Landslide Susceptibility Prediction Adaptive to Triggering Events

by

Ikechukwu Daniel Adebi

S.B. Computer Science and Engineering
Massachusetts Institute of Technology 2022

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

© 2023 Ikechukwu Daniel Adebi. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide, irrevocable, royalty-free license to exercise any and all rights under copyright, including to reproduce, preserve, distribute and publicly display copies of the thesis, or release the thesis under an open-access license.

Authored by: Ikechukwu Daniel Adebi
Department of Electrical Engineering and Computer Science
May 12, 2023

Certified by: Marcela Angel
MIT Environmental Solutions Initiative
Thesis Supervisor

Accepted by: Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Landslide Susceptibility Prediction Adaptive to Triggering Events

by

Ikechukwu Daniel Adebi

Submitted to the Department of Electrical Engineering and Computer Science
on May 12, 2023, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

Landslide detection and susceptibility prediction are valuable tools for disaster prevention. Despite there being many various solutions to accomplish these tasks, they all generally depend on topographic features of the environment. However, there are not many solutions that can adapt to triggering events such as hurricanes, earthquakes, or volcanic eruptions. This lack of adaptability can greatly limit the performance of the algorithms designed to solve these problems, which, in turn, makes it difficult for emergency managers and responders in the area to prepare for these events appropriately. This work experiments with various kinds of machine learning models and analyzes the effects of incorporating dynamic features based on triggering events in the training process. Ultimately, the final versions of the best performing models produced in this thesis will be deployed as a part of a landslide monitoring system to be used in Mocoa, Colombia. This system is being adapted and developed for the Drones/UAVs for Equitable Climate Change Adaptation (DECCA) project run by MIT's Environmental Solutions Initiative.

Thesis Supervisor: Marcela Angel Lalinde
Title: MIT Environmental Solutions Initiative

Acknowledgments

I would like to thank Sean Anklam for directly advising me throughout this project, as well as Marcela Angel and Professor John Fernandez for taking the time to supervise me during the DECCA project. Without these people, none of the work here would be possible, and I'm grateful to be able to participate in a project that is this impactful.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

A landslide is a term that is generally used to describe the down-slope movement or mass wasting of soil, rock, and organic materials under the effects of gravity, as well as the land-form that results from such movement [11]. Landslides can be incredibly destructive events that can wipe out critical infrastructure in communities and towns, cost governments billions of dollars, and destroy many lives. This is where the problem of landslide susceptibility, or landslide probability prediction becomes vital, as even though landslides are typically unpreventable, knowing how likely one is to occur in a given region can help communities and local authorities respond appropriately and take the proper actions in order to minimize losses these natural disasters may cause. This can include evacuating areas that are predicted to be high risk at a certain point in time, or not building infrastructure in certain spaces where the terrain may be unstable.

There has been significant progress in developing solutions to the landslide susceptibility prediction problem, and most of the best solutions utilize machine learning models trained on specific static features of the environment. These features are generally extracted from the topography of the region being studied, and they include attributes such as altitude, curvature, and slope of the land that is being studied. While these models have produced high accuracy predictions for landslide susceptibility based on these factors alone [31], they struggle to adapt to triggering events. Nonetheless, factors related to weather, seismology, volcanic activity, vegetation, and even human activity can play a major role in determining whether or not landslides are likely to occur in a given area [18], but most of the current methods don't account for these factors, or if they do, they play a minimal role. Because of this, these models often produced very inaccurate outputs, either because the topography would not be changed significantly enough to make the model change its prediction, or the topography would be changed so significantly that the previous patterns the model had learned before would become useless. The inability for these models to adapt to these trigger events can prevent local governments and emergency managers in the community from responding appropriately, which can not only cost billions of dollars, but also destroy many lives.

This work seeks to improve upon current machine learning based solutions by considering both the static topographical features and dynamic event related features in order to produce models that are more adaptive to sudden environmental changes. Dynamic features will be derived from sources that can give us more information about potential triggering events, including weather stations, geology databases, and vegetation and land cover changes. We will analyze some of these dynamic factors to determine which ones have the greatest influence on landslide susceptibility, specifically precipitation and land surface temperature. These factors were picked based on the availability of the data we have to work with, as well as prior research that shows how these factors influence how landslides occur [12] [3]. We will then use this information to train and test a model that will provide landslide susceptibility predictions predicated on both the static and dynamic features of the environment. We aim to test the hypothesis that including dynamic features can greatly improve performance, which can allow authorities to take the appropriate actions to minimize losses in their communities including more timely and targeted disaster response, direct landscape intervention and remediation, and better informed long-term urban planning.

It is also worth mentioning that other natural events such as earthquakes [16] and volcanic eruptions [21] can potentially increase landslide risk. However, we did not have the means to obtain sufficient data to perform analysis on other dynamic features outside of precipitation and temperature, as there were no ground sensors established in our area of interest to collect any seismic information for us to use.

The ultimate focus region for the landslide susceptibility modelling effort is Mocoa, Colombia. This region is being studied due to its mountainous terrain, heavy rainfall, and other features that cause landslides to occur there very frequently. However, due to acquisitions and logistical delays involved with the data collecting sensor and platform it was decided to begin landslide susceptibility modelling in Puerto Rico with the intent of transferring the model to work with data collected in Mocoa, Colombia as soon as it becomes available. Puerto Rico was selected as a landslide susceptibility modelling site to transfer our analysis from because it experiences frequent landslides and similar mass wasting events induced by hurricanes, flooding, and earthquakes. Furthermore, Puerto Rico shares many environmental features with Mocoa, Colombia including tropical and sub-tropical climate zones and montane cloud forest ecosystem dynamics. Additionally, not only does Puerto Rico have topographic landforms and geologic stratigraphy similar to Mocoa, the complexity of Puerto Rico's environment also allows the methods we are using to be transferable to any other region of the world, as the problem of modelling landslide susceptibility is most difficult with an environment like Puerto Rico which has tree canopies, mountainous terrain, and significant amounts of cloud cover. Finally, datasets similar to those intended to be collected for Mocoa, Colombia had already been collected for Puerto Rico and were available to utilize including foliage penetrating 3d lidar point clouds, multispectral and thermal imagery, and geology and weather data. All of these factors coalesced to make Puerto Rico a near ideal substitute candidate for developing and testing a landslide susceptibility model in anticipation of datasets that will be collected for Mocoa, Colombia in the near-term future.

For this thesis, as a part of the Drones/UAVs for Equitable Climate Change Adaptation (DECCA) project led by MIT's Environmental Solutions Initiative (ESI), in collaboration with MIT Lincoln Laboratory and the Colombian Regional Environmental Authority Corpoamazonia, I aim to contribute to producing an effective and robust landslide monitoring system that combines data collected by Unmanned Aerial Vehicles (UAVs) with the development of algorithms using machine learning and artificial intelligence to predict landslide susceptibility risk levels. I hypothesize that the incorporation of more dynamic features of the environment, specifically precipitation and temperature for this project, will allow the models to make more accurate predictions and analyses in determining landslide susceptibility. The results and insights gained from this thesis will be utilized in further developing and perfecting the models that will be deployed in Mocoa.

1.2 Related Work

1.2.1 Topography Data Based Solutions

Previous solutions to landslide detection and prediction problems have primarily relied on learning information from topographical data. This topographical data can be derived from Digital Terrain Models (DTMs), a form of Digital Elevation Model (DEM) consisting of 3 dimensional representations of the bare earth surface, disregarding above ground objects such as trees and buildings [11]. DTMs can be derived from many sources of data including topographic surveys, optical stereo photogrammetry [32], Synthetic-Aperture Radar (SAR) [29], and Light Detection and Ranging (lidar) [8]. We decided on using Lidar because of the balance it has in penetrating the earth to get high-resolution data, while also being able to avoid cloud cover, which is common in both the Puerto Rico and Mocoa regions. DTMs are an extremely useful tool for performing landslide analysis because of the rich topographic features the data contains, but current methods tend to only focus on second order features derived from these DTMs, rather than other more dynamic factors, which limits their performance.

1.2.2 Landslide Detection

While landslide susceptibility prediction will be our main focus for our project, landslide detection is still a valuable tool that has been used to determine what factors play the biggest roles in causing landslides to occur, and it is also one of the various forms of analysis you can perform once you have acquired a DTM. Interferometric SAR (InSAR) data in particular has been used to investigate land deformations and mass wasting events in different areas [13] [9].

