Student thesis series INES nr xx

Assessing the role of precipitation in logistic regression landslide susceptibility models in the Blue Ridge Mountains of North Carolina

Tom Samzelius Valentin

2026

Department of
Earth and Environmental Sciences
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

Tom Samzelius Valentin (2026)
Assessing the role of precipitation in logistic regression landslide susceptibility models in the Blue Ridge Mountains of North Carolina
Utvärdering av nederbördens roll i logistiska regressionsmodeller för jordskredsbenägenhet i Blue Ridge Mountains i North Carolina
Bachelor degree thesis, 15 credits in *Physical Geography and Ecosystem Analysis*
Department of Physical Geography and Ecosystem Science, Lund University

Level: Bachelor of Science (BSc)

Course duration: *January* 2026 until *June* 2026

Disclaimer

This document describes work undertaken as part of a program of study at the University of Lund. All views and opinions expressed herein remain the sole responsibility of the author, and do not necessarily represent those of the institute.

Assessing the role of precipitation in logistic regression landslide susceptibility models in the Blue Ridge Mountains of North Carolina

---

Tom Samzelius Valentin

Bachelor thesis, 15 credits, in
Physical Geography and Ecosystem Analysis

Thomas Pugh
Supervisor

Examiner name:
Andreas Persson

Acknowledgements

Firstly, I wish to thank my supervisor, Tom Pugh, for his sound advice and ability to answer panicked emails from a confused and nervous thesis student on short notice. Thanks, Big T. Additionally, I would like to extend my thanks to the teaching staff within the Physical Geography Department as a whole for making my time studying extremely engaging. I also want to thank my friends, who have made attending university a fantastic experience.

Furthermore, I wish to thank my family, specifically my mom, dad, and sister; whose undying support have pushed me to achieve my dream of graduating university. Finally, I want to express my endless gratitude to Julia, whose selfless encouragement and unwavering belief I would be nowhere without.

Abstract

Landslide susceptibility models (LSMs) are essential to mitigate the danger of landslides, and with a changing climate posing potential risks to model accuracy, it is vital to understand the role that precipitation plays within LSMs and the effect shifting climatic conditions may have on model accuracy. This study produced a logistic regression LSM in the Blue Ridge Mountain region of North Carolina for the 2001-2020 period in order to understand the significance of precipitation as a predictor, observe the change in landslide occurrence an increase in precipitation intensity could cause, and analyze the change in accuracy between models trained on one precipitation regime when applied to differing climate conditions. The model exhibited high accuracy, with an average AUC score of 0.85, with no signs of overfitting. Precipitation was the second most significant predictor after slope, with extreme statistical significance ($p<0.0001$) and a 58% susceptibility rise per standard deviation increase. Additionally, a precipitation sensitivity analysis found that a change in precipitation in the study area led to a significant change in the model's landslide predictions. The results of a decadal cross-validation showed an 8% decrease in overall accuracy and a 9% rise in Type I errors when applied to a decade with differing precipitation conditions. The decadal results did, however, also highlight a potential landslide inventory bias, which could have held undue influence over the model. Overall, the results of the study suggested that precipitation is a vital predictor within the LSM, and that changes in climate could potentially lead to decreased accuracy within landslide models. Further research is encouraged into this vital issue.

Table of Contents

1 Introduction

Landslides are one of the most destructive natural hazards in the world, and were responsible for a global death toll of roughly 160,000 from 1995 to 2014 (Haque et al., 2019). In addition to the tragic loss of life, the direct economic loss from landslides is estimated at $20 billion annually worldwide (Sim et al., 2022). Used as a general term for large-scale, downhill movement of rock, soil, and debris, landslides are considered most dangerous at higher velocities, although slower movements can still cause major damage to buildings and infrastructure (Hetley, 2010). Precipitation is a major trigger of landslides, with 79% of non-seismic fatal landslide events being triggered by rainfall in the period 2004 to 2016 (Froude & Petley, 2018). With climate change increasing the frequency and intensity of precipitation in many parts of the world (Tahroudi, 2025; Blenkinsop et al., 2021), it is estimated that the risk of landslides will similarly increase in affected areas (Gariano and Guzzetti, 2016).

Landslide susceptibility models (LSMs) are a vital part of modern hazard mapping, and are used to highlight areas considered vulnerable to landslides based on local conditions. The types of models used for susceptibility mapping can be broadly divided into five groups: heuristic or opinion-driven models, statistical models, physical-based or deterministic models, machine learning models, and neural network or deep learning models (Reichenbach et al., 2018; Merghadi et al., 2020). The models in these groups vary in their specific functionality, but their general idea is the same: use data that describe landslide conditioning factors, or predictors, present at locations where landslides have and have not occurred in the past, and apply this understanding to a study area. This means that although LSMs can assess where a risk of landslide exists, they cannot predict when a landslide will occur. According to Nocenti et al., (2023), LSMs are not temporal assessments of landslide predictions, but are more limited to a spatial modelling of landslide risk based on historical data. This highlights one of the most significant limitations of susceptibility models: the issue of temporal transferability.

Because LSMs are built and trained on historical landslide data, the characteristics of the past determine the model's understanding of the conditions that result in landslides. When those conditions, such as precipitation, change over time, the model may lack the understanding of how this affects the susceptibility of an area. This point is echoed in a study about the effect of climate change on landslides: “The use of a spatially established relationship to predict changes through time… relies on the spectrum of reconditions occurring through time matching those that influenced the original spatial relationship” (Cozier, 2010). The quasi-static assumptions regarding predictors are generally acceptable when considering predictors that change in centennial or millennial-scale, such as slope, geology, or soil type. However, this may present a problem when considering predictors that change over a smaller timescale, such as precipitation. Although there are emerging attempts to address temporal issues regarding precipitation using event-based or dynamic models, many studies developing LSMs use static, long-term precipitation variables, such as multi-decade annual averages (Ling et al., 2022; Shahabi and Hashim, 2015; Sun et al., 2018), which do not consider precipitation intensity, its short-term variations, or its potential rise. With global landslide susceptibility expected to rise in high-emission scenarios (Duan et al., 2025), continued research into the effect changing precipitation will have on current day susceptibility models is vital.

This study will explore the issue of temporal transferability by assessing the role of precipitation in an LSM in the Blue Ridge Mountain area of North Carolina. Specifically, this study will develop and train a statistical based, logistic regression landslide susceptibility model based on a series of predictor variables for a 20-year study period, 2001 to 2020, and thereafter examine the role of precipitation in the model, its strength within it, and its effect on susceptibility predictions across several precipitation scenarios. This study seeks to address three main objectives: How significant is precipitation as a predictor to an LSM for this study area? Would an increase in precipitation cause a similar increase in the predicted quantity of landslides in the study area? Is the accuracy of an LSM trained on one set of historical precipitation data significantly worsened when predicting landslides with different precipitation conditions? Answering these questions will lead to a clearer understanding of an important issue within landslide susceptibility modelling, and hopefully shed light on the potential need for future research and study within the topic.

1.1 Study Area

The area chosen for this study was the Blue Ridge Mountains of North Carolina, a part of the Southern Appalachian Highlands, itself a part of the larger Appalachian Mountains. The study area was defined as the North Carolinian counties within the Appalachian Regional Commission's (ARC) official map of Appalachia. The region was chosen for its history of landslide activity, as well as the considerable role precipitation plays in their occurrence. The southern Blue Ridge province, mainly situated in southeastern Tennessee and western North Carolina, features some of the highest elevations in the United States east of the Rocky Mountains, and experiences high mean annual precipitation due to movement of occluded low-pressure fronts and hurricane remnants from the warm, moist air of the Gulf Stream. These topographic and climatic conditions combine to cause frequent and recurring rapid mass wasting processes in the region (Clark, 2008). Wooten et al. (2016) corroborates these claims, finding that storms in the region trigger hundreds of debris flows roughly every 9 years. Furthermore, cyclonic storms from the Gulf Stream have triggered hundreds to thousands of debris flows in various surges since 1916.

The Appalachian Mountains are a well-known geographical feature of the eastern United States. They were, along with its sub-regions such as the Blue Ridge Mountains, created roughly 270 million years ago, when the African continental plate collided with the North American continental plate (Clark, 2001). Today, the geology of the Blue Ridge Mountains are largely dominated by metamorphosed sandstone and shale, which has through dynamic chemical weathering produced generally acidic soils supporting varying species of flora (Pittillo et al., 1998). The Appalachian-Blue Ridge forests are among the most biodiverse forests in North America, being home to 158 species of trees, with low elevations generally home to deciduous trees, transitioning to coniferous spruce-fir forests beyond roughly 400 meters altitude (Bement, 2024). The climate of the region prevents the creation of an alpine zone tree-line, thus the forests generally extend all the way to the mountaintops. Although building density has increased throughout the Southern Appalachian Mountains in recent decades, it is still generally dominated by forest cover (Turner et al., 2003).

Figure 1: A map of the study area, the Blue Ridge Mountain region of North Carolina.

2 Methods

The work undertaken in this study can be described in several distinct steps. They will be broadly described here before going into relevant detail in the following methodology sections. Firstly, an examination and scrutinization of multiple landslide databases was performed. This, and many other functions within this study, was undertaken by the use of Esri's ArcGIS Pro program. This procedure entailed a review of each database's accuracy and relevance before a final inventory was selected. Secondly, an initial investigation of the study area was undertaken; involving examining precipitation trends and landslide occurrence within the area. A study period was also defined: limiting the scope of the study to a 20-year period from 2001 to 2020. Thirdly, there was a broad ranging literature search, mainly to understand the various landslide conditioning factors considered by studies into the topic. This allowed the finalization of a list of predictor data to be used in the model. The fourth stage was data acquisition and harmonization. This process involved exploring regional, national, and international databases to find relevant geographical data. This data was then downloaded and examined to ensure its relevance and accuracy to the study. If found to be a relevant data source, it was then harmonized to the predetermined cell size, extent, and coordinate system (NAD 1983 UTM Zone 17N). Fifth, the data was extracted for each layer and combined into one dataset based on a set of historical landslide points and generated non-landslide points to be used as the training and validation dataset within the model. The end result was a dataset with a total 1406 points, each with a value from every predictor raster from where the point intersected with them. Sixth, a logistic regression model was designed and coded in Python. This model was put through several rounds of testing to determine the most accurate combination of predictors for a refined, final model. Seventh, the finalized model was used to project a series of precipitation scenarios, where precipitation in the validation set was altered stepwise from 25% to 200% of its original values to see the impact of increased precipitation on the model's landslide predictions. This data was then analyzed to understand the potential impact of increased precipitation within the study area according to the model. Finally, the dataset was divided by decade and used for two models, which were trained and cross-validated. The results were then used to analyze the potential change in accuracy due to a change in precipitation within the datasets.

2.1 Landslide Inventory

The most vital part of the data needed for this study was the database of known historical landslide points. The dataset was chosen based on a set of criteria including scope, accuracy, and total landslide inventory. Landslide Inventories across the United States ver. 2.0 (Belair et al., 2022) was chosen due to its initial inventory containing over 200,000 known landslide points throughout the United States, a large number of those being within the study area and period, as well as the inclusion of a confidence metric. The confidence metric was a semi-objective criteria calculated for each landslide point based on the methods of landslide identification, such as media reporting or geohazards mapping, and the accuracy of the data used, such as high resolution lidar or coarser alternatives. This confidence metric expressed the authors' certainty of both the temporal and spatial accuracy (B. Mirus, personal communication, April 17, 2026). All landslide points with less than high confidence were therefore removed from consideration to ensure that the model received only the highest quality dataset available.

There were other considerations before finalizing the landslide points as well. There are a multitude of subclassifications for landslides, with the general movement types like rockfalls, topples, and spreads each having their own subclassifications depending on the material, such as rock or soil (Hungr et al., 2014). For the purpose of this study, all types of landslides were considered regardless of classification. Additionally, as seen with the subset of original landslide points in Figure 2, there were potential issues of oversampling in one section of the study area. This was due to a large storm event in 2018, which seemingly triggered a substantial number of landslides in a southern portion of the study area, between Asheville and Spartanburg. This area therefore featured considerable quantity points within very close proximity to one another (<100 meters). In order to limit the effect of autocorrelation and possible skewness in the dataset caused by the extreme clustering, the points were subjected to a minimum distance filter within ArcGIS, wherein any group of landslide points within 100 meters of each other was limited to a single point, deleting the rest from the dataset. This was to ensure that the model was given a more balanced dataset which reduced, but did not eliminate, the effects of the clustered landslide points. After the reductions in sample size, the final landslide inventory included 703 landslides, with roughly 170 from the 2018 landslide event.

