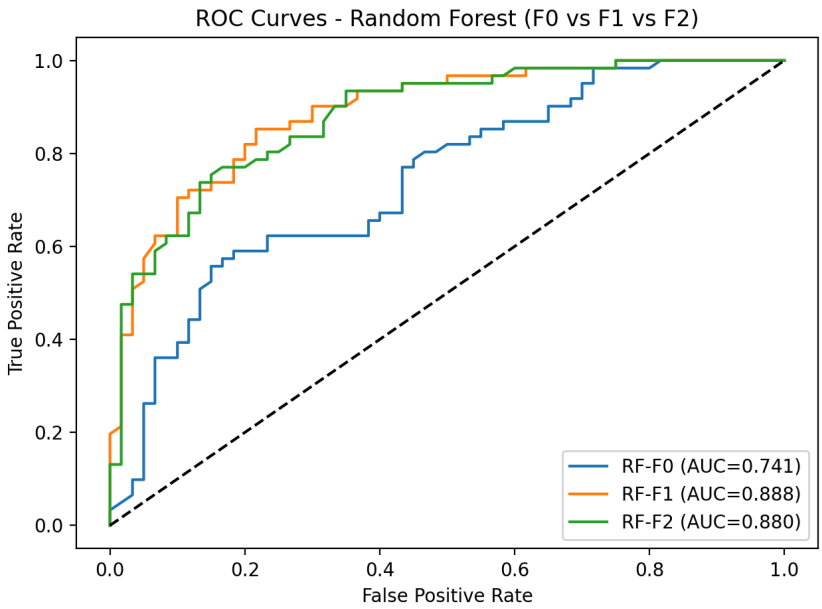

Figure 8: ROC Curves for the three models F0, F1, and F2 of Random Forest. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

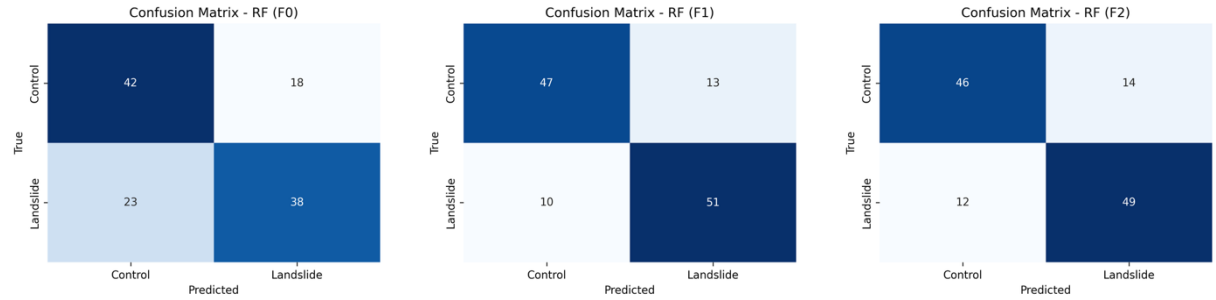

Figure 9-11: Confusion matrix of the three models F0, F1, and F2 of Random Forest.

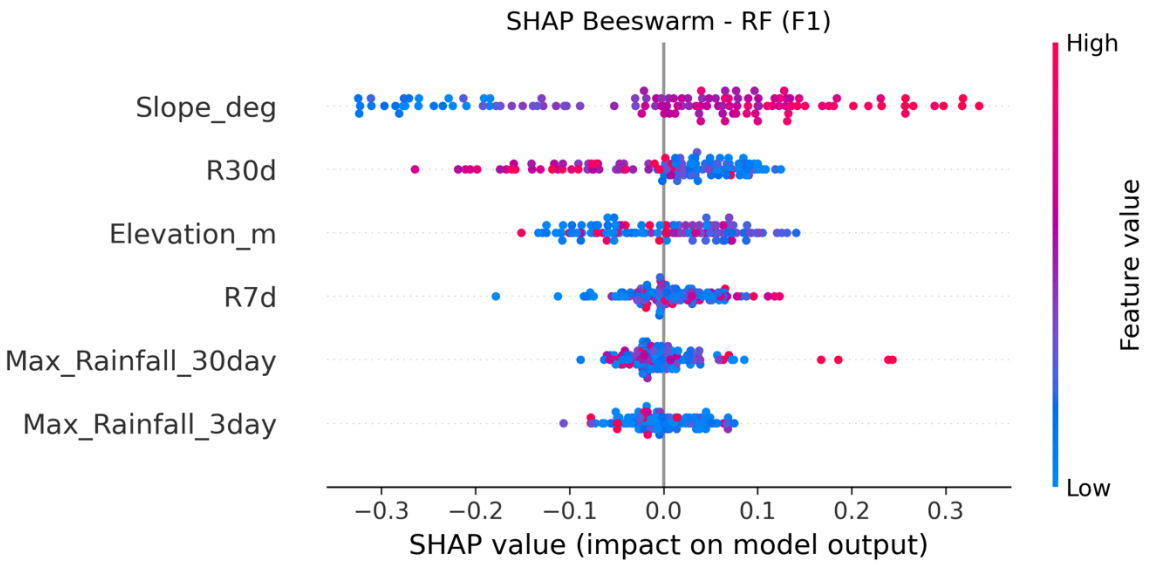

Figure 12: SHAP Beeswarm Chart for the F1 Random Forest Model.

The best-performing model, out of all three algorithms, was the Random Forest (RF) model. They were consistent across varying combinations of features (Table 5; Figure 8). Although the accuracy of the F0 model was 0.66 and the AUC 0.74—still notably higher than the accuracy of the Logistic Regression (LR) model—the F1 model setup, incorporating rainfall data, gave an accuracy of 0.81 and an AUC of 0.89, demonstrating the best performance of all. But the introduction of the soil feature in F2 detracted from the performance to an accuracy of 0.79 and an AUC of 0.88, departing from the trends of the LR. The slight performance decrease when soil depth is added (F1 to F2) may reflect the coarse spatial scale of SSURGO polygons relative to the 40-m DEM grid, introducing noise rather than meaningful stratification. Because most mapped soils in Buncombe County are uniformly deep (>200 cm), the limited variability may dilute stronger predictors such as slope and short-term rainfall.

Considering F1's superior performance, feature analysis was only conducted on the F1 model, focusing on SHAP figures (Figure 12). The figure confirmed that slope is the primary predictor, exerting the most influence on the model. Subsequently, it was followed by *R30d*, *elevation*, and short-term rainfall features such as *R7d*, *Max_Rainfall_30day*, and *Max_Rainfall_3day*. According to the SHAP beeswarm plot (Figure 12), slope indicated a positive relationship with the risk of landslides, while elevation showed a bidirectional influence. Short-term bursts of

rain, along with extreme maxima, exerted tremendously powerful effects in increasing the probability; conversely, cumulative rains over a 30-day period revealed a negative outcome, which diminished the likelihood of landslide occurrence.

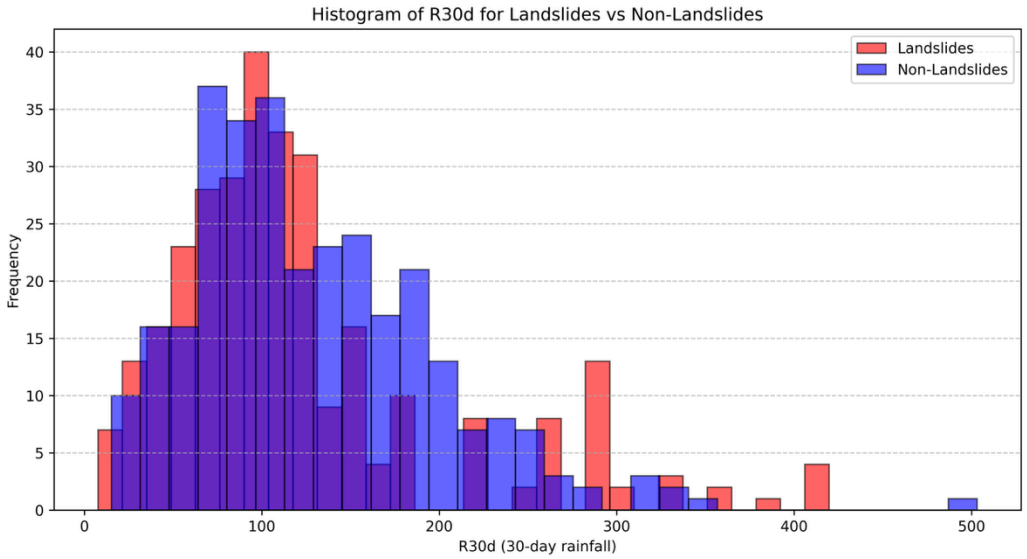

Figure 13: Distribution of 30-Day Cumulative Rainfall (R30d) for Landslide and Non-Landslide Events. The darker bars represent landslide events, and the lighter bars represent non-event controls. All rainfall values are measured in millimeters (mm).

As shown in the R30d histogram (Figure 13), both distributions of rainfall values were comparable, confirming that rainfall data were similarly distributed across both categories, thus strengthening the model's credibility. These findings underscore the capability of Random Forest to capture both anticipated and unexpected dynamics regarding landslide susceptibility. Intriguing is the discovery of the negative correlation between long-term cumulative rains and the model's predicted likelihood. This suggests that long-term rain accumulation, when not accompanied by brief extremes, may actually contribute to stabilizing the slope by facilitating gradual infiltration and percolation before pore pressurization persists long enough to initiate brief shallow failures. The results derived match past works citing landslide behavior on highly stepped landscapes as being mostly due to short-term high-magnitude events instead of being entirely the product of prolonged wet spells (e.g., Crozier, 2010; Tiranti & Rabuffetti, 2010; Bogaard & Greco, 2018). In addition, the decreased performance for F2 implies soil depth had

little bearing on enhancing predictive performance at the particular resolution utilized in this study. In conclusion, the results from applying the Random Forest model indicate that the landslide hazard in Buncombe is primarily caused by a combination of a highly stepped landscape on slope discontinuities and short periods of increased rainfall, rather than long periods of rainfall acting as an instantaneous trigger.

XGBoost

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.711	0.719	0.683	0.701	0.814	0.179
F1	0.802	0.757	0.883	0.815	0.879	0.141
F2	0.802	0.765	0.867	0.813	0.885	0.138

Table 6: XGBoost model performance across feature sets (F0, F1, F2).

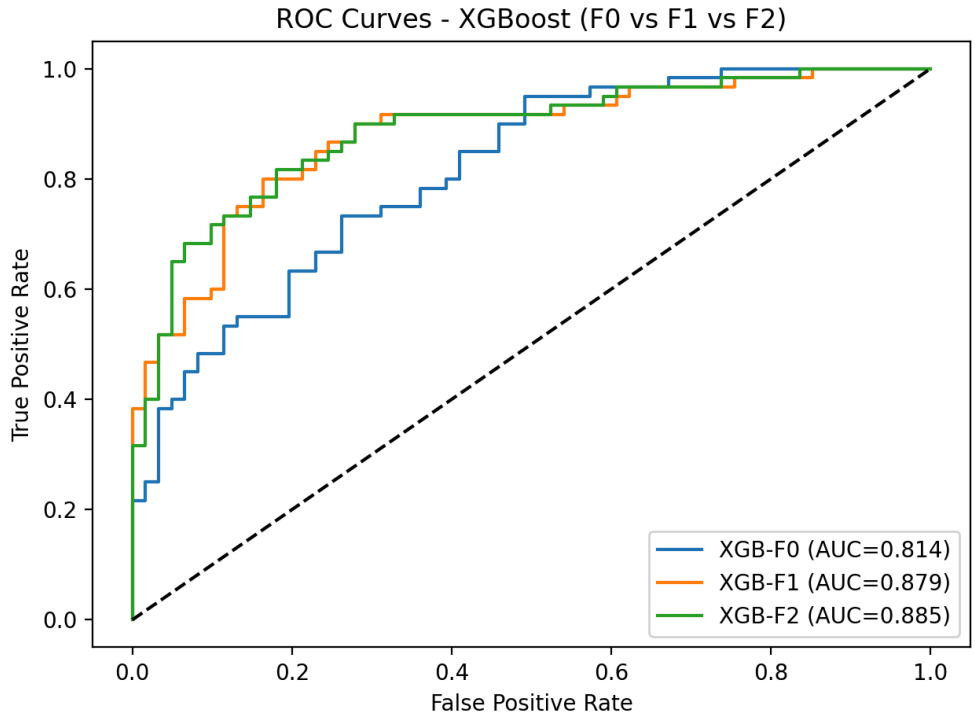

Figure 14: ROC curves of the three XGBoost models, F0, F1, and F2. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

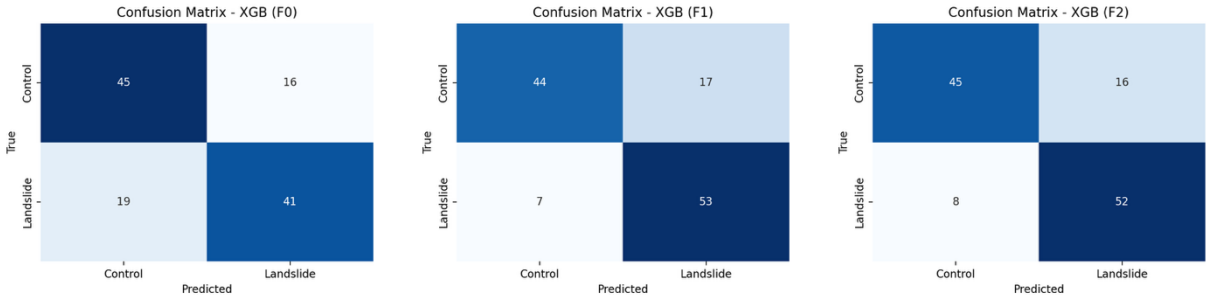

Figure 15-17: Confusion Matrix of the three models F0, F1, and F2 of XGBoost.

Surprisingly, XGBoost did not outperform the random forest model—instead, its F1 and F2 models showed similar prediction accuracy to the Random Forest Model (Table 6; Figure 14). The F0 XGBoost architecture excelled in model strength through an accuracy of 0.74 and an AUC of 0.81 in comparison to the other two F0 methods: Logistic Regression (LR) (accuracy 0.62, AUC 0.67) and Random Forest (RF) (accuracy 0.66, AUC 0.74). As expected, adding rainfall features improved the accuracy to 0.80 and AUC to 0.88, still better than the LR model, yet similar to the performance of the RF model (accuracy: 0.81, AUC: 0.89). Likewise, these trends appear to the full model (F2) equivalently. The full model (F2), which incorporated soil depth, produced the highest overall performance with an accuracy of 0.82 and an AUC of 0.89. The confusion matrix for F2 (Figure 17) highlights its balance, correctly classifying 45 controls and 52 landslides.