One project which relied more on using lidar data was performed on a region in Hong Kong [29]. This work utilized multiple machine learning methods, including convolutional neural networks (CNNs), random forest models, logistic regression, boosting, and support vector machines (SVMs). This work was able to be significantly more robust than previous methods, and we hope to build on their studies, in order to improve the state of art of landslide detection and susceptibility modelling.

1.2.3 Landslide Susceptibility Prediction

There has been significant progress in the development of solutions to the landslide susceptibility prediction problem, a lot of which has been enabled by recent advancements in machine learning [23], as well as technological advances of the sensors which enable machine learning models to collect higher-resolution data. An example of this progress can be seen in a project performed for a region in Malaysia [15]. This project uses more dynamic variables such as precipitation, but they only use two topographical features to train their models, which limits overall performance. Another project was performed in Romania which also used machine learning to predict landslides from SAR data and DTM analysis [4]. This provides valuable information regarding which topographical factors play the biggest roles in predicting landslides for that region, however, the focus was placed on the static features of the environment. I will go into more detail on specific methodologies used in Chapter 3.

The ability to efficiently store and compute significant quantities of data, in addition to access to a greater diversity of data on triggering factors of dynamic landslides such as weather and temperature, allowed us to develop landslide susceptibility models that can better adapt to triggering events. In the following chapters, I will describe how this work was accomplished.

Chapter 2

Dataset Overview and Preparation

2.1 Data Collection

2.1.1 Study Areas

The study area for the experimentation performed in this thesis is located on the border of the Utuado and Jayuya municipalities in Puerto Rico, as can be seen in the map in Figure A-1. This region was selected because it has thoroughly developed landslide inventories produced by the Federal Emergency Management Agency. Also, the data we collected was captured in the aftermath of Hurricane Maria in October 2017, a hurricane that caused severe damage to much of Puerto Rico, especially in the municipalities of Utuado and Jayuya, producing sufficient landslide information for which to extract and perform analysis.

The final models will be trained and deployed in Mocoa, Colombia, which can be seen on a map in Figure A-2. For this region, we will have more detailed time series data in order for us to perform predictions based on the state of the features at previous time steps, rather than just simply performing uni-temporal analysis as compared to the work in the Puerto Rico region. The models trained for this study area will be deployed as part of the DECCA project later in 2023 to help with risk prevention efforts within the region.

Data collection and preparation methods are universal across both regions unless otherwise specified.

2.1.2 3D Data Collection Methodology

Light detection and ranging (lidar), interferometric synthetic aperture radar (InSAR), and optical stereo photogrammetry are all sensing modalities and data sources that can be used to generate robust, accurate, and high resolution 3d digital elevation models (DEMs) for terrestrial regions of the earth including digital terrain models (DTMs) which focus on the topography of underlying terrain. While optical stereo photogrammetry has the advantage of being the cheapest and most common method for generating remotely sensed 3d model, it cannot be utilized from satellite platforms for the focus regions of this thesis because of the near perpetual cloud cover. Furthermore, even if flying beneath the cloud deck on a low altitude aerial platform such as an unmanned aerial vehicle (UAV), optical sensors cannot penetrate dense tree canopies in order to reveal the underlying terrain. Synthetic aperture radar seems like a good sensing modality to utilize from spaceborne platforms in order to penetrate cloud cover and generate 3d digital terrain models using across-track interferometric data collection and processing methods. However, my particular area of interest is densely tree covered, and the C-band synthetic aperture radar data we have available to us experiences rapid phase decay in regions that are densely tree covered. Also, C-band SAR data has extremely limited ability to penetrate foliage in order to reveal underlying terrain. Lidar sensing was found to occupy a good middle ground in that it can be flown beneath the cloud deck on low flying UAVs while offering excellent foliage penetrating capabilities at sufficiently high spatial resolution in order to assess the land form of the underlying terrain.

2.1.3 Features

After selecting our study areas, we selected various kinds of features with which to perform landslide susceptibility analysis and prediction. All of these features were selected based on the work of Deepankar Gupta, more specifically, his thesis work called "Interpretable Machine Learning Methods for Landslide Analysis" submitted in February 2022. These features were selected because they all have been scientifically shown to be landslide conditioning factors [11], and the results in Gupta landslide prediction analysis demonstrates that these features have significant predictive power.

All of these datasets have been resampled to 1 m resolution in ground sample distance (GSD). Below is a list of all features that were selected[1]:

- **Topographic Data:** In order to extract all of the topographic features of our study areas, we first had to generate DTMs. These DTMs were created from 3D lidar point cloud datasets collected by the Airborne Optical Systems Testbed (AOSTB) from MIT Lincoln Laboratory. With these DTMs, we derived the following features [2]:
  - ***Altitude*:** One of the most important features for determining landslide susceptibility is altitude, as this has effects on temperature, slope, aspect, vegetation type, among other environmental factors. In both study areas, altitude values are measured from mean sea level.
  - ***Slope*:** Slope describes the degree of incline of a surface. Slope can have a major influence on landslide severity, especially in regions where it may vary significantly.
  - ***Aspect*:** Aspect describes the compass direction that a slope is facing, with 0 degrees representing the northernmost direction. For sloped surfaces aspect can govern other parameters such as vegetation growth and vigor and evapotranspiration with the effect becoming more pronounced at more northerly or southerly latitudes.
  - ***Curvatures*:** For this feature, we extracted two different types of curvatures: plan and profile. Plan curvature, which is short for planform curvature, describes the curvature that is perpendicular to the direction of maximum slope. In other words, when planform curvature is positive,

[1] All features are continuous unless otherwise specified.

this means that the surface is laterally convex, and if it is negative, it is laterally concave [5]. In contrast, profile curvature describes the curvature that is parallel to the direction of maximum slope, which affects how flow accelerates or decelerates on a sloped surface. So a positive curvature in this setting would result in an upwardly concave surface, accelerating flow, while a negative curvature results in an upwardly convex surface, which decelerates flow [2, 5]. We also combined gathered general curvature values, (i.e. a combination of both planform and profile curvatures), and used this as a feature as well.

- ***Flow Direction*:** Flow direction is a categorical feature which describes the direction of water flow based on the maximum change in elevation between a given cell (which is simply a square section of space in a map) and its eight bordering cells. There are 8 possible flow directions: north, northeast, east, southeast, south, southwest, west, and northwest.
- ***Flow Accumulation*:** Flow accumulation of a given region provides a measure of total flow accumulated based on upstream catchment basins.
- ***Stream Network*:** The stream network is simply a map of all watersheds in a region, as moving water can influence landslide probability in a given area. Strahler Stream order was also calculated to get a measure of branching complexity of different streams within a network [5].
- ***Topographic Wetness Index (TWI)*:** TWI is an index used to describe how wet the region can be, which is useful for gathering information about hydrological processes that occur in the terrain [26].
- ***Sediment Transport Index (STI)*:** STI is an index used to describe the level of erosion in an area due to sediment movement, which can have a significant impact on how land moves when natural disasters occur.
- ***Stream Power Index (SPI)*:** SPI is an index used to describe the level of erosion in an area due to water flow. This factor can have a great effect on landslide susceptibility as well, especially after floods, hurricanes, or

other water based natural disasters.

- **Static Multispectral Data:** The visible and near-infrared (VNIR) multispectral data was collected by the Pleiades-1a satellite for calculating spectroradiometric indices for describing local vegetation and moisture conditions. Long-wave infrared (LWIR) image data was also acquired from the Landsat-8 satellite in order to calculate land surface temperature, but this is a dynamic feature that will be discussed later.

  – ***Normalized Difference Vegetation Index (NDVI)*:**

  $$NDVI = \frac{NIR - Red}{NIR + Red} \tag{2.1}$$

  NDVI is a spectroradiometric index used to quantify the presence of green vegetation by reading the amount of reflected red visible light (Red) and near-infrared light (NIR) [6]. Equation 2.1 demonstrates how the NDVI value is calculated. Values closer to 1 likely represent vigorous green vegetation, whereas values closer to -1 represent other surface materials such as open water which absorb near-infrared light.

  – ***Normalized Difference Wetness Index (NDWI)*:**

  $$NDWI = \frac{Green - NIR}{Green + NIR} \tag{2.2}$$

  NDWI is a spectroradiometric used to quantify moisture content in a region [1] by comparing the amount of reflected green light (Green) and near-infrared light (NIR). Equation 2.2 demonstrates how the NDWI value is calculated. Values closer to 1 represent bodies of open water, and values closer to -1 represent very dry areas, or non-water surfaces.