An inventory of known landslide points is vital to a landslide susceptibility model, but just as vital are the inputs which describe the conditions where landslides did not occur. This combination of landslide and non-landslide points was used by the model to understand the conditions where landslides were likely to occur and not occur in. Here, a decision must be made regarding the ratio of landslide to non-landslide points. Higher ratios, such as 1:1 or 1:2, may lead to overfitting of the model (Tang et al., 2023), but lower ratios, such as 1:50 or 1:100, may lead the model to understand that predicting every point as 'no landslide' will be the most mathematically accurate no matter the inputs. With these facts in mind, a ratio of 1:1 was decided to be the most fitting for the study. For every year where a landslide happened, an equal number of non-landslide points were generated, creating a total of 1406 landslide and non-landslide data points for the model to consider. When data extraction occurred, precipitation values were taken from the corresponding year the landslide and non-landslide points belonged to, in order to present as temporally sensitive precipitation data as possible. It is important to note that all other predictors were static and unchanging year-to-year, as they were deemed not as prone to a reliable and relevant shift as precipitation.

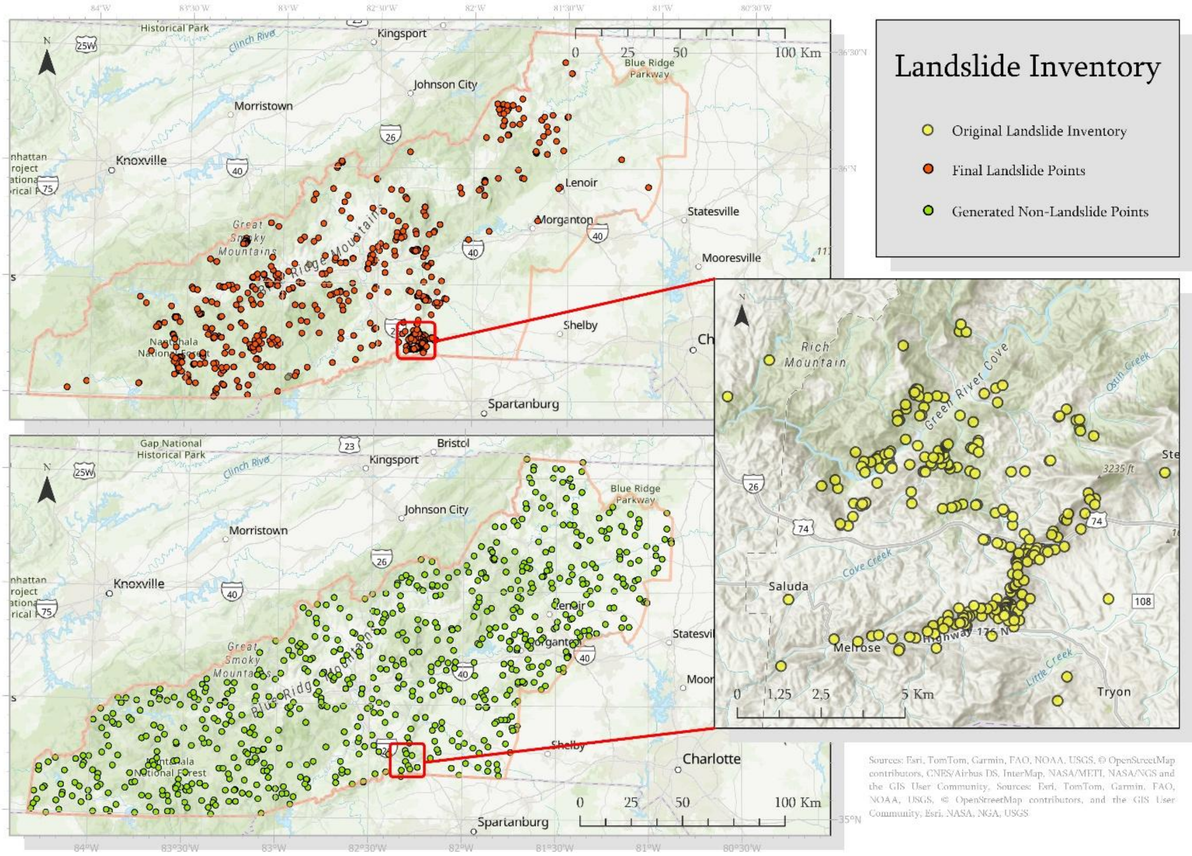

Figure 2: A map displaying the final landslide inventory (red), the generated non-landslide points (green), and a subset of the original landslide inventory for the area (yellow), highlighting the 2018 cluster.

2.2 Predictor Layers

In order to run the landslide susceptibility model, predictor layers would first need to be decided, acquired, and harmonized. The values from these layers would then be extracted by a multi-point extraction function, producing a final dataset of 1406 landslide and non-landslide points with predictor data for each point.

Table 1: A catalog of all predictors used in the study, including resolutions, database names, and sources.

Predictor	Resolution	Database	Source
Annual Precipitation	800m	PRISM	Northwest Alliance for Computational Science and Engineering (NASCE)
Daily Precipitation	1km	Daymet: Daily Surface Weather Data	National Aeronautics and Space Administration (NASA)
Elevation	1-arc second	The National Map (TNM)	United States Geological Survey (USGS)
Slope	-	-	Derived from Elevation
Aspect	-	-	Derived from Slope
Profile Curvature	-	-	Derived from Elevation
Plan Curvature	-	-	Derived from Elevation
Lithology	Vector-based	State Geologic Map Complation (SGMC)	Horton et al., 2017
Distance to Streams	Vector-based	NHDPlus High Resolution Hydrography Framework	Buto & Andersson, 2020
Distance to Roads	Vector-based	National Transportation Dataset	United States Census Bureau
NDVI	375m	EROS Visible Infrared Imaging Radiometer Suite (eVIIRS)	United States Geological Survey (USGS)
Land cover	30m	Annual NLCD Collection 1.0	United States Geological Survey (USGS)
Soil Type	Vector-based	Gridded Soil Survey Geographic (gSSURGO)	United States Department of Agriculture (USDA)

2.2.1 Precipitation

Because of the purpose of this study, to highlight the role that precipitation plays on landslide susceptibility models, the precipitation variable was the most vital predictor. Initially, annual precipitation was used, although later application within the model proved this metric to be of limited use as a predictor of landslides. Annual precipitation data was obtained through the PRISM Group, housed in the College of Engineering at Oregon State University, and part of the Northwest Alliance for Computational Science and Engineering (NACSE). The PRISM dataset contained total annual precipitation for the entire continental United States. The data contained 800m resolution data, a reasonably good resolution for precipitation data on a national scale. As with all predictor layers, this data was downloaded, projected into the correct cell size, extent, and coordinate system (NAD 1983 UTM Zone 17N).

As precipitation intensity is a more relevant metric to evaluate landslides (Guzzetti et al., 2008), maximum daily precipitation was also acquired and used within the model. This precipitation metric was deemed far more representative of landslide susceptibility than annual precipitation data, as many landslides are triggered by short-term extreme rainfall (Li et al., 2025). This data was obtained through the NASA web data source, the Application for Extracting and Exploring Analysis Ready Samples (AppEEARS). The daily precipitation itself belonged to a dataset called Daymet: Daily Surface Weather Data, from NASA's Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC). The maximum daily precipitation was calculated for each year, 2001 to 2020, with a resolution of 1km.

2.2.2 Elevation-based predictors

This study contained five elevation-based predictors. One DEM, and four layers derived from it: slope, aspect, and plan and profile curvature. Elevation was expected to be a relevant predictor for this region due to the mountainous landscapes, and was thus one of the first layers finalized. The DEM was obtained through the United States Geological Survey (USGS), through an online data collection application called The National Map (TNM) v2.0. This provided an elevation layer for the study area whose original coordinate system, NAD 1983, and cell resolution, 1-arc second, became the standard for the entire study. Five separate DEM layers were downloaded, combined, and clipped to the study area to produce the final DEM layer.

Slope was another predictor that was chosen, as the angle of a slope can be an extremely effective predictor for landslide susceptibility. The slope layer was derived from the DEM, as were the aspect and curvature layers. Aspect can play a role in the susceptibility of slope failure, due to increased evaporation and generally higher vegetation on southern-facing slopes (Çellek, 2021). Likewise, curvature can play a large role in the stability of a slope, with the shape of the horizontal slope (plan curvature) and the downhill slope (profile curvature) changing both the susceptibility, intensity, and accumulation of debris. In profile curvature, a convex curvature, or an outward bulge, is represented by negative numbers. Concave, or inward dips, are represented by positive numbers. The neutral value or zero means a flat, linear surface. For plan curvature, the relationship is reversed; a convex curvature is positive, while a concave, valley-like surface is described as negative (Çellek, 2024).

2.2.3 Lithology

The incorporation of lithology in LSMs is vital, as the geologic profile of an area can significantly affect its landslide susceptibility (Alcântara et al., 2025; Cui et al., 2025). For this study, the layer used was the State Geologic Map Compilation (SGMC) Geodatabase of the Conterminous United States by Horton et al., 2017. The data was imported as a polygon layer spanning the continental United States, containing various levels of geologic information. The level used within the study was a general lithology layer, which was subjected to harmonization and rasterization.

2.2.4 Distance to Streams/Rivers

Another predictor that was deemed important to the modelling was distance to streams, as proximity to hydrological features can affect the slope stability (Hamal et al., 2026). The database that was chosen for this area was the NHDPlus High Resolution Hydrography Framework for the Nation database (Buto & Andersson, 2020). There were a few issues with the data that needed addressing. Firstly, as the study area crossed several river basin maps, five separate databases were needed. The hydrographic data was combined into one layer and then clipped. The second issue was that the database was extensive, and included even minor ephemeral waterways. As the final layer would be a simple distance accumulation, all waterways would be treated equally by the model. This meant that even small, ephemeral streams would have just as much input as large, permanent rivers. Therefore, limitations were made to the data. Any watercourse that did not fit the description “streamriver” or “ArtificalPath” (to include manmade alterations to the hydrography), and did not have a geographical name were excluded from the layer. This decision was made to attempt to limit the layer to a reasonable level of hydrographic detail. A rasterization function followed by a distance accumulation function were then run to produce a final distance to streams/rivers predictor layer.

2.2.5 Distance to Roads

Distance to roads is an anthropogenic predictor that can have a significant impact on soil stability and landslide risk (He et al., 2025), and was therefore found to be a vital inclusion in the study. The raw data was acquired from the USGS National Transportation Dataset (NTD), created with data from the U.S. Census Bureau. Two major and minor road layers were downloaded for North Carolina, which were then combined, clipped, and rasterized. A distance accumulation function was then run to produce a final distance to roads layer.

2.2.6 NDVI

Vegetation plays a vital role in landslide susceptibility, with its inclusion integral in producing comprehensive and accurate models (Li and Duan, 2024). It was therefore paramount to include an accurate vegetation index within the model. The USGS EROS Visible Infrared Imaging Radiometer Suite (eVIIRS) was chosen as the dataset to represent this. Specifically, their set of Normalized Difference Vegetation Index (NDVI) data, which is used to highlight the health and density of vegetation at 375m resolution. This dataset was downloaded for the United States, and then subjected to harmonization procedures in order to produce a final NDVI layer for the study area.

2.2.7 Land Cover

Land cover is among the more fundamental predictors in landslide susceptibility, and has been found to influence both the size and mobility of landslides (Pratama et al., 2025). The land cover data used in this study was obtained through the Annual NLCD Collection 1.0, which is a land cover dataset provided by the USGS. The original dataset included four layers that together covered the study area, so they were combined and clipped to the desired coordinate system, cell size, and extent. This dataset suffered from some errors, however, including mislabeling and some restricted NoData issues, which were not discovered until the final model testing. Furthermore, the land cover group names were at the same time found to be problematic and often vague. Group 21, for example, was described in the source data as “Perennial Ice/Snow to Developed, High Intensity”. Similarly, group 81 was defined as “Developed, High Intensity to Open Water”. These issues in labeling would not affect the model functionality, as land cover was a categorical predictor which was divided into sub-predictors based on groups, however, did affect the interpretability of the results. When interpreting the results, a subjective decision was made as to what the group name best described. For example, group 21 was analyzed as Developed, High Intensity. Fortunately, the dataset was still usable within the model despite the limited errors, as labels did not matter for the final model (due to categorization), and the NoData issues were limited enough to be negligible, only affecting 37 out of 1406 points.

2.2.8 Soil Type

Soil type affects soil stability, which in turn affects the landslide susceptibility of an area. Therefore, soil type can be a crucial parameter in landslide modelling (Alcântara et al., 2025). The soil type database used in this study was the Gridded Soil Survey Geographic (gSSURGO) Database, provided by the United States Department of Agriculture (USDA). Though the database provided multiple raster layers and database tables, however, they were not initially linked within the database. Thereby, multiple corrections to the data needed to be applied before a harmonized version of the soil layer could be finalized. Unfortunately, the soil type layer experienced severe errors upon being extracted and imported into the model, likely due to the necessary joining of attribute tables. These errors were, unlike the land cover errors, severe enough that the final soil type layer could not be reasonably utilized by the model, and was left out of the final analysis due to time limitations.

