UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title

Applications of Artificial Intelligence (AI) Techniques on Remote Sensing Data for Ground Failure Detection, Mobility Assessment, and Infrastructure Monitoring

Permalink

https://escholarship.org/uc/item/6vx4j2q3

Author

Huang, Jhih-Rou

Publication Date

2024

Peer reviewed|Thesis/dissertation

Applications of Artificial Intelligence (AI) Techniques on Remote Sensing Data for
Ground Failure Detection, Mobility Assessment, and Infrastructure Monitoring

By

Jhih-Rou Huang

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Dimitrios Zekkos, Chair
Professor Kenichi Soga
Professor John Radke

Summer 2024

Applications of Artificial Intelligence (AI) Techniques on Remote Sensing Data for
Ground Failure Detection, Mobility Assessment, and Infrastructure Monitoring

Abstract

Applications of Artificial Intelligence (AI) Techniques on Remote Sensing Data for
Ground Failure Detection, Mobility Assessment, and Infrastructure Monitoring
by
Jhih-Rou Huang

Doctor of Philosophy in Engineering - Civil and Environmental Engineering
University of California, Berkeley
Professor Dimitrios Zekkos, Chair

This dissertation incorporates artificial intelligence (AI) techniques on remote-sensing data for ground failure detection, mobility assessment, and infrastructure monitoring. First, the use of AI for landslide detection is investigated. Although an increasing body of work is observed on this topic, a systematic investigation of the factors (input and algorithms) that affect the accuracy of the machine-learning co-seismic landslide detection model has not been attempted. This study leverages the state-of-the-art detailed 3D inventory of more than 700 landslides triggered by the M_w 6.5 Lefkada earthquake on November 17, 2015. The result highlights that feature selection is the most essential factor for successful landslide detection, but the number of features needed is not particularly high. The geospatial distribution and size of the training sample are also important. Input data resolution and machine learning algorithms are the secondary factors that influence detection accuracy. Geospatial distribution affects the training sample size needed to create an accurate landslide detection model, and a wider geospatial distribution of training samples generates a more precise landslide detection model. The work is expanded to consider the generality of the above results for two additional co-seismic landslide events, namely the 2016 Kaikōura earthquake and the 2021 Nippes earthquake, with the goal to identify the commonalities and differences in the success of the machine learning-based landslide detection model. It is found that although feature selection is the most vital factor in the landslide detection model, both topographic and spectral features are useful, with spectral features being most significant in two of the study areas due to their geologic and climatic setting. The input data resolution and training sample size similarly influence the model performance for the three earthquake events, but the importance of segmentation and machine learning algorithms varies across events. Next, a simple mechanistic-model that is based on the Voellmy friction law as incorporated in Rapid Mass Movement Simulation Debris Flow (RAMMS-DF) is tested against statistically significant observations of landslide runout for hundreds of mapped rock avalanches triggered by the M_w 6.5 Lefkada earthquake on November 17, 2015. It is found that the dry-Coulomb friction (μ) controls the simulation's performance, whereas the simulation is less sensitive to viscous-turbulent friction (ξ),

especially for large values of ξ. The simulation's accuracy positively correlates with landslide source area, height, and 3D travel distance. The model does not match very well landslides with small source areas ($<4,000$ m^2), but in these cases, it systematically overestimates landslide runout, i.e., it is inherently conservative. The fourth part of this dissertation leverages lessons learned from the damage observed along Highway 1 during the January 2021 atmospheric river event. A remote sensing-based methodology is developed for system-level monitoring and assessment following natural disasters. It is shown that remote sensing indicators of vegetation loss can detect the occurrence of debris flows and ground failure and indicate the severity of highway damage. Damage severity is correlated to increasingly broader distribution and a lower minimum value of the vegetation loss curve. The last part of this dissertation aims to develop a methodology for fully autonomous remote-sensing-based monitoring of mines. Specifically, the detection of mining instability using high-resolution satellite imagery for eight recent failure cases is considered: the 2022 Jagersfontein tailings dam failure, the 2022 Pau Branco iron ore mine landslide, the 2020 Carmen copper mine landslide, the 2020 Singrauli fly ash dam breach, the 2019 Córrego De Feijão tailings dam failure, the 2018 Cadia gold mine tailings dam failure, the 2014 Mount Polley mine tailing dam failure, and the 2013 Bingham Canyon copper mine landslide. The results show that remote sensing indexes can successfully detect mining failure. In summary, this dissertation demonstrates that new approaches that leverage Artificial Intelligence (AI) and remote-sensing data can be valuable for ground instability detection following natural hazards and can set the stage for fully-autonomous infrastructure monitoring in an expedited and efficient manner.

To my parents, Chi-Fang Huang and Yu-Hsuan Chou

Contents

List of Figures

List of Tables

Acknowledgments

I could not have completed my doctoral study without the support of many people in my life. First, I could not have asked for a better research advisor than Professor Dimitrios Zekkos. His experience and knowledge have guided me every step of the way. His patience, invaluable guidance, and continuous support throughout my doctoral study cannot be overstated. I am extremely grateful to be his student, and his mentorship has consistently motivated me to strive to be a better version of myself.

I would like to express my gratitude to my dissertation committee, Professor John Radke and Professor Kenichi Soga, for reviewing my dissertation and providing valuable feedback. In addition, I would like to sincerely thank my doctoral qualifying exam committee, Professor Jonathan Bray and Professor Marta Gonzalez, for giving me valuable insight during the early stage of my research.

I would like to express my deepest appreciation for the financial support from NASA through the Disasters program grant #80NSSC20K1032 and the Jane Lewis Fellowship for the research in this dissertation. Additionally, I am grateful to Professor Marin Clark for her assistance with the Lefkada earthquake inventory, Dr. Karantanellis for his assistance with the Kaikōura earthquake inventory, and Dr. Pukar Amatya for his guidance in setting up the Semi-Automatic Landslide Detection (SALaD) system. Also, I would like to gratefully acknowledge the support provided by Ryan Turner of Caltrans in providing historical information, cost information, and damage assessment photographs of Highway 1.

I was fortunate to meet many amazing and inspiring people during my doctoral journey. They are Weibing Gong, Parker Blunts, Drew Gomberg, Mirna Kassem, Jongchan Kim, YoungSeok Jo, Wonjun Cha, Xinyi Qian, Brittany Russo, Joel Given, Kil-Wan Ko and many more students and researchers at the University of California Berkeley. Last but not least, I could not have undertaken this journey without my family. I would like to thank my parents and sisters for always being there for me.

Chapter 1 Introduction

1.1   Background and Motivation

Natural hazards are detrimental events that occur beyond our control and are often indirectly made worse by human interventions (Alimohammadlou et al., 2013). Natural hazards have given rise to over 1.6 million fatalities globally since 1990 and have led to estimated economic losses at an average of around USD 260-310 billion annually (Ward et al., 2020). Landslides can lead to substantial economic losses and casualties and are the seventh largest killer among natural disasters (Herath and Wang, 2009; Kjekstad and Highland, 2009; Dai et al., 2002). In the United States, landslides lead to an estimated annual economic loss that ranges between 1 and 3.6 billion dollars and an average death rate of 25-50 people (Highland et al., 1998). Evaluating landslide risk at the global scale is challenging because the spatial extent of individual landslides is typically small, and the diversity of parameters affecting the landslide hazard susceptibility, such as topography, triggering factors, and material properties, is high (Ward et al., 2020). Once a landslide is triggered, it can have significant effects downhill. Thus, runout analyses are the key component of landslide risk assessment and mitigation. Natural disasters provide the opportunity to observe and document the occurrence and runout of landslides, and the observations can be used to calibrate simple empirical or mechanistic runout models, providing valuable site-specific insights.

Natural hazards and the instability of human-made infrastructure, such as mining facilities and highways, can result in significant economic loss and casualties. Resilient infrastructure is essential in maintaining community access and supporting economic activities, and it needs to be designed to resist human-induced and natural hazards over their lifetime (Little, 2002; Dong et al., 2022). Infrastructure is more vulnerable to natural hazards in a changing climate (Mühlhofer et al., 2023). Natural hazards such as earthquakes, landslides, and storms pose a massive risk to infrastructure systems (Little, 2002). Infrastructure failure can lead to enormous disruption of economic activity and economic loss (Kelly, 2015). Transportation infrastructures and road networks are essential to assure accessibility and social and

economic development of a country (Maltinti et al. 2012). The disruption of the transportation system significantly impacts the economy and society because of the interdependency and interconnectivity of the economic sectors (Sohouenou et al., 2020). System-level monitoring and assessment can improve the transportation system's operations and significantly reduce the cost of emergency repair by early detection and prioritization of the most impacted locations along the transportation system. Mining facilities such as tailings dams are some of the largest earth structures, and historical and recent tailings dam failures are reminders that these facilities remain vulnerable and are prone to catastrophic failures. (Azam and Li, 2010). Mining facilities' failure significantly negatively impacts people's lives, the economy, and surrounding properties (Draves and Fox, 1998; Lyu et al., 2020). Analyzing historical events, prediction methods, and forecasting approaches can enhance infrastructures' ability to withstand natural hazards and failure (Little, 2002).

The recently observed rapid proliferation of data and the increase in computational capabilities that can handle large data volumes allow the wide application of artificial intelligence (AI) theory (Guikema, 2020). Applying the AI method to natural hazards promises to enhance predictive accuracy and reduce the computational burden compared to more traditional engineering-based and physics-based approaches (Guikema, 2020). Wang et al. (2021a) presented a framework for the machine learning-based regional scale intelligent modeling of building information for natural hazard risk management. Pourghasemi et al. (2020) used machine learning to assess multi-hazard susceptibility. Yousefi et al. (2020) presented a machine-learning framework for multi-hazards modeling and mapping in a mountainous area. Accurate methods to efficiently detect landslides are needed to generate landslide inventories following natural disasters, such as earthquakes. Machine learning techniques are becoming popular in landslide mapping, detection, and spatial forecasting (Tehrani et al., 2022). However, the large dataset is critical to the success of the AI method, and acquiring a representative and sufficiently large set of training data for natural hazards is challenging, especially for rare events (Guikema, 2020).

Remote sensing gathers information about an environment from a distance, so remote sensing tools do not need human contact with the target and can be space-borne, air-borne, or ground-based (Whitehead and Hugenholtz, 2014; Casagli et al., 2023). The complex and widespread nature of hazards requires accurate measurement with extensive spatial coverage and high acquisition frequency to catch the considerable change in hazards (Casagli et al., 2023). Remote sensing can provide synoptic, systematic, and cost-effective data covering the entire world over a short period of time. Thus, it is well suited to managing natural disasters and infrastructure failures in a timely manner (Krishnamoorthi, 2016). Remote sensing has been deployed to support natural disaster mitigation, preparedness, response, and recovery including early warning systems, damage assessment, and resource allocation to monitor and predict natural hazards, evaluate their impact, and facilitate efficient response and recovery efforts (Twigg, 2004; Kucharczyk and Hugenholtz, 2021; Krichen et al., 2023; Alipouri et al., 2024; Mantovani et al., 1996). Satellite data has been utilized in natural hazard and infrastructure monitoring because of the advancement of algorithms

and the launch of satellites with higher data spatial resolution and acquisition frequency (Crosetto et al., 2016; Mondini et al., 2021; Casagli et al., 2017; Gagliardi et al., 2023; Huang and Zekkos, 2023b). Remote sensing data can be utilized to monitor the condition of infrastructure such as highways and operational and abandoned mining sites remotely and regularly.

1.2   Main Contributions

The main contribution of this dissertation is the implementation of AI techniques on remote sensing data to detect landslides, assess landslide mobility, and monitor ground instability that impacts infrastructure. First, although progress has been made in recent years in implementing AI in landslide detection, a more systematic assessment of the factors that affect the detection of landslides using machine learning algorithms is lacking. Thus, three different study areas at different geologic, climatic, and topographic settings where landslides have occurred during earthquakes are used to investigate systematically the influence of various factors (input and associated algorithms) that affect the results. The large data collected on co-seismic landslides is also used to develop landslide mobility prediction models that are mechanistic-based but are also grounded against field observations. Then, two examples of remote sensing-based frameworks for monitoring infrastructure are presented. The first case is the performance of a remote sensing-based monitoring system for Highway 1 during the January 2021 atmospheric rivers. The second case is developing a framework of an autonomous methodology for mining instability monitoring using remote sensing data.

The novel contributions of this dissertation are summarized as follows:

(a) It quantifies and ranks factors that affect the accuracy of machine learning algorithms in detecting co-seismic landslides in three different settings.

(b) It calibrates a physics-based model against quantified field observations to derive statistically-significant calibrated input parameters for hundreds of rock avalanches.

(c) It proposes a remote sensing-based monitoring system that can continuously track roadway infrastructure conditions on a large scale and follow natural hazards, such as storms and earthquakes.

(d) It presents a framework for an autonomous monitoring system to detect mining instability using high-resolution satellite imagery by evaluating the performance of different remote sensing indexes when failures occur.

1.3   Organization of the Dissertation

This dissertation has seven chapters and is organized below:

- This chapter (Chapter 1) presents this dissertation's background, motivation, contribution, and organization.
- Chapter 2 quantifies the influence of machine learning algorithms, the selection of features, the size and geospatial distribution of training samples, the resolution of input data, segmentation parameters, the spatial distance between non-landslide and landslide training samples, and the geometry of the training sample using the detailed 3D inventory of more than 700 landslides that occurred during the M_w 6.5 Lefkada earthquake on November 17, 2015.
- Chapter 3 expands on the work from Chapter 2 to assess the machine learning-based landslide detection model for the 2015 Lefkada earthquake, the 2016 Kaikōura earthquake, and the 2021 Nippes earthquake event to generate more generalized lessons about the importance of each factor in a range of settings.
- Chapter 4 deploys a Voellmy friction model and calibrates its model parameters against hundreds of mapped rock avalanches triggered by the M_w 6.5 Lefkada earthquake on November 17, 2015, to generate statistically significant observations of landslide runout.
- Chapter 5 classifies the damage observed along Highway 1, a scenic route along the California coastline system, from the January 2021 atmospheric river and develops a remote-sensing-based methodology for system-level monitoring and assessment following natural disasters.
- Chapter 6 presents an autonomous mining instability monitoring methodology by evaluating the performance of different remote sensing indexes from high-resolution satellite imagery and examines eight recent failure cases: the 2022 Jagersfontein tailings dam failure, the 2022 Pau Branco iron ore mine landslide, the 2020 Carmen copper mine landslide, the 2020 Singrauli fly ash dam breach, the 2019 Córrego De Feijão tailings dam failure, the 2018 Cadia gold mine tailings dam failure, the 2014 Mount Polley mine tailing dam failure, and the 2013 Bingham Canyon copper mine landslide.
- Chapter 7 summarizes the conclusions and future work of this dissertation.