- **Geology Data:** Geology data was collected from the United States Geological Survey (USGS). Geology data is valuable for landslide prediction due to the information it provides us regarding lithology, soil type, ground stability and erosion levels.

- ***Soil Classification*:** Soil classification is a categorical feature with values defined by the Unified Soil Classification System (USCS). It describes soils by their plasticity, liquid limits, and particle size and sorts soils into distinct classes of clays, gravels, sands, and silts.

- ***Lithology Classification*:** Lithology classification is a categorical feature that describes the type of rocks that are present in the region. The four lithology types that are in the classification for the Puerto Rico region are alluvium, amphibolite, granodiorite, and Robles Formation.

- **Impervious Surface Data:** The impervious surface data was collected from the Humanitarian OpenStreetMap Team [20]. It's a binary feature that identifies which areas are surfaces that are impervious to rainwater absorption, which includes buildings, paved roads, sidewalks, parking lots, and certain types of bare/exposed rocks.

- **Dynamic Features:** The dynamic features are the primary focus of this thesis. The two features being tested are the land surface temperature and precipitation.

  - ***Land Surface Temperature*:** Like the other multispectral features, the land surface temperature was collected by the Landsat 8 satellite. Land surface temperature is a critical component for deriving apparent evapotranspiration which in turn describes the portion of precipitation which is absorbed into the ground vs being evaporated back to the atmosphere. Temperature is included in the training process of these models because this factor is something that can change drastically under landslide triggering events.

  - ***Precipitation Intensity*:** Precipitation data is collected by the NASA Global Precipitation Measurement Mission. It measures the level of rainfall intensity per day in mm.

- **Historic Landslide Inventory:** The historical landslide inventory contains data about where and when landslides have occurred since 1985 for this specific region of Puerto Rico. The USGS and FEMA collected the landslide data for Puerto Rico.

2.2 Data Processing

All data was processed using ArcGIS Pro [22], an app that is designed to work with Geographic Information Systems (GIS), and Rasterio [7], a Python library used to handle rasterized files and data more easily.

2.2.1 Continuous Data

The majority of the features used in training have continuous values. Before training, all continuous features are normalized between 0 and 1. To fill in missing values, for each feature, I calculated the median to use as a placeholder to minimize any effects that could occur from skewing data. I made sure to calculate the percentage of missing values to make sure this would have as little effect on the training process as possible. When doing these calculations, the highest percentage I calculated was less than 0.05%, which gave me confidence to use my strategy of using median to handle these missing values.

2.2.2 Discrete Data

To handle missing values for the discrete data, for all features except the lithology and soil classifications, I calculated the mode in the feature's data and used the mode to fill in missing values. For the lithology and soil classification features, I added a new class for each of these categories signifying the fact that the statuses of the classifications are unknown.

Chapter 3

Methods

3.1 Machine Learning Strategies

For this thesis, I experimented with various machine learning methods, ranging from conventional machine learning methods, to deep neural networks, to ultimately determine the effect of adding dynamic features to the training process has in improving model performance. The conventional models were trained using scikit-learn, while the deep neural networks were trained using PyTorch.

For the experimentation process, using the historic landslide inventory as my target to predict, I employed two different strategies: regression and classification. For regression, I simply performed the normalization process I described in Section 2.2.1 to the landslide inventory, and used the methods described below to try to predict those target values directly, essentially creating a historic landslide density map. This essentially meant treating the target values as pseudo-probabilities, with 1 being the maximum probability of a landslide occurring, and 0 being the minimum. I only used conventional machine learning models for regression tasks due to the difficulty I had with demonstrating the relationship between landslide conditioning factors and landslide susceptibility when doing deep learning tasks.

For classification, however, I decided to divide up the historic landslide data into four susceptibility classes. These classes were numbered from 1 to 4, with 1 being the lowest risk of a landslide occurring, to 4 being the highest. To avoid having a skewed target distribution, I divided up the normalized landslide probabilities into quartiles, and used that to determine their classification values. It is important to note that a skewed target distribution of classes is not necessarily undesirable[2], but for the purposes of the experiments I am conducting, I decided it would be best to keep the four classes evenly distributed. Table B.1 provides information on what the precise quartile values were for the study area of Puerto Rico. I used both conventional machine learning and deep learning methods for classification.

For categorical features in both the regression and classification model training, I used binary encoding to embed these features before passing them into all of the models used.

3.1.1 Conventional Machine Learning Methods

Logistic Regression (LR)

Logistic regression is one of the most popular machine learning methods used to find relationships between variables, and this method has been consistently used in landslide prediction tasks for a long time [14]. Logistic regression is a supervised learning method that typically works well for classification problems, due to how it utilizes the concept of odds to assign probabilities to events.

Quantitatively, a logistic regression function for multiple variables can be represented as follows:

$$P(\mathbf{X}) = \frac{\exp(\boldsymbol{\beta} \cdot \mathbf{X})}{1 + \exp(\boldsymbol{\beta} \cdot \mathbf{X})} \tag{3.1}$$

where $\mathbf{X}$ represents the m landslide predictors of our problem, $\boldsymbol{\beta}$ is the set of learning parameters, and $P(\mathbf{X})$ is the probability of a landslide occurring given the predictors in $\mathbf{X}$.

For regression, we start by treating this as a binary classification problem at first, so that we only have to predict one probability value representing landslide susceptibility. But instead of picking which class has the higher likelihood, we simply

[2]The four classes do not have to be evenly distributed, depending on how much we want to emphasize certain landslide risk levels over others, as it may be desirable to have a model be more or less risk-averse.

use the probability value calculated as our output.

For classification, since we have four landslide susceptibility classes in total, for training, we essentially have to predict four probability values, one for each class, meaning there are four sets of learning parameters $\boldsymbol{\beta}_{1..4}$ we need to calculate to get our results. To predict what the correct classification is, we just take the probability with the maximum value.

Random Forest (RF)

Random forests is another popular method that has been used in landslide detection and other natural hazards prediction tasks [19]. It is an ensemble learning method based on combining the results of decision trees to give a consensus prediction. This method is very popular due to its robustness, and its ability to be used on very large datasets. This method also provides multiple parameters that we have direct control over, including the number of decision trees we want to use, as well as the branching factor and depth of each tree.

For regression, we replace our traditional decision trees with regression trees, and instead of determining a consensus prediction out of a discrete set of options, we average the results of each regression tree to get a final prediction.

For classification, nothing is fundamentally changed. Each decision tree just has to predict from four susceptibility classes.

Support Vector Machines (SVM)

Support vector machines are another machine learning method that can be used for classification tasks, and has been used for landslide prediction tasks in the past [28]. Support vector machines work by mapping the original input features into a higher dimensional feature space where a hyperplane can be found to act as a decision boundary. This mapping process is done by a kernel, which could be linear, polynomial, Gaussian, or some other form. The kernel used in this thesis is the radial basis function, due to the power it has as an approximator [5]. The radial basis function kernel is defined by the following equation:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\lambda \|\mathbf{x}_i - \mathbf{x}_j\|^2) \tag{3.2}$$

where $\mathbf{x}_i$ and $\mathbf{x}_j$ are feature vectors with all of the landslide predictor variables, and λ is a tuning parameter accounting for the smoothness of the decision boundary produced by the support vector machine.

For regression, instead of finding a hyperplane to act as a decision boundary, we look for more of a best fit line through the data to make continuous predictions. This method is known as support vector regression.

For classification, because there are four classes, we make four binary predictors to perform a multi-class classification task. Each binary predictor behaves by dividing one class of data from the other classes. This is just one common approach for perform multi-class classification tasks with support vector machines [30].

Adaptive Boosting (AB)

Adaptive Boosting, or AdaBoost, is another ensemble learning method which combines the performances of a bunch of weak classifiers to produce a final prediction [10]. Quantitatively, we can define the equation for a prediction with m weak classifiers and n data points as

$$F(x) = \arg\max_k \sum_{i=1}^{m} c_i \cdot \mathbb{I}(f_i(x) = k) \tag{3.3}$$

where $\mathbb{I}$ is an indicator function which is equal to 1 if a weak classifier f_i picks class k, and 0 otherwise, and $F(x)$ gives us our predicted class. The weak classifier used in this project is simply a decision tree, but other kinds of classifiers can also be used.

For regression, we replace our weak classifiers with regressors, and instead of using an $\arg\max$ operation, we simply take the weighted average of the results produced by each of our weak classifiers.

For classification, similarly to random forests, nothing fundamentally changes.