2.3 Logistic Regression Model

The model chosen for this study was based on an analysis of LSMs in current literature. Initially, a fuzzy logic-based design was considered, however, a logistic regression model was ultimately chosen due to a variety of factors. Firstly, although more advanced alternatives such as machine learning or neural network models have been found to be highly accurate, their interpretability is limited (Youssef et al., 2023). Logistic regression is a 'white box' model, whose inner functions and coefficients can be more easily interpreted than 'black box' models like neural networks. This was a better fit for a study more focused on the relative strength and impact of a potentially changing variable than raw model accuracy. Secondly, logistic regression models can be run through simple coding programs such as PyCharm, and can be implemented with minimal programming expertise. Logistic regression models allow for the relative simplicity of less advanced heuristic models like fuzzy logic or AHP models, while excluding the subjectivity of opinion-based weights in favor of mathematical objectivity (Okoli et al., 2023). In fact, some studies have found that although heuristic models are simple, the subjectivity introduced can be problematic for both accuracy and reproducibility (Reichenbach et al., 2018). Thirdly, the time restrictions of the study limited the scope at which models could realistically be designed, trained, and analyzed.

Logistic regression is a statistical model used for binary classification. In the case of landslide susceptibility, this is whether or not a location is likely to experience a landslide. As explained in Atkinson and Massari (1998), a logistic regression model first standardizes the data, then calculates a linear combination for all predictor variables, called the log-odds or logit (Equation 1). It does this by assigning a weight, or coefficient value, to each variable representing their importance and direction of effect through a process called maximum likelihood estimation. The logit value z, derived from Equation 1, can then be subjected to a logistic sigmoid function (Equation 2), compressing the outcomes into a probability range between 0 to 1. Once the coefficients are determined, they can be more easily interpreted by exponentiating them to produce odds ratios (Equation 3), which more easily highlight the size and direction of the predictor's influence. An odds ratio of more than 1 means an increase in landslide susceptibility per unit increase while an odds ratio of less than 1 means a decrease.

$$logit(p_i) = \beta^0 + \Sigma^{n}{}_{i}{}^{=1}\beta_i x_i \quad (1)$$

$$P\ (Y = 1) = 1/(1 + e^{-z}) \quad (2)$$

$$OR = e^{\beta_i} \quad (3)$$

The model was run using Python scripts which utilized the scikit-learn and statsmodels Python packages to run various combinations of model training splits, predictions, and applications. Artificial intelligence models (Google Gemini and Anthropic's Claude) were used for assistance in writing and understanding the code required to run the model, although all decisions about model design and functions were made independently, without AI assistance. The scripts are described in more detail in the following methods section.

2.4 Model Testing

2.4.1 Precipitation Trend Analysis

Before model testing could begin properly, precipitation trends in the study area were examined in order to contextualize the analysis and create a fundamental base for the study's objectives. Annual total and maximum daily precipitation were plotted to test for a significant increase through the study period. The values were taken from the average cell value of each metric's yearly precipitation layers. Additionally, maximum daily precipitation was divided into two decades, 2001-2010 and 2011-2020, so that the distributions in maximum daily precipitation per the decades could be compared and analyzed.

2.4.2 Input Data Analysis and Transformations

On top of trend analyses, an additional set of criteria were analyzed to ensure the data was ready for model testing. The first was a collinearity test, which analyses if multiple input data describe a similar trend. If two or more predictors are collinear, so-called multicollinearity, a logistic regression model would struggle to disentangle the affected predictors, and produce unreliable coefficient values, odds ratios, and p-values. The collinearity test can be interpreted through the VIF score, or Variance Inflation Factor, a measurement of how well an input's variance can be anticipated by another input. The higher the VIF for a predictor, the more redundant the predictor is for the model. A VIF score below 5 is acceptable, whereas above 5 indicates moderate collinearity. A predictor with a VIF score above 10, however, has high levels of collinearity, and any results derived from its use could be problematic. As multicollinearity within the model could pose an issue for the objective of this study, this test was run on all predictor data to diagnose potential issues pre-model testing. Decisions could then be made regarding the relevance and the potential exclusion of predictors found to be collinear.

The next analysis involved evaluation of predictor distributions. When a logistic regression model functions, it assumes a linear relationship between the continuous predictors and the log-odds of the outcome, meaning equal increments in the predictor should produce equal changes in the log-odds. If a predictor has a heavily skewed distribution, the model struggles to apply this predictor, as extreme values at the tail end of the distribution can significantly bias the calculations. Before the data can be introduced into the model, the distributions of each predictor were examined for skewness and, if found to be necessary, were transformed to a more balanced distribution.

2.4.3 Initial Model Testing

The first round of model tests were done to gain an initial understanding of the model's accuracy and the coefficient values and significance of each predictor. These coefficients, translated as odds ratios, would form the baseline understanding of how important each predictor was to the model. A coefficient value or odds ratio could not alone be used to analyze the predictor's strength within the model, however, as even a high coefficient value could be due to random chance within the model. These statistics would therefore be supplemented with p-values for all predictors, showing the statistical significance of each coefficient. A large coefficient with a low p-value (<0.05), for example, would indicate a strong and influential predictor within the model. Additionally, because of the standardization function performed by the model on all continuous data within the dataset, the effect of each continuous predictor could be measured by effect per standard deviation. Categorical data, conversely, would be transformed into a separate predictor per group within the model, effectively turning the lithology layer, for example, into several different predictors by group: sedimentary, metamorphic, etc. Their effect would then be measured not against a standard deviation, but as a direct comparison to the first group, the so-called reference group.

The initial model tests were run using a train-test split, which divided the dataset (the extracted landslide and non-landslide points and their associated predictor values) into a set of training and validation data, using a 70%/30% split. The training data would be used to fit the model and estimate the relationship between predictor data and landslide occurrence, while the validation data would be used to assess the predictive performance of the model. Numerous data and accuracy metrics were outputted to assess the model, including the aforementioned coefficient values, odds ratios, and p-values, as well as the AUC metric. The AUC, or Area Under the Curve, is a measurement of how well the model discriminates between two classes, in this case: landslide and non-landslide. Among the most commonly reported validation metrics in landslide susceptibility models (Reichenbach et al., 2018), AUC calculates the proportion of landslides correctly predicted against the proportion of non-landslides incorrectly predicted across all possible classification thresholds from 0 to 1, allowing for an evaluation of model accuracy.

Additionally, because the model would simultaneously conduct the requested test on both the validation data as well as the same training data it was fitted on, an analysis of potential over-fitting could also be done. Over-fitting could be inferred from a large discrepancy between the training and the validation data AUC scores, indicating the model simply memorized the training data instead of developing a true understanding of the relationships between predictors and landslides. To ensure that the results were not skewed by a potentially favorable random sampling, the script was run five times, each with a new random state within the Python code. These five iterations were then averaged to obtain one initial testing result, which was used to garner understanding of the initial model's coefficients and general accuracy. Additional measures of fit would be used to judge model accuracy, such as pseudo R2, a measurement of explained variance akin to the R2 metric in linear regression, and log-likelihood, a probability metric used to compare models (Allison, 2014).

2.4.4 AIC Testing

Once an analysis was conducted on the initial model, different combinations of predictors could then be tested to discern the most efficient model. Predictors were removed based on their coefficient scores from the initial testing, although different combinations even excluded several of the most significant predictors to ensure methodological soundness. These combinations were tested by removing the test-train split and allowing the model to train on the 100% of the dataset. The lack of a training split meant the output data (AUC, coefficients, p-values) would be optimistic, and thus invalid compared to the more fairly tested 70/30 model. This did not matter for this test, however, as the only metric used to judge the combinations was the AIC, or Akaike Information Criterion, a metric solely built to compare models, not accuracy. The AIC compares models based on how well they fit the observed data per predictor used by punishing the addition, and rewarding the removal, of predictors that do not add predictive power to the overall fit. By comparing the AIC scores of various combinations of predictors, a final model could be produced, with all unnecessary predictors removed. Judging which predictors remained in the final model could garner a further understanding of the significance of the predictors used in this study. Following this, the final model could be applied and its results evaluated to determine potential changes from the initial model's accuracy, coefficient values, and overall fit.

2.4.5 Landslide Susceptibility Map

With the final, most efficient model identified, a landslide susceptibility model could also be produced. This was done by applying the final model to every cell within the study area, a total of nearly 32 million. This was done to showcase the model's understanding of the area and provide evidence for the efficacy of the predictors used within the model.

2.4.6 Precipitation Sensitivity

In order to analyze the effect a change in the precipitation predictor could potentially have on the model's projection of landslide occurrence, the model was then tested using simulated maximum daily precipitation values. Using the standard 70%/30% train-test split, the precipitation data within the validation set was altered to simulate different precipitation scenarios. The values were altered stepwise by a multiplier: 25%, 50%, 75%, 100% (as a baseline comparison), 125%, 150%, 175%, and 200%. For each precipitation multiplier, a landslide prediction would be made for the validation dataset, with any point at 50% susceptibility or above being classified as a landslide. The predictions were made across five random states per multiplier, for a total of 40 tests. The number of points predicted to experience landslides were then averaged, plotted, and analyzed to assess the significance a changing precipitation variable would have on the model's landslide predictions.

2.4.7 Decadal Model Analysis

To attempt to answer one of the main objectives of the study, whether a changing precipitation variable can lead to a change in accuracy in LSMs, a decadal comparison of models was conducted. By sorting the dataset into two subsets based on year, two decadal models could be trained; Model A (2001-2010) and Model B (2011-2020). The models would be trained using a train-test split (70%/30%), and would be tested against its own training and validation data, as well as 100% of the other decade. Additionally, each model performed a landslide prediction of the other decade's dataset, using the same 50% classification as the sensitivity analysis. Comparing the results of each model's decadal cross-validation could therefore provide a basic understanding of how future changes in precipitation data could potentially affect an LSM trained on historical data within this study area.

3 Results

3.1 Precipitation Trend Analysis

Both total annual and maximum daily precipitation were tested for significant increases through the study area. This was done by taking the average raster value for every year, then applying a regression analysis to the trends. The tests revealed a significant increase ($p<0.05$) in annual precipitation within the study area in the period 2001 to 2020. Although there was variance within the trend, with an explained variance (R2) of 0.21, the test found a p-value of 0.04, indicating that the climate in the study area was significantly changing within the study period. The regression analysis did not, however, find a statistically significant ($p<0.05$) increase in the maximum daily precipitation across the study period. With a minimal explained variance of 0.006 due to extreme variance, the test produced a p-value of 0.76, confirming that there was no statistically significant increase in max daily precipitation in the study area from 2001 to 2020.

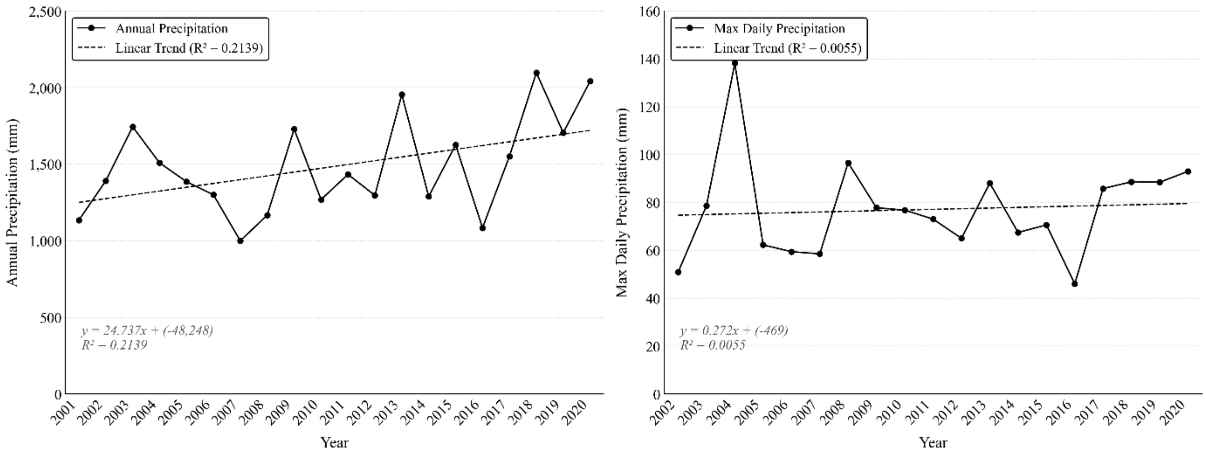

Figure 3: Graphs illustrating both the average annual and max daily precipitation across the entire study area for the years 2001 to 2020. Trend lines, equations, and R^2 values are included.

A comparison in maximum daily precipitation values between decades was also performed. As illustrated in Figure 4, Decade A (2001-2010) experienced increased variance and higher maximum precipitation values than Decade B (2011-2020). The result of a two-sample Kolmogorov-Smirnov test (Figure 5), a non-parametric statistical test which analyses the difference between cumulative distributions of samples, corroborates the significant difference between decades (D = 0.323, $p<0.001$). Although a minimal p value is expected when comparing samples of this size in a Kolmogorov-Smirnov test, the D statistic is still useful to quantify and confirm the visual difference seen between the distributions.