Chapter 2 Factors Affecting Machine Learning-based Landslide Detection for the 2015 Lefkada Earthquake

2.1 Introduction

Landslides remain among the most hazardous natural disasters. Between 1998 and 2017, landslides affected an estimated 4.8 million people and killed more than 18,000 individuals (Petrucci, 2022). Landslides result in an estimated economic loss of $20 billion annually (Klose, 2016). The 2017 Hurricane Maria triggered more than 70,000 landslides and damaged 90 % of roads in Puerto Rico, leaving millions of people without access to electricity, water, and emergency services (Lindley et al., 2023). The catastrophic landslides triggered by the 2008 Sichuan Earthquake caused 1381 fatalities, and the 2005 Kashmir earthquake-triggered landslides led to 600 deaths (Gomez et al., 2023). Strategies for landslide monitoring, prediction, and mitigation could mitigate the consequences of landslides (Sim et al., 2022). The creation of landslide inventories is essential for hazard and risk management (Galli et al., 2008). Presently, landslide mapping often relies mainly on visual interpretation, expert knowledge, and field surveys (Guzzetti et al., 2012). Visual interpretation of landslides is time-consuming and resource-intensive for large mapping areas. Also, landslide mapping is subjective and prone to human errors, especially when resources and time are constrained.

Remote sensing techniques such as optical and multispectral imagery, laser scanning, and ground-based interferometry can be utilized for the management of landslide risks (Casagli et al., 2023). Machine learning-based landslide detection using satellite imagery that integrates expert knowledge is expected to be particularly well suited for developing objective landslide inventories in a scalable and timely manner (Barlow et al., 2006; Martha et al., 2010). Nowadays, the availability of remote sensing

data, especially satellite imagery, can support monthly or even daily data collection with global coverage. Machine learning-based landslide detection has the potential to quickly produce a landslide inventory that, although not entirely accurate, can, at a minimum, assist as the first screening step for rapid disaster assessment and mitigation, as well as to improve more detailed mapping efficiency and accuracy.

Machine learning and deep learning algorithms have been applied to landslide detection. Table 2.1 shows some studies implementing ML on satellite imagery for landslide detection. Tehrani et al. (2021) successfully applied a random forest model for landslide detection. Piralilou et al. (2019) showed that the stacking method outperformed single machine learning algorithms such as logistic regression, neural networks, and random forest. Dou et al. (2020) showed that the support vector machine-boosting model outperformed the support vector machine-bagging model and support vector machine-stacking model. Ghorbanzadeh et al. (2022) concluded that the integration of Residual U-Net with the object-based image analysis can produce a better landslide detection model than pixel-based Residual U-Net. Wang et al. (2021b) pointed out that convolutional neural networks outperformed logistic regression, support vector machine, random forest, and boosting methods because of their strength in multi-dimensional data processing and feature extraction. Meena et al. (2022) showed that the U-Net model can detect landslides slightly better than the support vector machine, k-nearest neighbor, and random forest. However, the U-Net model's performance depends on the model's architecture and the complexity of geographical features in the imagery. Ghorbanzadeh et al. (2019) pointed out that deep-learning convolution neural networks do not always outperform support vector machine and random forests because their performance strongly depends on their network architecture design, such as layer depth, the size of the input window, and training strategies. A deep learning algorithm requires a large number of training samples to generate accurate results. A small training sample size is a significant drawback and limitation of the deep learning algorithm (Liu et al., 2020; Qi et al., 2020).

Manual landslide inventories can be used as training and testing samples in supervised machine learning (ML)-based landslide detection with the intent to train the ML model and assess accuracy (Feizizadeh et al., 2014). Training samples significantly impact landslide detection results, and the deficiencies in manual landslide inventory are the leading cause of errors in landslide detection (Pawluszek-Filipiak et al., 2020; Rodriguez-Galiano et al., 2012). The training sample size should be large enough to cover the variability of landslide conditioning factors (Heckmann et al. 2014). Most studies use a training-testing split ratio of 70/30, which is common in ML approaches (Chen et al., 2018; Hong et al., 2015; Huang et al., 2020; Tien Bui et al., 2016), but an investigation of the influence of the split ratio is not often conducted. Mohan et al. (2020) emphasized that the assessment of the model's performance with various training-testing sample ratios must be examined and explored further.

The quality and quantity of the manual landslide inventory are essential in the supervised machine learning-based landslide detection model. Piralilou et al. (2019) used the landslide inventory developed from GPS data and Planet satellite imagery.

Tehrani et al. (2021) collected landslide events from online resources, social media, and the Global Landslide Catalog (GLC) of NASA Goddard Space Flight Center. Dou et al. (2020) utilized the landslide inventory interpreted from post-disaster aerial photography provided by the Geospatial Information Authority of Japan. Ghorbanzadeh et al. (2022) mapped landslide inventory based on Sentinel-2 imagery. Stumpf and Kerle (2011) used the landslide inventory based on fieldwork and visual interpretation of aerial photography and satellite imagery. Tien Bui et al. (2018) used the landslide inventory generated based on satellite imagery, digital elevation model, aerial photos, and fieldwork. Meena et al. (2022) mapped landslides from RapidEye satellite imagery and field observations. Ghorbanzadeh et al. (2019) used the landslide inventory created based on GPS data from the field survey and satellite imagery.

The implementation of different steps in the ML procedures and data input affects the landslide detection model's performance. Piralilou et al. (2019), Tehrani et al. (2021), Dou et al. (2020), Ghorbanzadeh et al. (2022), Wang et al. (2021b), Meena et al. (2022), and Ghorbanzadeh et al. (2019) compared different machine learning models. Tehrani et al. (2021), Dou et al. (2020), and Stumpf and Kerle (2011) assessed the feature importance and selection. Piralilou et al. (2019), Tehrani et al. (2021), and Stumpf and Kerle (2011) evaluated the influence of segmentation. Stumpf and Kerle (2011) tested the landslide detection model with various sample datasets, including Quickbird, IKONOS, Geoeye-1, and aerial photographs. However, the study that comprehensively evaluates the effect of various procedures and data input in the ML landslide detection model is missing.

This study leverages a complete, fully three-dimensional, high-quality landslide inventory for the 2015 Lefkada Earthquake, developed by Zekkos and Clark (2020). The landslide inventory was manually mapped carefully based on high-resolution satellite imagery and three-dimensional models created by Unmanned Aerial Vehicles (UAVs), as shown in Figure 2.1. The inventory dataset includes the area, location, and volume of 716 landslides and separates the source and full area for each landslide. Leveraging this high-quality inventory as the "truth," a comprehensive evaluation of various machine learning algorithms, the features used, and the inputs are conducted with the intent to rank the importance and provide guidance on the factors that matter the most in the development of an ML methodology for landslide detection. In addition, this study explores the effect of the geospatial distribution of training samples, which has not been addressed in the literature.

2.2 Study Area

A M_w 6.5 earthquake occurred on November 17, 2015, in Lefkada, Greece, as shown in Figure 2.2, and triggered nearly 700 landslides along the island's west coast. Unmanned Aerial Vehicle (UAVs) equipped with ultra-high resolution optical cameras as well as satellite imagery were utilized to develop a detailed mapping that includes digital elevation models that were used in interpreting the geometric characteristics of landslides, such as area, height, and volume and separating the amalgamated landslides (Zekkos et al., 2017). Zekkos and Clark (2020) developed the landslide inventory triggered by the 2015 Lefkada earthquake according to the three-dimensional models and high-resolution satellite imagery. The landslide inventory has information on the location, area, and volume of 716 landslides, and the source and entire area for each landslide are generated separately.

Table 2.1: Machine learning algorithms applied in landslide detection

Machine Learning Algorithms	Reference
Neural network	Piralilou et al. (2019); Wang et al. (2021b); Ghorbanzadeh et al. (2019); Kumar et al. (2024); Qin et al. (2024); Wang et al. (2023)
Logistic regression	Piralilou et al. (2019); Wang et al. (2023)
Random forest	Piralilou et al. (2019); Tehrani et al. (2021); Stumpf and Kerle (2011); Wang et al. (2021b); Meena et al. (2022); Ghorbanzadeh et al. (2019); Wang et al. (2023)
K-means clustering	Tehrani et al. (2021); Meena et al. (2022)
Support vector machine	Dou et al. (2020); Tien Bui et al. (2018); Wang et al. (2021b); Meena et al. (2022); Ghorbanzadeh et al. (2019); Wang et al. (2023)
U-Net	Ghorbanzadeh et al. (2022); Yu et al. (2020); Meena et al. (2022); Das et al. (2023); Lu et al. (2023)
Boosting method	Wang et al. (2021b)
Stacking method	Piralilou et al. (2019); Dou et al. (2020)
FCN, U-net, DeeplabV3+, and MFFENet	Xu et al. (2024)

2.3 Methodology

Landslide detection is broadly a process involving extracting and classifying raster datasets. Success in landslide detection is affected by the following: (a) classification method and algorithm, (b) input data, (c) training and testing samples, and (d) metrics used for accuracy assessment. In this study, landslide detection is conducted using the image classification toolbox available in ArcGIS Pro 2.9.1 (Esri Inc., 2021), and the Semi-Automatic Landslide Detection (SALaD) system is used to evaluate the performance of different machine learning algorithms.
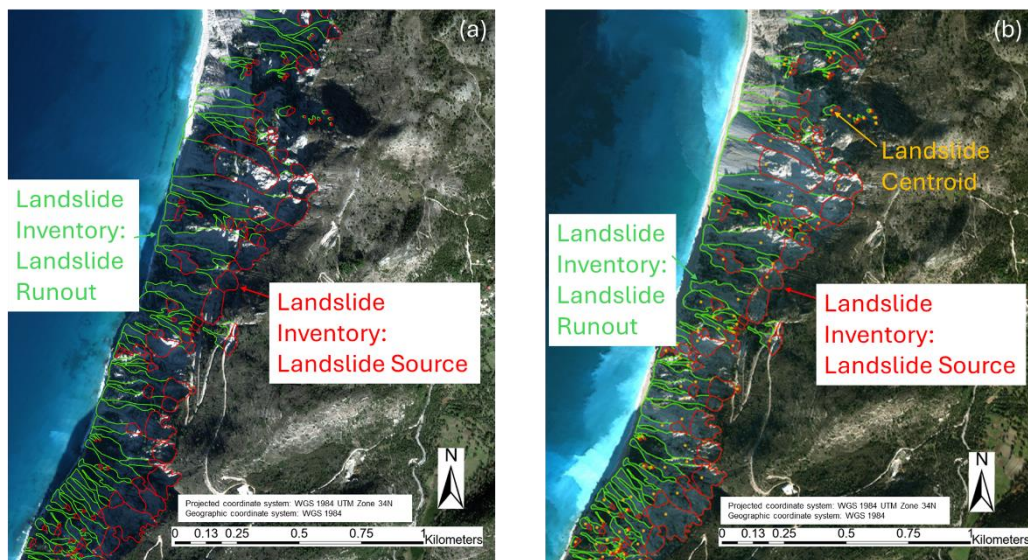
Figure 2.1: Satellite Imagery (a) before the 2015 Lefkada earthquake (b) after the 2015 Lefkada earthquake
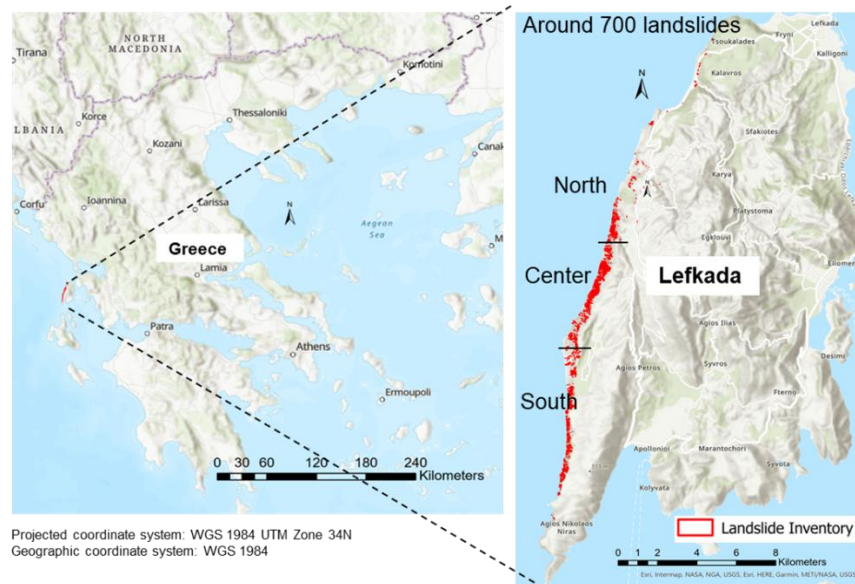
Figure 2.2: Location of study area

2.3.1 Classification Method and ML Algorithms

Object-based image analysis (OBIA) and pixel-based image analysis (PBIA) are the two main image classification approaches. PBIA does not consider neighboring pixels and classifies each pixel into individual classes. OBIA groups adjacent pixels with similar characteristics into objects through a segmentation process. The segmentation process integrates neighboring pixels' spectral, spatial, and textural information (Blaschke et al., 2014). In OBIA, landslides are ensembles of pixels instead of individual spatially unrelated pixels. Thus, OBIA eliminates the salt-and-pepper effect that is often observed in PBIA (Guzzetti et al., 2012; Lu et al., 2011) and allows for the extraction of landslide characteristics such as landslide count and size. In this study, OBIA, as implemented in ArcGIS Pro 2.9.1, considers spectral details, spatial detail, and minimum segment size in pixels. These three parameters define how pixels in the imagery are segmented into objects. Spectral detail and spatial detail range from 1 to 20. Spectral detail represents the importance of spectral differences between pixels, and spatial detail represents the proximity of pixels. Higher values of spectral detail provide better discrimination of different classes when pixels have similar characteristics. In contrast, a lower value contributes to smoothing results with less detail. A higher spatial value is more suitable if pixels of interest are clustered and small. A lower spatial value results in a spatially smoother result. Segments smaller than the minimum segment size in pixels are merged with best-fitting adjacent segments (Esri Inc., 2021). In this study, spectral details, spatial details, and minimum segment sizes have been parameterized to assess the parameters that produce the best landslide detection model.