Compared to Logistic Regression and Random Forest, XGBoost delivered higher accuracy and consistently stronger AUC across feature sets, confirming its robustness in handling complex nonlinear relationships. The results reinforce that steep slopes and short-term rainfall extremes are primary triggers, while long-term rainfall can act as a dampening factor rather than a direct trigger (Crozier, 2010; Bogaard & Greco, 2018). The improved performance of the full F2 model also suggests that soil depth, while secondary to slope, provides additional predictive power in boosted frameworks.

Accuracy improved across all models when rainfall was added to terrain predictors. Logistic Regression rose modestly from 0.62 (F0) to 0.70 (F2), Random Forest increased more sharply from 0.66 (F0) to 0.81 (F1) but dropped slightly with soil depth (0.79, F2), while XGBoost achieved the highest accuracies at every stage, from 0.74 (F0) to 0.82 (F2).

Conclusion

This study developed a localized near-real-time landslide prediction framework for Buncombe County, North Carolina, using machine learning models and environmental predictors. By compiling a balanced dataset of 302 landslides and 302 non-events, rainfall windows from PRISM, topography from USGS DEMs, and soil depth from SSURGO, we trained Logistic Regression, Random Forest, and XGBoost classifiers. The pipeline incorporated stratified and spatial-block cross-validation, allowing us to rigorously compare models while preserving spatial representativeness. SHAP analysis and intrinsic feature importance made it possible to understand how predictors shaped landslide susceptibility in this mountainous region.

Across models, Random Forest consistently achieved the highest predictive performance (accuracy ≈ 0.82, AUC ≈ 0.89), followed by XGBoost and Logistic Regression. In all frameworks, slope emerged as the dominant predictor, confirming prior work showing terrain steepness as the primary control on landslide occurrence (Zheng et al., 2025; Xiao et al., 2024). Short-term rainfall accumulations (1–7 days, maxima over 3 days) were strongly associated with failures, consistent with global findings that intense short-duration rainfall is the principal trigger for shallow landslides (Crosta, 2004; Pennington et al., 2014; Barthélemy et al., 2024). A notable theoretical contribution was the identification of long-term cumulative rainfall (R30d) as a *negative* predictor in several models, suggesting that extended wet periods may sometimes stabilize slopes by allowing gradual drainage and infiltration rather than immediate pore-pressure buildup (Crosta, 2004; Fan et al., 2020). Soil depth contributed marginally, adding value only in boosted frameworks, which echoes recent findings that subsurface factors are often secondary to slope–rainfall interactions (Lee et al., 2023).

These results underscore both the practical and scientific value of localized machine-learning–based landslide modeling. The framework demonstrates that integrating rainfall windows with terrain predictors can generate near-real-time susceptibility maps, offering actionable information for hazard managers after storms like Hurricane Helene. At the same time, the findings reinforce a broader theoretical consensus: steep slopes combined with short-term rainfall extremes are the key drivers of shallow failures, while long-term rainfall plays a preparatory but

not directly triggering role (Pennington et al., 2014; Gariano & Guzzetti, 2016). Limitations of this study include a simplified binary soil depth representation, a relatively small sample size, and a lack of landslide-type differentiation. A limitation of this study is the use of PRISM daily rainfall, which cannot capture short-duration, high-intensity bursts that often trigger shallow landslides. These sub-daily intensities (e.g., mm/hr) are important in operational thresholds, but they are smoothed in daily grids. Future work could incorporate higher-temporal-resolution rainfall from radar products such as MRMS or satellite–gauge datasets like IMERG to better capture intense triggering events. Because the current framework is based on observed rainfall up to the event date, it functions as a near-nowcast rather than a true lead-time forecast; integrating short-term precipitation forecasts would be necessary to produce actionable early-warning lead times. Ultimately, coupling such models with rainfall forecasts could enable truly real-time landslide "nowcasting," a direction increasingly emphasized in both local and global hazard research (Stanley et al., 2021; Khan et al., 2022).

References

Akosah, S., Gratchev, I., Kim, D.-H., & Ohn, S.-Y. (2024). Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review. *Remote Sensing*, *16*(16), 2947. https://doi.org/10.3390/rs16162947

Allstadt, K., McBride, S. K., Godt, J., Slaughter, S., Baxstrom, K., Sobieszczyk, S., & Stull, A. (2025). *Preliminary field report of landslide hazards following Hurricane Helene* (Report Nos. 2025–1028; Open-File Report). USGS Publications Warehouse. https://doi.org/10.3133/ofr20251028

Aydin, A. (2006). Stability of saprolitic slopes: Nature and role of field scale heterogeneities. *Natural Hazards and Earth System Sciences*, *6*(1), 89–96. https://doi.org/10.5194/nhess-6-89-2006

Barthélemy, S., Bernardie, S., & Grandjean, G. (2025). Assessing rainfall threshold for shallow landslides triggering: A case study in the Alpes Maritimes region, France. *Natural Hazards*, *121*(4), 4023–4049. https://doi.org/10.1007/s11069-024-06941-2

Bauer, J., Fuemmeler, S., Wooten, R., Witt, A., Gillon, K., Douglas, T., Eberhardt, E., Froese, C., Turner, K., & Lerouell, S. (2012). Landslide hazard mapping in North Carolina–Overview and improvements to the program. *Landslides and Engineered Slopes: Protecting Society through Improved*

Understanding. *11th International Symposium on Landslides and 2nd North American Symposium on Landslides*, 257–263. https://www.researchgate.net/publication/260077615_Landslide_hazard_mapping_in_North_Carolina-Overview_and_improvements_to_the_program

Bogaard, T., & Greco, R. (2018). Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences*, *18*(1), 31–39. https://doi.org/10.5194/nhess-18-31-2018

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brier, G. W. (1950). VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. *Monthly Weather Review*, *78*(1), 1–3. https://doi.org/10.1175/1520-0493(1950)078%3C0001:VOFEIT%3E2.0.CO;2

Burrough, P., & McDonnell, R. (1998). *Principle of Geographic Information Systems*. https://www.researchgate.net/publication/37419765_Principle_of_Geographic_Information_Systems

Chan, H.-C., Chen, P.-A., & Lee, J.-T. (2018). Rainfall-Induced Landslide Susceptibility Using a Rainfall–Runoff Model and Logistic Regression. *Water*, 10(10), 1354. https://doi.org/10.3390/w10101354

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Crozier, M. J. (2010). Deciphering the effect of climate change on landslide activity: A review. *Geomorphology*, 124(3–4), 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009

Fan, L., Lehmann, P., Zheng, C., & Or, D. (2020). Rainfall Intensity Temporal Patterns Affect Shallow Landslide Triggering and Hazard Evolution. *Geophysical Research Letters*, 47(1), e2019GL085994. https://doi.org/10.1029/2019GL085994

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fuemmeler, S. J. (2008, April 10). *Landslide Hazard Mapping Methodology and Examples from North Carolina*. The Geological Society of America:

Southeastern Section - 57th Annual Meeting. https://gsa.confex.com/gsa/2008SE/webprogram/Paper136714.html?utm_source=chatgpt.com

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Hatcher, R. D. (2010). The Appalachian orogen: A brief summary. In R. P. Tollo, M. J. Bartholomew, J. P. Hibbard, & P. M. Karabinos, *From Rodinia to Pangea: The Lithotectonic Record of the Appalachian Region*. Geological Society of America. https://doi.org/10.1130/2010.1206(01)

Horn, B. K. P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1), 14–47. https://doi.org/10.1109/PROC.1981.11918

Jaboyedoff, M., Michoud, C., Derron, M.-H., Voumard, jérémie, Leibundgut, G., Sudmeier-Rieux, K., Nadim, F., & Leroi, E. (2016). *Human-Induced Landslides: Toward the analysis of anthropogenic changes of the slope environment* (pp. 217–232).

Kang, J., Wan, B., Gao, Z., Zhou, S., Chen, H., & Shen, H. (2024). Research on machine learning forecasting and early warning model for

rainfall-induced landslides in Yunnan province. *Scientific Reports*, 14(1), 14049. https://doi.org/10.1038/s41598-024-64679-0

Khan, S., Kirschbaum, D. B., Stanley, T. A., Amatya, P. M., & Emberson, R. A. (2022). Global Landslide Forecasting System for Hazard Assessment and Situational Awareness. *Frontiers in Earth Science*, *Volume* 10-2022. https://doi.org/10.3389/feart.2022.878996

Khashchevskaya, D., Owen, L. A., Wegmann, K., Scheip, C., & Figueiredo, P. M. (2025). The characteristics and timing of multiphase major landslides along the Blue Ridge Escarpment of the southern Appalachians revealed by combined cosmogenic nuclide dating and Schmidt hammer rebound measurements. *Geomorphology*, 485, 109857. https://doi.org/10.1016/j.geomorph.2025.109857

Kirschbaum, D. B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., & Lerner-Lam, A. (2012). Advances in landslide nowcasting: Evaluation of a global and regional modeling approach. *Environmental Earth Sciences*, 66(6), 1683–1696. https://doi.org/10.1007/s12665-011-0990-3

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. *Earth's Future*, 6(3), 505–523. https://doi.org/10.1002/2017EF000715

Klose, M., Damm, B., & Terhorst, B. (2015). Landslide cost modeling for transportation infrastructures: A methodological approach. *Landslides*, 12(2), 321–334. https://doi.org/10.1007/s10346-014-0481-1

Klose, M., Maurischat, P., & Damm, B. (2016). Landslide impacts in Germany: A historical and socioeconomic perspective. *Landslides*, 13(1), 183–199. https://doi.org/10.1007/s10346-015-0643-9

Lee, S., Oh, S., Ray, Ram. L., Lee, Y., & Choi, M. (2023). Three-dimensional hydrological thresholds to predict shallow landslides. *Terrestrial, Atmospheric and Oceanic Sciences*, 34(1), 20. https://doi.org/10.1007/s44195-023-00052-4

Lin, S., Chen, S., Rasanen, R. A., Zhao, Q., Chavan, V., Tang, W., Shanmugam, N., Allan, C., Braxtan, N., & Diemer, J. (2024). Landslide Prediction Validation in Western North Carolina After Hurricane Helene. *Geotechnics*, 4(4), 1259–1281. https://doi.org/10.3390/geotechnics4040064

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic Information Science and Systems*. Wiley. https://books.google.com/books?id=C_EwBgAAQBAJ

Lundberg, S., & Lee, S.-I. (2017). A *Unified Approach to Interpreting Model Predictions* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Mondini, A. C., Guzzetti, F., & Melillo, M. (2023). Deep learning forecast of rainfall-induced shallow landslides. *Nature Communications*, 14(1), 2466. https://doi.org/10.1038/s41467-023-38135-y

NASA Earth Observatory. (2021, June 10). *Machine Learning Model Doubles Accuracy of Global Landslide 'Nowcasts'—NASA*. https://www.nasa.gov/missions/gpm/machine-learning-model-doubles-accuracy-of-global-landslide-nowcasts/

National Hurricane Center. (2025). *Tropical Cyclone Report: Hurricane Helene* (AL092024) (p. 17). https://www.nhc.noaa.gov/data/tcr/AL092024_Helene.pdf

North Carolina Department of Environmental Quality. (2024). *North Carolina Landslide Inventory Points* [Dataset]. NC OneMap. https://www.nconemap.gov/datasets/01965a193482438cb70332e5e524e38b_0/about

Office of State Budget and Management. (2024). *Hurricane Helene Damage and Needs Assessment*. https://www.osbm.nc.gov/hurricane-helene-dna/open?utm_source=chatgpt.com

Pennington, C., Dijkstra, T., Lark, M., Dashwood, C., Harrison, A., & Freeborough, K. (2014). Antecedent Precipitation as a Potential Proxy for Landslide Incidence in South West United Kingdom. In K. Sassa, P. Canuti, & Y. Yin (Eds.), *Landslide Science for a Safer Geoenvironment* (pp. 253–259). Springer International Publishing. https://doi.org/10.1007/978-3-319-04999-1_34

Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, 40(10), 927–930. https://doi.org/10.1130/G33217.1

Petley, D. (2025, January 6). Global fatal landslides in 2024. *Eos*. https://eos.org/thelandslideblog/fatal-landslides-in-2024

PRISM Climate Group. (2024). *PRISM Daily Precipitation Data (1981–present)* [Dataset]. Oregon State University. https://prism.oregonstate.edu/

Regorda, A., Lardeaux, J.-M., Roda, M., Marotta, A. M., & Spalla, M. I. (2020). How many subductions in the Variscan orogeny? Insights from numerical models. *Geoscience Frontiers*, *11*(3), 1025–1052. https://doi.org/10.1016/j.gsf.2019.10.005

Sidle, R. C., & Ochiai, H. (2006). *Landslides: Processes, Prediction, and Land Use* (Vol. 18). American Geophysical Union. https://doi.org/10.1029/WM018

Sim, K. B., Lee, M. L., Rasa Remenyte-Prescott, & Soon Yee Wong. (2022, August). *An Overview of Causes of Landslides and Their Impact on Transport Networks*. Advances in modelling to improve network resilience, https://op.europa.eu/en/publication-detail/-/publication/c81e8bc9-1469-11ed-8fa0-01aa75ed71a1/language-en. https://www.researchgate.net/publication/369437132_An_Overview_of_Causes_of_Landslides_and_Their_Impact_on_Transport_Networks

Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M., Medwedeff, W., & Clark, M. K. (2021). Data-Driven Landslide Nowcasting at the Global Scale. *Frontiers in Earth Science*, 9, 640043. https://doi.org/10.3389/feart.2021.640043

Thomas, M. A., Collins, B. D., & Mirus, B. B. (2019). Assessing the Feasibility of Satellite-Based Thresholds for Hydrologically Driven Landsliding. *Water Resources Research*, 55(11), 9006–9023. https://doi.org/10.1029/2019WR025577

Tiranti, D., & Rabuffetti, D. (2010). Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation. *Landslides*, 7(4), 471–481. https://doi.org/10.1007/s10346-010-0198-8

U.S. Census Bureau. (2022). *Cartographic Boundary Shapefiles – Counties* [Dataset]. https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html

U.S. Department of Agriculture, Natural Resources Conservation Service. (2024). *Soil Survey Geographic* (SSURGO) *Database* [Dataset]. https://websoilsurvey.nrcs.usda.gov/app/

U.S. Geological Survey. (2021). *National Land Cover Database* (NLCD) 2021 [Dataset]. https://www.usgs.gov/centers/eros/science/national-land-cover-database

U.S. Geological Survey. (2022). *3D Elevation Program* (3DEP), *1/3 arc-second DEM seamless products* [Dataset]. https://www.sciencebase.gov/catalog/item/627f3798d34e3bef0c9a3198

Watterson, N. A., & Jones, J. A. (2006). Flood and debris flow interactions with roads promote the invasion of exotic plants along steep mountain streams, western Oregon. *Geomorphology*, 78(1–2), 107–123. https://doi.org/10.1016/j.geomorph.2006.01.019

Wooten, R. M., Gillon, K. A., Witt, A. C., Latham, R. S., Douglas, T. J., Bauer, J. B., Fuemmeler, S. J., & Lee, L. G. (2008). Geologic, geomorphic, and meteorological aspects of debris flows triggered by Hurricanes Frances and Ivan during September 2004 in the Southern Appalachian Mountains of Macon County, North Carolina (southeastern USA). *Landslides*, 5(1), 31–44. https://doi.org/10.1007/s10346-007-0109-9

Xiao, X., Zou, Y., Huang, J., Luo, X., Yang, L., Li, M., Yang, P., Ji, X., & Li, Y. (2024). An interpretable model for landslide susceptibility assessment based on Optuna hyperparameter optimization and Random Forest. *Geomatics, Natural Hazards and Risk*, 15(1), 2347421. https://doi.org/10.1080/19475705.2024.2347421

Ye, C., Wu, H., Oguchi, T., Tang, Y., Pei, X., & Wu, Y. (2025). Physically Based and Data-Driven Models for Landslide Susceptibility Assessment: Principles, Applications, and Challenges. *Remote Sensing*, 17(13), 2280. https://doi.org/10.3390/rs17132280

Zheng, W., Fan, W., Cao, Y., Nan, Y., & Jing, P. (2025). Landslide Hazard Assessment Under Record-Breaking Extreme Rainfall: Integration of SBAS-InSAR and Machine Learning Models. *Remote Sensing*, 17(13), 2265. https://doi.org/10.3390/rs17132265

Author Biography —

[Author name redacted by Managing Editor] is a student researcher at [school name redacted by Managing Editor], where he focuses on data-driven modeling of rainfall-triggered landslides and the development of real-time hazard prediction systems for mountainous regions in Western North Carolina. His most recent work, "A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina," integrates PRISM rainfall data, USGS Digital Elevation Models, and SSURGO soil depth information using machine-learning frameworks such as Logistic Regression, Random Forest, and XGBoost. The study explores how dynamic rainfall windows and antecedent soil-moisture conditions shape landslide susceptibility, offering a near-nowcasting approach for operational hazard monitoring.