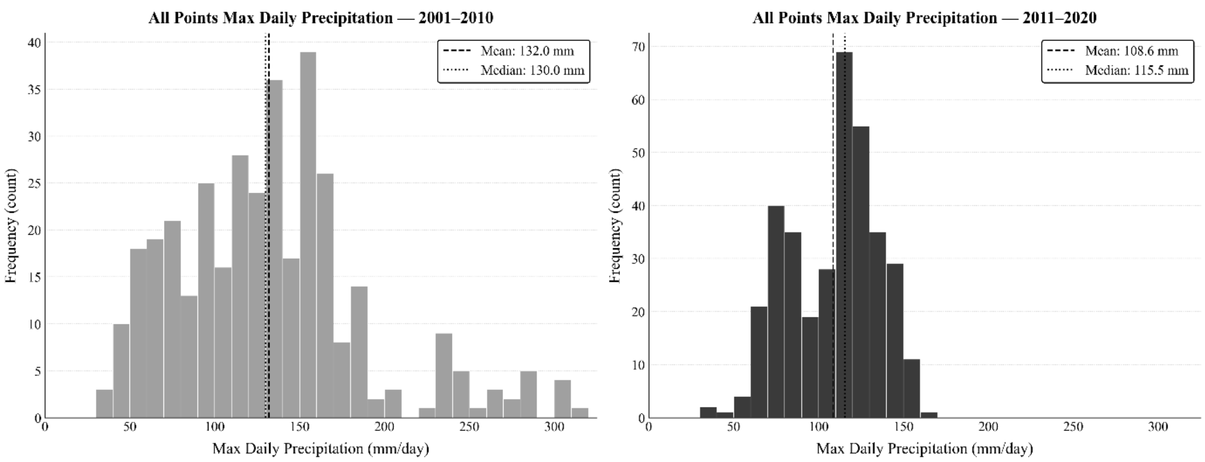

Figure 4: A histogram comparison between 2001-2010 and 2011-2020, highlighting the increased variance in the earlier decade.

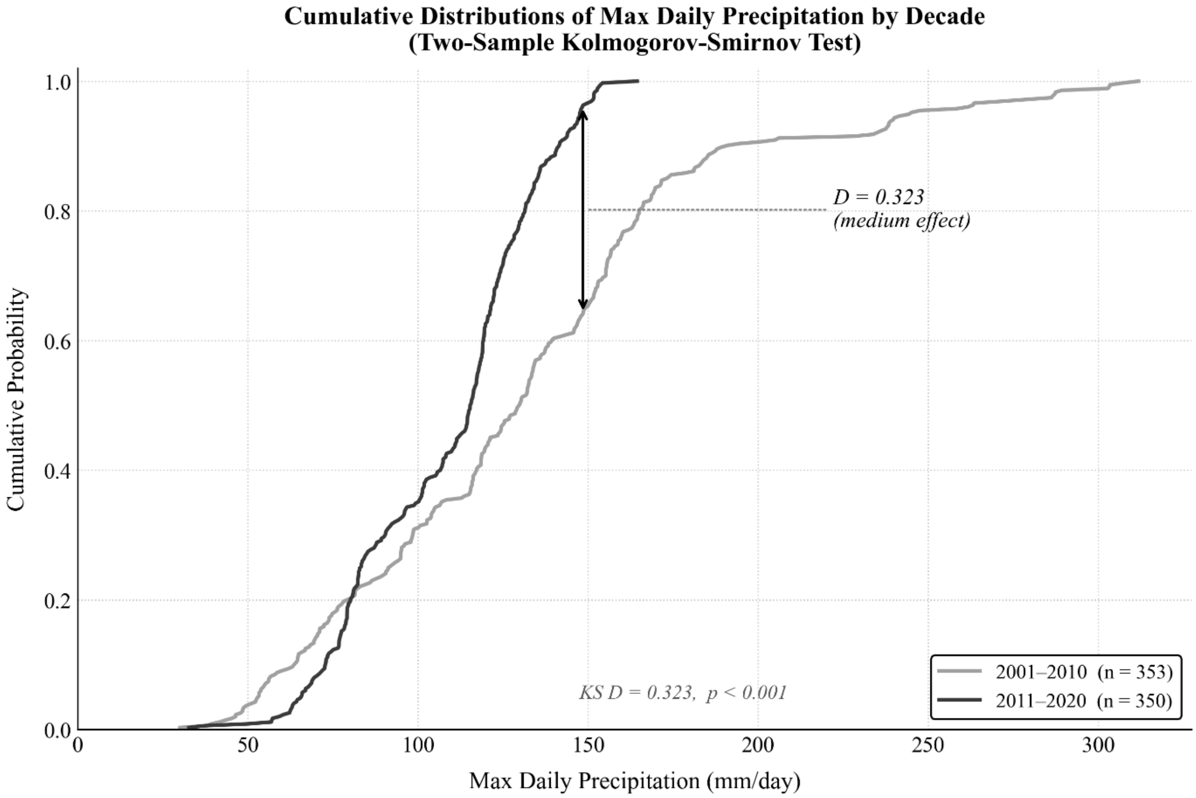

Figure 5: The results of a two-sample Kolmogorov-Smirnov test on the two sets of max daily precipitation data from each decade, highlighting the significant difference in precipitation variance. The D statistic and p-value of the test are included.

3.2 Input Data Analysis and Transformations

Before model testing, the input predictor data were tested for collinearity and skewness. Upon testing of the original dataset, significant multicollinearity was found. As presented in Table 2, both slope and DEM displayed moderate levels of collinearity, while the NDVI and the annual precipitation exhibited high levels of collinearity. A new test was run with an updated predictor dataset, removing the DEM predictor and replacing the precipitation variable with maximum daily precipitation instead. This new combination produced greatly improved collinearity levels, although a moderate collinearity between the new precipitation variable and NDVI was still observed.

Table 2: A table showing the result of the collinearity tests on the initial and updated predictor datasets, with DEM removed and annual precipitation replaced with max daily precipitation.

Initial Predictor Dataset			**Updated Predictor Dataset**		
Predictor Variable	**VIF Score**	**Collinearity**	**Predictor Variable**	**VIF Score**	**Collinearity**
Slope	6.1042	Moderate	Slope	4.8157	Acceptable
DEM	7.9294	Moderate	Aspect	3.9880	Acceptable
Aspect	4.0943	Acceptable	Profile Curvature	1.0580	Acceptable
Profile Curvature	1.0584	Acceptable	Plan Curvature	1.0457	Acceptable
Plan Curvature	1.0457	Acceptable	Lithology	2.3644	Acceptable
Lithology	2.4942	Acceptable	Soil Type	2.1384	Acceptable
Soil Type	2.3081	Acceptable	Distance to Streams/Rivers	2.6683	Acceptable
Distance to Streams/Rivers	2.7497	Acceptable	Land Cover	1.0505	Acceptable
Land Cover	1.0516	Acceptable	NDVI	8.9597	Moderate
NDVI	11.5899	High	Distance to Roads	1.3023	Acceptable
Distance to Roads	1.3117	Acceptable	Max Daily Precipitation	5.9743	Moderate
Annual Precipitation	12.5424	High			

VIF threshold: <5 acceptable, 5–10 moderate, ≥10 high collinearity.

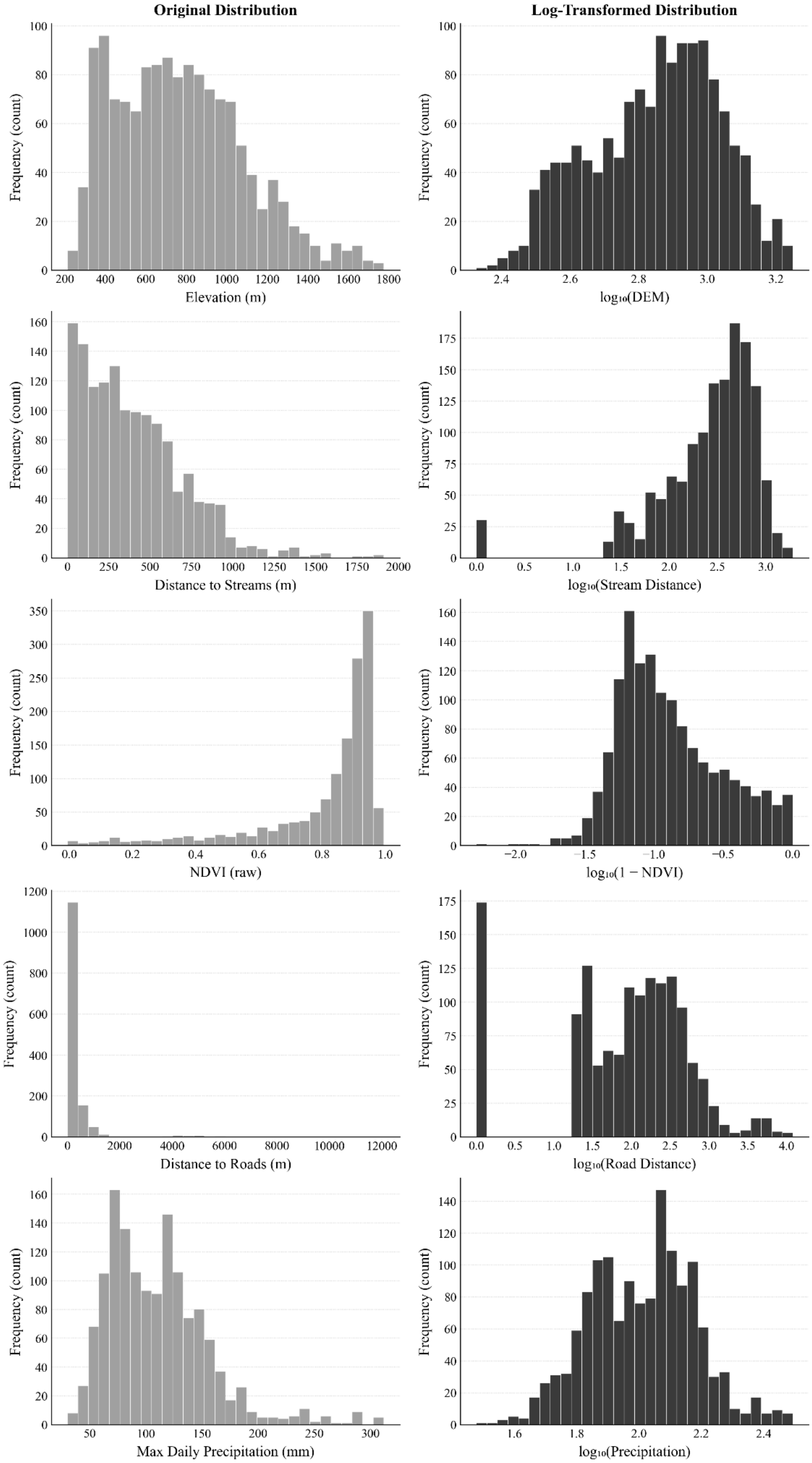

Figure 6: The predictor histograms for DEM, distance to streams/rivers and roads, NDVI, and maximum daily precipitation before and after log transformations were applied.

3.3 Initial Model Testing

Testing was done on five iterations of the initial model with all relevant predictors included, and then averaged to get a set of initial test results. The model results highlighted a few key initial findings. Seven of the predictors (nine if counting sub-groups) were found to be statistically significant ($p<0.05$), with four exhibiting particularly extreme statistical significance with miniscule p-values ($p<0.0001$): slope, max daily precipitation, distance to roads, and land cover 43 (Grassland/Herbaceous). Of these, slope demonstrated the strongest effect on the model. With a coefficient score of 1.49, meaning an odds ratio of 4.45, slope had the effect of a 345% increase in landslide susceptibility per standard deviation (SD) increase of slope. Maximum daily precipitation also exhibited significant influence over the model, with a 59% increase in susceptibility per standard deviation increase. Distance to roads showed a 51.2% decrease per SD increase, or that landslide susceptibility decreased the further from a road a point was. Other predictors showcased statistically significant influence as well, for example profile curvature. In the initial model testing, one SD increase in profile curvature, or an increase in concavity, increased landside susceptibility by 24%. Aspect and NDVI also both showed roughly 19% decrease in susceptibility per SD increase.

Of the categorical predictors, lithology proved statistically insignificant to the model, with no lithology type proving more effectual than the reference. Land cover, however, did show relevant differences between the predictor groups, with land cover 22, or Barren Land, indicating a near 400% increase compared to the reference layer of Developed, High Intensity. Moreover, land cover groups 41 (Mixed Forest) and the aforementioned 43 (Grassland/Herbaceous), constituted a significant decrease in susceptibility when compared to the reference layer.

Table 3: The coefficient scores, odds ratios, effect sizes, and significance for all predictors used within the initial model. Slope is highlighted as the most significant predictor, alongside maximum daily precipitation and distance to roads.