The Semi-Automatic Landslide Detection (SALaD) system is an open-source landslide detection model in a Python environment developed by Amatya et al. (2021) using OBIA and random forests (RF). This study modifies the SALaD system with supervised machine learning (ML) models from the Python Scikit-learn packages (Pedregosa et al., 2011) to examine the performance of the following ML algorithms: random forests (RF), AdaBoost (AB), gradient boosting (GB), support vector machine (SVM), nearest centroid (NC), naive Bayes (NB), maximum likelihood (MLE). RF is an ensemble machine learning method comprised of multiple decision trees developed by Breiman (2001). For a single decision tree, a small change in training would cause high variance and lead to low accuracy (Breiman, 1996). RF lets each decision tree decide which class the data belongs to and assigns a class based on the majority of the class from decision trees, so RF is less susceptible to overfitting problems in a complex dataset than decision trees (Stumpf and Kerle, 2011; Ghorbanzadeh et al., 2019). Furthermore, RF can handle missing values and resist outliers (Youssef et al., 2016). AB is an iterative ensemble boosting classifier that combines multiple weak classifiers and sets modified weights to incorrectly classified instances to improve the model's performance (Freund and Schapire, 1997). GB is an ensemble classifier that combines weak classifiers and optimizes the classifiers' weights based on the errors of previous iterations to minimize the bias error and enhance the model's accuracy (Friedman, 2001). SVM is a binary classifier based on statistical learning theory (Vapnik, 1999). SVM aims to find the optimal hyperplane by maximizing the margin between two classes (Kavzoglu et al., 2014). SVM can reduce model complexity and works effectively with high-dimensional and non-separable datasets (Mountrakis et al., 2011). NC is a simple classifier in which data is classified into the class with the nearest centroid (Levner, 2005). NB is one of the most simple and effective classifiers based on Bayes's theorem (Murphy, 2006). MLE is a parametric classification relying on the statistical probability density function (Paola et al., 1995). MLE assumes each class has a statistically significant number of normally distributed data and builds a probability density function for each class (Sun et al., 2013). Each data point is assigned to a specific class according to the relative likelihood of that data in the probability density function for each class (Hagner et al., 2007).

2.3.2 Data

The occurrence of landslides can be observed from several features, including topographic and surface texture characteristics, as well as their change before and after a landslide. In this study, Worldview© satellite imagery that has a panchromatic band of 30 to 50 centimeters and multispectral bands of 1.84 m resolution is used. Considering the cloud coverage and imagery quality, the pre-event Worldview© satellite imagery on January 15, 2015, and the post-event Worldview© satellite imagery on December 28, 2015, are collected. The multispectral Worldview© satellite imagery has eight bands: coastal blue, blue, green, yellow, red, red edge, near-infrared (NIR), and near-infrared2 (NIR2). The Worldview© imagery is the same imagery used

for the manual mapping developed by Zekkos and Clark (2020).

The eight bands in Worldview© satellite imagery are used to calculate indexes such as the Red-Edge Normalized Difference Vegetation Index (NDVIre), Red-Edge Chlorophyll Index (CIre), Red-Edge Triangulated Vegetation Index (RTVICore), Global Environmental Monitoring Index (GEMI), Green Chlorophyll Index (CIg), Green Normalized Difference Vegetation Index (GNDVI), Visible Atmospherically Resistant Index (VARI), Modified Soil Adjusted Vegetation Index (MSAVI), Simple Ratio (SR), Red-Edge Simple Ratio (SRre), Modified Triangular Vegetation Index (MTVI2), Enhanced Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Iron Oxide (IO), and mean brightness.

NDVIre, CIre, RTVICore, GEMI, CIg, GNDVI, VARI, MSAVI, SR, SRre, MTVI2, EVI, and NDVI are vegetation indexes. NDVIre calculated from Equation 2.1 considers NIR and red edge band and is effective in estimating crop health (Dong et al., 2019). CIre in Equation 2.2 uses the ratio of NIR and red edge band to estimate the chlorophyll content in leaves (Clevers et al., 2013). RTVICore, in Equation 2.3, estimates the biomass and leaf area utilizing NIR, red edge, and green bands (Dong et al., 2019). GEMI in Equation 2.4 is a nonlinear vegetation index less sensitive to atmospheric effects (Pinty et al., 1992). CIg, as shown in Equation 2.5, estimates chlorophyll content in leaves using NIR and green band (Wu et al., 2012). GNDVI estimates photosynthetic activity in the plant, as shown in Equation 2.6 (Wang et al., 2007a). VARI in Equation 2.7 utilizes visible spectrum, red, green, and blue bands to estimate vegetation (Stow et al., 2005). MSAVI in Equation 2.8 was developed to reduce the effect of bare soil on the vegetation index (Qi et al., 1994). SR in Equation 2.9 and SRre in Equation 2.10 can minimize the influence of topography and atmosphere on the estimation of vegetation (Chen, 1996). MTVI2 in Equation 2.11 uses red, green, and NIR bands to estimate the leaf chlorophyll content (Nguy-Robertson, 2013). EVI in Equation 2.12 uses NIR, red, and blue bands to evaluate vegetation and is less sensitive to background noise (Matsushita et al., 2007). NDVI in Equation 2.13 is computed from the red and NIR bands (Huang et al., 2021). NDWI in Equation 2.14 is a remote sensing index monitoring the change in surface water content (McFeeters, 1996). IO in Equation 2.15 is the red and blue band ratio because the limonitic iron oxide and limonitic-bearing phyllosilicates can lead to the reflectance of the red band and absorption in the blue band (Segal, 1982). Mean brightness is the average value of the red, green, and blue bands.

A 2-m digital elevation model (DEM) before November 17, 2015, is used as the pre-event DEM, and a 2-m DEM on December 28, 2015, is used as the post-event DEM. Both pre-event and post-event DEM are derived from the UAV and satellite imagery. The pre-event and post-event DEM are used to extrapolate terrain characteristics, such as slope inclination, curvature, aspect, flow direction, and hillshade. The slope inclination represents the steepness of the surface, which is the angle between each surface and the horizontal reference point (Yilmaz et al., 2012; Dehnavi et al., 2015). Curvature describes the shape of the slope. A positive curvature value means the surface is upwardly convex, and a negative value indicates the surface

is upwardly concave at that cell. A zero value means the surface is flat (Constantin et al., 2011). Aspect describes the direction of the slope surface (Teillet et al., 1982). Flow direction is created using the D8 flow method from each cell to its neighbors (Tarboton, D. G., 1997). Hillshade visualizes the terrain surface by the slope, aspect of the elevation surface, and light source (Eeckhaut et al., 2005).

The feature difference between the pre-event and post-event values is also calculated and may indicate the change caused by the 2015 Lefkada earthquake or other non-earthquake-related changes that may have occurred between the dates that the data is collected. The differences associated with spectral and terrain features are calculated using Equation 2.16. The pre-event value, post-event value, and feature difference value for each feature are analyzed separately in an effort to quantify the significance of each feature. Additionally, features are composited together, and the performance of multi-feature models is evaluated.

The influence of various satellite imagery is also assessed by considering two more satellite imagery datasets. RapidEye Orthorectified imagery (REOrtho) has five bands (blue, green, red, red edge, and NIR) with a resolution of 3.125 m. The pre-event REOrtho imagery on November 11, 2015, and post-event REOrtho imagery on May 13, 2016, are collected. Landsat 8 has eleven bands (coastal aerosol, blue, green, red, NIR, swir1, swir2, pan, cirrus, tir1, and tir2) with a resolution of 30 m. The pre-event Landsat 8 imagery on November 12, 2015, and post-event Landsat 8 imagery on April 4, 2016, are collected. The imagery difference that subtracts the pre-event imagery from the post-event imagery for Worldview©, REOrtho, and Landsat 8 imagery is utilized to examine the effect of datasets in the landslide detection model.

$$\text{NDVIre} = \frac{\text{NIR} - \text{Red Edge}}{\text{NIR} + \text{Red Edge}} \tag{2.1}$$

$$\text{CIre} = \left(\frac{\text{NIR}}{\text{Red Edge}} - 1\right) \tag{2.2}$$

$$\text{RTVICore} = \left(100(\text{NIR} - \text{Red Edge}) - 10(\text{NIR} - \text{Green})\right) \tag{2.3}$$

$$\text{GEMI} = \text{eta} \, (1 - 0.25 \, \text{eta}) - \left(\frac{\text{Red} - 0.125}{1 - \text{Red}}\right) \tag{2.4}$$

$$\text{where eta} = \frac{2\left(\text{NIR}^2 - \text{Red}^2\right) + 1.5\text{NIR} + 0.5\text{Red}}{\text{NIR} + \text{Red} + 0.5}$$

$$\text{CIg} = \left(\frac{\text{NIR}}{\text{Green}}\right) - 1 \tag{2.5}$$

$$\text{GNDVI} = \frac{\text{NIR} - \text{Green}}{\text{NIR} + \text{Green}} \tag{2.6}$$

$$VARI = \frac{Green - Red}{Green + Red - Blue} \tag{2.7}$$

$$MSAVI = \frac{1}{2}\left(2(NIR + 1) - \sqrt{(2NIR + 1)^2 - 8(NIR - Red)}\right) \tag{2.8}$$

$$SR = \frac{NIR}{Red} \tag{2.9}$$

$$SRre = \frac{NIR}{Red\ Edge} \tag{2.10}$$

$$MTVI2 = 1.5\left(1.2(NIR - Green) - 2.5(Red - Green)\right)\sqrt{(2NIR + 1)^2 - \left(6NIR - 5\sqrt{Red}\right) - 0.5} \tag{2.11}$$

$$EVI = \frac{2.5(NIR - Red)}{(NIR + 6Red - 7.5Blue + 1)} \tag{2.12}$$

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{2.13}$$

$$NDWI = \frac{Green - NIR}{Green + NIR} \tag{2.14}$$

$$IO = \frac{Red}{Blue} \tag{2.15}$$

$$Feature\ Difference = Feature_{post-event} - Feature_{pre-event} \tag{2.16}$$

2.3.3 Training and Testing Samples

All supervised machine learning algorithms require training samples to train the algorithm and testing samples to assess the model's accuracy. The landslide detection model in the current study is a binary classification with two classes: landslide and non-landslide. The manually developed landslide inventory is divided into training and testing samples, and landslides and non-landslide samples are created accordingly. Equalized stratified random sampling is used in this study to ensure enough representation of each class and reduce the variability issue in random sampling (Acharya et al., 2013). Equalized stratified random sampling is a sampling method when samples are randomly distributed within each class, and each class has the same amount of samples (Dhakal et al. 2000). Class imbalance and overlap can also adversely affect the results (Stumpf and Kerle 2011), and thus, balanced and non-overlapping training samples are built to be randomly distributed in each class to prevent bias. In assessing the effect of factors such as segmentation parameters, feature selection, and different satellite imagery, the commonly used training-testing splitting ratio of 70/30 is used. For the evaluation of the influence of the geospatial distribution and size of the training sample, various sizes of manual landslide inventory at different locations are set up as training samples. This study defines the training-testing ratio as the percentage of training and testing samples over the entire manual landslide inventory. As such, a metric is more meaningful in understanding the training and testing needed to achieve specific results for an entire inventory. However, often following a natural disaster, a complete landslide inventory is not available for extended periods of time, especially right after the event. Therefore, expressing the training and testing samples as a percentage of an entire raster dataset is a more practical definition that can be applied anytime after a landslide event.

For sampling, the inventory polygons of the landslides are used, and they are also converted into circles around the landslide centroid to compare the impact of the sampling shape on the results, as shown in Figure 2.1b. The effect of the spatial distance between the landslide and the non-landslide training samples is also assessed.