3.1.2 Deep Learning Methods

Up to this point, I have mainly focused on describing more conventional machine learning methods that I used in my experimentation process. While conventional machine learning models have the benefit of being simple to use and implement, they tend to struggle to learn more complex patterns in data. Because of this, I decided to experiment with deep neural networks as well, to determine the potential performance differences when incorporating more dynamic features into the training process. More specifically, I experimented with utilizing segmentation models provided by PyTorch's segmentation model library, as segmentation models have been proven to be successful in landslide detection tasks in the past [25].

Data Pre-processing

In order to pass the features into the segmentation model, there was some pre-processing that needed to be done. Experiments were performed dividing the data into 64×64 patches and 128×128 patches. The patches were also padded with 4 pixels on each side, and 8 pixels on each side, respectively, in order to make sure the map did not appear too discretized when produced.

Model Architecture

The framework architecture I am using as the segmentation model is called U-Net [24]. In a nutshell, U-Net uses an encoder-decoder structure to perform predictions, using convolutional layers to both compress the original input features, and expand this compression to produce an image with classified pixels. This architecture is especially good at performing segmentation tasks because of how the encoder part of the network is able to learn how to classify pixels, and how the decoder part of the network can recover the exact spatial location of those pixels, making this a very powerful method in practice.

U-Net also provides us with the freedom to change the complexity of the encoder itself, and for this project, two different encoders were used. Both of these encoders are

ResNet architectures [27], which are deep residual networks that are known to perform well on image classification tasks. The two different versions of this architecture used are ResNet-18 and ResNet-50. These architectures were used to observe if there were any significant changes in model performance if the deep neural network was given more layers, potentially making the addition of dynamic features in the training process more significant.

Loss Function

The loss function used for this segmentation model is Dice loss [33]. This loss function was picked due to how well it performs in segmentation tasks under situations where there is major data imbalance. The dice loss function can be written as follows:

$$\mathcal{L}_{Dice} = 1 - \frac{2\sum_{i=1}^{n}\sum_{j=1}^{c} y_{ij}p_{ij} + \epsilon}{\sum_{i=1}^{n}\sum_{j=1}^{c}(y_{ij} + p_{ij}) + \epsilon} \tag{3.4}$$

where y_{ij} represents whether or not data point i has target class j, p_{ij} represents whether or not data point i is predicted to be class j, and ϵ is a constant used to prevent the denominator from being 0. The minimum (and optimal) Dice loss value is 0, and the maximum is 1.

Chapter 4

Experiments and Results

4.1 Experiments

The ultimate goal of our experiments is to determine the effects of adding dynamic features in the training process of landslide susceptibility models.

For this study area, we train all of the machine learning models described in Chapter 3, using the features described in Section 2.1.3. For the conventional machine learning models, the data is randomly split into training, validation, and test sets, with 70% of the data in the training set, 10% of the data in the validation set, and the remaining 20% of the data in the test set. For the deep neural networks, instead of dividing up the data by individual data points, we would divide by the 64×64 and 128×128 patches described in Section 3.1.2.

4.2 Evaluation Metrics

4.2.1 Classification

To evaluate the models in the case of classification, we use accuracy, precision, and recall, along with the Area-Under-Curve (AUC) score given by the receiver operating characteristic (ROC) curve. All of these equations use four variables: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). The equations for the first three metrics can be described as follows:

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{FN} + \text{TP} + \text{FP}} \tag{4.1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{4.2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{4.3}$$

The ROC curve simply plots the True Positive Rate (or the recall) vs False Positive Rate (FPR), where FPR is defined by the following equation:

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{4.4}$$

With the curve, the area under it gives us an aggregate measure of how accurate our predictions of each class are, with respect to all of our other classes.

4.2.2 Regression

To evaluate models in the case of regression, we simply take the root mean squared error between the predicted landslide probability and the ground truth landslide probability generated in Chapter 3.

4.3 Effects of Adding Dynamic Features

4.3.1 Feature Importances

The process of analyzing the effects of adding dynamic features to the training process begins with examining how much influence these features have compared with all of the other features. Figure A-3 displays a list of feature importance information based on Gini importance values extracted from training some of the best performing conventional machine learning models based on accuracy, and averaging these results over all models obtained [17].

From this feature importance plot, we can see that precipitation is the highest valued feature by all of the best performing predictors by a significant margin. Static features such as altitude, lithology classification, and soil classification are also valued significantly by most of the predictors. Both precipitation and land surface temperature are top 5 in importance metrics, which shows that adding these features to the training process is likely to have significant influence on the performance of the models.

4.3.2 Regression Model Results

Table B.2 contains the root mean square error results produced by all of the regression models. For these regression models, we can observe that in general, the addition of dynamic features improves model performance. In fact, it improved performance for all conventional models used except for AdaBoost, although the AdaBoost algorithm did not perform optimally in this environment anyway.

Figure A-8 shows maps produced for each conventional model except the support vector machines, as the training time complexity for this model is extremely large. From these maps, we can see that the models were biased towards predicting very low risk values across the board, regardless of whether or not we used dynamic features in the training process. This is likely due to the fact that the distribution of ground truth risk values is skewed towards lower values, as can be seen from Table B.1, therefore the predictions end up imitating this distribution to minimize error, and the addition of new features is unlikely to change this behavior significantly. To mitigate this problem in future studies, one could try transforming the distribution of ground truth values in such a way that the predictions will not be as biased towards predicting low risk values.

4.3.3 Classification Model Results

Tables B.3 and B.4 contain all of the classification based evaluation metrics for the conventional models and segmentation models, respectively. Figures A-4 and A-5 contain a histogram showing the proportions of each landslide susceptibility class predicted for all models trained without and with the use of dynamic features.

For the conventional models, we can see from Table B.3 that the addition of dynamic features consistently improves model performance in every metric utilized in this experimentation process. Considering that both dynamic features were top 5 in feature importances, it is reasonable to conclude that the addition of dynamic features in the training process improved model performance.

From Figure A-4 and the generated susceptibility maps in Figure A-6, we can see that the conventional models are more likely to predict very high risk values, regardless of whether or not dynamic features are utilized in the training process, however, the addition of dynamic features actually increased the likelihood of predicting the lowest risk class across the board. This shows that the models trained with dynamic features are less risk-averse than the models trained without these dynamic features.

For the segmentation models, we actually notice some different patterns arise. For starters, the addition of dynamic features actually decreases model performance for all models except the ResNet-18 model trained with 128×128 patch size. This difference in performance was more prominent as the models increased in size from 18 layers to 50 layers.

Also, when looking at Figure A-5 and the generated susceptibility maps in Figure A-7, we can see that models trained with a patch size of 64×64 is much more likely to predict a susceptibility class value of "Very High" when dynamic features are added to the training process, but with a patch size of 128×128, the distribution is bimodal for models trained without dynamic features, and more spread out amongst predicting other susceptibility classes (except the "High" class) when dynamic features are added to the training process. So while the raw evaluation metrics values show that the performance is better without the use of dynamic features in the training process, the addition of dynamic features can create more balanced distributions, which can be beneficial in creating more useful landslide susceptibility prediction models.

4.3.4 Other Visual Observations

For the classification conventional models, the addition of dynamic features in the training process appear to produce the higher susceptibility risk regions more accurately. For the regression conventional models, there was not much of a significant difference, except for the AdaBoost case, where the addition of dynamic features did allow for the model to predict some higher probability values.

For the segmentation models, similar trends were visible. Higher susceptibility regions in the ground truth map were predicted to have higher landslide susceptibility values, although, the predictions would often either drastically overestimate or underestimate the landslide risk, meaning that many "Medium" level areas would be predicted as "Low" risk, whereas many "High" level areas would be predicted as "Very High" risk.

Chapter 5

Conclusion

The results presented in this thesis show that the addition of dynamic features to the training process can definitely impact the performance of landslide susceptibility models, although simply adding features like temperature or precipitation will not guarantee that predictions will be more accurate. In the case of conventional models, performance improved across the board, regardless of the model used. However, as the models got more complex, improvement of performance was not guaranteed.

Other observations worth noting include that the regression models all bias their predictions towards lower susceptibility levels, while the classification models generally biased their predictions towards either higher risk, or predicting more low and very high risk value classes, ignoring the medium risk options. Even though I chose to have the susceptibility classes be equally distributed across the region for the purposes of eliminating potential bias, it is possible to engineer a more optimal distribution of susceptibility classes that could eliminate, or potentially emphasize, the biases observed in these experiments, depending on what effects would be more desired by those evaluating landslide risk for any given region.