Predictor	Coefficient	Odds Ratio	P-Value	Significance	Effect Size
Slope	1.4907	4.4491	0.0000	Sig. Increases	+344.9% per SD increase
DEM	-0.0357	0.9667	0.6160	Not Significant	—
Aspect	-0.2072	0.8136	0.0289	Sig. Decreases	-18.6% per SD increase
Profile Curvature	0.2169	1.2443	0.0271	Sig. Increases	+24.4% per SD increase
Plan Curvature	0.1051	1.1122	0.3741	Not Significant	—
Distance to Streams/Rivers	-0.0136	0.9866	0.8513	Not Significant	—
NDVI	-0.2126	0.8093	0.0251	Sig. Decreases	-19.1% per SD increase
Distance to Roads	-0.7173	0.4885	0.0000	Sig. Decreases	-51.2% per SD increase
Max Daily Precipitation	0.4603	1.5871	0.0000	Sig. Increases	+58.7% per SD increase
SGMC_2	0.3533	1.4300	0.0949	Not Significant	—
SGMC_3	-1.1589	0.3395	0.0560	Not Significant	—
SGMC_4	0.7351	2.0996	0.2836	Not Significant	—
SGMC_8	-0.1343	0.8835	0.6180	Not Significant	—
SGMC_999	-0.8673	0.4257	0.0688	Not Significant	—
Land Cover_22	1.5458	4.9078	0.0184	Sig. Increases	+390.8% vs reference
Land Cover_41	-0.9442	0.3932	0.0044	Sig. Decreases	-60.7% vs reference
Land Cover_42	-0.8311	0.4745	0.3292	Not Significant	—
Land Cover_43	-1.6072	0.2014	0.0000	Sig. Decreases	-79.9% vs reference
Land Cover_81	-0.5726	0.5687	0.2207	Not Significant	—
Land Cover_999	0.1473	1.1837	0.7707	Not Significant	—

Sig. = Significantly

When examining general model accuracy, the AUC scores can be assessed. The mean validation AUC score was 0.84 with only a 0.01 standard deviation across five iterations of the model. Similarly, the training AUC across five iterations was 0.86 with a 0.0005 SD, meaning an average AUC gap of only 0.02. Additionally, a pseudo R2 value of 0.35 falls well within the scale of acceptable accuracy (0.2 to 0.4), while a log-likelihood of -444.8 can be used to compare to the later finalized model.

Table 4: A table presenting accuracy metrics for the initial model. Notable is the high accuracy and low AUC gap between the validation and training data, indicating no overfitting.

Metric	Mean	Std Dev	Min	Max
Validation AUC (30% holdout)	0.8436	0.0156	0.8142	0.8595
Training AUC (70% training set)	0.8682	0.0050	0.8613	0.8764
AUC Difference (gap)	0.0246	—	—	—
No. Observations (training)	984	—	—	—
Pseudo R-squared	0.3476	0.0098	0.3347	0.3641
Log-Likelihood	-444.87	6.76	-453.78	-433.45
N Runs	5	—	—	—

Results averaged across 5 runs (random states 41–45), 70/30 train/validation split.

3.4 AIC Testing

In order to solidify a final model, an AIC test was performed to determine the most efficient combination of predictors. The first test involved a model including all predictor data, which exhibited an AIC score of 1336.13 (Table 5). This score was bettered by the next version, Model 2, which removed the distance to streams/rivers layer. The additional removal of both the DEM and plan curvature layers led to a model with the lowest AIC score, 1331.48 for Model 4 (DrDPl model). The further removal of various predictors were found to instead increase the AIC score, meaning a decrease in fit per predictor. For methodological soundness, additional variations were run, including the removal of highly significant predictors, however no other predictor combination reached an equal or lower AIC score than Model 4. Notable was the large increase in AIC when removing either the max daily precipitation or the distance to roads predictors from Model 4, with an increase of 35.4 (Model 13) and 76.07 (Model 14) respectively. Additionally, the removal of the lithology predictor increased AIC, despite insignificance during initial testing. Following the selection of the optimal model configuration, the finalized model was then tested using the same methods as the initial model. Upon testing, the final model displayed only a marginal increase in AUC, predictor significance, and overall fit (Appendix A), highlighting that the removal of insignificant predictors had minimal impact on overall model performance.

Table 5: The results of the AIC testing for all predictor combinations, with model number and names, AIC scores, and layers removed compared to the initial model. The final model combination is bolded.

#	Model	AIC Score	Removed Predictors
1	Initial Model	1336.13	None — full model
2	Model −Dr	1334.15	Dist. to Rivers
3	Model −DrD	1332.97	Dist. to Rivers, Elevation (DEM)
4	**Model −DrDPl***	**1331.48**	**Dist. to Rivers, Elevation (DEM), Plan Curvature**
5	Model −DrDPlL	1342.29	Dist. to Rivers, Elevation (DEM), Plan Curvature, Lithology
6	Model −DrDPlLLc	1389.13	Dist. to Rivers, Elevation (DEM), Plan Curvature, Lithology, Land Cover
7	Model −DrDPlLLcA	1399.13	Dist. to Rivers, Elevation (DEM), Plan Curvature, Lithology, Land Cover, Aspect
8	Model −DrDPlLLcAPr	1407.76	Dist. to Rivers, Elevation (DEM), Plan Curvature, Lithology, Land Cover, Aspect, Profile Curvature
9	Model −DrDPlA	1340.02	Dist. to Rivers, Elevation (DEM), Plan Curvature, Aspect
10	Model −DrDPlAPr	1347.35	Dist. to Rivers, Elevation (DEM), Plan Curvature, Aspect, Profile Curvature
11	Model −DrDPlN	1339.70	Dist. to Rivers, Elevation (DEM), Plan Curvature, NDVI
12	Model −DrDPlIRd	1407.55	Dist. to Rivers, Elevation (DEM), Plan Curvature, Roads
13	Model −DrDPlPm	1366.88	Dist. to Rivers, Elevation (DEM), Plan Curvature, Max Daily Precipitation
14	Model −DrDPlS	1669.24	Dist. to Rivers, Elevation (DEM), Plan Curvature, Slope

** Final model (lowest AIC score)*

3.5 Landslide Susceptibility Map

In order to test the applicability of the model, the trained and fitted model was used to produce a landslide susceptibility map of the study area. As illustrated in Figure 7, the lowland regions along the eastern part of the study area around Lenoir and Morganton are characterized by consistently low susceptibility. A clear elevation change can be seen on the susceptibility map as it extends northeast across the study area from the south, dividing the lowland regions from the Blue Ridge Mountain chain. Similarly, the valley region stretching along Highway 26 from the south-central portion of the study area into the central region, through Asheville, experiences a similar level of consistent low susceptibility. The southwest region of the study area, along the borders to Tennessee and South Carolina, are also associated with generally low susceptibility.

The most pronounced area of high susceptibility, conversely, is the southwest-central region stretching from the Nantahala National Forest to Asheville. A similarly high level of susceptibility can be seen north of Lenoir in the northeast portion of the study area. A general pattern can be gleaned from the map, with the regions along the Blue Ridge Mountains facing a considerably higher risk of landslides than lowland, mainly eastern areas.

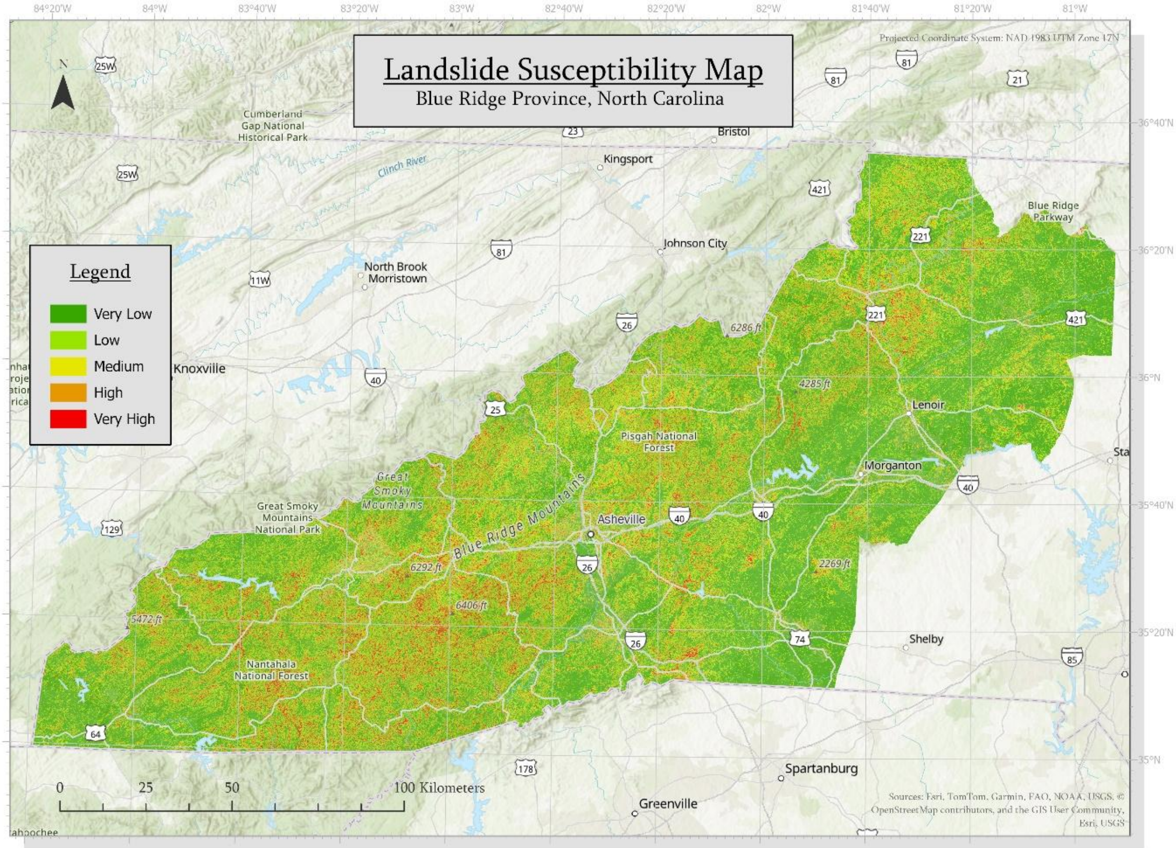

Figure 7: A landslide susceptibility map of the study area, highlighting landslide-prone areas based on 20% intervals of landslide susceptibility. Notable is the low susceptibility within the eastern lowland areas, as well as the high susceptibility found in the mountainous areas.

3.6 Precipitation Sensitivity

A sensitivity analysis was also conducted for various precipitation scenarios, testing stepwise multipliers of 25% to 200% of the original max daily precipitation values. The results, presented in Table 6, can be assessed to develop a deeper understanding of the strength that precipitation has in the model, as well as provide basic estimation of how changes in precipitation in the study area could potentially change the area's landslide susceptibility. As seen in TABLE, stepwise increases in maximum daily precipitation led to an increasing landslide prediction rate within the model. With a 50% increase in maximum daily precipitation, the model's landslide prediction increases by 10%, whereas a doubling of precipitation leads to a 16% increase according to the model. Larger changes are seen in the opposite way, with a 50% decrease resulting in a 14% decrease in landslide predictions. When plotted, the data can be seen as a logarithmic relationship, matching the log transformation applied to the precipitation data pre-model.

Table 6: A table showing the results of the precipitation sensitivity analysis, with average counts for the predictions given five iterations, as well as changes from baseline and stepwise increases. A logarithmic increase in landslide occurrence is observed per stepwise multiplier.

Metric	**0.25x**	**0.5x**	**0.75x**	**1.0x**	**1.25x**	**1.5x**	**1.75x**	**2.0x**
Predicted Landslide (avg count)	93.2	142.8	181.4	202.8	226.6	243.0	257.8	270.2
Predicted No Landslide (avg count)	328.8	279.2	240.6	219.2	195.4	179.0	164.2	151.8
Predicted Landslide % (avg)	**22.09%**	**33.84%**	**42.99%**	**48.06%**	**53.70%**	**57.58%**	**61.09%**	**64.03%**
From Baseline (%)	-25.97%	-14.22%	-5.07%	Baseline	+5.64%	+9.52%	+13.03%	+15.97%
Stepwise Increase in % Points	—	+11.75	+9.15	+5.07	+5.64	+3.88	+3.51	+2.94

Baseline = 1.0x (actual precipitation). Susceptibility is the modelled proportion of points predicted as landslide.