Beyond this project, [Author name redacted by Managing Editor] has engaged in a broad range of quantitative and technical pursuits, including robotics programming, geospatial modeling, and environmental data analysis. His previous work includes coding real-time rainfall ingestion pipelines, building geospatial visualizations, and assisting peers in advanced mathematics and computer science as a tutor and leader.

[Author name redacted by Managing Editor] academic interests extend across environmental engineering, machine learning, and earth-systems science. He hopes to continue developing applied data-science tools that improve community resilience, especially in regions facing increased hydrological extremes due to climate change.

Acknowledgements

[Mentor name redacted by Managing Editor] provided valuable guidance during the formative stages of this research project. Her mentorship was instrumental in helping refine the initial research direction, shape the core questions, and clarify the conceptual framework connecting hydrologic thresholds, soil characteristics, and geomorphology. Through early discussions, she offered thoughtful feedback on how to situate the study within broader landslide-risk research and helped the author identify a focused and meaningful gap in existing modeling approaches. [Mentor name redacted by Managing Editor] also advised on interpreting long-term rainfall patterns and incorporating terrain-based predictors, ensuring that the project maintained coherence between its scientific objectives and methodological decisions. Her insights strengthened the logical flow of the manuscript, particularly

in the development of the introduction and the explanation of the environmental context.

[Mentor name redacted by Managing Editor] contributed additional guidance related to data acquisition and modeling design. His expertise in environmental prediction and machine-learning applications helped refine the selection and processing of key datasets, including PRISM rainfall windows, SSURGO soil depth, and the North Carolina landslide inventory. His feedback supported clearer alignment between the statistical methods and the goals of real-time hazard assessment.

The mentors' roles were limited to conceptual and structural guidance; all data processing, analysis, coding, interpretation, and writing were conducted independently by the author. Throughout the process, their support encouraged methodological precision, deeper critical thinking, and a clearer connection between the research objectives and final findings, providing an invaluable foundation for the study's development.

AI Disclosure

AI tools (ChatGPT) were used solely for grammar and formatting edits; all analysis and interpretation are original to the author.

A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina

Abstract

Rainfall-triggered landslides commonly occur across the Southern Appalachians, yet effective, localized “nowcasts” that fuse dynamic precipitation with terrain characteristics remain limited. We develop a near-real-time landslide risk model for Buncombe County, North Carolina, integrating daily PRISM precipitation windows with static hydrological data. The novel dataset containing 302 mapped landslides (1981–2024) was compiled by integrating USGS Digital Elevation Model (DEM) and Soil Survey Geographic Database (SSURGO) soil data set. Across models, slope was the dominant feature, while short-duration rainfall (1–7 days and 3-day maxima) most strongly increased landslide probability. In particular, 30-day cumulative rainfall often exhibited a negative association with predicted risk, suggesting that prolonged rainfall without intense surges may reduce immediate triggering. The resulting model enables county-scale, data-driven susceptibility updates based on forecasted precipitation data, offering a practical foundation for near-nowcasting and a landslide alarm system during and after intense storms.v

Keywords: Rainfall-triggered landslides, Nowcasts, Buncombe County, Random Forest, XGBoost, PRISM precipitation, Landslide susceptibility modeling

[Repository link redacted for review; will be included upon acceptance]
[Images that are under 300 dpi can be found in higher quality in this google docs]

Introduction

Landslides are devastating natural hazards that take place worldwide and cause wide-ranging damage to modern society, such as human casualties, economic losses, and infrastructure damage (Froude & Petley, 2018; Kirschbaum & Stanley, 2018; Petley, 2012). In the year 2024 alone, 766 instances of landslides took place around the world, killing 4,933 individuals (Petley, 2025). Landslides across the world incur an economic cost of US$20 billion annually, taking 17% of the yearly mean expenses caused by all natural disasters worldwide between 1980 and 2013 (Klose et al. 2016). Social infrastructures, especially transportation and communication, are highly susceptible to landslides (Sim, Lee, Remenyte-Prescott,

& Wong, 2022). Rural areas are likely to experience more impairments due to landslides, where these infrastructures are scarcer and scattered (Klose et al. 2015).

In search of effective landslide prediction methods, researchers have proposed various theories on the sources of these disasters, ranging from geological activity such as volcanic eruptions and earthquakes to human-induced sources such as modification of slopes and obstruction of hydrological flows (Sidle & Ochiai, 2006; Jaboyedoff et al., 2016). Among those factors, precipitation is regarded as the leading cause of landslides worldwide; extreme rainfall triggered 70% of all landslide events across the world between 2004 and 2010 (Petley, 2012). Fittingly, the role of precipitation in landslides will only continue to grow with climate change; Gariano and Guzzetti (2016) predict that landslides triggered by rainfall will be increasingly catastrophic and recurrent as climate change increases the frequency of short, intense storm episodes. The predicted increase in landslide prevalence forebodes the urgency of an effective rainfall-induced landslide prediction model.

The Western North Carolina region has historically been dominated by rainfall-induced shallow landslides due to geographical and meteorological factors (Wooten et al., 2008). The Blue Ridge Mountains, as a part of the Southern Appalachian mountain range, are characterized by steep slope gradients (Khashchevskaya et al., 2025). Its bedrock layer mostly comprises metamorphic and igneous rocks, which, when decomposed, create saprolite, which constitutes the top layer of the horizons (Hatcher, 2010; Watterson & Jones, 2006). The saprolite-heavy soil absorbs water easily, which increases pore pressure, resulting in a higher probability of shallow landslides (Aydin, 2006). The region's geomorphic susceptibility has been called for focus again after Hurricane Helene's extensive landslide damage in Western North Carolina in 2024 (Lin et al., 2024). Allstadt et al. (2025) state that the hurricane caused unprecedented geographic disruptions in the region, particularly in Buncombe, Henderson, Rutherford, and Yancey counties. Freshwater flooding contributed to 95 of the 176 direct deaths that Helene caused, most of which were landslide-related; around 2,000 independent cases of landslides were reported to be related to Helene, with the majority located over western North Carolina (National Hurricane Center, 2025). Helene significantly damaged the communications infrastructure in western North Carolina, considering WNC's rural characteristic, which exacerbated the situation (Office of State Budget and Management, 2024). Helene's unexpected landslide damage calls

for a real-time probabilistic forecasting model, localized for Western North Carolina.

At present, there is no near-nowcast model for Western North Carolina that captures dynamic changes in precipitation. In fact, nearly all of the current landslide hazard maps use minimal or no precipitation inputs. Specifically, the USGS Landslide Hazard Mapping Program's landslide susceptibility map identifies areas with elevated risk for landslips under extreme rainfall conditions (Fuemmeler et al., 2008); these, however, do not take account of real-time changes in precipitation since they were based on past correlations of landslides with geographical elements (Bauer et al., 2012). Academically, even though there have been numerous attempts at developing a physical/statistical model for the prediction of landslides (e.g., an ML-based model by Lin et al. (2024) that combines a variety of predictors such as the NC landslide database, soil surveys, digital elevation models, and water body locations), very few, if any, utilized dynamic rainfall as their primary predictor. To this end, an effective localized model of Western North Carolina for rainfall-induced landslides has proven necessary, as seen in the extensive damage Helene caused in the region.

Thus, to address the absence of a localized prediction model that incorporates dynamic rainfall variability, the proposed research will address the question: How does incorporating static geographic data (elevation and soil depth) with the variability of rainfall data affect the predictability of shallow landslides in the Western North Carolina region? Ultimately, this study hopes to create an efficient prediction model in near real-time (rainfall data are updated daily) for the area, developing an early warning system for residents of Western North Carolina, specifically, Buncombe County. Because the model relies on rainfall windows that end on the event day, it produces a near-nowcast rather than a true forecast with a defined lead time. Generating operational lead-time predictions would require integrating short-term rainfall forecasts, which is beyond the scope of the present study but represents a clear direction for future work.

Global models based on rainfall-triggered landslides have been developed and utilized for several years, with generally positive results. They are considered ‘nowcasts,’ because they incorporate both static geological characteristics and dynamic observations of rainfall to provide nearly real-time forecasts (Kirschbaum et al., 2012). For example, NASA's LHASA v2 model is the first operational global

system to utilize dynamic satellite rainfall data along with static geological and environmental characteristics (Stanley et al., 2021). It is based on a Logistic Regression algorithm and examines static predictor variables as well as IMERG precipitation (1 km resolution) to generate probabilistic hazard labels. The new model is purported to be two-fold more accurate than the previous LHASA v1 system, which utilized a standard threshold algorithm (Kirschbaum & Stanley, 2018; NASA Earth Observatory, 2021).

Only recently has the ~~the number of statistical models predicting landslides increased,~~ volume of published landslide susceptibility studies grown substantially, from 31 in 2016 to 219 in 2024 (Ye et al., 2025). Supervised machine learning (ML) was a popular choice for its ability to capture nonlinear relationships between predictor types and landslide probability, resulting in a higher level of accuracy than conventional linear statistical models (Regorda et al., 2020). Mondini et al. (2023) developed a time-dependent landslide model based on Italy. Exploiting the strength of Recurrent Neural Networks (RNNs) in sequential data, they used continuous rainfall data for a window of 30 days, excluding geological or environmental factors. Similarly, Chan et al. (2018) built a logistic regression-based model aimed at predicting landslides caused by a typhoon's heavy rain in the southern region of Taiwan. The model mainly used runoff flow depth, not pure precipitation data, taking into account soil saturation, which reached an accuracy of 80~85%. Kang et al. (2024) recently constructed a localized model for Yunnan Province, China, with its Random Forest model reaching an accuracy level of 0.906. All the above ML-based models center around rainfall data. Their high accuracy levels affirm the need for dynamic rainfall data in a localized landslide model.

Focusing on North America, Thomas et al. (2019) challenged if satellite rainfall data can replace in situ hydrological data to evaluate the soil saturation threshold for a slope failure, reflecting a recent trend of increasing remote geospatial data-based models (Akosah et al., 2024). The findings from this localized model of California demonstrated that rainfall (specifically satellite-measured rainfall data) cannot be relied upon to predict landslides and encourage a hydrogeological gauge that calibrates rainfall data with on-site hydrologic data to develop a soil wetness index. Other studies that support the aforementioned findings include a comprehensive model (Lee et al., 2023) blending precipitation duration/intensity and normalized soil moisture capacity, which dropped false alarms (FA) from ~26 to

3. Both studies emphasize the importance of contextualizing raw rainfall data with local hydrologic features in evaluating the soil saturation level.

The contribution of this study is twofold. We build a localized, near-real-time landslide prediction model. For the modeling, an equal number of non-events were randomly chosen based on the landslide inventory of Buncombe County (North Carolina Department of Environmental Quality, 2024). To avoid pattern biasing in specific constellations of events, stratification and a spatial block cross-validation approach were utilized. Three different ML algorithms were tested overall, with three models per algorithm (F0, F1, and F2) based on feature type. The assessment of performance and feature analysis was interpreted using evaluation metrics, SHAP plots, and permutation importance.

To fit and evaluate the model, we designed a hydrologic/environmental dataset that collates daily precipitation data (PRISM Climate Group, 2024), a topography map (USGS DEM), and soil depth (USDA SSURGO). The dataset is open and available for further research at the specified repository.

The next section describes the datasets and methodology of the research. Section 3 illustrates the results of the findings. Section 4 delivers the conclusion of the research.

Data and Methodology

Dataset

This research aimed to create a dynamic prediction model localized in Buncombe County, North Carolina, geographically located within the Blue Ridge Mountains (Figure 2a). The inventory of landslide events was extracted from the North Carolina Landslide Points dataset (via NC OneMap; North Carolina Department of Environmental Quality, 2024). The dataset included event geometry (points) and traceability fields: *IsLandslide* (event flag), *Event_Date*, *Sort_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, and *Source_Period*. The inventory was both a temporal anchor (*Event_Date*) for rainfall-window calculations referred to by dates and a spatial anchor (X, Y) for topography and soil data extraction. Additionally, an equal number of non-event samples were created for control, randomly selecting the same time (January 1981–December 2021) and same region

of interest (Buncombe County) as those in the landslide dataset. Ultimately, equal numbers of events and non-events were included in the final dataset used for modeling.

Gridded precipitation (4 km resolution) was extracted from PRISM daily products (PRISM Climate Group, 2024) throughout the entire analysis period (spanning from January 1981 to December 2024). Using the reference dates of the events and non-events—the event date (*Event_Date*) or the control (*Random_Date*)—historical sums over windows of 1-day (*R1d*), 3-days (*R3d*), 7-days (*R7d*), and 30-days (*R30d*) were computed. Other temporal windows included maximum rainfall sums calculated for both 3-day (*Max_Rainfall_3day*) and 30-day (*Max_Rainfall_30day*) intervals. All precipitation measurements were in units of millimeters, and all the calculation windows' end dates coincided with the reference date of the observation (Table 1 contains the definition and sources of all features).

Topography was taken from a DEM (Digital Elevation Model) encompassing Buncombe County published by the U.S. Geological Survey (2022). The DEM already provided elevation (*Elevation_m*), while slope (*Slope_deg*) was computed from Horn's (1981) 3×3 finite difference gradients and converted into degrees (Figure 2; Equations 1, 2, and 3).