With the addition of precipitation and temperature into the training process, landslide susceptibility models can potentially better adjust to triggering events like hurricanes and floods. However, there is work to be done still to determine how these triggering events may influence landslide risk in specific areas of regions. The models in this study are also prone to underfitting, which is potentially due to a combination of the models being too small, and the data being insufficient to get a better picture of what the landslide risks truly are in different parts of the region. More work can be done in the future to expand this analysis to larger models and datasets.

When transferring this work to Mocoa, it is important to take into account the fact that these models can be biased towards very high risk or very low risk predictions, leaving out intermediate risk level prediction values. This issue can be magnified when creating a real-time landslide monitoring system, especially if the predictions vary frequently between time steps. It may be necessary to adjust the distributions of susceptibility risk classes when training these models for the Mocoa region, so that they are more likely to predict intermediate classes to make real-time updates in the monitoring system smoother.

It is also worth mentioning that precipitation and temperature are only two of many potential dynamic features of an environment that can influence landslide susceptibility. Earthquakes, volcanic eruptions, and river velocity are just a few other factors that could potentially increase landslide risk, and with the appropriate data, more experiments can be done to determine how much of an influence these factors have in improving landslide susceptibility prediction.

Appendix A

Figures

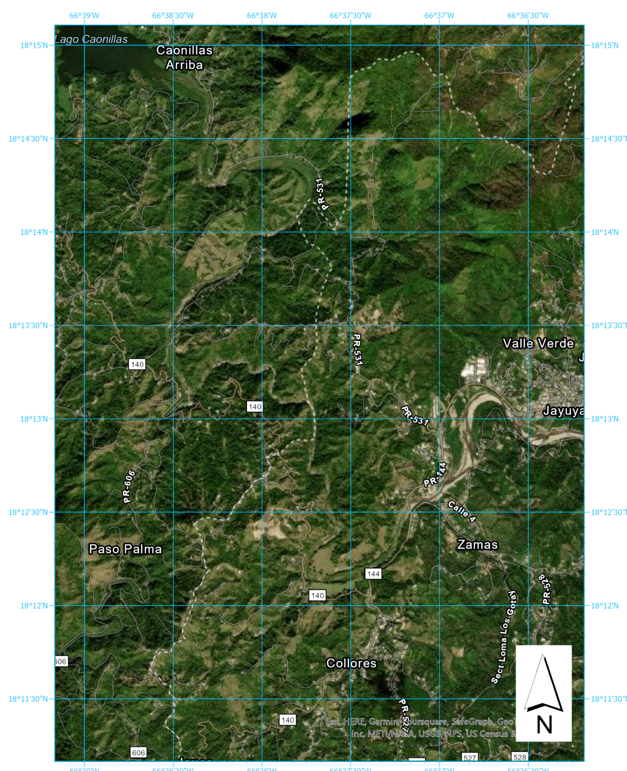

Figure A-1: Map of the southern part of the border between the Utuado and Jayuya municipalities of Puerto Rico. Initial training and testing of machine learning models was performed here for data analysis spanning approximately 17 km^2 of the region.

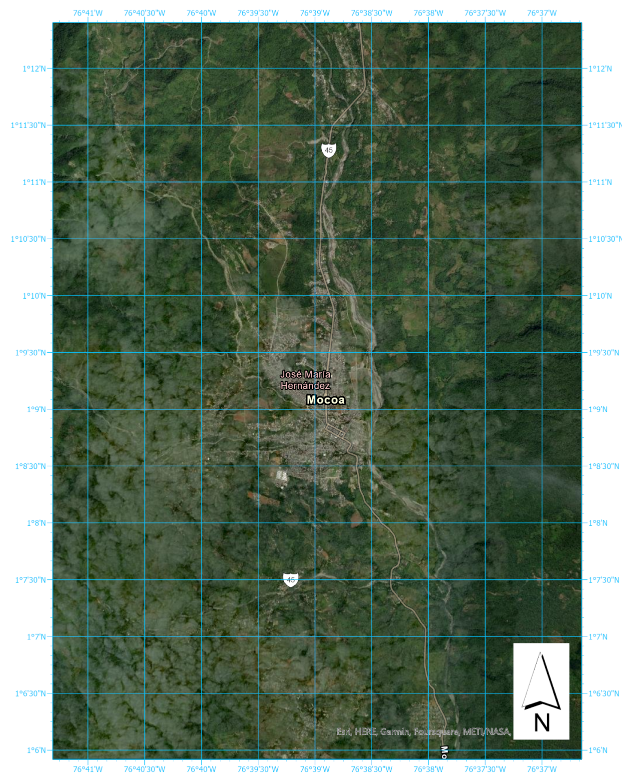

Figure A-2: Map of the region of Mocoa, Colombia. The final versions of the model will be trained later in the year, after the completion of this thesis.

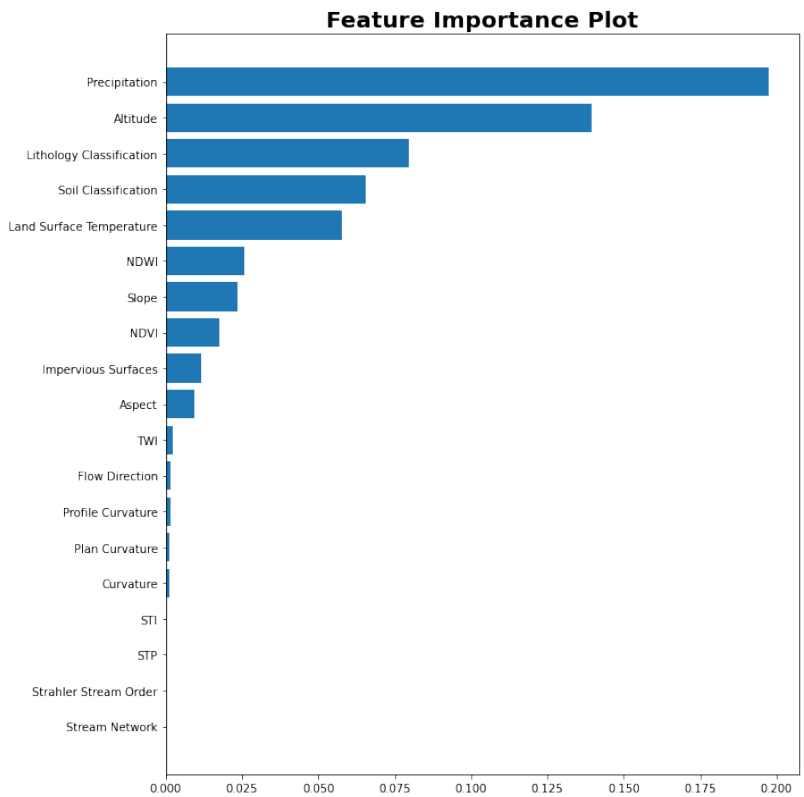

Figure A-3: A bar plot measuring the average feature importances from the best performing conventional machine learning models trained using both static and dynamic features.