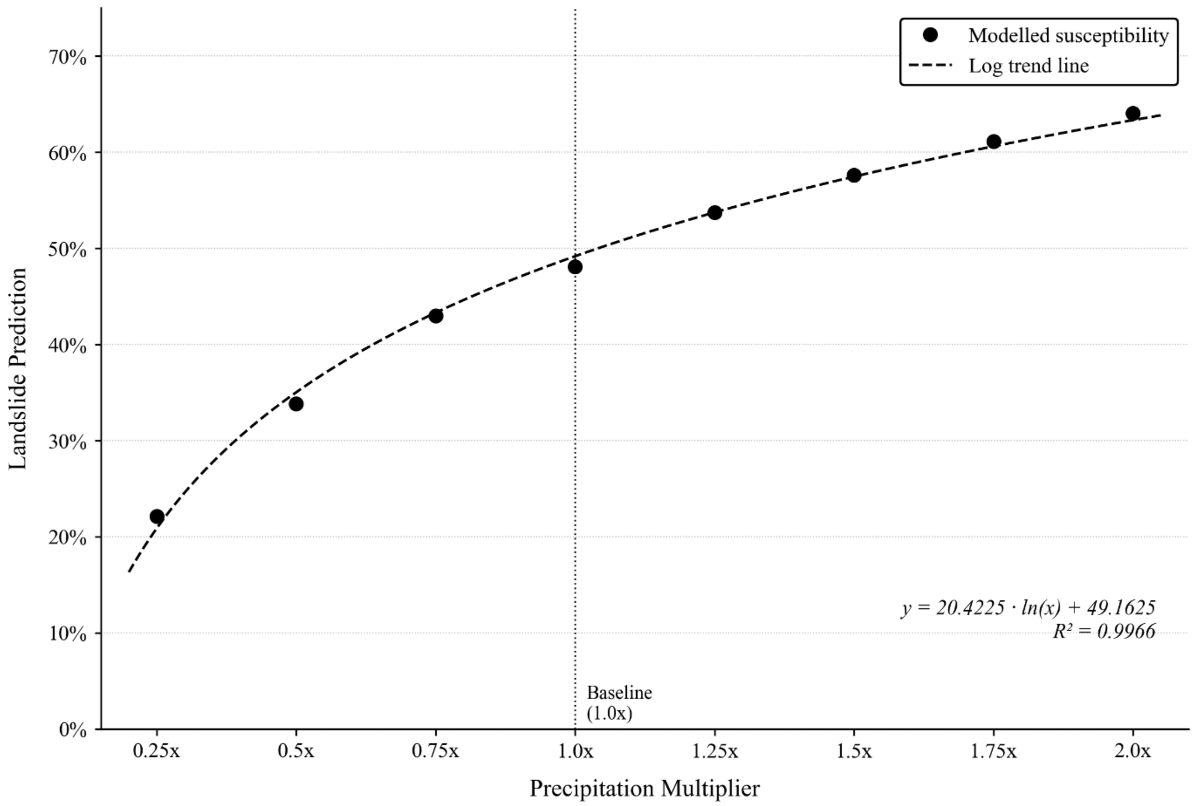

Figure 8: Graph illustrating the logarithmic increase in landslide prediction per stepwise increase in the precipitation multiplier.

3.7 Decadal Model Comparison

The results of the decadal cross-validation indicated a decrease in accuracy for both models when predicting cross-decade compared to the accuracy within their own decade. The overall accuracy of Model A, based on precipitation data from the decade with the more extreme precipitation values (2001-2010), dropped around 4%, while the AUC declined from 0.86 to 0.81 when cross-validating into the less variable precipitation decade (2011-2020). Model B, trained on the more stable decade, experienced a larger drop of roughly 8% overall accuracy and 0.08 decline in AUC when validating into the more extreme decade.

The overall accuracy was not the only metric measured, however. As seen in FIGURE, Model A experienced a sharp decline in sensitivity, or landslides correctly identified, when comparing within-decade to cross-decade (77% to 65%). Correspondingly, Model B's specificity, or non-landslides correctly identified, decreased by over 17% (77% to 60%) when comparing within to cross-decade. This is again evidenced when examining the confusion matrix. Model A's cross-validation yielded a decrease in proportional true positives while an increase in false negatives (Type II errors). Conversely, Model B demonstrated a marked decline in true negatives and an increase in false positives (Type I errors).

Table 7: The results of the decadal model cross-validation, with accuracy and prediction changes. The percentage values in the confusion matrix are the proportion of predictions within the same column. Notable is the decrease in accuracy for both decadal cross-validations, as well as the changes in sensitivity and specificity.

Metric	**Model A (2001–2010) Within-Decade**	**Model A → B (Cross-Decade 2011–2020)**	**Model B (2011–2020) Within-Decade**	**Model B → A (Cross-Decade 2001–2010)**
AUC				
AUC — Training (70%)	0.876	—	0.877	—
AUC — Validation	**0.856**	**0.813**	**0.879**	**0.802**
Accuracy & Classification				
Overall Accuracy (%)	77.36%	73.60%	79.62%	71.33%
Sensitivity / Recall (%)	76.60%	64.63%	82.10%	82.66%
Specificity (%)	78.11%	82.57%	77.14%	60.00%
Precision (Landslide)	0.778	0.788	0.783	0.674
F1 Score (Landslide)	0.772	0.710	0.802	0.743
Confusion Matrix (avg counts)				
True Positives (avg)	81.2 (38.3%)	226.2 (32.3%)	86.2 (41.0%)	291.8 (41.3%)
True Negatives (avg)	82.8 (39.1%)	289.0 (41.3%)	81.0 (38.6%)	211.8 (30.0%)
False Positives (avg)	23.2 (10.9%)	61.0 (8.7%)	24.0 (11.4%)	141.2 (20.0%)
False Negatives (avg)	24.8 (11.7%)	123.8 (17.7%)	18.8 (9.0%)	61.2 (8.7%)

Bold = key performance metric. Cross-decade validation tests transferability of each model to the opposite decade.

The coefficient analysis of the models highlights notable differences. Although Model B's coefficients look similar to both the initial and final full dataset models, Model A finds significance in only the distance to roads and slope predictors, although both exhibit a marked increase in effect size. Notable are profile curvatures and land cover, both of which were close to statistically significant ($p<0.05$), as well as the insignificance of the max daily precipitation predictor, showing the potential impact of the high variance of precipitation within this decade.

Table 8: Showing the coefficients, odds ratios, and significance of the decadal models. Notable is the lack of significance in most predictors in Model A, indicating proportionally higher influence of the landslide points from the second decade.

	Model A (2001–2010)				Model B (2011–2020)			
Predictor	**Odds Ratio**	**P-Value**	**Significance**	**Effect Size**	**Odds Ratio**	**P-Value**	**Significance**	**Effect Size**
Slope	5.281	0.000	Sig. Increases	+428.0% per SD	4.081	0.000	Sig. Increases	+308.1% per SD
Aspect	0.864	0.250	Not Significant	—	0.748	0.032	Sig. Decreases	-25.2% per SD
Profile Curvature	1.311	0.051	Not Significant	—	1.085	0.488	Not Significant	—
NDVI	1.048	0.578	Not Significant	—	0.674	0.004	Sig. Decreases	-32.6% per SD
Distance to Roads	0.309	0.000	Sig. Decreases	-69.1% per SD	0.636	0.002	Sig. Decreases	-36.4% per SD
Max Daily Precipitation	1.274	0.108	Not Significant	—	2.346	0.000	Sig. Increases	+134.6% per SD
Land Cover 41	0.453	0.066	Not Significant	—	0.267	0.009	Sig. Decreases	-73.3% vs ref
Land Cover 43	0.376	0.056	Not Significant	—	0.128	0.001	Sig. Decreases	-87.2% vs ref
Land Cover 81	1.888	0.326	Not Significant	—	—	—	Not in model	—
Land Cover 999	0.957	0.669	Not Significant	—	0.540	0.254	Not Significant	—
SGMC 2	1.499	0.189	Not Significant	—	1.566	0.168	Not Significant	—
SGMC 3	—	—	Not in model	—	0.581	0.394	Not Significant	—
SGMC 8	0.889	0.644	Not Significant	—	1.711	0.267	Not Significant	—
SGMC 999	0.803	0.381	Not Significant	—	0.321	0.069	Not Significant	—

Sig. = Significantly. Odds ratios and p-values are means across 5-fold cross-validation runs. Effect sizes per SD increase (continuous) or vs. reference category (categorical).

4 Discussion

4.1 Precipitation Trend Analysis

An observed increase in annual precipitation in the study area contextualized the objectives of analyzing changes in accuracy of LSMs across shifts in precipitation. The insignificant trend seen in maximum daily precipitation, however, limited the ability of the study to examine the effect a yearly change would have on a landslide susceptibility model's accuracy. Nevertheless, notable differences were observed in the variance and distribution of extreme precipitation values when comparing the two decades of the study period. This disparity provided the basis for a comparative analysis between decadal models.

4.2 Input Data Analysis and Transformations

As stated in the results, multiple predictors exhibited moderate to high collinearity in the initial analysis. The collinearity of slope and elevation was anticipated, given slope is derived from the elevation layer. Nevertheless, the collinearity meant the DEM was excluded from initial tests. Notably, the annual precipitation and NDVI predictors displayed exceptionally high collinearity. This was likewise unsurprising, given increased growth rates often accompany higher precipitation rates. This did, however, further highlight the limited use of annual precipitation as a landslide predictor. These findings, alongside literature identifying extreme rainfall as a stronger trigger of landslides than cumulative rainfall (Zhao et al., 2025), informed the decision to replace annual precipitation within the model with the intensity-based maximum daily precipitation. This replacement, and the exclusion of the DEM layer, lowered collinearity to more acceptable levels for data testing. Notable still was the moderate collinearity between maximum daily precipitation and NDVI, which would have to be considered when analyzing model results.

4.3 Initial Model Analysis

The most significant predictor within the initial model was determined to be slope, with an effect nearly six times larger than the nearest predictor. This finding was consistent with expectations given the predictor's direct impact on soil stability (Hamal et al., 2026) and the prominence of the mountainous terrain within the study area. It was further corroborated by a study examining the landslide susceptibility of a smaller part of the study area, which likewise found slope to be the most significant predictor within the model (Agboola et al., 2024). The second strongest predictor was maximum daily precipitation, a result that directly addressed the first major objective within the study; identifying the significance of precipitation as a predictor within the model. The extreme significance of precipitation ($p<0.0001$) indicated that it was a vital predictor for landslide susceptibility modelling within the study area, aligning with prior literature on the subject (Bezerra et al., 2025). Distance to roads was similarly identified as an especially significant predictor, with susceptibility decreasing considerably the larger the distance from a road. This is consistent with past literature on the subject, which found that road construction could disturb the equilibrium of natural slopes and play a crucial role in their failure (Kaushal et al., 2025). This result could also, however, reflect a potential sampling bias within the landslide inventory. Because landslides occurring near roads are more likely to be reported and documented accurately, the confidence metric applied by the authors of the inventory could disproportionately favor more accessible and clearly reported incidents. Therefore, it is possible that the use of only high-confidence landslides within the study could have inadvertently biased the samples toward more road-adjacent landslides.

Both the vegetation index and the land cover sub-predictors mixed forest and grassland/herbaceous highlighted similar relationships, with an increase in vegetative cover associated with decreased landslide susceptibility. This reaffirms the consensus in literature, which finds vegetation increases soil stability and shear strength (Lann et al., 2024). Notably, according to the model, a shift from high density developed to barren land would see a dramatic increase in landslide susceptibility. Although urbanized areas are often associated with high landslide susceptibility (Rohan et al., 2023), some studies have found susceptibility in both developed and barren land (Hamal et al., 2026). This suggests the substantial difference in susceptibility found in the results may be more reflective of potential inconsistencies within the original land cover classification than a genuine geophysical relationship.

Model accuracy was also calculated, demonstrating that the initial model was highly accurate, with high stability over five random iterations. Additionally, the minimal AUC gap between training and validation sets confirmed there were no indications of overfitting within the model. These metrics illustrated the model's capability of accurately calculating susceptibility of landslides within the study area.

4.4 AIC Testing

AIC testing revealed the most efficient configuration of predictors to be Model 4, with the distance to streams/rivers, DEM, and plan curvature removed. Given these predictors displayed limited value in the initial model testing, the minimal change in accuracy between the initial and final model was an expected outcome. Notable was that the removal of the otherwise insignificant lithology predictor reduced predictive ability in the model. This can be potentially explained by the handling of categorical predictors in the model, as each group within the layer is assigned to a unique sub-predictor. Although these may not present significant predictive power on their own, their combined influence may be enough to significantly affect the model when removed entirely. Additionally, the removal of both the precipitation as well as distance to roads predictors resulted in a considerable drop in predictive ability, again demonstrating the significance of the precipitation variable within the model. Notably, however, the removal of the distance to roads layer produced a larger increase in AIC than the removal of max daily precipitation, despite the larger effect size of the former. This is likely explained by the moderate collinearity between maximum daily precipitation and NDVI. Because the AIC score considers predictive power per distinct variable, predictors with moderate levels of collinearity may be considered as providing less independent predictive power than their coefficient and effect size may suggest. The low collinearity of the distance to roads predictor likely therefore meant its unique predictive contribution was more fully reflected in the AIC score than maximum daily precipitation.

4.5 Landslide Susceptibility Map

The spatial patterns analyzed in the final landslide susceptibility map are consistent with relationships identified in model testing. As the most significant predictor, slope's influence is reflected in the distribution of susceptibility across the map, with mountainous regions exhibiting increased susceptibility and lowland areas considerably less affected. Additionally, the potential orographic effect of the Blue Ridge Mountains, driven by its sharp increases in elevation, likely affects the precipitation regime within the study area. As the southerly movement of moist, warm air from the Gulf of Mexico pushes across the mountains, orographic lifting induces release of precipitation (Baldwin, 2007), a process likely represented by the increased susceptibility observed in the higher elevation areas southeast of Asheville. The final landslide susceptibility map highlighted similar patterns to existing literature, with Lin et al. (2024) producing a multi-hazard earthquake and landslide susceptibility map with comparable distributions. This lends support to the validity and relevance of the predictors employed within the model.