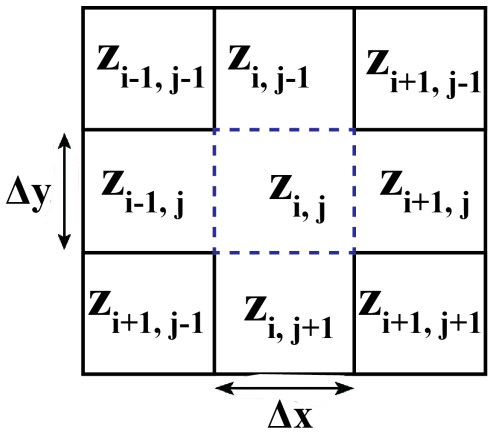

Figure 1. Visualization of Horn's (1981) 3×3 finite difference gradient calculator. $Z_{i,j}$ symbolizes the position of the elevation of a particular cell. Δx and Δy are the horizontal and vertical grid distances of the cells.

$$p = [(z_{i+1,j-1} + 2z_{i+1,j+1} - z_{i-1,j-1}) - (z_{i+1,j+1} + 2z_{i-1,j} + z_{i-1,j+1})]/8\Delta x \quad (1)$$

$$q = [(z_{i-1,j+1} + 2z_{i,j+1} - z_{i+1,j+1}) - (z_{i-1,j-1} + 2z_{i-1,j} + z_{i-1,j-1})]/8\Delta y \quad (2)$$

$$Slope = tan^{-1}(\sqrt{p^2 + q^2}) \quad (3)$$

The values p and q in Equations 1 to 3 represent the rate of elevation change in the east-west direction and the north-south direction, respectively. The DEM cell resolution for Buncombe County showed an equal width and height of 40 meters. To avoid unit discrepancy across different datasets, all datasets were under a consistent projected coordinate reference system (NAD83/North Carolina; EPSG:32119) and maintained a resolution of 40m.

Soil data originated from USDA NRCS Soil Survey Geographic Database (SSURGO) map units (U.S. Department of Agriculture, Natural Resources Conservation Service, 2024). A point-polygon comparison was used to match MUKEY and MUSYM to each sample point, linking depth information from the related tables. Soil depth was converted to a numeric variable (*Soil_Depth_cm*) by extracting data from raw strings (*Soil_Depth_cm_raw*) and converting the variable into a binary flag (*Soil_Depth_Deep200_Flag*), where entries larger than 200 were changed to 1, while any other entries were changed to 0. These soil properties were viewed as static covariates that reflect conditions of regolith associated with instability triggered by rainfall.

For map compilation and masking, administrative boundaries by county were employed to clip rasters and to define the sampling window for controls with a modest buffer to attenuate edge effects (U.S. Census Bureau, 2022). Land cover from NLCD was retained for descriptive mapping and voluntary sensitivity tests but was not included in the baseline prediction set (U.S. Geological Survey, 2021).

Gridded variables such as rainfall windows, elevation, and slope were sampled by bilinear interpolation at coordinates defined by points (X, Y) (Burrough & McDonnell, 1998). Soil data, specifically soil depth, came from the MUKEY and MUSYM area codes. Units of measurement have been standardized with precipitation in millimeters, elevation in meters, and slope measured in degrees. We

verified the rainfall windows to check that the end dates coincide with the event reference date. Data instances with missing values were removed from the dataset.

The dataset created is a linked dataset, combining information from PRISM daily rainfall data, the USGS DEM, and the USDA SSURGO map specific to Buncombe county, and can be used for further analyses related to landslide risk. Linking information from the various datasets was not trivial, as it required gathering chronological information and summarizing it for the selected data points before linking it to the main dataset (Longley et al., 2015; Burrough & McDonnell, 1998). Other challenges included incompatible reference systems that had to be reconciled. The complete dataset includes the outcome variable (*IsLandslide*) and full fields for past rainfall, topography, and soil depth with comprehensive metadata for provenance and audit purposes, which comprises *Event_Date/Random_Date*, X, Y, *County*, *GlobalID*, *OBJECTID*, *Data_Type*, *Source_Period*, *MUKEY*, *MUSYM*, and *Soil_Depth_cm_raw*.

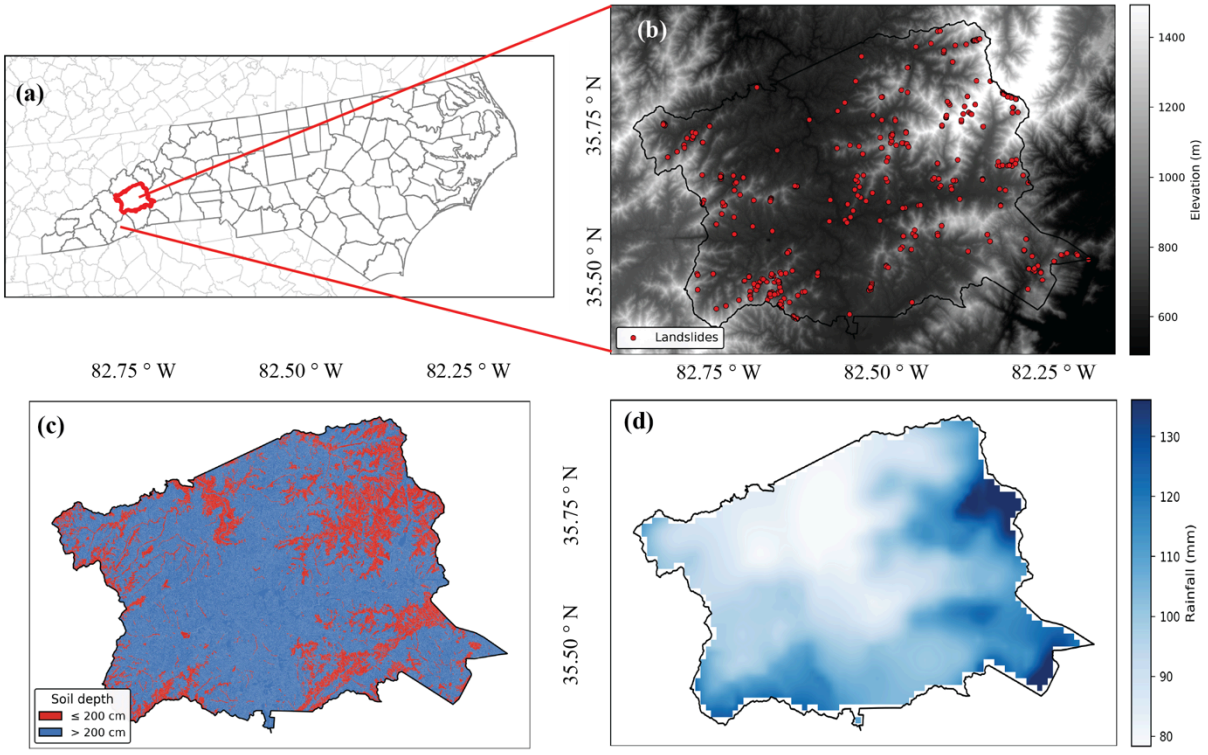

Figure 2: (a) Geographical location of the study area, (b) elevation, (c) soil depth, (d) 30 year average rainfall.

Symbol (unit)	How it was calculated	Source
R1d (mm)	Sum of daily PRISM rainfall for 1 day	PRISM (800m)
R3d (mm)	Sum of daily PRISM rainfall for 3 days	PRISM (800m)
R7d (mm)	Sum of daily PRISM rainfall for 7 days	PRISM (800m)
R30d (mm)	Sum of daily PRISM rainfall for 30 days	PRISM (800m)
Max_Rainfall_3day (mm)	Rolling maximum of 3-day totals in past 30 days	PRISM (800m)
Max_Rainfall_30day (mm)	Rolling maximum of 30-day totals in past 90 days	PRISM (800m)
Elevation_m (m)	Extracted directly from DEM	USGS DEM (40m)
Slope_deg (°)	Derived from DEM using slope algorithm	USGS DEM (40m)
Soil_Depth_Deep200_Flag (—)	Flag for soils deeper than 200 cm	SSURGO Soil

Table 1: Features definitions (Rainfall time-frames are from the reference point of the corresponding event)

All data collection, processing and model development steps were done in Python 3.9 in the PyCharm IDE. Libraries, including NumPy, Pandas, SciPy, Matplotlib, and Seaborn, were used for scientific computing; GeoPandas, Shapely, and Rasterio were used for data processing; and scikit-learn and XGBoost models combined with SHAP were chosen for model development and assessment. The models were run on a Mac Apple M2 processor.

Modeling

To assess which rainfall variables specifically influenced landslide probability, the temporal accumulation/maximum rainfall predictors were analyzed with all non-empty possible combinations. A primary concern was minimizing overfitting due to the large number of rainfall predictors (*R1d*, *R3d*, *R7d*, *R30d*, *Max_Rainfall_3day*, *Max_Rainfall_30day*). Thus, a training–test split was implemented via a stratified five-fold cross-validation approach using a constant

seed and data point shuffling. In addition, the evaluation was repeated for every rainfall predictor subset. The model that showed the most accurate predictions on the test data was included in the final comparison across other types of models. The above processes were independently conducted for each machine learning (ML) algorithm model.

Three separate ML algorithms were employed for the actual study—Logistic Regression, Random Forest and XGBoost (Breiman, 2001; Chen & Guestrin, 2016). Due to logistic regression's simple and linear design, the model was effective in serving as a benchmark to compare with nonlinear models. Receiver Operating Characteristic (ROC) curves plot the true positive rate against the false positive rate across probability thresholds. The Area Under the Curve (AUC) summarizes model discrimination. The diagonal dashed line represents the no-skill baseline, where the model performs no better than random chance. The performance of the model was quantified through ROC curves, confusion tables, and summary measures: accuracy, precision, recall, ROC–AUC, and the Brier score (Brier, 1950). For logistic regression, the coefficient magnitudes summarized how each variable influenced the model prediction, while standard errors, p-values and confidence intervals measured the degree of statistical significance of that prediction.

A Random Forest algorithm was used as the baseline nonlinear model. Interpretation of the model involved a combination of intrinsic feature importance as well as SHAP analysis. Intrinsic feature importance was determined as the mean decrease in impurity, indicating the effectiveness of each predictor in splitting the data. Shapley values for each variable and data point were illustrated via SHAP bar plots, SHAP beeswarm plots and SHAP waterfall plots per sample (Lundberg & Lee, 2017). The bar plot visualizes the average contribution of each feature to the overall prediction, while the beeswarm plots show both the distribution as well as the direction of the predictor contributions. SHAP waterfall plots give local explanations regarding the most likely landslide samples. A full dataset, including SHAP values as well as both the imputed and the original feature values, expected values, probabilities, and true labels, in addition to original metadata, was also provided so that both global and case-specific insights could be obtained. Model discrimination as well as calibration were also assessed as part of ROC curves and confusion matrices, as well as the same summary metrics that were used as part of Logistic Regression.

In answering the hypothesis, three separate models were created for each type of ML algorithm. Model F0 used just slope and elevation, aiming to assess the predictive potential based on the static predictor: terrain alone. Model F1 added rainfall accumulation intervals, alongside slope and elevation, as the basis to also evaluate the short- as well as longer-term rainfall triggers. Following on from F1, Model F2 also included soil depth so that the effect of the subsurface interacting with both rainfall and the terrain factors could be determined. Each algorithm was then trained across the three feature settings, ultimately leaving us with parallel models that allowed a direct comparison.

The F0, F1, and F2 configurations for each model (Logistic Regression, Random Forest, and XGBoost) were trained and tested on the same sample and therefore any discrepancies in results were not driven by sampling but rather variable differences.

Comparison analysis conducted on XGBoost involved classification accuracy, showing ROC curves and confusion matrices, as well as the complete set of metrics, including accuracy, precision, recall, F1 score, and ROC–AUC, as well as the Brier score. Interpretable techniques, such as SHAP dependence plots or permutation importance, are only presented for the Random Forest model.

Results and Discussion

The compiled landslide dataset contained 9,092 events recorded within North Carolina between 1940 and 2024, of which 398 occurred within the study area of Buncombe County (Figure 3). Because the PRISM Weather daily rainfall data before 1981 was unavailable for the public, the inventory was limited to the time frame between 1981 and 2024, limiting the data points to 302 landslide events. Corresponding to these landslide points, an equal number of control points were created by sampling sites and dates randomly within the same county and time frame. As such, the final dataset contained 604 total entries, split evenly between events and controls, thus reducing potential biases for certain variables and increasing its robustness toward new data.

All the records had a collection of hydrological covariates. Rainy features comprised both cumulative windows ($R1d$, $R3d$, $R7d$, $R30d$) as well as highest

intensities (*Max_Rainfall_3day*, *Max_Rainfall_30day*). Geological features included elevation as well as slope at 40 m resolution. Soil depth was included as an added predictor expressed as a binary flag (>200 cm vs. ≤200 cm), as the vast majority of the sites had extremely deep soils.

The final dataset merged the rainfall, terrain, and soil characteristics for each observation. Its well-balanced design (302 events + 302 controls) ensured the risk of class bias was kept small, and the range of covariate variability enabled the models to incorporate long-term conditioning factors (e.g., slope, soil depth) as well as short-term triggers (e.g., extreme rains).

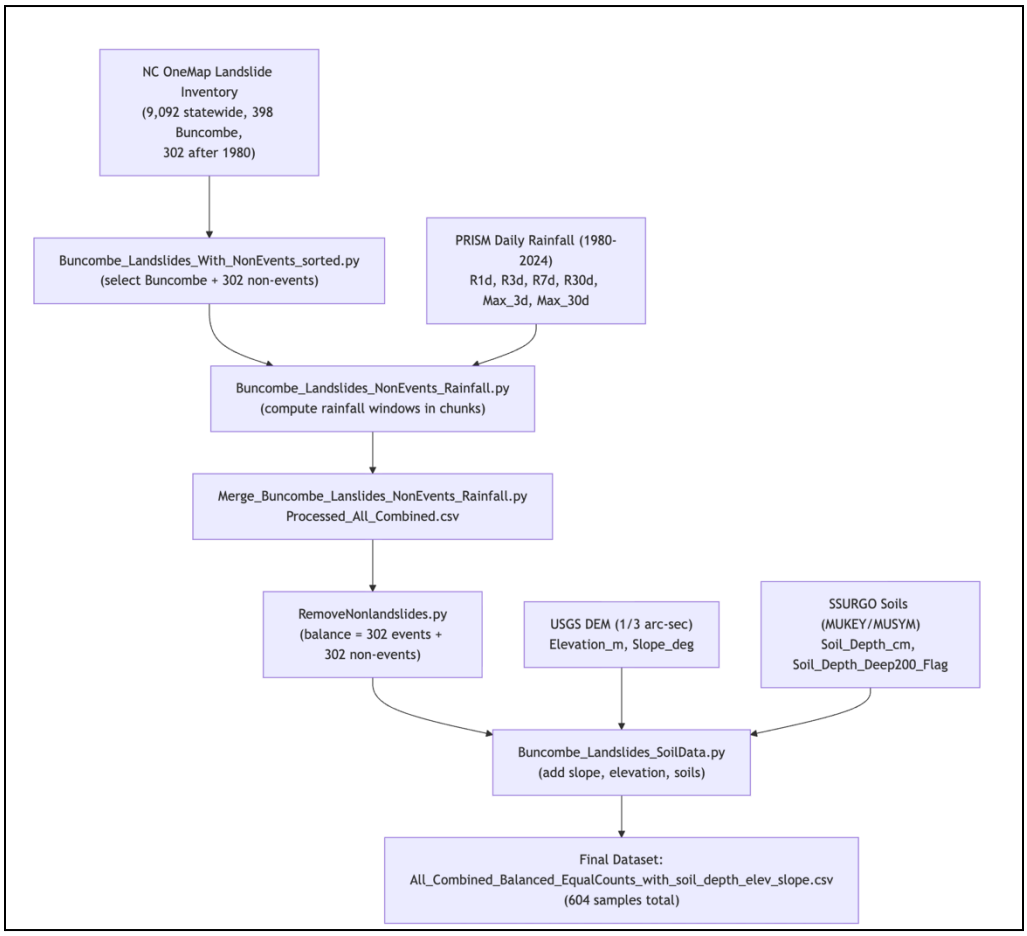

Figure 3: Flowchart of landslide events and non-events data processing

Feature selection:

Feature selection indicated consistency across the nonlinear and the linear models. In the Logistic Regression category, the optimum subset of rainfall data comprised *Max_Rainfall_30day*, *R3d*, *R7d*, and *R30d*, achieving an accuracy of 0.593

as well as an ROC–AUC of 0.614 (see Table 2). Short-period rainfall accumulation, including three-day as well as seven-day buildup, was frequently shown in the top-ranking combinations, showing how these predictors have a strong relationship with landslide risks. However, long-term predictors such as *Max_Rainfall_30day* and *R30d* also showed up occasionally, albeit not as frequently as short-term variables.