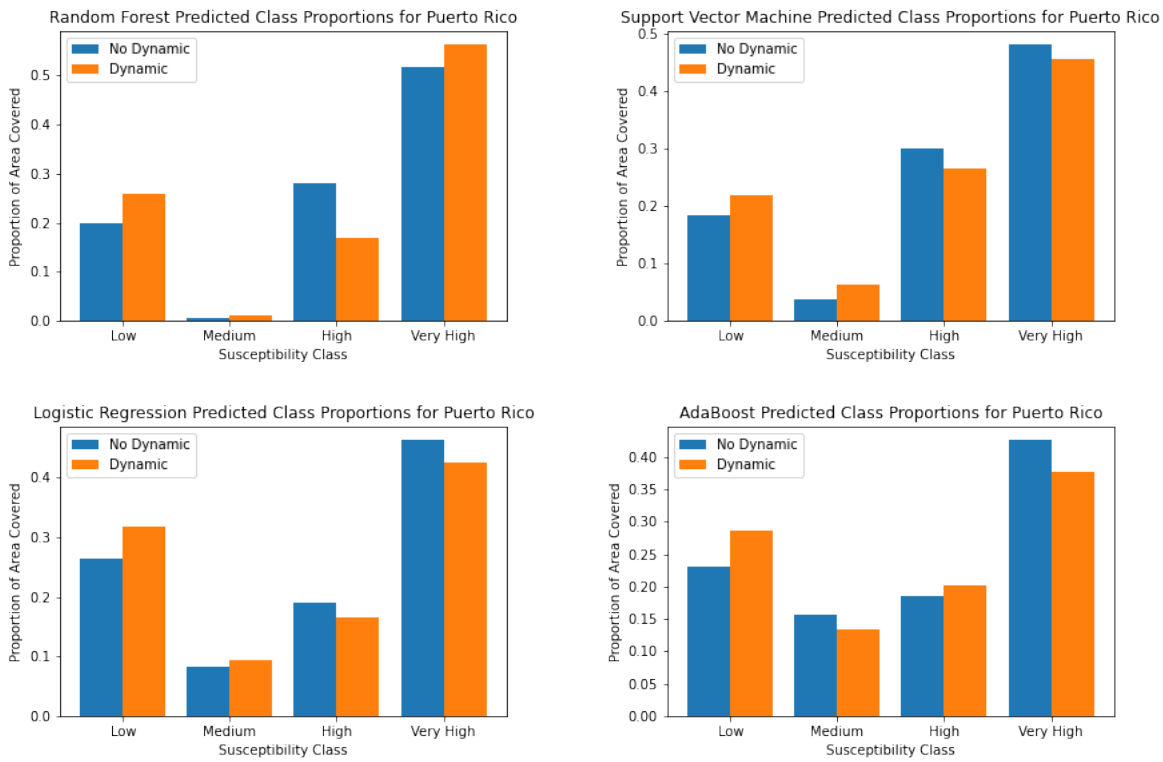

Figure A-4: Plots showing percentages of susceptibility classes predicted for each conventional model for the Puerto Rico region, comparing the results of the models with and without the use of dynamic features in the training process.

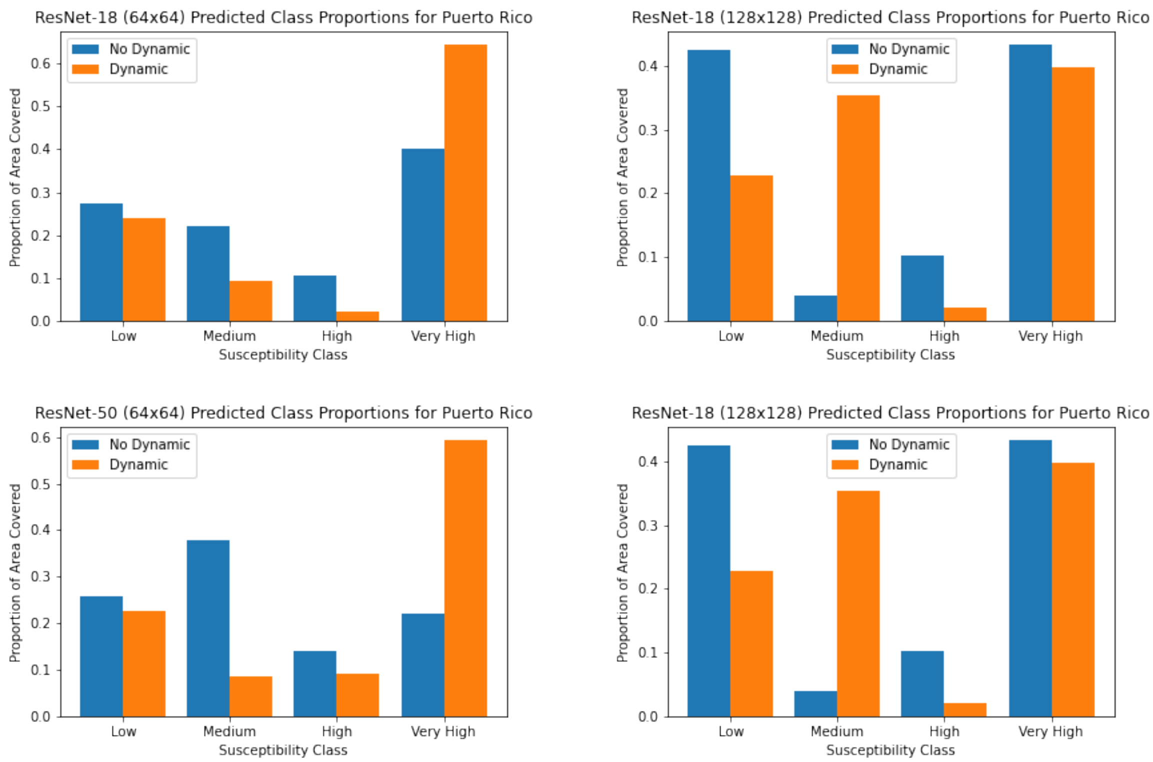

Figure A-5: Plots showing percentages of susceptibility classes predicted for each segmentation model for the Puerto Rico region, comparing the results of the models with and without the use of dynamic features in the training process.

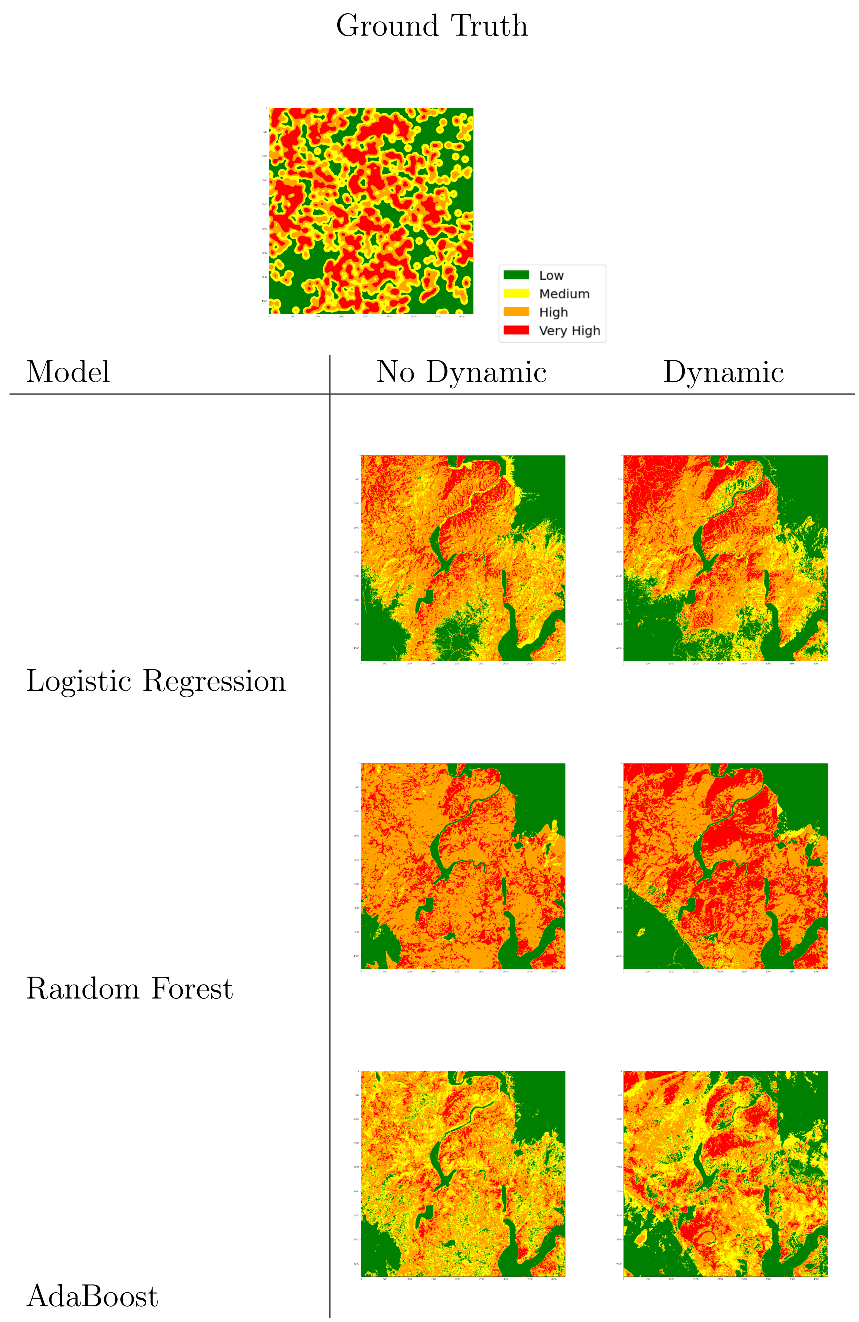

Figure A-6: Visualizations of discrete classification susceptibility maps produced by each of the conventional models for the Puerto Rico region, excluding the support vector machines. No map was produced for the support vector machines due to the large time complexity of training.