4.6 Precipitation Sensitivity Analysis

The stepwise increase in predicted landslide occurrence across precipitation multipliers demonstrates the model's predictions are directly affected by changes in precipitation. The results in this analysis indicate that according to this model, a shift in maximum daily precipitation could lead to a significant change in the number of predicted landslides within the study area. The observed logarithmic relationship of the data is to be expected given the log transformation applied to the original data, however the strength in prediction response indicates precipitation's vital role as a landslide predictor within the model. Furthermore, the results found within this sensitivity analysis suggest that LSMs trained on one climate regime may potentially exhibit reduced accuracy under differing precipitation patterns, a relationship that is further explored in the following section.

4.7 Decadal Model Comparison

A decrease in both overall accuracy as well as AUC in both decadal cross-validations suggested the accuracy of the models were affected but not entirely compromised. Notably, Model B, trained on the decade exhibiting more stable precipitation values, experienced a proportionally larger drop in accuracy and AUC when validating into the more climatically variable precipitation regime of Decade A, than Model A's cross-validation. This could indicate that models trained on more stable climatic conditions may struggle when applied to periods characterized by more extreme precipitation patterns. More specifically, the change in sensitivity and specificity within each model's cross-validation is of particular note. Model A's drop in sensitivity during cross-validation, and concurrent proportional rise in false negatives, indicated that the model underestimated landslide occurrence within the more stable precipitation regime of Decade B, possibly a consequence of the model associating more extreme precipitation values with elevated landslide probability. Conversely, Model B exhibited a marked increase in false positives when cross-validating, suggesting a systemic overprediction of landslide occurrence when applied to a more extreme precipitation regime. This more cautious approach is likely a result of a model trained on more stable precipitation interpreting increased precipitation variability with increased landslide susceptibility.

The coefficient values of each model, however, reveal a potential issue with the decadal comparison methodology. While Model B displays predictor significance broadly similar to the full dataset model (2001-2020), Model A identifies significance in only two predictors: slope and distance to roads. Although several predictors exhibit p-values approaching significance, the insignificance of the precipitation predictor may invalidate Model A's cross-validation. Furthermore, the marked difference in coefficient values between the decadal models may point to an issue discussed earlier in the report: the landslide inventory. The close alignment between Model B and the full dataset model's coefficient values may suggest the majority of the full model's predictive power is derived from the 2011-2020 period. Given a substantial proportion of this period's landslide points originate from the 2018 landslide cluster highlighted in the landslide inventory section, this raises the possibility of a systemic bias within the model towards the conditions associated with that event. Although this bias may limit the relevance of Model A's cross-validation results, Model B's cross validation still retains analytical value. Model B's observed decline in predictive ability when validating across a differing precipitation regime suggests that the accuracy of an LSM trained on one set of historical precipitation data is indeed affected when predicting landslide susceptibility in contrasting precipitation conditions, directly addressing the third objective within the study.

4.8 Limitations and Suggestions for Further Research

Possibly the most relevant limitation within the study was the issue of potential landslide inventory bias, discussed in the previous section. Due to the prominence of the 2018 landslide event within the inventory, approximately 170 points of the 703 total, the model could be heavily influenced to primarily represent the conditions present there. Despite measures implemented to reduce the influence of the event, including a 100-meter minimum distance filter, the limited predictor significance identified in the 2001-2010 decadal model suggest a disproportionate influence of the 2011-2020 decade, and thus likely the 2018 cluster within it. Further research into the subject is encouraged to encompass a larger, more varied landslide inventory to avoid potential biases caused by single landslide events. Moreover, errors and inconsistencies identified in the later stages of the study limited the applicability and interpretability of several predictor layers. Most notably, the soil type layer was excluded from analysis following identification of critical processing errors which left the layer invalid. Similarly, the broad classification found within the land cover layer reduced interpretability of analysis results, limiting the understanding that could be gained regarding land cover types and landslide susceptibility. Additionally, reprojection required due to the extreme variation in spatial resolutions of acquired predictor layers introduced potential interpolation errors. An additional potential bias includes the distance to roads predictor, which was derived using a simple distance accumulation function. As the influence of distance to roads on landslide susceptibility reasonably becomes irrelevant beyond a certain distance, a simple accumulation function may not accurately represent this relationship, potentially inflating the significance of the distance to roads predictor. This concern may be unwarranted, however, as the similarly derived distance to streams/rivers predictor was among the most consistently insignificant predictors within the study. This observation suggests the accumulation function alone may not be responsible for the significance of the distance to roads predictor, though the potential limitation is still noteworthy.

Another notable limitation was the use of a logistic regression-based model instead of more advanced model alternatives. Logistic regression models assume a linear relationship between the predictors and the log-odds, something that is rarely the case in the natural world. Additionally, as the collinearity tests highlighted, the assumed independence of the predictors may also be faulty, as many predictors have complex, interdependent relationships not considered within the model. Other alternatives, such as random forest or neural network models, are more able to capture the predictor's complex, non-linear relationships, and may therefore be more valid models to apply. Finally, logistic regression models may be sensitive to outliers within the data, and despite efforts to reduce this limitation with log transformations, it is still noteworthy in the case of potential extreme values. Given the limits of logistic regression outlined above, future research into the topic would benefit from the use of more advanced modelling techniques, such as neural networks or deep learning models, which are more capable of capturing the complex interdependencies between landslide conditioning factors.

A further significant limitation concerns the precipitation variable used within the study. While annual precipitation was identified as an ineffective landslide predictor and was subsequently replaced with the more relevant maximum daily precipitation, daily interval precipitation data is still an imperfect predictor of landslide occurrence. Bezerra et al. (2025), for example, found that 48-hour metrics were the ideal metric for landslide susceptibility. Alternatively, effective antecedent precipitation (EAP), a metric that tracks the effective cumulative rainfall in the days leading up to a landslide event, has been found to have a much stronger correlation to landslide occurrence than other precipitation indices (Cheng et al., 2025). Additionally, a study using compound temporal precipitation, a multitemporal predictor constructed by jointly considering both long- and short-term precipitation data, found a significant increase in AUC across multiple experiments (Wang et al., 2026). Further research into the subject would benefit from the incorporation of more advanced precipitation metrics, such as EAP or compound temporal precipitation, which could more accurately observe the role of precipitation predictors within landslide susceptibility models.

5 Conclusion

This study developed and evaluated a logistic regression landslide susceptibility model for the Blue Ridge Mountain region of North Carolina, addressing three key objectives concerning the role of precipitation within the model and temporal transferability. The developed model demonstrated consistently high predictive accuracy across all random iterations, with the final full dataset model exhibiting an average AUC score of 0.85 with minimal variance, and displayed no signs of overfitting to the training data. Slope was identified as the most significant predictor within the model, followed closely by maximum daily precipitation, addressing the first major objective of the study. The extreme statistical significance of precipitation ($p<0.0001$) and its substantial effect size confirmed precipitation as a critical predictor within the landslide susceptibility model. The precipitation sensitivity analysis, meanwhile, demonstrated a relationship between increasing precipitation and predicted landslide occurrence within the model, addressing the second objective of the study. The stepwise increase in precipitation and the subsequent rise in predicted landslides within the dataset suggested a strong relationship between precipitation and landslide occurrence. Due to the logarithmic transformation of the raw precipitation data, however, diminishing returns were observed in landslide occurrence per stepwise increase in precipitation.

The decadal cross-validation addressed the third and most central research objective, with results of the study suggesting that the predictive accuracy of an LSM trained on one precipitation regime is affected when validating against a differing set of precipitation conditions. Particularly noteworthy was the proportionally larger drop in accuracy in Model B, trained on the more stable decade, when cross-validating into the more variable decade. Additionally, the marked increase in false positive errors within the same cross-validation suggested models based on more stable climatic conditions may experience an increasingly cautious overprediction of landslide occurrence when faced with more variable precipitation regimes. Noteworthy, however, was the potential landslide inventory bias associated with the over-representation of the 2018 landslide event, which may have disproportionately influenced model behavior and largely invalidated the decadal cross-validation for the first decade.

Collectively, this study's findings indicate that maximum daily precipitation plays a vital role as a predictor for landslide susceptibility, and that change in the variable causes a significant response in the accuracy and predicted occurrence of landslides for logistic regression LSMs within the study area. More broadly, these findings highlight the potential issue of temporal transferability within landslide modelling, and that further research is encouraged into the potential effect of a changing climate on LSM accuracy in the near future.

6 References

Agboola, G., Beni, L.H., Elbayoumi, T., & Thompson, G. (2024) Optimizing landslide susceptibility mapping using machine learning and geospatial techniques. Ecological Informatics, 81, Article 102583. https://doi.org/10.1016/j.ecoinf.2024.102583

Alcântara, E., Baião, C.F., Guimarães, Y.C., Mantovani, J.R., & Marengo, J.A. (2025) Machine learning reveals lithology and soil as critical parameters in landslide susceptibility for Petrópolis (Rio de Janeiro State, Brazil). Natural Hazards Research, 5(3), 539-553. https://doi.org/10.1016/j.nhres.2025.01.008

Allison, P.D. (2014) Measures of Fit for Logistic Regression [Conference paper]. SAS Global Forum 2014, Washington, D.C., United States. https://support.sas.com/resources/papers/proceedings14/1485-2014.pdf

Appalachian Regional Commission. (2026). About the Appalachian Region. https://www.arc.gov/about-the-appalachian-region/

Atkinson, P.M. & Massari, R. (1998) Generalised linear modelling of susceptibility to landsliding in the central Apennines, Italy. Computers & Geosciences, 24(4), 373-385. https://doi.org/10.1016/S0098-3004(97)00117-9

Baldwin, W. (2007). An Analysis of Key Aspects of Warm and Cool Season Flash Flooding in the Southern Appalachians [Master's thesis, Western Kentucky University]. TopSCHOLAR. https://digitalcommons.wku.edu/theses/961/

Belair, G.M., Jones, E.S., Slaughter, S.L., & Mirus, B.B. (2022) Landslide Inventories across the United States (Version 2.0, June 2022) [Dataset]. https://www.sciencebase.gov/catalog/item/61f326dfd34e622189b93308

Bement, H. (2024) Appalachian-Blue Ridge Forests. EBSCO Research Starters. https://www.ebsco.com/research-starters/forestry/appalachian-blue-ridge-forests

Bezerra, A.P., dos Santos, C.A.C., Santos, C.A.G., Gonçalves, W.A., & de Oliveira, G. (2025) Assessment of landslide susceptibility triggered by precipitation in the Metropolitan Region of Recife, Brazil. Journal of South American Earth Sciences, 168, Article 105812. https://doi.org/10.1016/j.jsames.2025.105812

Blenkinsop, S., Muniz Alves, L., & Smith, A.J.P. (2021). ScienceBrief Review: Climate change increases extreme rainfall and the chance of floods. In C. Le Quéré, P. Liss & P. (Eds.) ForsterCritical Issues in Climate Change Science. https://doi.org/10.5281/zenodo.4779119

Buto, S.G. & Anderson, R.D. (2020). NHDPlus High Resolution (NHDPlus HR)---A hydrography framework for the Nation [Dataset]. U.S. Geological Survey. https://doi.org/10.3133/fs20203033

Çellek, S. (2024) Effect of the curvature parameter and its classification on landslides. Mühendislik Bilimleri ve Tasarım Dergisi, 12(1), 49-63. https://doi.org/10.21923/jesd.1391818

Çellek, S. (2022) The Effect of Aspect on Landslide and Its Relationship with Other Parameters. In Y. Zhang & Q. Cheng (Eds.) Landslides. IntechOpen.