Logistic Regression (Top 5)	Accuracy	ROC–AUC
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.593	0.614
Max_Rainfall_30day, *Max_Rainfall_3day*, *R3d*, *R7d*	0.589	0.621
Max_Rainfall_30day, *R1d*, *R3d*, *R7d*, *R30d*	0.589	0.613
Max_Rainfall_30day, *R7d*, *R30d*	0.584	0.600
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.584	0.619

Table 2: Top-performing subsets of Rainfall Features for Logistic Regression

	Accuracy	ROC–AUC
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*	0.808	0.888
Max_Rainfall_30day, *R1d*, *R7d*, *R30d*	0.806	0.882
Max_Rainfall_30day, *R3d*, *R7d*, *R30d*	0.805	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R3d*, *R7d*	0.801	0.890
Max_Rainfall_30day, *Max_Rainfall_3day*, *R1d*, *R7d*, *R30d*	0.801	0.883

Table 3: Top-performing subsets of Rainfall Features for Random Forest

In the Random Forest output, the best-performing subset included *Max_Rainfall_30day*, *Max_Rainfall_3day*, *R1d*, and *R3d*, yielding exceptional accuracy up to 0.808 and ROC-AUC reaching as high as 0.888 (see Table 3). Moreover, subsets including R3d and R7d also appeared among the very best, a point also supported by the results yielded by the Logistic Regression (see Table 2).

However, the Random Forest model also brought to light that short-term rainfall predictors like *R1d* as well as the *Max_Rainfall_3day* significantly improved classification, bringing to the forefront nonlinear interactions. These findings indicated that although both nonlinear and linear prototypes invariably recognized the paramount significance of short-term rainfall, the Random Forest technique, in addition, captured the additional nuances of the lengths of the rainfall affecting the probabilities of the landslide occurrences (refer to Table 3).

Actual Models:

Logistic Regression:

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.620	0.609	0.650	0.629	0.673	0.229
F1	0.702	0.688	0.733	0.710	0.771	0.193
F2	0.702	0.682	0.750	0.714	0.775	0.191

Table 4: Logistic Regression model performance across feature sets (F0, F1, F2).

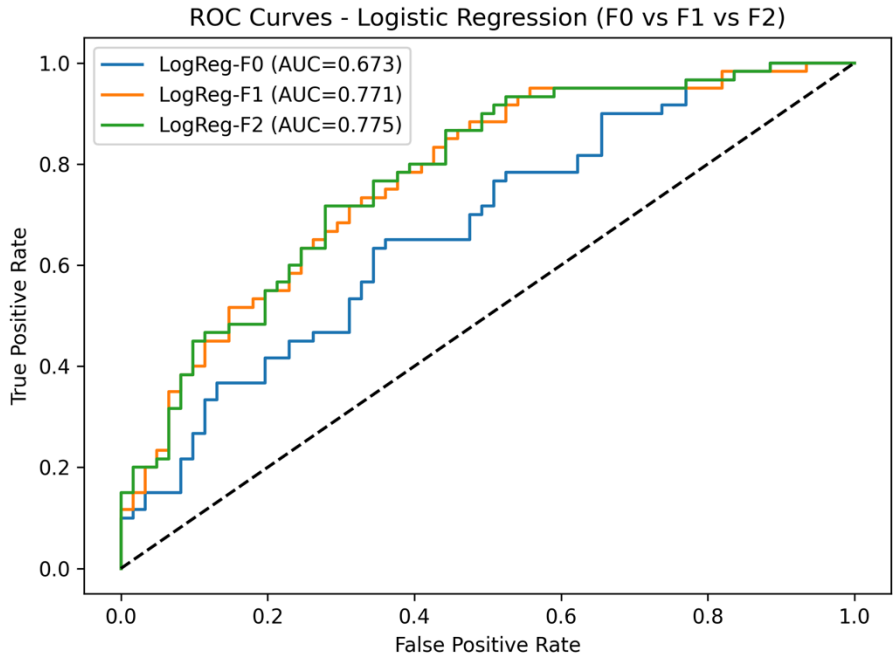

Figure 4: ROC Curves for the three Logistic Regression models F0, F1, and F2. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

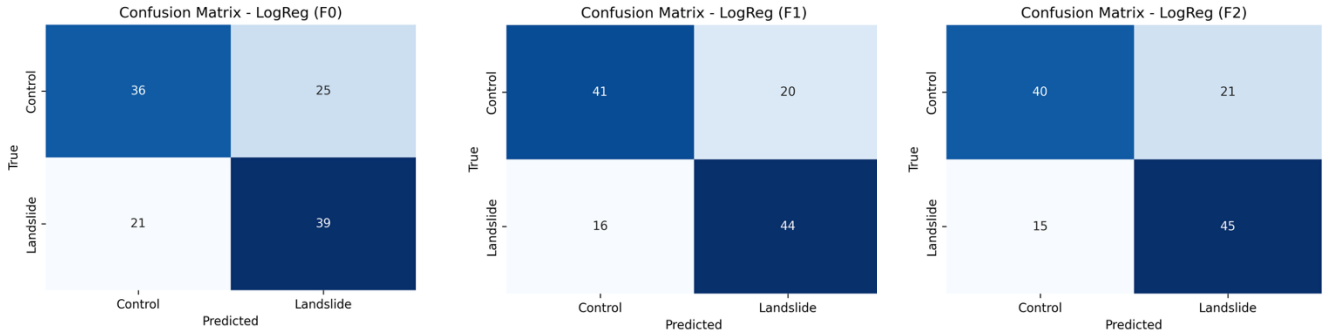

Figure 5-7: Confusion Matrices of F0, F1, F2 Logistic Regression Models

Feature	**F0**	**F1**	**F2**
Slope_deg	0.7150	1.3574	1.3813
Max_Rainfall_30day	-0.7287	0.5666	0.5871
R30d	0.2208	-0.5968	-0.6327
R7d	-0.1018	0.2192	0.2136
Elevation_m	-0.0287	-0.1886	-0.1099
R3d		-0.0571	-0.0620
Soil_Depth_Deep200_Flag			0.1943

Table 5: Regression coefficients for each feature of F0, F1, F2 Logistic Regression Models. Note that rainfall-related coefficients are intentionally blank for the F0 model, as F0 includes only static terrain predictors (slope and elevation) and therefore does not estimate coefficients for any rainfall variables.

Performance for Logistic Regression got progressively better as more predictors were added (Table 4). In the terrain-only version F0, accuracy was 0.62 with an AUC (Area Under Curve evaluates how well the model distinguishes between cases and controls) of 0.67. Inclusion of the rainfall predictors in F1 got the accuracy up to 0.70 and the AUC up to 0.77 (Figure 4). Inclusion of the full set of predictors in the complete version F2 got the best accuracy at 0.70 as well as the best AUC at 0.78. The confusion matrices further corroborated this ranking of models (Figures 5-7): F2 got 40 correct controls as well as 45 correct landslides,

compared to 36 and 39 cases for the F0 model and 41 and 44 cases for the F1 model respectively.

No formal statistical test (e.g., McNemar's test or bootstrap confidence intervals) was conducted to evaluate whether these accuracy differences are statistically significant, and the improvements should therefore be interpreted cautiously.

Model interpretation relied on the regression coefficients (Table 5). The slope was the strongest predictor variable, followed closely by *Max_Rainfall_30day*, both having a positive effect on the risk of landslides, except for the F0 model. R30d represents the total cumulative rainfall in the 30 days preceding the event, whereas Max_Rainfall_30day represents the single wettest 30-day period within the past 90 days. These variables capture different hydrologic processes: R30d reflects gradual wetness buildup, while Max_Rainfall_30day captures past extreme episodes. Their differing definitions explain why one may be positive and the other negative in regression coefficients. Here, the cumulative 30-day rainfall (*R30d*) tended to have a negative coefficient, meaning long-term accumulation did not significantly increase the probability of slope failure after accounting for short-term accumulation or extreme events. These results strengthen the conclusion that the best explanation for landslides lies in the interaction between the region's steep slope and short-term, high-rate bursts of rain, shown in the Brier Scores of Table 4.

Random Forest:

FeatureSet	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**ROC_AUC**	**Brier**
F0	0.661	0.679	0.623	0.650	0.741	0.216
F1	0.810	0.797	0.836	0.816	0.888	0.139
F2	0.785	0.778	0.803	0.790	0.880	0.144

Table 5: Random Forest model performance across feature sets (F0, F1, F2).

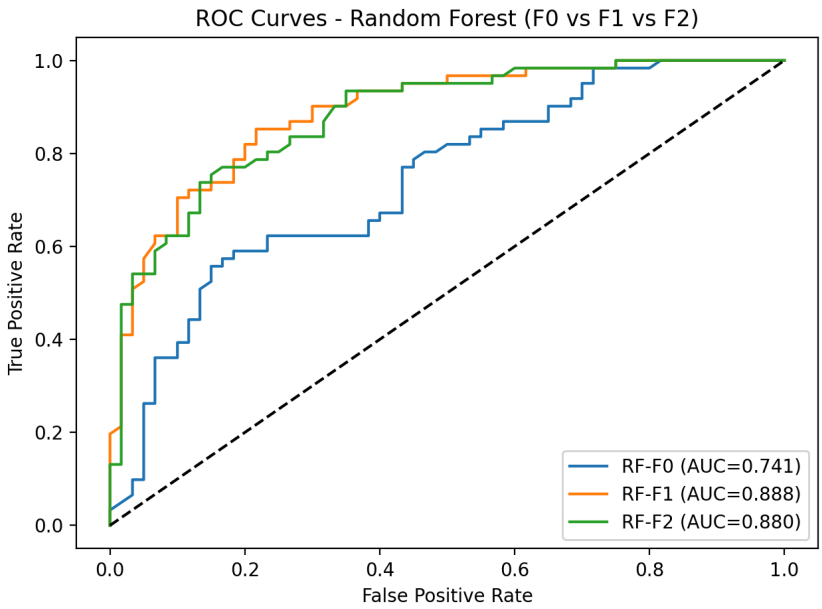

Figure 8: ROC Curves for the three models F0, F1, and F2 of Random Forest. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

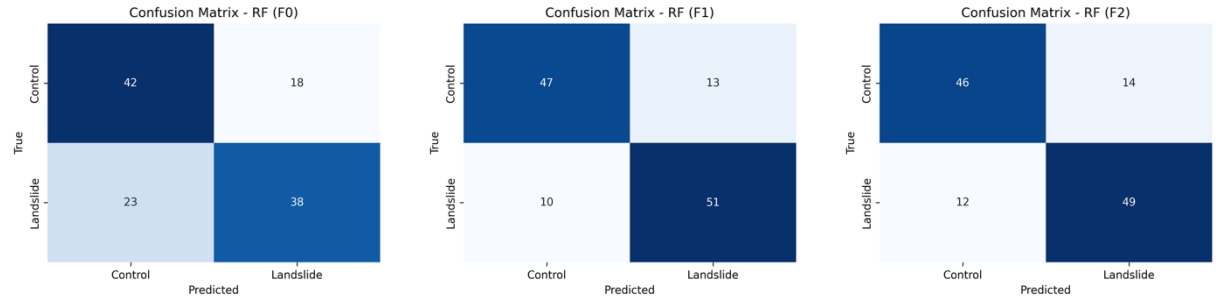

Figure 9-11: Confusion matrix of the three models F0, F1, and F2 of Random Forest.

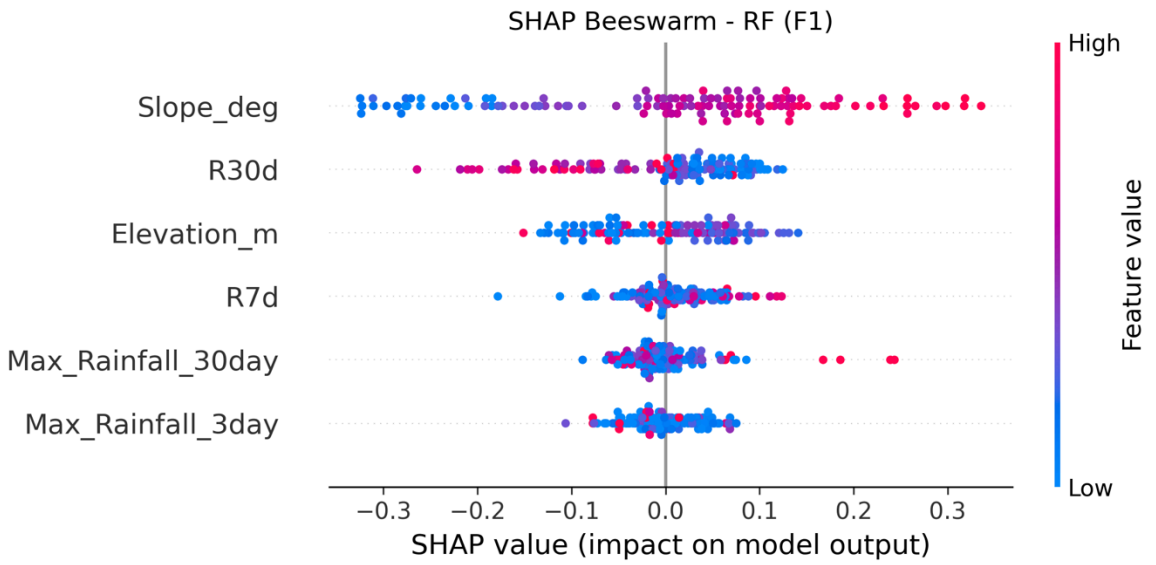

Figure 12: SHAP Beeswarm Chart for the F1 Random Forest Model.