2.3.4 Accuracy Assessment

In reporting landslide detection results, a positive result means landslides, and a negative means non-landslide. True indicates the result is correct, but false means the result is wrong. Positive predictive value (PPV) in Equation 2.17 quantifies the proportion of true positive samples among all predictive positive samples, while negative predictive value (NPV) evaluates the fraction of true negative samples among all predictive negative samples in Equation 2.18. True positive rate (TPR) measures the proportion of true positive samples among all actual positive samples in Equation 2.19. True negative rate (TNR) is the ratio of actual negatives, which are correctly classified in Equation 2.20. Accuracy and F1 score have been extensively used to evaluate binary classification performance. Accuracy is defined as the ratio between the number of correctly classified samples and the entire number of samples in Equation 2.21 (Wang et al., 2007b). F1 score focuses on one class only and is independent of samples correctly classified as negative (Powers 2015) in Equation 2.22. The F1 score ranges from 0 to 1 and takes increasing values for improving results. However, accuracy and F1 scores are biased by the majority class and are unreliable as the dataset is imbalanced (Bekkar et al., 2013). Matthews correlation coefficient (MCC) can overcome the class imbalance problem. MCC considers positive and negative samples, as shown in Equation 2.23, and ranges between -1 and 1. MCC value of 1 means the perfect classification, 0 represents the classification is no better than random, and -1 represents a complete disagreement between classification and ground truth. MCC has high scores only if all four confusion matrix categories (true positive, false negative, true negative, and false positive) are correctly classified (Chicco and Jurman 2020). The accuracy, F1 score, and MCC are all proportional to the model performance, so in this study, a combined index, overall accuracy (OA), is used to quantify the model's overall performance, as displayed in Equation 2.24. The minimum accuracy, F1 score, and MCC values are 0 %, 0, and -1, respectively. The maximum accuracy, F1 score, and MCC values are 100 %, 1, and 1, respectively. A random classifier has a value of accuracy of 50 %, an F1 score of 0.5, and an MCC of 0. The minimum value of OA is -33.33 %, and the maximum value of OA is 100 %. The random classifier has an OA of 33.33 %, calculated in Equation 2.24.

$$\text{Positive Predictive Value (PPV)} = \frac{TP}{TP+FP} \tag{2.17}$$

$$\text{Negative Predictive Value (NPV)} = \frac{TN}{TN+FN} \tag{2.18}$$

$$\text{True Positive Rate (TPR)} = \frac{TP}{TP+FN} \tag{2.19}$$

$$\text{True Negative Rate (TNR)} = \frac{TN}{TN+FP} \tag{2.20}$$

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \ (\%) \tag{2.21}$$

$$\text{F1 score} = \frac{\text{TP}}{\text{TP}+\frac{1}{2}(\text{FP+FN})} \tag{2.22}$$

$$\text{Matthews correlation coefficient (MCC)} = \frac{\text{TP·TN}-\text{FP·FN}}{\sqrt{(\text{TP+FP})(\text{TP+FN})(\text{TN+FP})(\text{TN+FN})}} \tag{2.23}$$

$$\text{Overall Accuracy (OA)} = \left(\frac{\text{Accuracy}}{100} + \text{F1 score} + \text{MCC}\right) \text{x} \ \frac{100}{3} \ (\%) \tag{2.24}$$

where TP=Ture Positive, TN=True Negative, FP=False Positive, FN=False Negative

2.4 Results and Discussion

2.4.1 Feature Importance

A total of 92 features, listed in Table 2.2, that capture topographic as well as spectral aspects of the landslides are considered as input in a Random Forest (RF) model. As discussed subsequently, RF is shown to have the best results compared to the other machine learning (ML) models and is thus used in the majority of this study. For each feature, the pre-event, post-event, and the difference between the post-event and pre-event values are used as input in the landslide detection model. Classification results using a single feature are shown in Table 2.2. In Table 2.2, topographic features are highlighted in gray, and spectral features are highlighted in white. Among all features, the most important features are the post-slope and pre-slope, and have the highest overall accuracy (OA), around 80 %, followed by pre-event vegetation indexes, pre-NDVI, pre-NDVIre, pre-CIre, and pre-RTVICore. Post-elevation, pre-elevation, aspect difference, curvature difference, and GEMI difference have the lowest OA (between 13 % and 19 %). Figure 2.3a is the histogram of post-slope for the landslide and non-landslide areas in the training samples. Figure 2.3b is the histogram of curvature difference value for the landslide and non-landslide areas in the training samples. Figure 2.3a shows that the post-slope values of landslide and non-landslide in training samples are significantly different. Non-landslide training samples have a mode of post-slope value around 16°, while landslide training samples have a mode of post-slope value around 36°. Non-landslide training samples tend to have a lower post-slope than the landslide training samples. In contrast, landslide and non-landslide training samples have similar curvature differences. Both landslide and non-landslide training samples have a mode of curvature difference value around zero, which is why the curvature difference is not a vital feature in helping the model detect landslides and non-landslides.

Results improve if multiple features are used, which is expected. Thus, using multiple features for landslide detection is advantageous compared to any model that is based on a single feature listed in Table 2.2. This is shown in Figure 2.4, where the performance of multi-feature models using the top 5, 10, 20, 30, 40, 50, 60, 70, and 80 features with best performance in Table 2.2 is shown. The top 5 features are post-slope, pre-slope, pre-NDVI, pre-NDVIre, and pre-CIre. Two of the top 5 features are topographic features, and three of the top 5 features are spectral features. The top 10 features are post-slope, pre-slope, pre-NDVI, pre-NDVIre, pre-CIre, pre-RTVICore, NDVI difference, post-flow direction, pre-CIg, and post-GEMI. Three of the top 10 features are topographic features, and seven of the top 10 features are spectral features. This shows that both topographic and spectral features play an important role in the successful detection of the Lefkada landslides.

Table 2.2: Feature importance

Ranking	Feature	OA (%)	Ranking	Feature	OA (%)	Ranking	Feature	OA (%)
1	Post Slope	79.9	41	EVI Difference	50.6	81	Post NDVIre	31.9
2	Pre Slope	79.3	42	Iron Oxide Difference	49.9	82	Pre SR	28.6
3	Pre NDVI	73.8	43	MSAVI Difference	49.9	83	NIR2 band Difference	27.4
4	Pre NDVIre	70.6	44	Post NIR	49.0	84	Pre SRre	25.4
5	Pre CIre	70.0	45	Post NIR2 band	48.9	85	Post Iron Oxide	24.2
6	Pre RTVICore	69.1	46	NDWI Difference	48.8	86	Post SR	22.8
7	NDVI Difference	68.5	47	CIre Difference	47.2	87	Post Coastal band	21.5
8	Post Flowdirection	66.9	48	Pre MSAVI	47.0	88	Post DEM	19.3
9	Pre CIg	66.4	49	Pre MTVI2	46.9	89	Pre DEM	18.7
10	Post GEMI	66.3	50	Pre Flowdirection	46.4	90	Aspect Difference	16.8
11	Post Aspect	66.2	51	Pre Aspect	46.2	91	Curvature Difference	16.2
12	Pre Yellow band	65.4	52	Post CIre	45.7	92	GEMI Difference	12.9
13	CIg Difference	64.8	53	Post CIg	45.6			
14	Post NDVI	64.6	54	Post SRre	45.2			
15	Post NDWI	64.2	55	NIR Difference	45.0			
16	Pre NIR	63.4	56	Pre VARI	44.2			
17	Pre NDWI	63.1	57	Pre Blue band	44.1			
18	GNDVI Difference	62.6	58	VARI Difference	43.8			
19	Pre GEMI	61.6	59	Pre Iron Oxide	43.7			
20	Post VARI	60.6	60	Pre Mean Brightness	43.6			
21	Post Red band	60.5	61	RTVICore Difference	43.3			
22	Post Hillshade	59.7	62	NDVIre Difference	42.2			
23	Pre GNDVI	59.3	63	Post Green band	42.0			
24	Pre Red band	59.2	64	Post RTVICore	41.3			
25	Post Yellow band	58.2	65	Hill Shade Difference	41.1			
26	Pre RedEdge band	58.0	66	MTVI2 Difference	40.8			
27	Pre Green band	57.6	67	Post RedEdge band	40.7			
28	Post Mean brightness	55.5	68	Pre Hillshade	40.1			
29	Coastal band Difference	55.5	69	Pre Coastal band	39.8			
30	Mean Brightness Difference	55.3	70	Post EVI	39.3			
31	SR Difference	55.0	71	NDVI Difference	39.1			
32	Post MTVI2	54.9	72	Yellow band Difference	38.4			
33	Pre NIR2 band	53.7	73	NDWI Difference	37.3			
34	Blue band Difference	52.2	74	Green band Difference	37.2			
35	Red band Difference	51.7	75	SRre Difference	37.1			
36	Flow Direction Difference	51.3	76	Post GNDVI	35.9			
37	Post MSAVI	51.2	77	Post Curvature	35.0			
38	RedEdge band Difference	50.8	78	Slope Difference	34.8			
39	Post Blue band	50.7	79	Pre Curvature	32.7			
40	Pre EVI	50.6	80	DEM Difference	32.4			
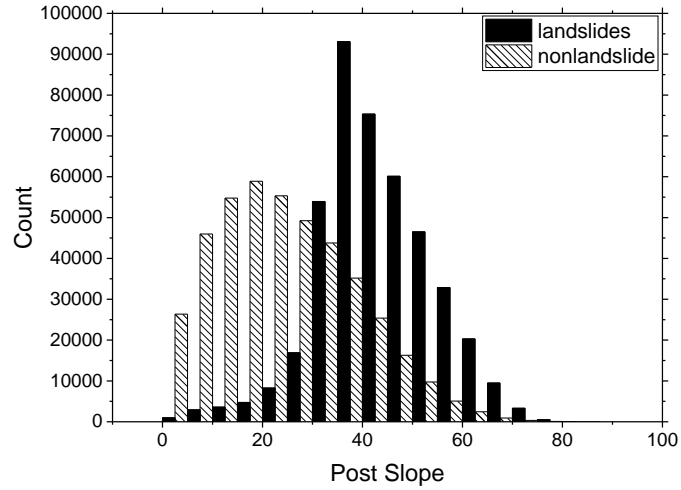
(a) Post Slope
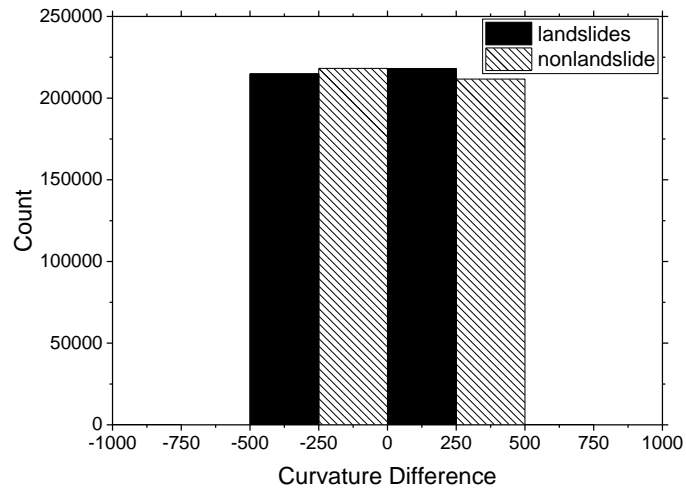
(b) Curvature Difference

Figure 2.3: Histogram of feature

The OA is 83.4 % with the top 5 features and increases to 89 % for the top 10 features. As shown in Figure 2.4, the single-feature models have OA between 12.9 % and 79.9 %. However, the multiple-feature model with the top 5 features has an OA of 83.4 %, and the model performance remains generally the same when additional features are included. The models with the top 20, 30, 40, 50, 60, 70, 80, and 92 features all have OA near 90 %, indicating no benefit of incorporating many additional features. Thus, in subsequent analyses to assess the influence of sample size and spatial distribution, the top 10 features are included in the model. Figure 2.5 shows that the model with the 10 best features can reveal the landslide location well, as highlighted in true positive. However, the result is overestimated compared to the landslide inventory, as shown in false positives. A few landslide areas are detected as non-landslides, as demonstrated by false negatives.
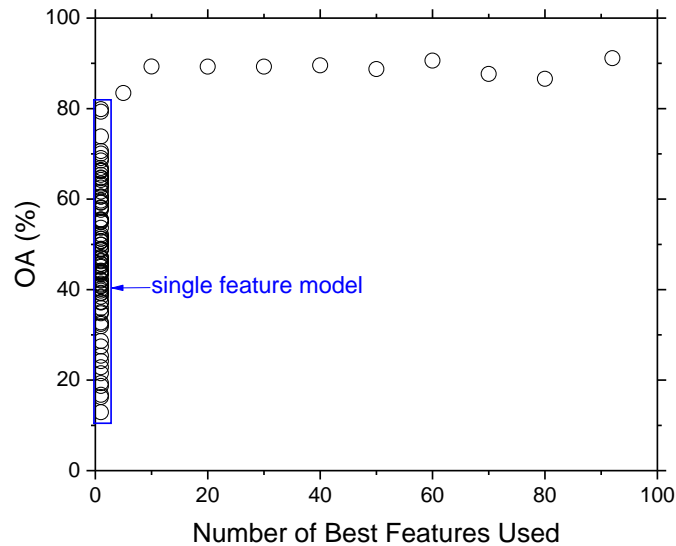
Figure 2.4: The performance of single and multi-feature models
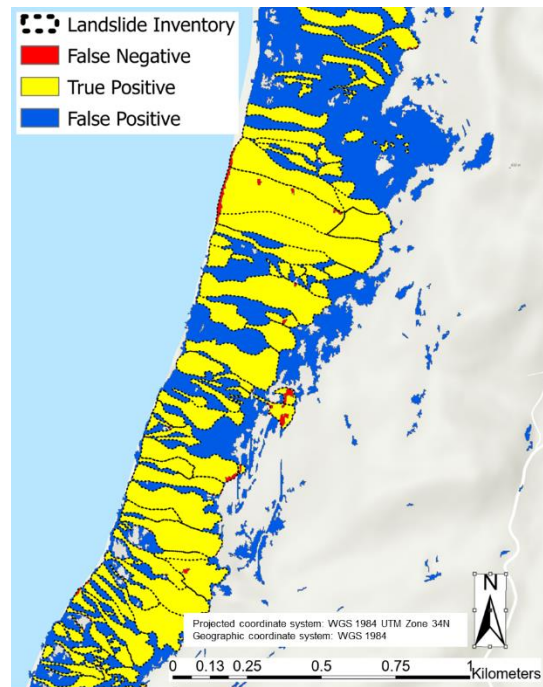
Figure 2.5: The landslide detection result with the 10 best features

2.4.2 Effect of Geospatial Distribution and Size of Training Sample

As mentioned earlier, the 2015 Lefkada earthquake triggered nearly 700 landslides concentrated on the steep western coastline of the Lefkada island. The landslide inventory is divided geographically into the north, center, and south regions of the coastline in Figure 2.2, and the different number of landslides from each region is considered as training samples to assess the impact of the size and geospatial distribution of the training sample. The results are compared against a testing sample that is evenly distributed along the entire coastline.