Cheng, C., Li, Y., Zhu, D., Liu, Y., Wu, Y., Lin, D., & Guo, H. (2025) Rain-Induced Shallow Landslide Susceptibility Under Multiple Scenarios Based on Effective Antecedent Precipitation. Applied Sciences 15(11), Article 6241. https://doi.org/10.3390/app15116241

Clark, G.M. (2008) Debris slides and flows in the Southern Appalachians - Historical, Process, and Environmental Geomorphology. Vignettes: Key Concepts in Geomorphology. https://serc.carleton.edu/vignettes/collection/25557.html

Clark, S.H.B. (2001) Birth of the Mountains: The Geologic Story of the Southern Appalachian Mountains. United States Geological Survey, United States Department of the Interior. https://pubs.usgs.gov/gip/birth/birth.pdf

Cozier, 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.
Cui, W., Chen, K., Qing, Z., Zheng, Z., Zhang, W., Zhou, Z., Han, P., Zhang, J., Zhu, Z. & Sun, C. (2025) Interpretable machine learning incorporating major lithology for regional landslide warning in northern and eastern Guangdong. npj Natural Hazards 2, Article 89. https://doi.org/10.1038/s44304-025-00146-8
Duan, Y., Ding, M., He, Y., Zheng, H., Delgado-Téllez, R., Sokratov, S., Dourado, F., & Fuchs, S. (2025) Global projections of future landslide susceptibility under climate change. Geoscience Frontiers, 16(4), Article 102074. https://doi.org/10.1016/j.gsf.2025.102074
Ehsan, M., Anees, M.T., Bakar, A.F.B.A., & Ahmed, A. (2025) A review of geological and triggering factors influencing landslide susceptibility: artificial intelligence-based trends in mapping and prediction. International Journal of Environmental Science and Technology, 22, 17347-17382. https://doi.org/10.1007/s13762-025-06741-6
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
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
Guzzetti, F., Peruccacci, S., Rossi, M., & Stark, C.P. (2008) The rainfall intensity–duration control of shallow landslides and debris flows: an update. Landslides, 5, 3-17. https://doi.org/10.1007/s10346-007-0112-1
Haque, U., da Silva, P.F., Devoli, G., Pilz, J., Zhao, B., Khaloua, A., Wilopo, W., Andersen, P., Lu, P., Lee, J., Yamamoto, T., Keellings, D., Wu, J.H., & Glass, G.E. (2019) The human cost of global warming: Deadly landslides and their triggers (1995–2014). Science of the Total Environment, 682, 673-684. https://doi.org/10.1016/j.scitotenv.2019.03.415.
Hamal, S., Dhakal, S., Shahi, B., Budha, P.B., Alkhuraiji, W.S., & Zhran, M. (2026) Assessing landslide susceptibility using frequency ratio and logistic regression models in the Khudi Watershed. Geoscience Letters, 13, Article 19. https://doi.org/10.1186/s40562-026-00467-0
He, Q., Wu, S., Zhao, X., Hui, Z., Wang, Z., Tsangaratos, P., Ilia, I., Chen. W., Chen, Y., & Hao, Y. (2025) Evaluation of landslide susceptibility of mountain highway based on RF and SVM models. Scientific Reports, 15, Article 24991. https://doi.org/10.1038/s41598-025-08774-w
Hetley, D. (2010) Chapter 6: Landslide hazards. In I. Alcántara & A. Goudie, (Eds.), Geomorphological hazards and disaster prevention (pp. 63-74). Cambridge University Press.
Horton, J.D. (2017). The State Geologic Map Compilation (SGMC) Geodatabase of the Conterminous United States [Dataset]. U.S. Geological Survey. https://doi.org/10.3133/ds1052
Hungr, O., Leroueil, S., & Picarelli, L. (2014) The Varnes classification of landslide types, an update. Landslides, 11, 167.194. https://doi.org/10.1007/s10346-013-0436-y
Kaushal, S., Thanveer, J., ul Islam, S.M., Negi, A., Dhanshyan, A., Beetan, Y., Subramanian, S.S., Sajinkumar, K.S., & Yunus, A.P. (2025) Unveiling the amplifying impact: Anthropogenic activities and the two-fold surge in landslides in the Lesser Himalayas. Catena, 250, Article 108771. https://doi.org/10.1016/j.catena.2025.108771.
Lann, T., Bao, H., Lan, H., Zheng, H., Yan, C., & Peng, J. (2024) Hydro-mechanical effects of vegetation on slope stability: A review. Science of the Total Environment, 926, Article 171691. https://doi.org/10.1016/j.scitotenv.2024.171691
Li, Y. & Duan, W. (2024) Decoding vegetation's role in landslide susceptibility mapping: An integrated review of techniques and future directions. Biogeotechnics, 2(1), Article 100056. https://doi.org/10.1016/j.bgtech.2023.100056

Li, Y., Yan, E., & Xiao, W. (2025) Study on Shallow Landslide Induced by Extreme Rainfall: A Case Study of Qichun County, Hubei, China. Water, 17(4), Article 530. https://doi.org/10.3390/w17040530

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

Ling, S., Zhao, S., Huang, J., & Zhang, X. (2022) Landslide susceptibility assessment using statistical and machine learning techniques: A case study in the upper reaches of the Minjiang River, southwestern China. Frontiers in Earth Science 10. https://doi.org/10.3389/feart.2022.986172

Merghadi, A., Yunus, A.P., Dou, J., Whiteley, J., ThaiPham, B., Tien Bui, D., Avtar, R., & Abderrahmane, B. (2020) Machine learning methods for landslide susceptibility studies: A comparative overview of algorithm performance. Earth-Science Reviews, 207, Article 103225. https://doi.org/10.1016/j.earscirev.2020.103225

Okoli, J., Nahazanan, H., Nahas, F., Kalantar, B., Shafri, H.Z.M., & Khuzaimah, Z. (2023) High-Resolution Lidar-Derived DEM for Landslide Susceptibility Assessment Using AHP and Fuzzy Logic in Serdang, Malaysia. Geosciences, 13(2), Article 34. https://doi.org/10.3390/geosciences13020034

Pittillo, D.J., Hatcher, R.D., Jr., & Buol, S.W. (1998) Introduction to the environment and vegetation of the southern Blue Ridge Province. Castanea, 63(3), 202-216.

Pratama, G.M., Gomi, T., Noviandi, R., Ritonga, R.P., Fathani, T.F., & Wilopo, W. (2026) Land Cover and Land Use Controls on Landslide Morphometry and Occurrence in a Heterogeneous Mountain Watershed. GeoHazards, 7(1), Article 31. https://doi.org/10.3390/geohazards7010031

PRISM Climate Group. (2026). Prism gridded climate data (2001-2026) [Dataset]. Northwest Alliance for Computational Science and Engineering, Oregon State University. https://prism.oregonstate.edu/data/

Reichenbach, P., Rossi, M., Malamud, B.D., Mihir, M., & Guzzetti, F. (2018) A review of statistically-based landslide susceptibility models. Earth-Science Reviews, 180, 60-91. https://doi.org/10.1016/j.earscirev.2018.03.001

Remondo, J., Sánchez-Díaz, M., & Cuesta-Albertos, J.A. (2026) An Increasing Trend of Landslides as a Consequence of the Global Change. Earth Systems and Environment, 10, 1461-1474. https://doi.org/10.1007/s41748-025-00685-0

Rohan, T., Shelef, E., Mirus, B., & Coleman, T. (2023) Prolonged influence of urbanization on landslide susceptibility. Landslides, 20, 1433-1447. https://doi.org/10.1007/s10346-023-02050-6

Shahabi, H. & Hashim, M. (2015) Landslide susceptibility mapping using GIS-based statistical models and Remote sensing data in tropical environment. Scientific Reports, 5, Article 9899. https://doi.org/10.1038/srep09899

Sim, K.B., Lee, M.L., & Wong, S.Y. (2022) A review of landslide acceptable risk and tolerable risk. Geoenviron Disasters 9(3), Article 3. https://doi.org/10.1186/s40677-022-00205-6

Soil Survey Staff. (n.d.). Gridded Soil Survey Geographic (gSSURGO) Database for the United States of America [Dataset]. U.S. Department of Agriculture, National Resources Conservation Service. https://gdg.sc.egov.usda.gov/

Sun, X., Chen, J., Bao, Y., Han, X., Zhan, J., & Peng, W. (2018) Landslide Susceptibility Mapping Using Logistic Regression Analysis along the Jinsha River and Its Tributaries Close to Derong and Deqin County, Southwestern China. International Journal of Geo-Information, 7(11), Article 438. https://doi.org/10.3390/ijgi7110438.

Tahroudi, M.N. (2025) Comprehensive global assessment of precipitation trend and pattern variability considering their distribution dynamics. Scientific Reports, 15, Article 22458. https://doi.org/10.1038/s41598-025-06050-5

Tang, L., Yu, X., Jiang, W., & Zhou, J. (2023) Comparative study on landslide susceptibility mapping based on unbalanced sample ratio. Scientific Reports, 13, Article 5823. https://doi.org/10.1038/s41598-023-33186-z

Thornton, M.M., Shrestha, R., Wei, Y., Thornton, P.E., & Kao, S.-C. (2022). Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1 (Version 4.1) [Dataset]. ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/2129

Turner, M.G., Pearson, S.M., Bolstad, P., & Wear, D.N. (2003) Effects of land-cover change on spatial pattern of forest communities in the Southern Appalachian Mountains (USA). Landscape Ecology, 18, 449-464. https://doi.org/10.1023/A:1026033116193

U.S. Geological Survey. (2023). USGS National Transportation Dataset (NTD) [Dataset]. U.S. Department of the Interior. https://data.usgs.gov/datacatalog/data/USGS:ad3d631d-f51f-4b6a-91a3-e617d6a58b4e

U.S. Geological Survey, Earth Resources Observation and Science (EROS) Center. (2021). USGS EROS Archive - Vegetation Monitoring - EROS Visible Infrared Imaging Radiometer Suite (eVIIRS) [Dataset]. U.S. Department of the Interior. https://www.usgs.gov/centers/eros/science/usgs-eros-archive-vegetation-monitoring-eros-visible-infrared-imaging#overview

U.S. Geological Survey, Earth Resources Observation and Science (EROS) Center. (2024). Annual National Land Cover Database [Dataset]. U.S. Department of the Interior. https://www.usgs.gov/centers/eros/science/annual-national-land-cover-database

U.S. Geological Survey, National Geospatial Technical Operations Center. (2024). 1 Arc-second Digital Elevation Models (DEMs) - USGS National Map 3DEP [Dataset]. U.S. Department of the Interior. https://data.usgs.gov/datacatalog/data/USGS:35f9c4d4-b113-4c8d-8691-47c428c29a5b

Wang, J., Wu, J., Fang, H., & Wang, M. (2026) Incorporating compound temporal precipitation dynamics to enhance landslide susceptibility modeling. npj Natural Hazards 3, Article 18. https://doi.org/10.1038/s44304-026-00181-z

Wooten, R.M., Witt, A.C., Miniat, C.F., Hales, T.C., & Aldred, J.L. (2016) Frequency and magnitude of selected historical landslide events in the southern Appalachian Highlands of North Carolina and Virginia: relationships to rainfall, geological and ecohydrological controls, and effects. In C.H. Greenberg & B.S. Collins (Eds.), Natural Disturbances and Historic Range of Variation (203–262). Springer International Publishing.

Youssef, K., Shao, K., Moon, S. & Bouchard, L.-S. (2023) Landslide susceptibility modeling by interpretable neural network. Communications Earth & Environment, 4, Article 162. https://doi.org/10.1038/s43247-023-00806-5

Zhao, B., Zhang, L., Gu, X., Luo, W., Yu, Z., & Yuan, L. (2025) How is the occurrence of rainfall-triggered landslides related to extreme rainfall? Geomorphology, 475, Article 109666. https://doi.org/10.1016/j.geomorph.2025.109666.

7 Appendix

7.1 Appendix A

This appendix contains the results of the final model test, which displayed little change from the initial model. Small increases in AUC can be noted, as well as slight changes in effect sizes for predictors, although the changes were deemed minimal enough to not be included in the results.

Predictor	Coefficient	Odds Ratio	P-Value	Significance	Effect Size
Slope	1.4734	4.3691	0.0000	Sig. Increases	+336.9% per SD increase
Aspect	-0.2092	0.8119	0.0272	Sig. Decreases	-18.8% per SD increase
Profile Curvature	0.2324	1.2642	0.0148	Sig. Increases	+26.4% per SD increase
NDVI	-0.2155	0.8068	0.0194	Sig. Decreases	-19.3% per SD increase
Distance to Roads	-0.7140	0.4901	0.0000	Sig. Decreases	-51.0% per SD increase
Max Daily Precipitation	0.4547	1.5783	0.0000	Sig. Increases	+57.8% per SD increase
Land Cover_22	1.5601	4.9628	0.0165	Sig. Increases	+396.3% vs reference
Land Cover_41	-0.9540	0.3893	0.0036	Sig. Decreases	-61.1% vs reference
Land Cover_42	-0.8044	0.4809	0.3315	Not Significant	—
Land Cover_43	-1.6125	0.2003	0.0000	Sig. Decreases	-80.0% vs reference
Land Cover_81	-0.5827	0.5624	0.2099	Not Significant	—
Land Cover_999	0.1813	1.2274	0.7544	Not Significant	—
SGMC_2	0.3557	1.4337	0.0933	Not Significant	—
SGMC_3	-1.1456	0.3407	0.0523	Not Significant	—
SGMC_4	0.6947	2.0214	0.3068	Not Significant	—
SGMC_8	-0.1533	0.8586	0.5517	Not Significant	—
SGMC_999	-0.8792	0.4221	0.0671	Not Significant	—

Sig. = Significantly

Metric	Value	Std Dev	Min	Max
Validation AUC (30% holdout)	0.8451	0.0136	0.8204	0.8603
Training AUC (70% training set)	0.8677	0.0048	0.8614	0.8759
AUC Difference (gap)	0.0226	—	—	—
No. Observations (training)	984	—	—	—
Pseudo R-squared	0.3466	0.0095	0.3342	0.3629
Log-Likelihood	-445.53	6.62	-454.12	-434.26
N Runs	5	—	—	—

Results averaged across 5 runs (random states 41–45), 70/30 train/validation split.

7.2 Appendix B

The histograms of predictors which were log transformed were highlighted in the results section, while the remaining histograms are displayed here. The remaining predictors were found to not need log transformations due to a relative normality in their distributions or the categorical nature of the predictor making transformations invalid.