The best-performing model, out of all three algorithms, was the Random Forest (RF) model. They were consistent across varying combinations of features (Table 5; Figure 8). Although the accuracy of the F0 model was 0.66 and the AUC 0.74—still notably higher than the accuracy of the Logistic Regression (LR) model—the F1 model setup, incorporating rainfall data, gave an accuracy of 0.81 and an AUC of 0.89, demonstrating the best performance of all. But the introduction of the soil feature in F2 detracted from the performance to an accuracy of 0.79 and an AUC of 0.88, departing from the trends of the LR. The slight performance decrease when soil depth is added (F1 to F2) may reflect the coarse spatial scale of SSURGO polygons relative to the 40-m DEM grid, introducing noise rather than meaningful stratification. Because most mapped soils in Buncombe County are uniformly deep (>200 cm), the limited variability may dilute stronger predictors such as slope and short-term rainfall.

Considering F1's superior performance, feature analysis was only conducted on the F1 model, focusing on SHAP figures (Figure 12). The figure confirmed that slope is the primary predictor, exerting the most influence on the model. Subsequently, it was followed by *R30d*, *elevation*, and short-term rainfall features such as *R7d*, *Max_Rainfall_30day*, and *Max_Rainfall_3day*. According to the SHAP beeswarm plot (Figure 12), slope indicated a positive relationship with the risk of landslides, while elevation showed a bidirectional influence. Short-term bursts of

rain, along with extreme maxima, exerted tremendously powerful effects in increasing the probability; conversely, cumulative rains over a 30-day period revealed a negative outcome, which diminished the likelihood of landslide occurrence.

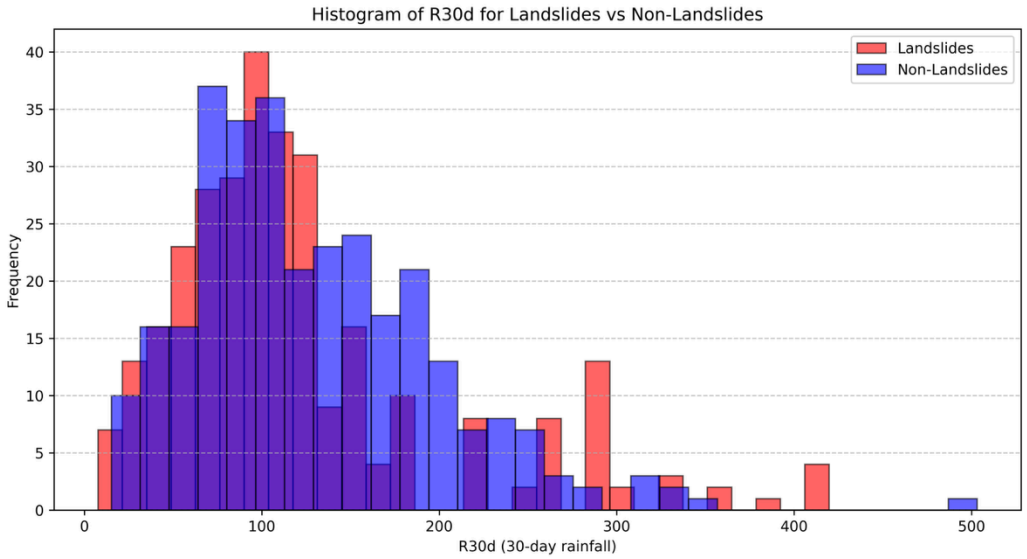

Figure 13: Distribution of 30-Day Cumulative Rainfall (R30d) for Landslide and Non-Landslide Events. The darker bars represent landslide events, and the lighter bars represent non-event controls. All rainfall values are measured in millimeters (mm).

As shown in the R30d histogram (Figure 13), both distributions of rainfall values were comparable, confirming that rainfall data were similarly distributed across both categories, thus strengthening the model's credibility. These findings underscore the capability of Random Forest to capture both anticipated and unexpected dynamics regarding landslide susceptibility. Intriguing is the discovery of the negative correlation between long-term cumulative rains and the model's predicted likelihood. This suggests that long-term rain accumulation, when not accompanied by brief extremes, may actually contribute to stabilizing the slope by facilitating gradual infiltration and percolation before pore pressurization persists long enough to initiate brief shallow failures. The results derived match past works citing landslide behavior on highly stepped landscapes as being mostly due to short-term high-magnitude events instead of being entirely the product of prolonged wet spells (e.g., Crozier, 2010; Tiranti & Rabuffetti, 2010; Bogaard & Greco, 2018). In addition, the decreased performance for F2 implies soil depth had

little bearing on enhancing predictive performance at the particular resolution utilized in this study. In conclusion, the results from applying the Random Forest model indicate that the landslide hazard in Buncombe is primarily caused by a combination of a highly stepped landscape on slope discontinuities and short periods of increased rainfall, rather than long periods of rainfall acting as an instantaneous trigger.

XGBoost

FeatureSet	Accuracy	Precision	Recall	F1 Score	ROC_AUC	Brier
F0	0.711	0.719	0.683	0.701	0.814	0.179
F1	0.802	0.757	0.883	0.815	0.879	0.141
F2	0.802	0.765	0.867	0.813	0.885	0.138

Table 6: XGBoost model performance across feature sets (F0, F1, F2).

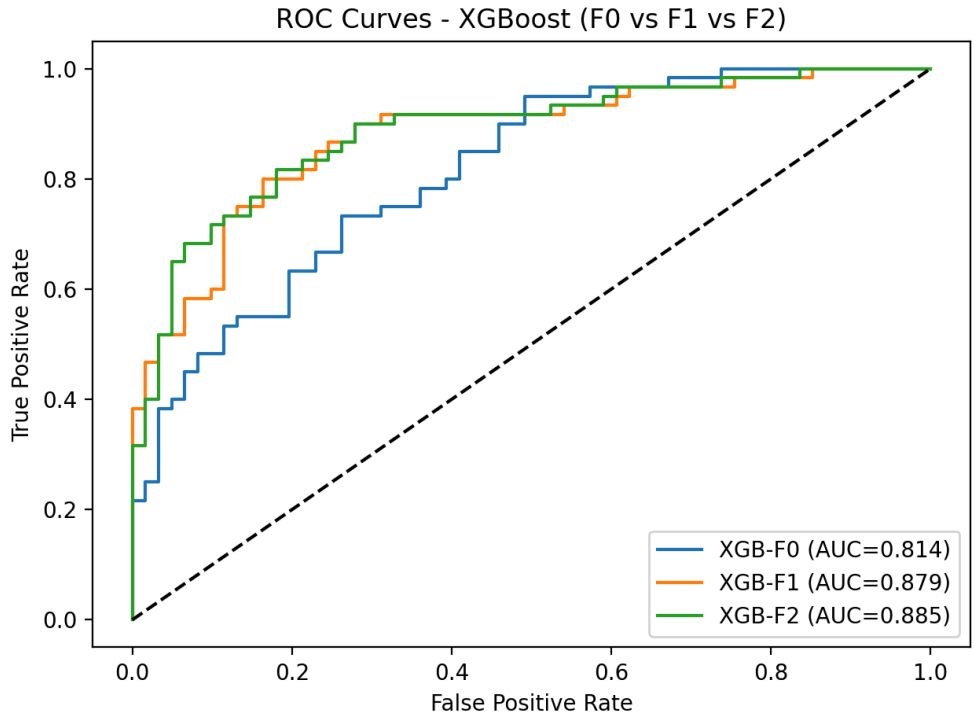

Figure 14: ROC curves of the three XGBoost models, F0, F1, and F2. Colors may be difficult to distinguish for some readers; interpretation should rely on line style and the AUC values.

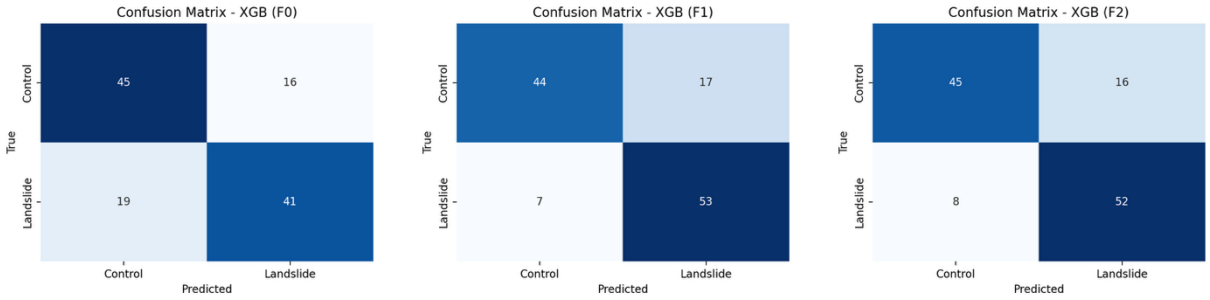

Figure 15-17: Confusion Matrix of the three models F0, F1, and F2 of XGBoost.

Surprisingly, XGBoost did not outperform the random forest model—instead, its F1 and F2 models showed similar prediction accuracy to the Random Forest Model (Table 6; Figure 14). The F0 XGBoost architecture excelled in model strength through an accuracy of 0.74 and an AUC of 0.81 in comparison to the other two F0 methods: Logistic Regression (LR) (accuracy 0.62, AUC 0.67) and Random Forest (RF) (accuracy 0.66, AUC 0.74). As expected, adding rainfall features improved the accuracy to 0.80 and AUC to 0.88, still better than the LR model, yet similar to the performance of the RF model (accuracy: 0.81, AUC: 0.89). Likewise, these trends appear to the full model (F2) equivalently. The full model (F2), which incorporated soil depth, produced the highest overall performance with an accuracy of 0.82 and an AUC of 0.89. The confusion matrix for F2 (Figure 17) highlights its balance, correctly classifying 45 controls and 52 landslides.

Compared to Logistic Regression and Random Forest, XGBoost delivered higher accuracy and consistently stronger AUC across feature sets, confirming its robustness in handling complex nonlinear relationships. The results reinforce that steep slopes and short-term rainfall extremes are primary triggers, while long-term rainfall can act as a dampening factor rather than a direct trigger (Crozier, 2010; Bogaard & Greco, 2018). The improved performance of the full F2 model also suggests that soil depth, while secondary to slope, provides additional predictive power in boosted frameworks.

Accuracy improved across all models when rainfall was added to terrain predictors. Logistic Regression rose modestly from 0.62 (F0) to 0.70 (F2), Random Forest increased more sharply from 0.66 (F0) to 0.81 (F1) but dropped slightly with soil depth (0.79, F2), while XGBoost achieved the highest accuracies at every stage, from 0.74 (F0) to 0.82 (F2).

Conclusion

This study developed a localized near-real-time landslide prediction framework for Buncombe County, North Carolina, using machine learning models and environmental predictors. By compiling a balanced dataset of 302 landslides and 302 non-events, rainfall windows from PRISM, topography from USGS DEMs, and soil depth from SSURGO, we trained Logistic Regression, Random Forest, and XGBoost classifiers. The pipeline incorporated stratified and spatial-block cross-validation, allowing us to rigorously compare models while preserving spatial representativeness. SHAP analysis and intrinsic feature importance made it possible to understand how predictors shaped landslide susceptibility in this mountainous region.

Across models, Random Forest consistently achieved the highest predictive performance (accuracy ≈ 0.82, AUC ≈ 0.89), followed by XGBoost and Logistic Regression. In all frameworks, slope emerged as the dominant predictor, confirming prior work showing terrain steepness as the primary control on landslide occurrence (Zheng et al., 2025; Xiao et al., 2024). Short-term rainfall accumulations (1–7 days, maxima over 3 days) were strongly associated with failures, consistent with global findings that intense short-duration rainfall is the principal trigger for shallow landslides (Crosta, 2004; Pennington et al., 2014; Barthélemy et al., 2024). A notable theoretical contribution was the identification of long-term cumulative rainfall (R30d) as a *negative* predictor in several models, suggesting that extended wet periods may sometimes stabilize slopes by allowing gradual drainage and infiltration rather than immediate pore-pressure buildup (Crosta, 2004; Fan et al., 2020). Soil depth contributed marginally, adding value only in boosted frameworks, which echoes recent findings that subsurface factors are often secondary to slope–rainfall interactions (Lee et al., 2023).

These results underscore both the practical and scientific value of localized machine-learning–based landslide modeling. The framework demonstrates that integrating rainfall windows with terrain predictors can generate near-real-time susceptibility maps, offering actionable information for hazard managers after storms like Hurricane Helene. At the same time, the findings reinforce a broader theoretical consensus: steep slopes combined with short-term rainfall extremes are the key drivers of shallow failures, while long-term rainfall plays a preparatory but

not directly triggering role (Pennington et al., 2014; Gariano & Guzzetti, 2016). Limitations of this study include a simplified binary soil depth representation, a relatively small sample size, and a lack of landslide-type differentiation. A limitation of this study is the use of PRISM daily rainfall, which cannot capture short-duration, high-intensity bursts that often trigger shallow landslides. These sub-daily intensities (e.g., mm/hr) are important in operational thresholds, but they are smoothed in daily grids. Future work could incorporate higher-temporal-resolution rainfall from radar products such as MRMS or satellite–gauge datasets like IMERG to better capture intense triggering events. Because the current framework is based on observed rainfall up to the event date, it functions as a near-nowcast rather than a true lead-time forecast; integrating short-term precipitation forecasts would be necessary to produce actionable early-warning lead times. Ultimately, coupling such models with rainfall forecasts could enable truly real-time landslide "nowcasting," a direction increasingly emphasized in both local and global hazard research (Stanley et al., 2021; Khan et al., 2022).