First, the landslide detection model's performance improves significantly as the training sample size increases from 1 % to approximately 20 %, as shown in Figure 2.6a-d for the accuracy, MCC, F1 score, and OA, respectively. Small improvements are observed beyond 20 %, but no significant model improvement is observed when the training sample size exceeds 40 %.

Figure 2.7 illustrates radar plots based on PPV, NPV, TPR, and TNR. The polygon in the radar plot becomes bigger as the PPV, NPV, TPR, and TNR have higher values, indicating better model performance. In Figure 2.7a, the training samples are selected from the entire coastline and have the best performance, as illustrated by the largest polygon. The model's performance worsens for regionally concentrated training

samples, specifically samples restricted to the north, south, and center, with only 3 %
of the landslide inventory as the training sample. When the training sample size
increases to 20 %, the training samples in the north, south, and entire coastline lead to
similar performance, as shown in Figure 2.7b. The model has the worst performance
when the training samples are in the center of the coastline only, even when 20 % of
the landslide inventory is used as training samples. In short, selecting landslides and
non-landslide samples from the entire coastline contributes to the best landslide
detection ability, even for the small amount of training samples (<3 %) of the landslide
inventory. In contrast, training samples from the center of the coastline only do not
produce a comparably good landslide detection result, even with 20 % of inventory as
training samples. When the training sample size reaches 40 percent of the inventory,
the geospatial distribution of training samples does not affect the model performance.
As displayed in Figure 2.7c, training samples in the center and the entire coastline have
a value of 90 for PPV, NPV, TPR, and TNR. PPV, NPV, TPR, and TNR remain
constant when the training sample size increases from 40 % to 80 % with the
distribution of training samples over the entire coastline.

(a)

(b)

(c)

(d)
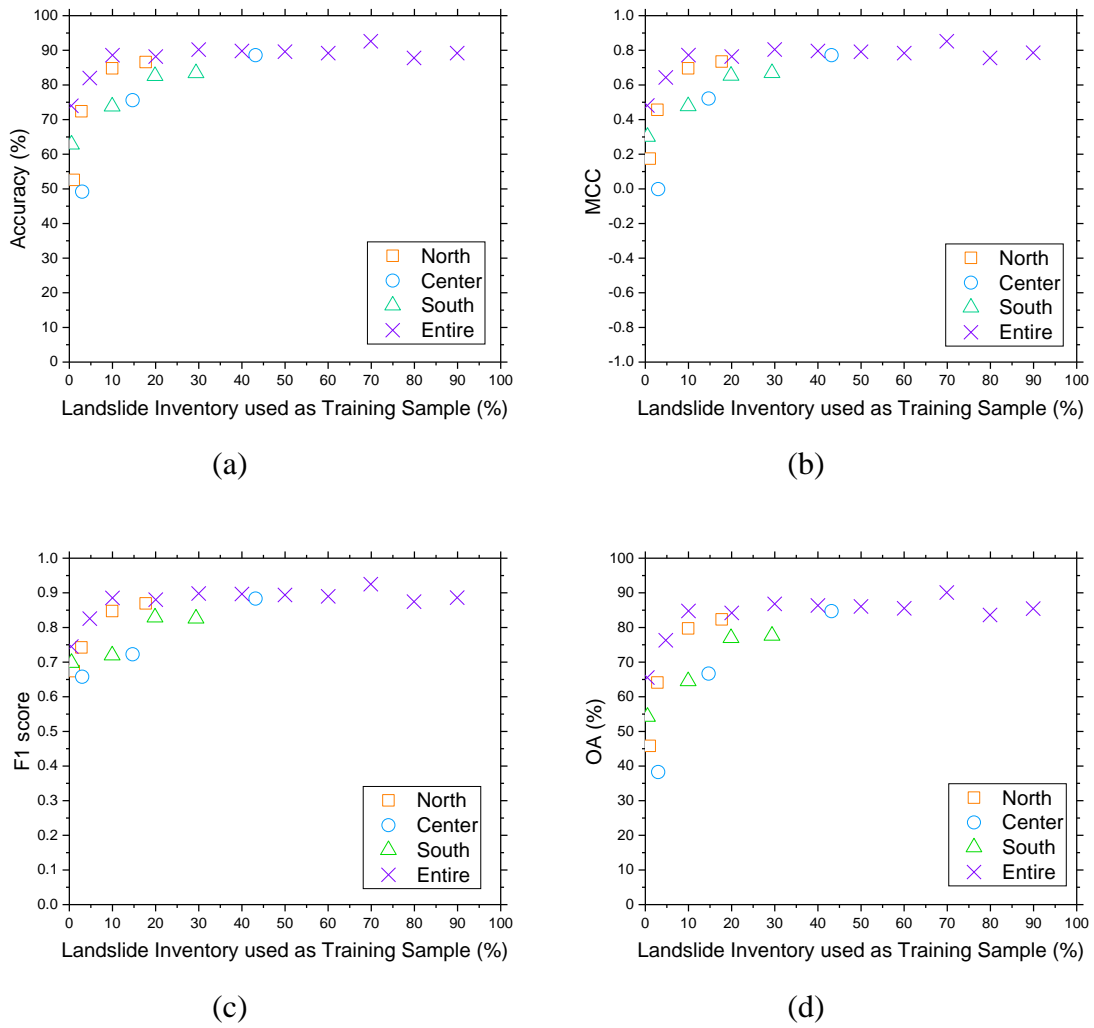
Figure 2.6: Model performance- geospatial distribution and size of training sample
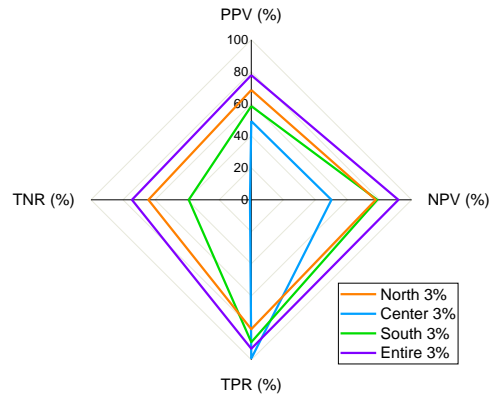
(a)
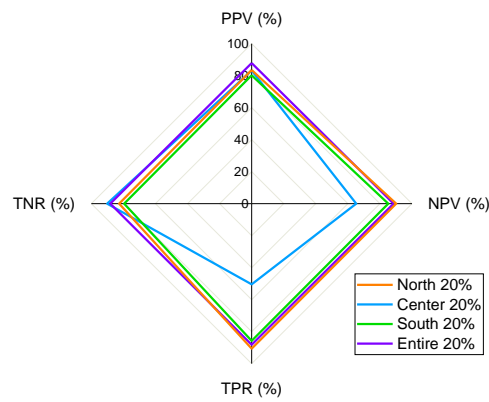
(b)
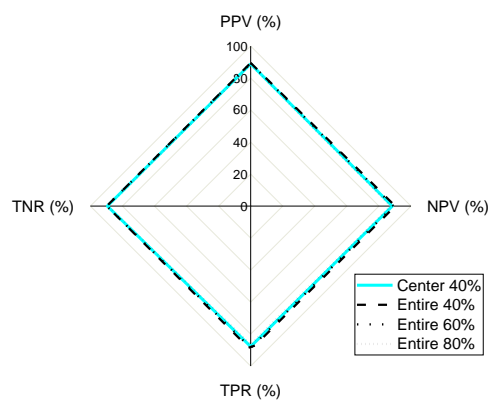
(c)

Figure 2.7: Model performance- training sample size (a) 3 % (b) 20 % (c) 40 %, 60 %
and 80 % of landslide inventory

Mohan et al. (2020) concluded that the small training sample size is a major drawback in landslide detection. Pawluszek-Filipiak et al. (2020) and Rogan et al. (2008) also showed that the landslide detection ability decreases as the training sample size decreases, as a small training sample may not be enough to capture the spatial variability of features. The landslide detection model results become more reliable as the training sample size increases. However, the large training sample size may affect the model's independent classification. 70/30 is the most commonly used training-testing splitting ratio in the supervised landslide detection and susceptibility model (Chen et al., 2018; Hong et al., 2015; Huang et al., 2020; Tien Bui et al., 2016), but evaluation of the effect of the different split ratios, as done in this study, is lacking.

Shirzadi et al. (2019) indicated that 60/40 and 70/30 splitting ratios provided the highest prediction accuracy in 10 m resolution data, and 80/20 and 90/10 splitting ratios had the best prediction accuracy in 20 m resolution raster data in the landslide susceptibility model. In the current study, the performance of the landslide detection model is significantly improved when the training sample size increases from 0 to 20 percent of landslide inventory, as shown in Figure 2.6. As the amount of training sample achieves 40 percent or above, the model performance remains stable.

This study highlights that the geospatial distribution of training samples is essential in landslide detection. Training samples throughout the event area can capture the characteristics of the entire dataset well and provide the best landslide detection result compared to training samples restricted in geographic regions such as the north, center, and south of the coastline. Even with a small training sample size, training samples covering the entire event area can still produce reliable landslide detection results. As training samples are distributed across the entire coastline, the model's accuracy can reach 75 % even with only 1 % of landslide inventory as training samples in Figure 2.6a. When the size of training samples reaches 10 % of the landslide inventory all over the entire coastline, the model accuracy reaches 90 %. The model is stable and has reached its accuracy limit as the size of training samples increases above 10 % of inventory over the entire coastline. The landslide detection model with training samples from localized regions such as the coastline's north, center, and south parts needs a larger amount of training samples to generate an accurate model.

As shown in Figure 2.6a, the training samples from the north, center, or south part of the coastline individually can result in 50 % to 70 % accuracy, with the training samples size below 5 % of the landslide inventory. When the training sample size reaches 10 percent or above, the accuracy can be 70 % or above, even if the training samples are in a localized area. Among the north, center, and south parts of the coastline, the training samples in the center have the worst performance, especially with a small training sample size. The reason is that the 2015 Lefkada earthquake triggered new landslides, mainly in the north and south of the Lefkada coastline. Zekkos et al. (2017) showed that most landslides in the center of the Lefkada coastline are re-activated, which already failed on the August 14, 2003 earthquake and failed again in 2015. So, using landslides in the center of the coastline as training samples results in the worse detection performance, possibly because re-activated landslides are misclassified as new landslides. The small size of training samples from a localized

part of the event area may provide biased information and generate an unreliable landslide detection model. This has implications for mapping: manual mapping should start from landslides with a sparse but wide geospatial distribution throughout the entire dataset rather than densely concentrated landslides in a localized event area to efficiently get an accurate landslide detection model.

Figure 2.8 compares the landslide area in the detection result to the landslide area in the inventory. As the training sample size increases, the detected landslide area approaches the mapped landslide inventory. Note that the entire mapped landslide area is 2.29 km^2 out of the event area of 26 km^2, i.e., it represents a small part of the entire area. When the training sample size is below 10 % of the landslide inventory, the ratio of the detected landslide area to the mapped landslide area ranges between 3 and 11, i.e., the model over-predicts significantly the landslides. When the training sample size is above 30 %, the ratio of detected landslides to the mapped landslides decreases to 2.5 and becomes stable. This indicates that even in the best model predictions, the detected landslides are greater than the mapped landslides. This is not surprising and has been shown by others (Ghorbanzadeh et al., 2022; Stumpf and Kerle, 2011). This is due to several factors such as input resolution, data quality, etc., but most importantly, landslides are always a very small part of an affected area.

The effort is also made not only to analyze landslides in terms of area coverage but also in terms of the amount and the size of individual landslide in the detection model. Landslide amalgamation represents a challenge in this case, as clusters may include more than one landslide. Flow direction is introduced to separate amalgamated landslides better in the detection result. This effort shows that the model overestimation is particularly due to the over-prediction of small landslides. The resolution of the input raster in this study is 2 m, so landslides smaller than 4 m^2 are considered unreliable predictions and are ignored. Figure 2.9 shows that the landslide detection model detects around 47,000 landslides with the landslide area < 5000 m^2, and the landslide inventory has around 590 landslides with the landslide area < 5000 m^2. The landslide detection result has a similar amount of landslides with the landslide area > 5000 m^2 to the landslide inventory. So, the landslide detection model captures larger landslides better than smaller landslides. The OA for landslides larger than 1600 m^2 is 1 % higher than those larger than 36 m^2.
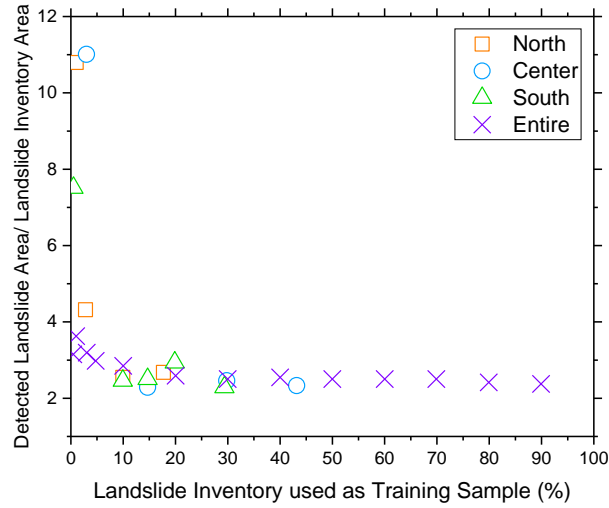
Figure 2.8: Comparison between detected landslide area and manual landslide inventory
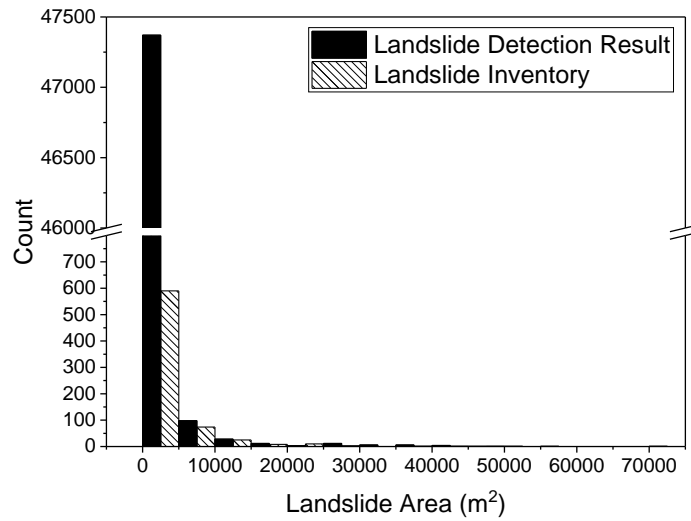
Figure 2.9: Histogram of landslide area in the landslide detection result and inventory

2.4.3 Effect of Machine Learning Models

A number of machine learning algorithms are also investigated, specifically random forests (RF), AdaBoost (AB), gradient boosting (GB), support vector machine (SVM), nearest centroid (NC), naive Bayes (NB), and maximum likelihood (MLE). It is found that the OA for RF, AB, GB, SVM, NC, NB, and MLE is 82.9 %, 78.3 %, 79.1 %, 71.2 %, 67.6 %, 68.4 %, and 66.2 %, respectively. Ensemble methods (RF, AB, and GB) generate the best detection result among all machine learning algorithms. RF, in particular, has the highest OA in the landslide detection model for the 2015 Lefkada earthquake event. This is because RF is able to learn complicated geospatial characteristics, and an ensemble model can improve machine learning performance (Chen et al., 2018). MLE has the lowest OA because MLE is sensitive to the number of samples and needs a statistically significant number of samples to create a reliable classification (Belgiu et al., 2016). MLE is suitable for unimodal data but not for multimodal data since the MLE classifier assumes the data is normally distributed (Liu et al., 2011), which is not the case for some input data. For this reason, RF is used to investigate the influence of input data and sampling strategies in the landslide detection model.