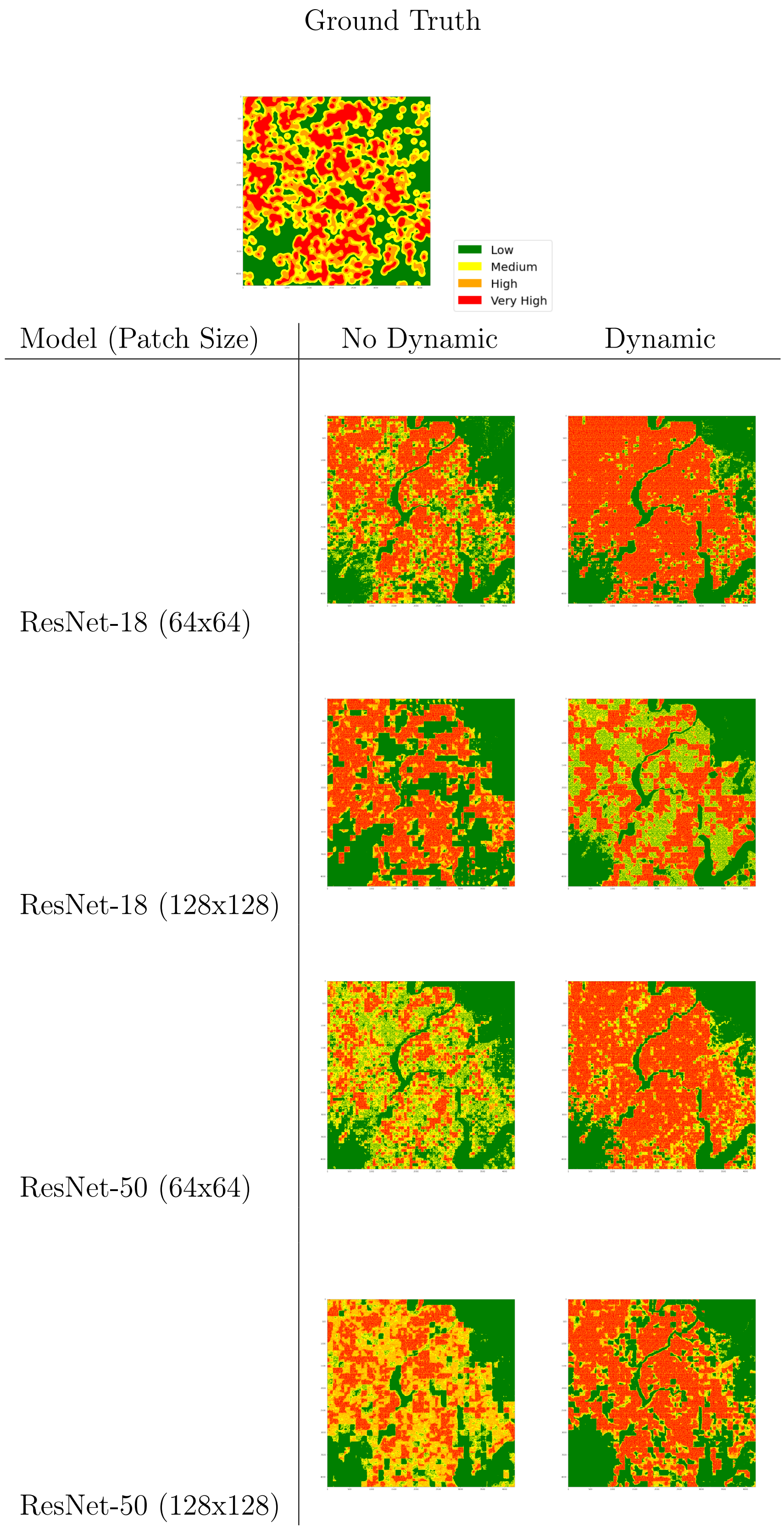

Figure A-7: Visualizations of discrete classification susceptibility maps produced by each of the segmentation models for the Puerto Rico region.

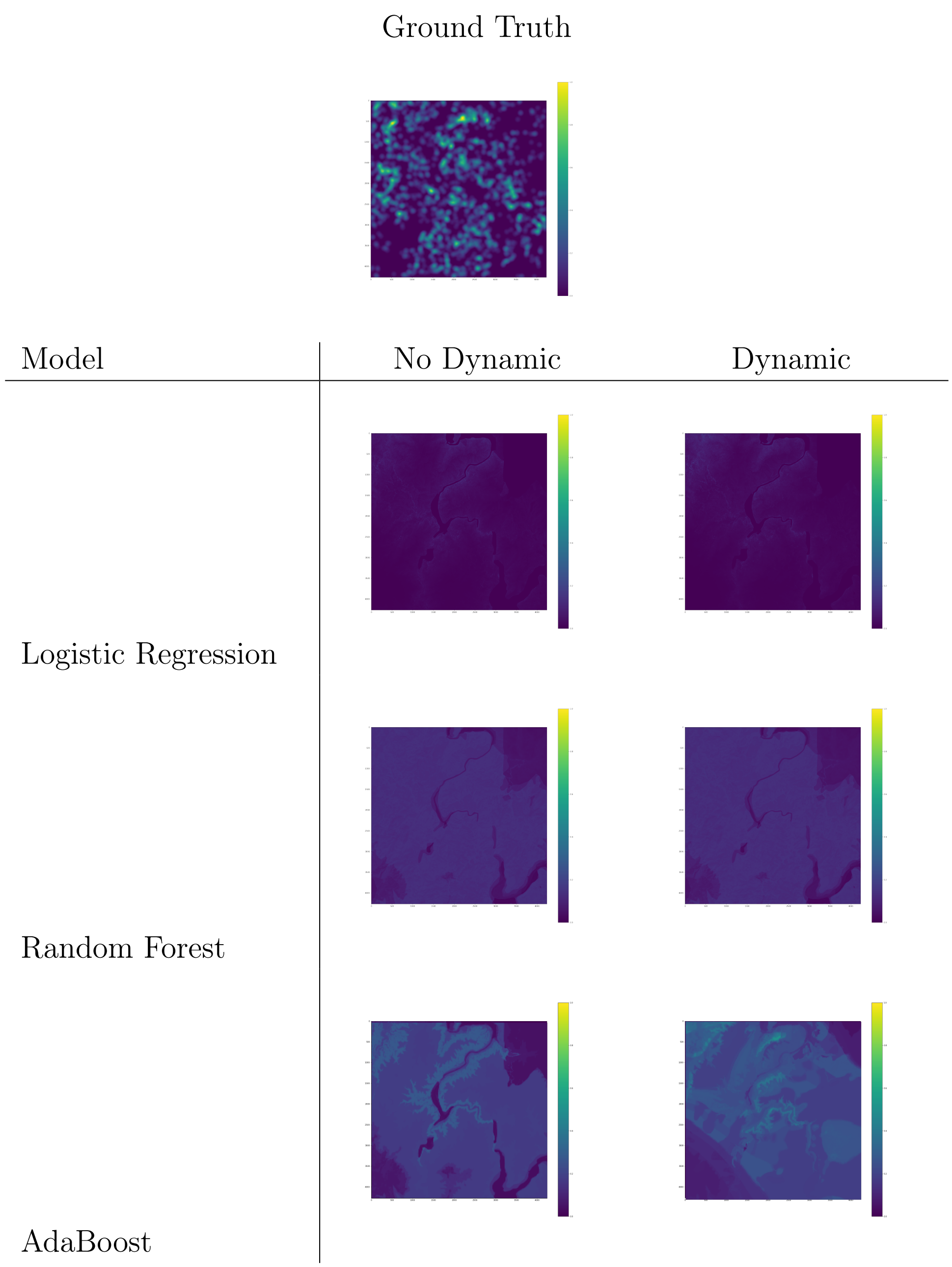

Figure A-8: Visualizations of maps produced by each of the continuous predictions for the Puerto Rico region, excluding the support vector machines. No map was produced for the support vector machines due to the large time complexity of training.

Appendix B

Tables

Table B.1: Landslide density quartile values for Puerto Rico region.

Min	Q1	Median	Q3	Max
0	6.204×10^{-3}	7.446×10^{-2}	1.762×10^{-1}	1

Table B.2: Root mean squared error values for conventional models trained with and without dynamic features for the Puerto Rico region.