References

Akosah, S., Gratchev, I., Kim, D.-H., & Ohn, S.-Y. (2024). Application of Artificial Intelligence and Remote Sensing for Landslide Detection and Prediction: Systematic Review. *Remote Sensing*, *16*(16), 2947. https://doi.org/10.3390/rs16162947

Allstadt, K., McBride, S. K., Godt, J., Slaughter, S., Baxstrom, K., Sobieszczyk, S., & Stull, A. (2025). *Preliminary field report of landslide hazards following Hurricane Helene* (Report Nos. 2025–1028; Open-File Report). USGS Publications Warehouse. https://doi.org/10.3133/ofr20251028

Aydin, A. (2006). Stability of saprolitic slopes: Nature and role of field scale heterogeneities. *Natural Hazards and Earth System Sciences*, *6*(1), 89–96. https://doi.org/10.5194/nhess-6-89-2006

Barthélemy, S., Bernardie, S., & Grandjean, G. (2025). Assessing rainfall threshold for shallow landslides triggering: A case study in the Alpes Maritimes region, France. *Natural Hazards*, *121*(4), 4023–4049. https://doi.org/10.1007/s11069-024-06941-2

Bauer, J., Fuemmeler, S., Wooten, R., Witt, A., Gillon, K., Douglas, T., Eberhardt, E., Froese, C., Turner, K., & Lerouell, S. (2012). Landslide hazard mapping in North Carolina–Overview and improvements to the program. *Landslides and Engineered Slopes: Protecting Society through Improved*

Understanding. 11th International Symposium on Landslides and 2nd North American Symposium on Landslides, 257–263. https://www.researchgate.net/publication/260077615_Landslide_hazard_mapping_in_North_Carolina-Overview_and_improvements_to_the_program

Bogaard, T., & Greco, R. (2018). Invited perspectives: Hydrological perspectives on precipitation intensity-duration thresholds for landslide initiation: Proposing hydro-meteorological thresholds. *Natural Hazards and Earth System Sciences*, *18*(1), 31–39. https://doi.org/10.5194/nhess-18-31-2018

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brier, G. W. (1950). VERIFICATION OF FORECASTS EXPRESSED IN TERMS OF PROBABILITY. *Monthly Weather Review*, *78*(1), 1–3. https://doi.org/10.1175/1520-0493(1950)078%3C0001:VOFEIT%3E2.0.CO;2

https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Burrough, P., & McDonnell, R. (1998). *Principle of Geographic Information Systems*.

https://www.researchgate.net/publication/37419765_Principle_of_Geographic_Information_Systems

Chan, H.-C., Chen, P.-A., & Lee, J.-T. (2018). Rainfall-Induced Landslide Susceptibility Using a Rainfall–Runoff Model and Logistic Regression. *Water*, 10(10), 1354. https://doi.org/10.3390/w10101354

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Crozier, M. J. (2010). Deciphering the effect of climate change on landslide activity: A review. *Geomorphology*, 124(3–4), 260–267. https://doi.org/10.1016/j.geomorph.2010.04.009

Fan, L., Lehmann, P., Zheng, C., & Or, D. (2020). Rainfall Intensity Temporal Patterns Affect Shallow Landslide Triggering and Hazard Evolution. *Geophysical Research Letters*, 47(1), e2019GL085994. https://doi.org/10.1029/2019GL085994

Froude, M. J., & Petley, D. N. (2018). Global fatal landslide occurrence from 2004 to 2016. *Natural Hazards and Earth System Sciences*, 18(8), 2161–2181. https://doi.org/10.5194/nhess-18-2161-2018

Fuemmeler, S. J. (2008, April 10). *Landslide Hazard Mapping Methodology and Examples from North Carolina*. The Geological Society of America: Southeastern Section - 57th Annual Meeting. https://gsa.confex.com/gsa/2008SE/webprogram/Paper136714.html?utm

Gariano, S. L., & Guzzetti, F. (2016). Landslides in a changing climate. *Earth-Science Reviews*, 162, 227–252. https://doi.org/10.1016/j.earscirev.2016.08.011

Hatcher, R. D. (2010). The Appalachian orogen: A brief summary. In R. P. Tollo, M. J. Bartholomew, J. P. Hibbard, & P. M. Karabinos, *From Rodinia to Pangea: The Lithotectonic Record of the Appalachian Region*. Geological Society of America. https://doi.org/10.1130/2010.1206(01)

Horn, B. K. P. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE*, 69(1), 14–47. https://doi.org/10.1109/PROC.1981.11918

Jaboyedoff, M., Michoud, C., Derron, M.-H., Voumard, jérémie, Leibundgut, G., Sudmeier-Rieux, K., Nadim, F., & Leroi, E. (2016). *Human-Induced Landslides: Toward the analysis of anthropogenic changes of the slope environment* (pp. 217–232). https://doi.org/10.1201/b21520-20

Kang, J., Wan, B., Gao, Z., Zhou, S., Chen, H., & Shen, H. (2024). Research on machine learning forecasting and early warning model for rainfall-induced landslides in Yunnan province. *Scientific Reports*, 14(1), 14049. https://doi.org/10.1038/s41598-024-64679-0

Khan, S., Kirschbaum, D. B., Stanley, T. A., Amatya, P. M., & Emberson, R. A. (2022). Global Landslide Forecasting System for Hazard Assessment and Situational Awareness. *Frontiers in Earth Science*, *Volume* 10-2022. https://doi.org/10.3389/feart.2022.878996

Khashchevskaya, D., Owen, L. A., Wegmann, K., Scheip, C., & Figueiredo, P. M. (2025). The characteristics and timing of multiphase major landslides along the Blue Ridge Escarpment of the southern Appalachians revealed by combined cosmogenic nuclide dating and Schmidt hammer rebound measurements. *Geomorphology*, 485, 109857. https://doi.org/10.1016/j.geomorph.2025.109857

Kirschbaum, D. B., Adler, R., Hong, Y., Kumar, S., Peters-Lidard, C., & Lerner-Lam, A. (2012). Advances in landslide nowcasting: Evaluation of a global and regional modeling approach. *Environmental Earth Sciences*, 66(6), 1683–1696. https://doi.org/10.1007/s12665-011-0990-3

Kirschbaum, D., & Stanley, T. (2018). Satellite-Based Assessment of Rainfall-Triggered Landslide Hazard for Situational Awareness. *Earth's Future*, 6(3), 505–523. https://doi.org/10.1002/2017EF000715

Klose, M., Damm, B., & Terhorst, B. (2015). Landslide cost modeling for transportation infrastructures: A methodological approach. *Landslides*, 12(2), 321–334. https://doi.org/10.1007/s10346-014-0481-1

Klose, M., Maurischat, P., & Damm, B. (2016). Landslide impacts in Germany: A historical and socioeconomic perspective. *Landslides*, 13(1), 183–199. https://doi.org/10.1007/s10346-015-0643-9

Lee, S., Oh, S., Ray, Ram. L., Lee, Y., & Choi, M. (2023). Three-dimensional hydrological thresholds to predict shallow landslides. *Terrestrial, Atmospheric and Oceanic Sciences*, 34(1), 20. https://doi.org/10.1007/s44195-023-00052-4

Lin, S., Chen, S., Rasanen, R. A., Zhao, Q., Chavan, V., Tang, W., Shanmugam, N., Allan, C., Braxtan, N., & Diemer, J. (2024). Landslide Prediction Validation in Western North Carolina After Hurricane Helene. *Geotechnics*, 4(4), 1259–1281. https://doi.org/10.3390/geotechnics4040064

Longley, P. A., Goodchild, M. F., Maguire, D. J., & Rhind, D. W. (2015). *Geographic Information Science and Systems*. Wiley. https://books.google.com/books?id=C_EwBgAAQBAJ

Lundberg, S., & Lee, S.-I. (2017). A *Unified Approach to Interpreting Model Predictions* (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1705.07874

Mondini, A. C., Guzzetti, F., & Melillo, M. (2023). Deep learning forecast of rainfall-induced shallow landslides. *Nature Communications*, 14(1), 2466. https://doi.org/10.1038/s41467-023-38135-y

NASA Earth Observatory. (2021, June 10). *Machine Learning Model Doubles Accuracy of Global Landslide 'Nowcasts'—NASA*. https://www.nasa.gov/missions/gpm/machine-learning-model-doubles-accuracy-of-global-landslide-nowcasts/

National Hurricane Center. (2025). *Tropical Cyclone Report: Hurricane Helene* (AL092024) (p. 17). https://www.nhc.noaa.gov/data/tcr/AL092024_Helene.pdf

North Carolina Department of Environmental Quality. (2024). *North Carolina Landslide Inventory Points* [Dataset]. NC OneMap. https://www.nconemap.gov/datasets/01965a193482438cb70332e5e524e38b_0/about

Office of State Budget and Management. (2024). *Hurricane Helene Damage and Needs Assessment*. https://www.osbm.nc.gov/hurricane-helene-dna

Pennington, C., Dijkstra, T., Lark, M., Dashwood, C., Harrison, A., & Freeborough, K. (2014). Antecedent Precipitation as a Potential Proxy for Landslide Incidence in South West United Kingdom. In K. Sassa, P. Canuti, & Y. Yin (Eds.), *Landslide Science for a Safer Geoenvironment* (pp. 253–259). Springer International Publishing. https://doi.org/10.1007/978-3-319-04999-1_34

Petley, D. (2012). Global patterns of loss of life from landslides. *Geology*, 40(10), 927–930. https://doi.org/10.1130/G33217.1

Petley, D. (2025, January 6). Global fatal landslides in 2024. *Eos*. https://eos.org/thelandslideblog/fatal-landslides-in-2024

PRISM Climate Group. (2024). *PRISM Daily Precipitation Data (1981–present)* [Dataset]. Oregon State University. https://prism.oregonstate.edu/

Regorda, A., Lardeaux, J.-M., Roda, M., Marotta, A. M., & Spalla, M. I. (2020). How many subductions in the Variscan orogeny? Insights from numerical models. *Geoscience Frontiers*, *11*(3), 1025–1052. https://doi.org/10.1016/j.gsf.2019.10.005

Sidle, R. C., & Ochiai, H. (2006). *Landslides: Processes, Prediction, and Land Use* (Vol. 18). American Geophysical Union. https://doi.org/10.1029/WM018

Sim, K. B., Lee, M. L., Rasa Remenyte-Prescott, & Soon Yee Wong. (2022, August). *An Overview of Causes of Landslides and Their Impact on Transport Networks*. Advances in modelling to improve network resilience, https://op.europa.eu/en/publication-detail/-/publication/c81e8bc9-1469-11ed-8fa0-01aa75ed71a1/language-en. https://www.researchgate.net/publication/369437132_An_Overview_of_Causes_of_Landslides_and_Their_Impact_on_Transport_Networks

Stanley, T. A., Kirschbaum, D. B., Benz, G., Emberson, R. A., Amatya, P. M., Medwedeff, W., & Clark, M. K. (2021). Data-Driven Landslide Nowcasting at the Global Scale. *Frontiers in Earth Science*, 9, 640043. https://doi.org/10.3389/feart.2021.640043

Thomas, M. A., Collins, B. D., & Mirus, B. B. (2019). Assessing the Feasibility of Satellite-Based Thresholds for Hydrologically Driven Landsliding. *Water Resources Research*, 55(11), 9006–9023. https://doi.org/10.1029/2019WR025577

Tiranti, D., & Rabuffetti, D. (2010). Estimation of rainfall thresholds triggering shallow landslides for an operational warning system implementation. *Landslides*, 7(4), 471–481. https://doi.org/10.1007/s10346-010-0198-8

U.S. Census Bureau. (2022). *Cartographic Boundary Shapefiles – Counties* [Dataset]. https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html

U.S. Department of Agriculture, Natural Resources Conservation Service. (2024). *Soil Survey Geographic (SSURGO) Database* [Dataset]. https://websoilsurvey.nrcs.usda.gov/app/

U.S. Geological Survey. (2021). *National Land Cover Database (NLCD) 2021* [Dataset]. https://www.usgs.gov/centers/eros/science/national-land-cover-database

U.S. Geological Survey. (2022). *3D Elevation Program (3DEP), 1/3 arc-second DEM seamless products* [Dataset]. https://www.sciencebase.gov/catalog/item/627f3798d34e3bef0c9a3198

Watterson, N. A., & Jones, J. A. (2006). Flood and debris flow interactions with roads promote the invasion of exotic plants along steep mountain

streams, western Oregon. *Geomorphology*, 78(1–2), 107–123. https://doi.org/10.1016/j.geomorph.2006.01.019

Wooten, R. M., Gillon, K. A., Witt, A. C., Latham, R. S., Douglas, T. J., Bauer, J. B., Fuemmeler, S. J., & Lee, L. G. (2008). Geologic, geomorphic, and meteorological aspects of debris flows triggered by Hurricanes Frances and Ivan during September 2004 in the Southern Appalachian Mountains of Macon County, North Carolina (southeastern USA). *Landslides*, 5(1), 31–44. https://doi.org/10.1007/s10346-007-0109-9

Xiao, X., Zou, Y., Huang, J., Luo, X., Yang, L., Li, M., Yang, P., Ji, X., & Li, Y. (2024). An interpretable model for landslide susceptibility assessment based on Optuna hyperparameter optimization and Random Forest. *Geomatics, Natural Hazards and Risk*, 15(1), 2347421. https://doi.org/10.1080/19475705.2024.2347421

Ye, C., Wu, H., Oguchi, T., Tang, Y., Pei, X., & Wu, Y. (2025). Physically Based and Data-Driven Models for Landslide Susceptibility Assessment: Principles, Applications, and Challenges. *Remote Sensing*, 17(13), 2280. https://doi.org/10.3390/rs17132280

Zheng, W., Fan, W., Cao, Y., Nan, Y., & Jing, P. (2025). Landslide Hazard Assessment Under Record-Breaking Extreme Rainfall: Integration of

SBAS-InSAR and Machine Learning Models. *Remote Sensing*, 17(13), 2265. https://doi.org/10.3390/rs17132265

Mentor Disclosure

The author conducted this research independently.

AI Disclosure

AI tools (ChatGPT) were used solely for grammar and formatting edits; all analysis and interpretation are original to the author.

Review- A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina

The authors develop a model to predict landslides with better information, accuracy, and timeliness. They do an excellent job articulating the importance and value of this topic, and their review of existing literature is extensive and very well written. Their methodology and approach to developing their models and evaluating them to identify the unique contributions of rainfall above and beyond other more static factors.

I just have two minor critiques:

1. If dynamic rainfall data is the primary contribution of this paper, why not have F0 and F1 models be the static variables and the F2 model be the addition of dynamic variables? It would be more coherent and clearly show the contribution of rainfall above and beyond the static factors.

We kept F0 as terrain-only, F1 as terrain plus rainfall, and F2 as terrain, rainfall, and soil depth because this structure makes the incremental contribution of rainfall clearer. We clarified these definitions in the Methods section.

2. The sample size is pretty small, especially for a powerful model like XGBoost. I do wonder if it may outperform Random Forest with a larger, more complex dataset, but don't think this necessarily detracts from the overall value and impact of this paper

We added a brief note in the Conclusion acknowledging that the limited sample size may constrain XGBoost's performance and that larger inventories may allow it to outperform Random Forest in future work.

Overall, this is an excellent paper and I am happy to recommend it for publication!

Recommendation: Accept as is

Review of Paper: “A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina”

This study addresses an important natural hazard – landslides – which are often deemed geologic risks, but as this study addresses, they may interact, particularly non-linearly, with various hydrologic factors, including rain accumulation and rain rates on various time scales, the latter of which are often triggers beyond particular USGS (United States Geological Survey) thresholds depending on context and other factors. This study introduces a landslide-risk model with PRISM gridded precipitation data, as well as static USGS Digital Elevation Model data and SSURGE (acronym not defined) soil depth data using various Machine-Learning (ML) methods. While not an inherent forecast model per se, the implications of this study, as is stated, “lay the foundation for near-nowcasting,” critical for gauging near-term risk and therefore enabling the potential to create alarms during and after hydrological triggers of concern. The primary conclusions of this study are that while the primary static feature – namely the slope, is the dominant correlate with landslides, shorter-term rain rates add significant skill in predicting landslides, whereas the accumulation of rainfall over a month or longer, particularly precipitation that is not intense but is able to percolate into soils, makes areas less susceptible for landslides – these non-linear precipitation interactions seem to be at the heart of the novelty of this study, and may be consistent or even corroborative with past work, though the framework here potentially allows new separation of variables in a real-time way given fairly high-resolution observational data and sufficiently dense static data from the USGS.