2.4.4 Effect of Satellite Imagery

The imagery difference of Worldview©, REOrtho, and Landsat 8 imagery has an OA of 77.11 %, 75.39 %, and 63.49 %, respectively. Worldview©, REOrtho, and Landsat 8 have several differences, including the different number of bands, the timing of available pre-event and post-event imagery, and imagery resolution. Worldview© imagery has eight bands with a resolution of 1.84 m, REOrtho imagery has five bands with a resolution of 3.125 m, and Landsat 8 imagery has eleven bands with a resolution of 30 m. The pre-event Worldview© imagery, REOrtho imagery, and Landsat 8 are collected on January 15, 2015, November 11, 2015, and November 12, 2015, respectively, and the post-event Worldview© imagery, REOrtho imagery, and Landsat 8 are collected on December 28, 2015, May 13, 2016, and April 4, 2016, respectively.

Among the three imagery datasets, the Worldview© imagery generates the best landslide detection model, and the REOrtho imagery also leads to a good detection model. Worldview© imagery contributes to a more accurate landslide detection model than REOrtho and Landsat 8 imagery because the landslide inventory developed by Zekkos and Clark (2020) was mapped on the Worldview© imagery. However, the REOrtho imagery can still generate a good landslide detection model with a similar performance as the Worldview© imagery.

2.4.5 Effect of Data Resolution

The raster data composited of the 10 best features in Table 2.2 is resampled into 5 m, 10 m, 30 m, and 50 m resolutions to examine the effect of data resolution, excluding other factors such as the number of bands, the date of pre-event and post-event imagery from different satellite imagery datasets. Figure 2.10 shows that the finer resolution contributes to a better detection model. The OA difference between the 5 m raster and 50 m raster data is 21.14 %.
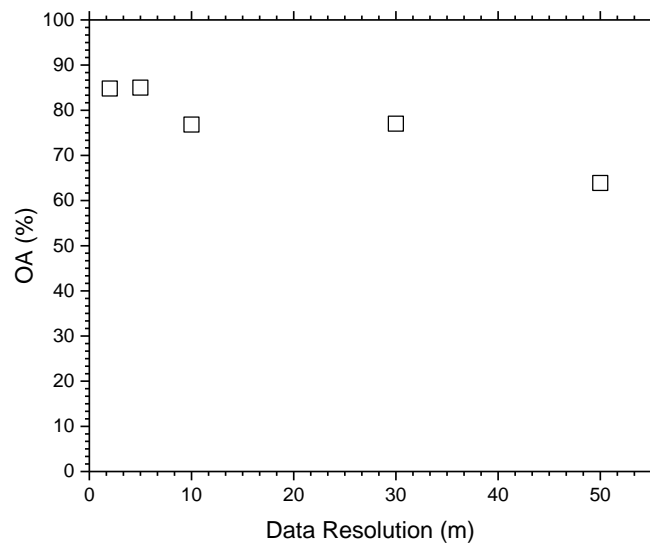
Figure 2.10: Model performance- data resolution

2.4.6 Effect of Spatial Distance Between Landslide and Non-Landslide Training Samples

The influence of the spatial distance between landslide and non-landslide training samples is also investigated to determine the best strategy for generating non-landslide training samples. For this investigation, the commonly used training-testing sample ratio of 70/30 is utilized, and the same number of non-landslide samples with different distances from landslides is developed accordingly. Non-landslide training samples are selected to be located within 18 m, between 70 m and 100 m, and more than 100 m from landslides, as well as completely random distribution, and the results are compared. OA varied from 75.8 % to 89.3 %. The random distribution of non-landslide training samples has the best landslide detection performance, indicating that the location of non-landslide training samples does affect the landslide detection performance. Besides random distribution, OA increases as the non-landslide training samples are further away from the landslide area, as shown in Figure 2.11.
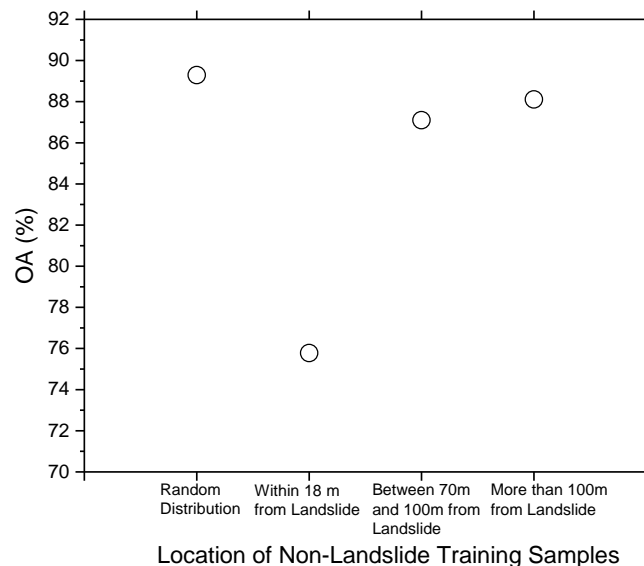
Figure 2.11: Model performance- the spatial distance between non-landslide samples and landslide samples

The sampling strategy is vital in the landslide detection model. Lin et al. (2022) showed that the landslide susceptibility model with non-landslide samples in the mountainous area outperformed the non-landslide samples in the entire event area and concluded that the spatial extent of non-landslide samples influences the landslide susceptibility model. This study shows that distances between landslides and non-landslide training samples affect the performance of landslide detection with non-landslide samples randomly distributed throughout the event area, giving the best

classification result. However, non-landslide training samples further away from the landslide area have better detection results than the closer-distance ones.

2.4.7 Effect of Segmentation

Segmentation is known to influence the performance of OBIA (Piralilou et al., 2019; Tehrani et al., 2021; Stumpf and Kerle, 2011; Huang and Zekkos, 2023a). In this study, three parameters, spectral detail, spatial detail, and minimum segment size in pixels, are applied to group neighboring pixels into objects. The spectral detail and spatial detail range between 1 and 20. The spectral detail and spatial detail values of 1, 10, and 20 are tested to see the influence of spectral detail and spatial detail on the model. The model with the spectral detail of 1, 10, and 20 has an OA of 87.4 %, 89.6 %, and 92.2 %, respectively. The model with the spatial detail of 1, 10, and 20 has an OA of 91.6 %, 89.6 %, and 90.4 %, respectively. Varying the minimum segment size in pixels from 1 to 10, then 20, resulted in similar OA (OA was 89.2 %, 89.6 %, and 89.6 % respectively). The results show that higher spectral details can yield a better landslide detection model. However, the variation in spatial detail and the minimum segment size in pixels does not significantly affect the model's performance.

2.4.8 Summary Ranking of Importance for Different Factors in Landslide Detection

Since a machine learning-based landslide detection process includes many algorithm steps and inputs, an effort is made to objectively quantify the importance of each of the parts for the 2015 Lefkada earthquake event. One way to quantify the importance of each step of the process is to calculate the difference in OA for the best and worst results when that process is evaluated. The results are summarized in Table 2.3. Among all factors, feature selection governs the landslide detection performance the most. For example, the post-slope feature can produce the landslide detection model with an OA of 79.9 %, but the GEMI Difference feature can only generate the model with an OA of 12.9 %. This results in a difference in OA of 67 %.

The size and geospatial distribution of the training sample are the second essential factors in the detection model. The OA difference is 51.82 % between the training sample size, 3 %, and 70 %. When the size of the training samples is 20 % of the landslide inventory, the OA difference is 33.05 % between the center of the coastline and the entire coastline.

The data resolution affects the model's performance. The OA difference between the 5 m raster and 50 m raster data is 21.14 %. The OA difference generated by different algorithms is 16.69 %, which is significantly lower than the previous factors. RF achieves the best performance, but MLE has the lowest OA. Satellite imagery and the spatial distance between non-landslide and landslide training samples similarly influence the detection result. The OA difference between the Worldview©

and the Landsat 8 is 13.62 %. Different spatial distances between landslide and non-landslide training samples can lead to an OA difference of 13.51 %. Segmentation parameters, the geometry of the training sample, and landslide size have the minimum effect with an OA difference of around or less than 10 %. Segmentation affects the OBIA performance, and the difference between the minimum and maximum OA caused by different segment parameters is 10.07 %. The OA difference between using landslide inventory polygons as training samples and circles buffering from centroids of landslides is 1.62 %. The landslide detection model detects larger landslides better than smaller landslides and can lead to an OA difference of 1 %.

Table 2.3: Ranking of importance for different factors in landslide detection

	OA Difference % (Maximum OA - Minimum OA)
Feature	67
Size of Training Sample	51.82
Geospatial Distribution of Training Sample	33.05
Data Resolution	21.14
Machine Learning Algorithm	16.69
Satellite Imagery	13.62
Spatial Distance Between Landslide and Non-Landslide Training Samples	13.51
Segmentation Parameters	10.07
Geometry of Training Sample	1.62
Size of Landslides	1

2.5 Conclusions

This study investigates and quantifies the importance of different settings and inputs of the machine learning-based landslide detection model for the 2015 Lefkada earthquake event. The type of features used, their geospatial distribution, and the size of training samples are the most essential factors in detecting landslides. Both topographic and spectral features are important, but just using the top 10 features is similar in detection success to using 91 features. The geospatial distribution of training samples matters. A wider geospatial distribution of training samples can better capture the data characteristics and generate a more reliable detection model, especially when a small training sample is used. Data resolution and the type of machine learning algorithms also affect the results, but not as much as the previous factors. Segmentation parameters, the geometry of the training sample, and landslide size are less critical.

Chapter 3 Lessons Learned on the Efficacy of Machine Learning-based Landslide Detection Following Three Recent Earthquake Events

3.1 Introduction

Landslides represent a major geologic hazard and can cause damage to natural and social environments (Pardeshi et al., 2013). Hundreds to thousands of landslides can be triggered by earthquakes or storm events (Keefer, 2002). Manual mapping requires a high degree of skill and experience in visual interpretation, has limitations in reproducibility, and is time and resource-intensive (Ramli et al., 2010). The application of Machine learning (ML) or deep learning algorithms using remote sensing data to detect landslides semi-automatically or automatically has been the focus of several studies in recent years (Ghorbanzadeh et al., 2019; Ullo et al., 2021; Ghorbanzadeh et al., 2022a; Ye et al., 2019; Nava et al., 2021; Meena et al., 2022; Sameen et al., 2019; Cai et al., 2021). Ghorbanzadeh et al. (2019) evaluated the performance of the artificial neural network, support vector machine, random forest, and deep-learning convolution neural networks with Rapid Eye satellite imagery and topographic factors for the landslide detection in the Rasuwa district in Nepal. They concluded that the deep learning convolution neural networks do not outperform other algorithms because of the network architecture design and data fusion. Ullo et al. (2021) reported that the ResNet-101 can produce an accurate model with few landslide photographs as training samples. Ghorbanzadeh et al. (2022a) integrated the pixel-based deep learning model, ResU-Net, with the object-based image analysis (OBIA) to detect landslides triggered by the 2009 Morakot typhoon in Taiwan and showed that the ResU-Net-OBIA has better performance than ResU-Net. On the other hand, Ye et al. (2019) used the deep

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

36

learning framework with constraints to detect landslides in Sichuan, China, and showed that deep learning can generate more accurate models than machine learning models. Nava et al. (2021) used the deep learning convolutional neural networks to compare the performance of the landslide detection model using optical images from Sentinel-2 and synthetic aperture radar (SAR) images from Sentinel-1 for the coseismic landslides triggered by the 2018 Hokkaido earthquake in Japan and revealed that SAR data might help the landslide detection task during storms or under dense cloud cover. Meena et al. (2022) evaluated the performance of U-Net, support vector machine, k-nearest neighbor, and the random forest to detect landslides in Nepal using RapidEye satellite imagery and Advanced Land Observing Satellite Phased Array type L-band Synthetic Aperture Radar (ALOS-PALSAR) derived topographical data and revealed that the digital elevation model could help the model differentiate human settlement area and river sand bars. Sameen et al. (2019) used the residual networks (ResNet) with spectral and topographic information to detect landslides in Cameron Highlands, Malaysia, and found that ResNet can generate faster and better models. Cai et al. (2021) used the dense convolutional networks (DenseNets) with satellite imagery, geological, topographic, hydrological, and land cover factors to detect landslides in the Gorges reservoir area in China. They showed that environmental factors and DenseNets can improve the model's accuracy.