Model	No Dynamic	Dynamic
LR	0.1617	**0.1612**
RF	0.1284	**0.1229**
SVM	0.1255	**0.1249**
AB	**0.1390**	0.1563

Table B.3: Classification results for conventional models trained on the Puerto Rico region (ND means no dynamic features were used in the training process, while D means dynamic features were used).

Model	Accuracy		Precision		Recall		AUC Score	
	ND	D	ND	D	ND	D	ND	D
LR	0.3589	**0.3760**	0.3447	**0.3514**	0.3589	**0.3760**	0.6171	**0.6337**
RF	0.3603	**0.3826**	0.3789	**0.3917**	0.3603	**0.3826**	0.6318	**0.6599**
SVM	0.3535	**0.3737**	0.3538	**0.4073**	0.3219	**0.3730**	0.6123	**0.6398**
AB	0.3743	**0.4087**	0.3667	**0.3947**	0.3743	**0.4087**	0.6308	**0.6613**

Table B.4: Classification results for segmentation models trained on the Puerto Rico region (ND means no dynamic features were used in the training process, while D means dynamic features were used). All models were trained for 25 epochs.

Encoder	Patch Size	Accuracy		Precision		Recall		AUC Score	
		ND	D	ND	D	ND	D	ND	D
ResNet-18	64x64	**0.4270**	0.4034	**0.4008**	0.3794	**0.4273**	0.4033	**0.6892**	0.6746
	128x128	0.4017	**0.4253**	0.3527	**0.4187**	0.4028	**0.4255**	0.6584	**0.6819**
ResNet-50	64x64	**0.4227**	0.4059	**0.4301**	0.3930	**0.4234**	0.4057	**0.6934**	0.6751
	128x128	**0.4606**	0.4371	**0.4527**	0.4016	**0.4611**	0.4377	**0.7144**	0.7050

Bibliography

[1] Normalized difference water index: NDWI formula and calculations, Jul 2022.

[2] Alkhasawneh, M. S., Ngah, U. K., Tay, L. T., Mat Isa, N. A., and Al-batah, M. S. Determination of important topographic factors for landslide mapping analysis using mlp network. *The Scientific World Journal 2013* (2013).

[3] Avdan, U., and Jovanovska, G. Algorithm for automated mapping of land surface temperature using landsat 8 satellite data. *Journal of Sensors 2016* (Feb 2016), 1480307.

[4] Brock, J., Schratz, P., Petschko, H., Muenchow, J., Micu, M., and Brenning, A. The performance of landslide susceptibility models critically depends on the quality of digital elevation models. *Geomatics, Natural Hazards and Risk 11*, 1 (2020), 1075–1092.

[5] Di Napoli, M., Marsiglia, P., Di Martire, D., Ramondini, M., Ullo, S. L., and Calcaterra, D. Landslide susceptibility assessment of wildfire burnt areas through earth-observation techniques and a machine learning-based approach. *Remote Sensing 12*, 15 (2020), 2505.

[6] Frampton, W. J., Dash, J., Watmough, G., and Milton, E. J. Evaluating the capabilities of sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing 82* (2013), 83–92.

[7] Gillies, S., et al. Rasterio: geospatial raster i/o for Python programmers, 2013–.

[8] Gillin, C. P., Bailey, S. W., McGuire, K. J., and Prisley, S. P. Evaluation of lidar-derived dems through terrain analysis and field comparison. *Photogrammetric Engineering Remote Sensing 81*, 5 (2015), 387–396.

[9] Gong, W., Thiele, A., Hinz, S., Meyer, F., Hooper, A., and Agram, P. Comparison of small baseline interferometric sar processors for estimating ground deformation. *Remote Sensing 8* (04 2016), 330.

[10] Hastie, T., Rosset, S., Zhu, J., and Zou, H. Multi-class adaboost. *Statistics and its Interface 2*, 3 (2009), 349–360.

[11] Highland, L., and Bobrowsky, P. *The Landslide Handbook – A Guide to Understanding Landslides.* 01 2008.

[12] Johnston, E. C., Davenport, F. V., Wang, L., Caers, J. K., Muthukrishnan, S., Burke, M., and Diffenbaugh, N. S. Quantifying the effect of precipitation on landslide hazard in urbanized and non-urbanized areas. *Geophysical Research Letters 48*, 16 (Aug. 2021).

[13] Lanari, R., Mora, O., Manunta, M., Mallorqui, J., Berardino, P., and Sansosti, E. A small-baseline approach for investigating deformations on full-resolution differential sar interferograms. *IEEE Transactions on Geoscience and Remote Sensing 42*, 7 (2004), 1377–1386.

[14] Lee, S. Application of logistic regression model and its validation for landslide susceptibility mapping using gis and remote sensing data. *International Journal of Remote Sensing 26*, 7 (2005), 1477–1491.

[15] M;, S. H. Landslide susceptibility mapping using gis-based statistical models and remote sensing data in tropical environment.

[16] Mainsant, G., Larose, E., Brönnimann, C., Jongmans, D., Michoud, C., and Jaboyedoff, M. Ambient seismic noise monitoring of a clay landslide: Toward failure prediction. *Journal of Geophysical Research: Earth Surface 117*, F1 (Mar. 2012), n/a–n/a.

[17] Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., and Hamprecht, F. A. A comparison of random forest and its gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC bioinformatics 10*, 1 (2009), 1–16.

[18] Meten, M., PrakashBhandary, N., and Yatabe, R. Effect of landslide factor combinations on the prediction accuracy of landslide susceptibility maps in the blue nile gorge of central ethiopia. *Geoenvironmental Disasters 2*, 1 (Mar 2015), 9.

[19] Nhu, V.-H., Mohammadi, A., Shahabi, H., Ahmad, B. B., Al-Ansari, N., Shirzadi, A., Geertsema, M., R. Kress, V., Karimzadeh, S., Valizadeh Kamran, K., Chen, W., and Nguyen, H. Landslide detection and susceptibility modeling on cameron highlands (malaysia): A comparison between random forest, logistic regression and logistic model tree algorithms. *Forests 11*, 8 (2020).

[20] OpenStreetMap contributors. Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org , 2017.

[21] Parker, R. N., Rosser, N. J., and Hales, T. C. Spatial prediction of earthquake-induced landslide probability. *Natural Hazards and Earth System Sciences Discussions 2017* (2017), 1–29.

[22] Redlands, C. E. S. R. I. Arcgis desktop: Release 10, 2011.

[23] Reichenbach, P., Rossi, M., Malamud, B. D., Mihir, M., and Guzzetti, F. A review of statistically-based landslide susceptibility models. *Earth-Science Reviews 180* (2018), 60–91.

[24] Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (2015), Springer, pp. 234–241.

[25] Soares, L. P., Dias, H. C., Garcia, G. P. B., and Grohmann, C. H. Landslide segmentation with deep learning: Evaluating model generalization in rainfall-induced landslides in brazil. *Remote Sensing 14*, 9 (2022), 2237.

[26] Sörensen, R., Zinko, U., and Seibert, J. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. *Hydrology and Earth System Sciences 10*, 1 (2006), 101–112.

[27] Targ, S., Almeida, D., and Lyman, K. Resnet in resnet: Generalizing residual architectures. *arXiv preprint arXiv:1603.08029* (2016).

[28] Tien Bui, D., Pradhan, B., Lofman, O., and Revhaug, I. Landslide susceptibility assessment in vietnam using support vector machines, decision tree, and naive bayes models. *Mathematical problems in Engineering 2012* (2012).

[29] Wang, H., Zhang, L., Yin, K., Luo, H., and Li, J. Landslide identification using machine learning. *Geoscience Frontiers 12*, 1 (2021), 351–364.

[30] Weston, J., and Watkins, C. Multi-class support vector machines. Tech. rep., Citeseer, 1998.

[31] Zbiciak, A., and Markiewicz, T. A new extraordinary means of appeal in the polish criminal procedure: the basic principles of a fair trial and a complaint against a cassatory judgment. *Access to Justice in Eastern Europe 6*, 2 (Mar. 2023), 1–18.

[32] Zhang, B. Intelligent photogrammetry for digital elevation model production. In *Automatic Target Recognition XXXI* (2021), R. I. Hammoud, T. L. Overman, and A. Mahalanobis, Eds., vol. 11729, International Society for Optics and Photonics, SPIE, pp. 1 – 13.

[33] Zhao, R., Qian, B., Zhang, X., Li, Y., Wei, R., Liu, Y., and Pan, Y. Rethinking dice loss for medical image segmentation. In *2020 IEEE International Conference on Data Mining (ICDM)* (2020), IEEE, pp. 851–860.