Overall, the content and the presentation seem quite compelling, with some degree of modification needed here and there (further elucidated in the specific comments below), though there are some assumptions that the writer may have in terms of *a priori* knowledge of the reader that may not be deemed to be common, general knowledge. While this is a modest hindrance of accessibility of the study overall, it could fairly straightforwardly be ameliorated with some additional description and definition of some of the statistical/machine learning terms including, but certainly not an exhaustive list - Logistic Regression, Random Forest, and XGBoost. Furthermore, several acronyms are used without being spelled out or necessarily defined. Eventually, a reader may be able to deduce how to interpret these, but as this is not a highly specialized journal, spelling out what key acronyms mean is critical so that nearly any reader can walk away with key takeaways of the study without extensive background reading first. In general, unless explicitly stated by a journal, most, if not all acronyms should be defined once before using them throughout.

Related to the lack of some of the key definitions utilized throughout the manuscript, a number of the figures, which though decipherable for someone in the field of certain areas of specialized statistics or ML, may not be as easy to interpret for a generalist in hydrology, geology, mathematics, atmospheric sciences, or even traditional statistics. These include the ROC curves (Figures 4 and 8), the SHAP beeswarm plot (Figure 12), and perhaps to a lesser extent the Confusion Matrices (which might be mislabeled in terms in terms of the legends and figure numbers – more on that below). A brief description and interpretation may be needed for each of these, rather than merely explaining the results. One reason I am bringing this up is that many readers may be most familiar with scatter plots relating two or more variables, with degree of goodness of fit, often presented in the form of a correlation coefficient, is measured in terms of closeness to the diagonal (at least for linear relationships), but the best performance for an ROC (Receiver Operating Characteristic) curve is towards the vertical (top left corner), and on an ROC curve the diagonal simply suggests a random classifier. As

someone not as well-versed in ROC curves, it took me some additional time to understand quantitively why model sets F1 and F2 are generally superior to F0 based purely on the figures alone. The text generally is satisfactory at explaining why rainfall metrics improve the models, but the characterization of some of the figures needs additional information in the manuscript.

Finally, there is some consolidation that can likely be done which might have the dual objective of saving some space, reducing the number of figures, and better allowing a more direct comparison among the different ML methods – and that would be including all of the ROC curves (from Logistic Regression, Random Forest, and XGBoost) onto one panel with perhaps different symbols on some of the curves to properly distinguish them. Alternatively, consider a multi-panel figure with all three models that save the reader from flipping to different pages to compare on different pages.

Overall though, while some improvement is possible with regards to accessibility, presentation, and at some points even some ambiguity of the interpretation of some of the rainfall variables, this paper should be accepted pending mostly satisfactory minor (and maybe some moderate) revision. With regards to grammar, the paper is generally well written, though there are a few instances in which tenses switch to past tense whereas present tense would be more appropriate. A thorough (though non-exhaustive) list of these technical and grammatical points are provided below.

Specific Points:

1) Lines 10-11: USGS DEM appears with being spelled out – in an abstract it should be written as "USGS Digital Elevation Model (DEM)". The same issue exists with SSURGO. It is not until line 183 that DEM is written out. If SSURGO is not written out anywhere else (I couldn't find it anywhere), it should be in the table of definitions (Table 1), and at the very least on line 11 add "depth" after "SSURGO soil".

SSURGO refers to the dataset, which is accessible online, and we use it multiple times in the paper, and it is irrelevant to the actual content of the paper.

2) Lines 90-92: Why is the question in quotes? Is this to bring additional emphasis, and are these words original to the author?

Removed quotes on line 90

3) Lines 93-94 and more generally: What lead time is being sought preferentially with regards to the proposed "early warning system?" How much lead time is necessary to potentially reduce mortality related to landslides under more extreme conditions?

I'm predicting the risk, thus providing the probability of the landslides, so there cannot be a specific lead for the predictions. Line 93-94

4) Lines 108-109: These numbers referred to in the citation (Ye et al., 2025) refer to the number of publications of landslide susceptibility studies, rather than the number of statistical models, it appears. Please clarify.

I updated the text to indicate that this represents the number of published studies, not unique model types.

5) Line 114: What does RNN stand for? It is not defined anywhere and only seems to appear once here. My best guess (after searching a bit) is Recurrent Neural Network, but please spell out if this is this case.

Revised the text to introduce the term as "Recurrent Neural Network (RNN)."

6) Lines 250-251, Table 1, and more generally: I notice that there are not maximum rain rates at shorter time scales (subdaily, even hourly), which might be important as well. Speaking of which, the USGS often has millimeter (mm)/hr or inches (in)/hr thresholds, which if breached, can be triggers of both landslides and mudslides, so wonder about the appropriateness of PRISM only data for very intense, short events? I am thinking here also more about the generalizability of some of the results of this study – including landslides/mudslides following burn areas on hillsides/mountains with steep slopes and denuded vegetation; 5-minute to hourly rain rates can be critical at determining the potential susceptibility of landslide hazards. Or, would high temporal resolution rain rates be beyond the scope of the research/model presented here? The latter could possibly come from radar data (https://mrms.nssl.noaa.gov/qvs/product_viewer/) or IMERG precipitation data, which is merged satellite-gauge data (https://gpm.nasa.gov/data/imerg).

I added a paragraph in the Conclusion noting that PRISM daily data smooths intense, short-duration events and Short-timescale rainfall (mm/hr) is critical for shallow failures. I explicitly state this limitation and frame it as a direction for follow-up studies

7) Line 262 and more generally: The first instance of ROC appears to occur here on line 262, but it is not defined, even though it is a heavily used metric/concept throughout the study, and also in multiple figures/tables. I believe that it refers to Receiver Operating Curve, but please confirm. ROC may not be a common acronym, so for the benefit of readers, I would suggest spelling it out during its first instance, and also perhaps defining a bit more, in particular additional characterization of the plots and how a non-specialist reader should interpret the ROC curves. This also applies to the acronym AOC (Area Under the Curve).

I now define both terms at first mention and provide a brief explanation of ROC interpretation, including the meaning of the "no-skill" diagonal. Figure captions have also been updated to clarify line meanings and accessibility.

8) Lines 288-289 and Lines 400-402: As the authors rightly note, including soil depth does not necessarily improve results beyond including rain rates/rain accumulation in explaining landslides – any improvement at times does not seem statistically significant. Why do you think this might be?

I added context explaining that SSURGO soil depth is coarse relative to 40-m DEM resolution, and most mapped soils are classified as very deep (>200 cm), reducing variability.

9) Lines 329 and beyond – More on the tenses switching to past when present would be more appropriate: It's best to be consistent and maintain present as much as possible. The exception to this is if an action was done at a particular point in time – e.g. "Data were collected on XX February XX", in which case past tense is correct.

I revised the Results and Methods sections to maintain present tense for interpretation and past tense only when describing actions already taken

10) Lines 329-337 and Results in Tables 2 & 3: The author emphasizes that "short-period rainfall accumulation" … "have a strong relationship with landslide risks", but "Max_Rainfall_30day" is such a prominent variable in all five of the top five logistic regression groups in Table 2. Perhaps a bit surprising to me, "Max_Rainfall_3day" only appears in one subset. The predictors for the Random Forest Table (Table 3) are somewhat different, with Max_Rainfall_3day making an appearance twice, although the 30-day equivalent is present in all five of the subsets. Indeed, when I look at all five subsets of Table 3, it seems that rainfall at the short-term (often known as synoptic scales of a few days up to two weeks) to monthly are broadly and nearly equally important.

I expanded the explanation by noting that Max_Rainfall_30day captures antecedent hydrologic conditioning and Short-term rainfall (R1d–R7d) captures direct triggering.

11) Figures 4 (and 8 and 14): Particularly for those with varying degrees of color blindness, I would highly suggest altering the styles of the lines to better distinguish them, or perhaps even adding symbols to one or more to help differentiate. Even for someone without color blindness, the blue and green (F0 and F2 curves, respectively) are a bit tricky to distinguish (a good way of checking how figures appear to someone with varying degrees of color blindness can be found here: https://www.color-blindness.com/coblis-color-blindness-simulator/)

    Somewhat relatedly, for more general readers, I would be sure to specify that the thick dashed line is the "No Skill" line – not everyone knows how to interpret these ROC curves without some additional guidance. This also needs to be included in the figure legends.

I updated figure captions to explicitly identify guidance for readers with color vision deficiencies.

12) Table 5: I am a little confused – why are rainfall parameters included in the F0 Logistic Regression Model – I thought only static slope/elevation parameters were used? What is the meaning of these coefficients in the context of F0?

I reviewed the table and confirmed that F0 contains only terrain predictors. I added a clarifying sentence that rainfall coefficients are blank for F0 because the F0 configuration intentionally excludes rainfall variables.

13) Lines 370-371: This accuracy (of the F2 model) is the same (0.70) as when F1 was included as well – I'm not quite sure I follow?

I added an explanation that soil depth offers minimal incremental signal in logistic regression due to coarse resolution and low variability, causing performance between F1 and F2 to converge.

14) Lines 373-374: While this statement sounds promising in terms of the improvement in correct identification of landslide cases correct from F0 versus F2, is this improvement statistically significant? Is there a way to formally test for this?

I have clarified in the manuscript that no formal statistical significance test was performed to assess whether the differences in model accuracy are statistically significant.

15) Table 5: The behavior of Max_Rainfall_30day for F1 and F2 is positive in terms of the sign of the regression coefficient, but strongly negative for F0 (again, with the caveat that I still do not know how to interpret the rainfall parameters for F0), but R30d is negative for F1/F2, which is a little difficult to digest. Max_Rainfall_30day is simply the rolling maximum rainfall of 30-days, is it not? Or, is it the maximum rainfall (of a day) within a 30-day period? Why is the behavior of R30d and Max_Rainfall_30day so different? Please help me understand what I am missing.

I added a clear distinction—R30d = cumulative rainfall over last 30 days, while Max_Rainfall_30day = wettest 30-day period in the last 90 days. These measure different hydrologic processes (gradual wetness vs. antecedent extremes). Their opposite coefficient signs therefore reflect different physical mechanisms.

16) Lines 398-399: Yes, I see this in the ROC curves as well – why might the inclusion of soil depth slightly degrade the performance of F2 in terms of the Random Forest model?

We don't know

17) Figure 13: I am confused about the two shades of blue/purple in this Figure 13 – what do the darker purple bars refer to? They are not defined in the legend or the caption. Also, no units are provided in Figure 13 – are they mm?

I have revised the caption for Figure 13 to clearly indicate what the two color shades represent and to specify the units.

18) Please be certain to check all the DOIs in the Works Cited/References section are correct – I caught at least one in which the URL did not work for the reference in lines 544-547. It should be: https://doi.org/10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2

Corrected the DOI for Brier (1950) as suggested. Checked all the other doi links, and replaced/deleted links that did not work.

19) Figure 12 and corresponding explanation in text: Please provide additional guidance for readers in how a SHAP Beeswarm chart is constructed and how to interpret it. Many non-specialists are not familiar with this type of plot.

SHAP explanation added and Basic interpretation included

20) Figures 5-7, 9-11, and 15-17: The three confusion matrices for each of the models do not require a separate figure number; they can simply be deemed a single figure, with multiple panels. I like these figures, but these so-called 9 figures really only need to be three. Furthermore, I might provide a little more narrative in the text about how to understand a Confusion Matrix, especially as many readers may not be that familiar with them.

Although these matrices could be combined into multi-panel figures, we present them individually to maintain readability and allow direct comparison across F0, F1, and F2 for each algorithm.

Grammatical Suggestions:

1) First line of abstract (line 5): Change “Rainfall-triggered landslides are commonly occurring across the Southern” to “Rainfall-triggered landslides commonly occur across the Southern”
2) Lines 121-122: “Incorporating every aspect of geographic, environmental and hydrological factors into eight variables” is an incomplete sentence. I am not sure if the author intended to convey an additional thought here, or if this sentence got truncated?
3) Line 259: Need a period after the parentheses.
4) Line 273: Change “visualized” to “visualizes”
5) Line 274: Change “showed” to “show”
6) Line 275: An example of where the tense should be present (not past): Change “SHAP waterfall plots gave” to “SHAP waterfall plots give” (there are also multiple other instances in this paragraph in which the tense of statements should be present).
7) Line 400: Awkward Transition: “Thus” would be a better way to start the new sentence than “Still”

Addressed all grammar mistakes.

Follow-Up Review of Paper: “A Dynamic Risk Prediction Model for Rainfall-Triggered Landslides in Buncombe County, North Carolina”

As this represents the second review, this time of a revised manuscript, this report is quite concise, but there remain just a few small items worth enumerating and important to address prior to publication. First and foremost, I appreciate the author’s thorough responses to concerns expressed in the previous review, and the additional text adds worthy description, clarification and context for readers where needed, so the latest iteration represents a fairly appreciable improvement. Spelling out some of the more esoteric acronyms, particularly in the Abstract (e.g. Soil Survey Geographic Database (SSURGO)), improves the accessibility of the paper, and furthermore, whatever is contained in the abstract usually represents both the topline results and most essential datasets used in the study, so being as clear as possible is critical.

With regards to the recommendation of different line styles for the line plots in Figures 4, 8, and 14, these figures appear to be the same as before albeit with enhanced characterization about how to interpret them, and admittedly, I ran them through a monochromatic simulator and while it is possible, with the updated guidance of using the AOC values appropriately, to discern the lines corresponding to models F0, F1, and F2 (with some effort), it might still be preferable (though I am not requiring this as a modification) for F1 (or F2) to have a different line style such as dots or short dashes.

While I appreciate the changes overall to streamline the tenses and to be consistent with regards to research being presented (largely present tense) versus actions taken (past tense), a final proofread by a professional editor may be helpful to shore up any remaining minor inconsistencies. For instance, after line 161, I might suggest in the Dataset section the following: “This research **aims** to create a dynamic prediction model …” (then on line 166): “The dataset **includes** event …”

All of this aside, these are very minor remaining comments, and I accept this manuscript for publication pending these modifications.

Remaining Minor Comments:

1) There’s an extra character “v” at the end of the abstract (line 18), which is likely inadvertent.
2) Lines 209: I would add “Map Unit Key” after MUKEY and “Map Unit Symbol” after MUSYM” for clarification for the reader.
3) Figure 12 interpretation: Elevation encompasses both negative and positive SHAP values, and it’s a little tricky to discern whether more purple/red dots populate negative or positive values, hence the mention of “bidirectional influence” on line 439. Are you able to expand the discussion around how elevation impacts landslides and how the reader should interpret this finding? This still remains slightly fuzzy to me.