Landslides in different locations or from different events occur in various climates and have different characteristics, such as vegetation conditions, geology, and topography. Most studies (Ghorbanzadeh et al., 2019; Ullo et al., 2021; Ghorbanzadeh et al., 2022a; Ye et al., 2019; Nava et al., 2021; Meena et al., 2022; Sameen et al., 2019; Cai et al., 2021; Huang and Zekkos, 2023) focus on evaluating the performance of various ML algorithms in the landslide detection model for a single event. Fewer studies used three to four landslide cases to validate their landslide detection model (Stumpf and Kerle, 2011; Ghorbanzadeh et al., 2022b). Specifically, Stumpf and Kerle (2011) used object-based image analysis with the random forest algorithm to map the landslides in Haiti, Italy, China, and France and showed that feature selection affects the model performance. Ghorbanzadeh et al. (2022b) evaluated eleven deep learning models: U-Net, ResU-Net, PSPNet, ContextNet, DeepLab-v2, DeepLab-v3+, FCN-8s, LinkNet, FRRN-A, FRRN-B, and SQNet with Sentinel-2 satellite imagery and digital elevation model and slope from ALOS-PALSAR for four landslide cases in Iburi, Kodagu, Gorkha, and Taiwan and revealed that ResU-Net outperformed other models in the landslide detection. However, a more systematic assessment of the importance of each factor for several events is still lacking.

This study aims to evaluate the importance of various factors included in the process of the landslide detection model for three recent seismic events that caused landslides. Specifically, the importance of features, training sample size, machine learning algorithms, segmentation parameters, and the resolution of the input data are quantified for each event, and similarities and differences in results are compared to provide improved guidance on which factors play more critical roles in a particular environment and need to be prioritized during the model development.

3.2 Study Areas & Seismic Events

The three different events considered are recent earthquake events that caused extensive landslides. Specifically, the 2015 Lefkada earthquake, the 2016 Kaikōura earthquake, and the 2021 Nippes earthquake are considered.

The M_w 6.5 Lefkada earthquake occurred on November 17, 2015, in Greece and triggered around 700 landslides along the west coast of Lefkada island (Zekkos and Clark, 2020). The M_w 7.8 Kaikōura earthquake occurred on November 14, 2016, in Kaikōura, New Zealand, and triggered approximately 29,000 landslides (Massey et al., 2020). On August 14, 2021, a M_w 7.2 earthquake hit the Tiburon Peninsula of western Haiti and triggered around 14,000 landslides (Amatya et al., 2023). As shown in Table 3.1, the climate of the Lefkada is Mediterranean, the climate of Kaikōura is subtropical, and the climate of Nippes is tropical. Figure 3.1 shows the satellite imagery for the three events, and Figure 3.2 shows the histogram of NDVI before the earthquake. Figure 3.2 shows the minimum, maximum, standard deviation, mean, and mode values of NDVI for the Lefkada, which are -0.98, 0.972, 0.248, 0.04, and 0.27, respectively. The minimum, maximum, standard deviation, and mean values of NDVI for the Kaikōura are -0.843, 0.624, 0.226, and 0.03, respectively. The NDVI density curve for the 2016 Kaikōura earthquake has two peaks: one is at -0.18, and the other is at 0.28. The minimum, maximum, standard deviation, mean, and mode values of NDVI for the Nippes are -0.258, 0.972, 0.113, 0.753, and 0.84, respectively. Nippes has the highest mean value of NDVI. Kaikōura and Lefkada have similar mean values of NDVI, but Lefkada has a lower minimum value of NDVI than Kaikōura. As indicated by the NDVI values and shown in the satellite imagery in Figure 3.1, the Nippes is the most vegetated case among the three cases.

Table 3.1: Earthquake events

	Event Year	Location	Earthquake Moment Magnitude	Climate	Imagery (Resolution)
Lefkada earthquake	2015	Greece	6.5	Mediterranean	Worldview© (1.84 m)
Kaikōura earthquake	2016	New Zealand	7.8	Subtropical	PSOrthoTile (3.125 m)
Nippes earthquake	2021	Haiti	7.2	Tropical	PSOrthoTile (3.125 m)

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

38
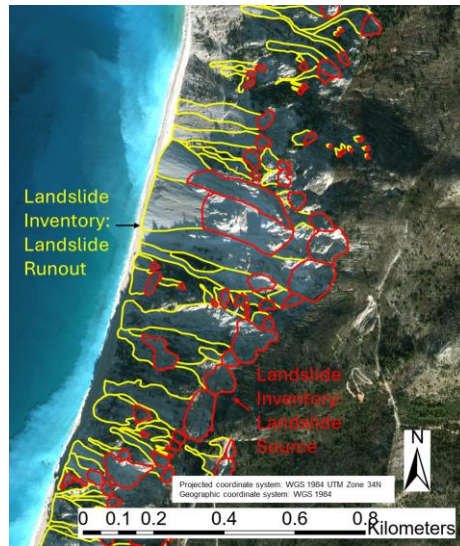
(a) 2015 Lefkada Earthquake
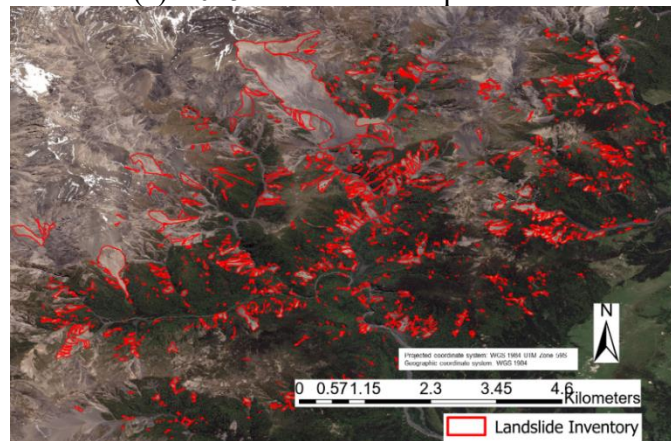
(b) 2016 Kaikōura Earthquake
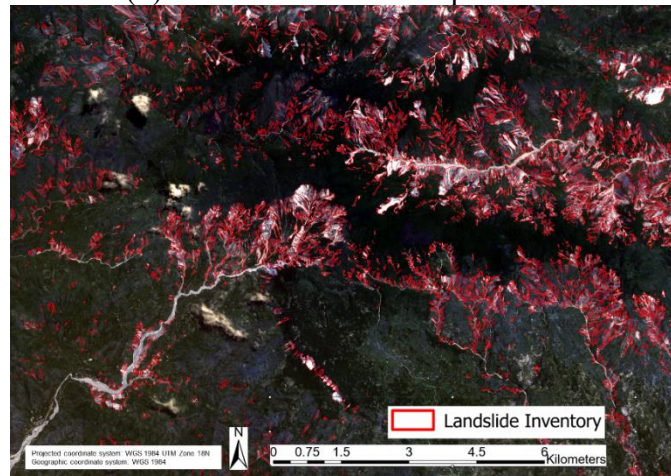
(c) 2021 Nippes Earthquake Event

Figure 3.1: Post-event imagery and landslide inventory

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

39
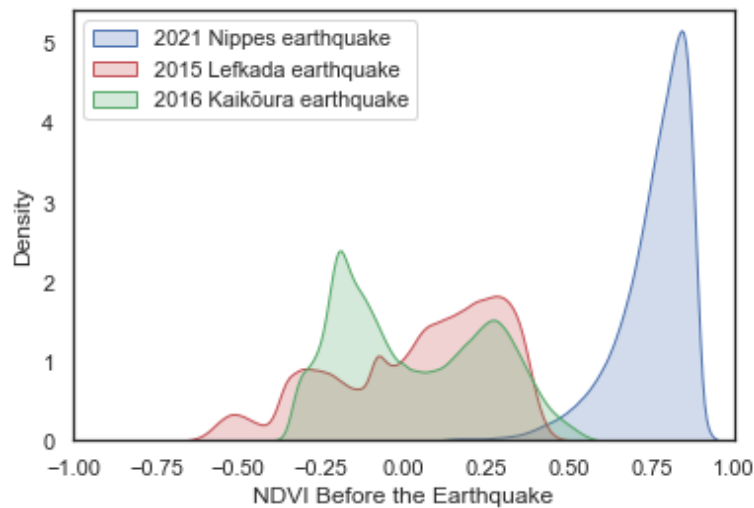
Figure 3.2: NDVI density curves before the earthquake

3.3 Data

As is often the case, the available data for each event varies. For the 2015 Lefkada
earthquake event, the pre-event Worldview© satellite imagery on January 15, 2015,
and the post-event Worldview© satellite imagery on December 28, 2015, are used.
The multispectral Worldview© satellite imagery has eight bands: coastal blue, blue,
green, yellow, red, red edge, near-infrared (NIR), and near-infrared2 (NIR2). Also, a
2-m digital elevation model (DEM) is used as the pre-event DEM, and a 2-m DEM on
December 28, 2015, is used as the post-event DEM. Both pre-event and post-event
DEM are derived from the UAV and satellite imagery. The eight bands in Worldview©
satellite imagery are used to calculate several indexes including the Red-Edge
Normalized Difference Vegetation Index (NDVIre), Red-Edge Chlorophyll Index
(CIre), Red-Edge Triangulated Vegetation Index (RTVICore), Global Environmental
Monitoring Index (GEMI), Green Chlorophyll Index (CIg), Green Normalized
Difference Vegetation Index (GNDVI), Visible Atmospherically Resistant Index
(VARI), Modified Soil Adjusted Vegetation Index (MSAVI), Simple Ratio (SR), Red-
Edge Simple Ratio (SRre), Modified Triangular Vegetation Index (MTVI2), Enhanced
Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI),
Normalized Difference Water Index (NDWI), Iron Oxide (IO), and mean brightness
as shown in Equation 2.1-2.15.

For the 2016 Kaikōura earthquake, the pre-event PSOrthoTile satellite imagery
on September 10, 2016, and the post-event PSOrthoTile satellite imagery on March
18, 2017, are collected. A 2-m digital surface model (DSM) before November 14,
2016, is used as the pre-event DSM, and a 2-m DSM in January 2017 is used as the

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

40

post-event DSM. The DSMs are derived from satellite imagery. For the 2021 Nippes Earthquake, the pre-event PSOrthoTile satellite imagery on August 10, 2021, and the post-event PSOrthoTile satellite imagery on December 18, 2021, are collected. Also, a 1.5-m digital terrain model (DTM) acquired in 2016 (The World Bank, 2021) is used as the pre-event topography data. The multispectral PSOrthoTile satellite imagery has four bands: blue, green, red, and near-infrared (NIR). The four bands in PSOrthoTile satellite imagery are used to calculate indexes, including the Global Environmental Monitoring Index (GEMI), Green Chlorophyll Index (CIg), Green Normalized Difference Vegetation Index (GNDVI), Visible Atmospherically Resistant Index (VARI), Modified Soil Adjusted Vegetation Index (MSAVI), Simple Ratio (SR), Modified Triangular Vegetation Index (MTVI2), Enhanced Vegetation Index (EVI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), and Iron Oxide (IO).

The pre-event and post-event topography data are used to extract terrain characteristics: slope inclination, curvature, aspect, flow direction, and hillshade. In addition to the pre-event and post-event features, the feature difference between the pre-event and post-event values is also calculated. This difference may indicate the change caused by the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, 2021 Nippes earthquake, or other non-earthquake-related changes that may have occurred between the dates that the data is collected. The differences associated with spectral and terrain features are calculated as shown in Equation 2.16. The pre-event, post-event, and feature difference values for each feature are analyzed separately to quantify each feature's significance. Additionally, features are composited together, and the performance of multi-feature models is evaluated.

3.4 Machine Learning Algorithms and Image Segmentation Process

Three machine learning (ML) algorithms are considered in the analyses: random forests (RF), support vector machine (SVM), and maximum likelihood (MLE). Object-based image analysis (OBIA), as implemented in ArcGIS Pro 2.9.1, is considered in this study. A crucial part of OBIA is the imagery segmentation process that defines the clusters or objects and considers spectral details, spatial details, and minimum segment size in pixels. Spectral detail and spatial detail range from 1 to 20. Segments smaller than the minimum segment size in pixels are merged with best-fitting adjacent segments (Esri Inc., 2021). In this study, spectral details, spatial details, and minimum segment sizes have been parameterized to assess the parameters that produce the best landslide detection model.

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

41

3.5 Training and Testing Samples

Supervised machine learning algorithms require training samples to train the algorithm and testing samples to assess the model's accuracy. The landslide detection models are a binary classification with two classes: landslide and non-landslide. The manually developed landslide inventory is divided into training and testing samples, and landslides and non-landslide samples are created accordingly.

Zekkos and Clark (2020) used high-resolution satellite imagery and three-dimensional models to develop the 2015 Lefkada earthquake landslide inventory, which contains the location, area, and volume of 716 landslides. Each landslide's source and entire area were mapped separately, as shown in Figure 3.1a. Karantanellis et al. (2023) developed the landslide inventory containing 1448 landslides for the 2016 Kaikōura earthquake using the WorldView© imagery and digital surface model, as shown in Figure 3.1b. Juang et al. (2019) developed the polygon-based manually mapped landslide inventory using satellite imagery from Pléiades, WorldView©, and Sentinel-2 for the landslides triggered by the 2021 Nippes earthquake, as shown in Figure 3.1c. The landslide inventory has a total number of 14,482 landslides. Both the 2015 Lefkada earthquake and the 2016 Kaikōura earthquake landslide inventory have the area and volume information for landslides, but the 2021 Nippes earthquake landslides only have the landslide area information.

As shown in Figure 3.3, most landslides in the inventory have a landslide area <2500 m^2 for all three earthquake events. In assessing the effect of factors such as segmentation parameters, feature selection, and data resolution, the commonly used training-testing splitting ratio of 70/30 is used. The training sample size from each manual landslide inventory is also varied to evaluate the influence of training sample size on the results.

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

42
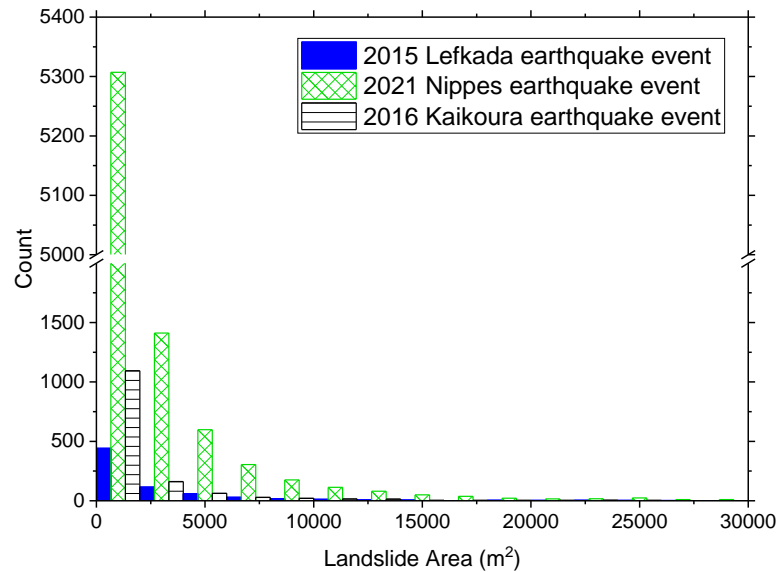
Figure 3.3: The landslide area in the landslide inventory

3.6 Accuracy Assessment

In reporting landslide detection results, a positive result means landslides, and a negative result means non-landslide. True indicates the result is correct, but false means the result is wrong. The overall accuracy (OA), as shown in Equation 2.24, is used to assess the model's performance. The minimum value of OA is -33.33%, and the maximum value of OA is 100%. The random classifier has an OA of 33.33%.

3.7 Results and Discussion

3.7.1 Feature Importance

Pre-event, post-event, and the feature difference for each feature (when available) are utilized as input in the landslide detection model. Random Forest (RF) has the best landslide detection performance than other ML models, as discussed subsequently, so RF is used to evaluate the influence of the selection of features, training sample size, data resolution, and segmentation on the detection model. Tables 2.2, Table 3.2, and Table 3.3 list each feature's model performance when considered individually. Terrain features are highlighted in gray, and spectral features are highlighted in white. There are 92 features for landslides triggered by the 2015 Lefkada earthquake event. The 2016 Kaikōura earthquake event has 63 features, and the 2021 Nippes earthquake has 51 features. All three study sites have considered the spectral features, including the

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

43

GEMI, CIg, GNDVI, VARI, MSAVI, SR, MTVI2, EVI, NDVI, NDWI, and IO. Because the 2015 Lefkada earthquake event has the red-edge band from the Worldview© satellite imagery, the NDVIre, CIre, RTVICore, and SRre are calculated additionally and included as spectral features. The 2015 Lefkada earthquake has the pre-event DEM and post-event DEM, the 2016 Kaikōura earthquake has the pre-event DSM and post-event DSM, and the 2021 Nippes earthquake has the pre-event DTM. The slope inclination, curvature, aspect, flow direction, and hillshade are derived from these terrain models and utilized as terrain features.

For the 2015 Lefkada earthquake event, post-slope and pre-slope have the highest OA, around 80 %, followed by pre-event vegetation indexes, pre-NDVI, pre-NDVIre, pre-CIre, and pre-RTVICore. Post-elevation, pre-elevation, aspect difference, curvature difference, and GEMI difference have the lowest OA between 13 % and 19 %. For the 2016 Kaikōura earthquake event, vegetation index difference, NDVI difference, and MSAVI difference individually generated the best two landslide detection models, with OA of 66.3 % and 64.7 %, respectively. Post vegetation index, post NDVI, SR difference, post MSAVI, post NDWI, and NDWI difference have the OA between 60.5 % and 64.2 %. All terrain features, such as post-slope and pre-DSM, have an OA of around or below 50 %. The lowest OA, which is 25.9 %, comes from post-DSM. For the 2021 Nippes earthquake event, the post-vegetation index, post-CIg, and post-SR have the highest OA of 62.1 % and 60.8 %, respectively, followed by pre-slope. Post-NDVI, GNDVI difference, post-NDWI, post MSAVI, and NDVI difference have the OA between 52.3 % and 57 %. The pre-DTM has an OA of 50 %. A random classifier has an OA of 33.33 %. EVI Difference, Pre SR, Pre blue band, Pre CIg, IO difference, and Pre EVI have worse model performance than a random classifier.

Generally, the pre-slope and post-slope terrain features are the most essential features in the landslide detection model for the 2015 Lefkada earthquake event. The vegetation index differences, NDVI difference, and MSAVI difference, are the two features with the highest OA for the 2016 Kaikōura earthquake event. The post-CIg and post-SR, the post vegetation index, are the two most critical features for the 2021 Nippes earthquake event. As indicated by the NDVI statistics for each area, shown in Figure 3.2, the 2021 Nippes earthquake event is in the most vegetated area, and the 2015 Lefkada earthquake event is in the least vegetated area among the three cases. Since Nippes is in a highly vegetated area, as shown in the post-event imagery in Figure 3.1c, the post-vegetation index plays a dominant role in generating a reliable landslide detection result. The vegetation condition of Kaikōura is between Lefkada and Nippes, so vegetation index differences, i.e., the difference between the pre-event and post-event vegetation indexes, are the most essential features. Lefkada is in the least vegetated area with overall low vegetation along the slope. Therefore, topographic characteristics, especially the slope angle rather than the vegetation index, are critical in improving the model's performance.

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

44
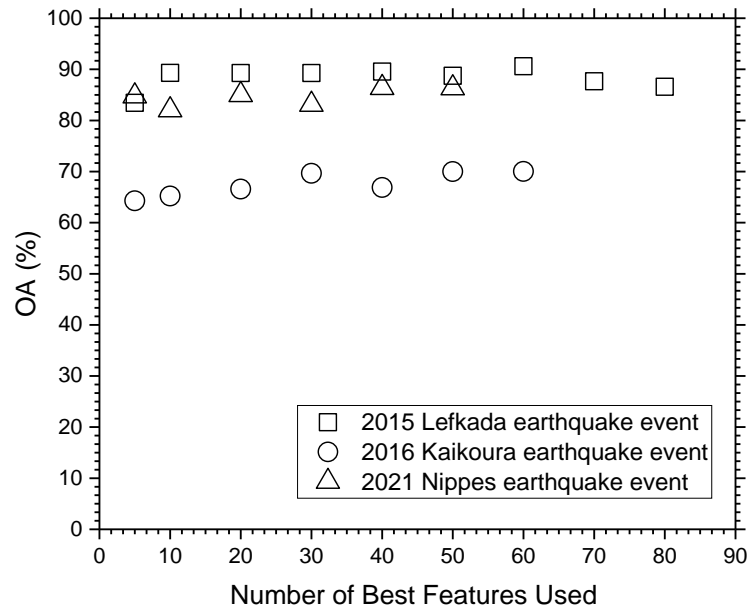
Figure 3.4: Model's performance using multiple features

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

45

Table 3.2: Kaikōura feature importance

Ranking	Feature	OA (%)	Ranking	Feature	OA (%)
1	NDVI Difference	66.3	41	Red band Difference	38.1
2	MSAVI Difference	64.7	42	Pre CIg	38.0
3	Post NDVI	64.2	43	Post Aspect	38.0
4	SR Difference	62.3	44	Curvature Difference	37.9
5	Post MSAVI	61.1	45	Post Flow Direction	37.5
6	Post NDWI	60.7	46	Pre VARI	37.4
7	NDWI Difference	60.5	47	Pre SR	37.1
8	Post GNDVI	60.0	48	Pre Aspect	35.9
9	Green band Difference	59.7	49	Pre Curvature	35.7
10	GNDVI Difference	58.1	50	Post Curvature	35.0
11	Mean Brightness Difference	58.1	51	Pre Flow Direction	34.5
12	Post CIg	56.6	52	Post EVI	34.2
13	Post GEMI	54.1	53	DSM Difference	33.6
14	MTVI2 Difference	51.8	54	Aspect Difference	33.3
15	Post Red band	51.4	55	Hillshade Difference	32.7
16	Post SR	51.3	56	Slope Difference	32.4
17	Post Slope	51.2	57	GEMI Difference	31.9
18	Pre DSM	50.0	58	Pre Red band	31.5
19	Pre EVI	50.0	59	Pre Mean Brightness	30.6
20	Pre Iron Oxide	50.0	60	Post Hillshade	29.0
21	EVI Difference	50.0	61	Flow Direction Difference	27.7
22	Iron Oxide Difference	50.0	62	Pre GEMI	26.7
23	Pre Slope	49.1	63	Post DSM	25.9
24	Post MTVI2	48.6			
25	CIg Difference	48.2			
26	Blue band Difference	47.3			
27	VARI Difference	47.0			
28	Pre Hillshade	46.0			
29	Post Green Band	45.6			
30	Pre Blue band	44.5			
31	Post VARI	44.2			
32	Pre GNDVI	44.1			
33	Pre NDWI	43.8			
34	Pre NDVI	42.7			
35	Post Mean Brightness	41.8			
36	Pre MSAVI	41.5			
37	Pre Green band	40.0			
38	Pre MTVI2	39.4			
39	Post Iron Oxide	39.0			
40	Post Blue band	38.6			

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

46

Table 3.3: Nippes feature importance

Ranking	Feature	OA (%)	Ranking	Feature	OA (%)
1	Post CIg	62.1	41	Pre VARI	36.6
2	Post SR	60.8	42	Pre Flow Direction	35.9
3	Pre Slope	60.3	43	GEMI Difference	35.8
4	Post NDVI	57.0	44	NIR band Difference	35.3
5	GNDVI Difference	56.9	45	Post EVI	34.7
6	Post NDWI	56.1	46	EVI Difference	30.0
7	Post MSAVI	53.6	47	Pre SR	28.6
8	NDVI Difference	52.3	48	Pre Blue band	28.3
9	Pre Curvature	52.2	49	Pre CIg	21.9
10	Post GNDVI	51.9	50	Iron Oxide Difference	20.0
11	Pre DTM	50.0	51	Pre EVI	16.0
12	Blue band Difference	49.9			
13	Pre Aspect	49.4			
14	NDWI Difference	49.3			
15	Pre Hillshade	48.5			
16	MTVI2 Difference	48.1			
17	Post GEMI	48.0			
18	Post VARI	47.4			
19	Pre GNDVI	47.1			
20	Red band Difference	46.4			
21	Green band Difference	45.7			
22	VARI Difference	45.2			
23	Post Red band	44.7			
24	CIg Difference	44.6			
25	Pre MSAVI	44.5			
26	MSAVI Difference	44.4			
27	Pre Green band	44.4			
28	Post Green band	44.1			
29	Pre NDVI	43.4			
30	Pre Iron Oxide	42.7			
31	Post MTVI2	42.7			
32	Post Iron Oxide	41.4			
33	Pre NIR band	41.4			
34	Pre Red band	40.8			
35	Post Blue band	40.5			
36	SR Difference	39.7			
37	Post NIR band	39.4			
38	Pre MTVI2	38.6			
39	Pre NDWI	37.5			
40	Pre GEMI	36.9			

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

47

Utilizing multiple features in the landslide detection model improves the model's performance compared to the single feature shown in Tables 2.2, Table 3.2, and Table 3.3. Figure 3.4 shows the multi-feature model's performance for the 2015 Lefkada earthquake event, the 2016 Kaikōura earthquake event, and the 2021 Nippes earthquake event using the top 5, 10, 20, 30, 40, 50, 60, 70, and 80 features that has the best performance in Tables 2.2, Table 3.2, and Table 3.3. For the 2015 Lefkada earthquake event, the top 10 features are post-slope, pre-slope, pre-NDVI, pre-NDVIre, pre-CIre, pre-RTVICore, NDVI difference, post-flow direction, pre-CIg, and post-GEMI. Three of the top 10 features are terrain features, and seven of the top 10 features are spectral features. For the 2016 Kaikōura earthquake event, the top 10 features are NDVI difference, MSAVI difference, post-NDVI, SR difference, post M SAVI, post-NDWI, NDWI difference, post-GNDVI, green band difference, and GNDVI difference. All of the top 10 features are spectral features. For the 2021 Nippes earthquake event, the top 10 features are post-CIg, post-SR, pre-slope, post-NDVI, GNDVI difference, post-NDWI, post MSAVI, NDVI difference, pre-curvature, and post-GNDVI. Two of the top 10 features are terrain features, and eight of the top 10 features are spectral features.

For the 2015 Lefkada earthquake event, the OA is 83.4 % using the top 5 features in the model and increases to 89 % for the top 10 features. The model's performance becomes stable as additional features are included. The models with the top 20, 30, 40, 50, 60, 70, and 80 features all have OA near 90%, showing that adding more features does not improve the model performance. In subsequent analyses, the top 10 features are used to assess the influence of sample size, machine learning algorithm, resolution, and segmentation. For the 2016 Kaikōura earthquake event, the OA is 64.3 % for the top 5 features and increases to 69.7 % for the top 30 features. The models with the top 40, 50, and 60 features all have OA near 70%. Hence, the top 30 features are used to evaluate the influence of sample size, machine learning algorithm, resolution, and segmentation in subsequent analyses. For the 2021 Nippes earthquake event, the multiple features model with the top 5 features has an OA of 84.7%, increasing to 85% for the top 20 features. The models with the top 40 and 50 features all have OA near 85%. As a result, the model with the top 20 features is used in subsequent analyses.

Figure 3.5a-c show that the landslide detection model can reveal the landslide location well, as highlighted in true positive. However, some non-landslide areas are detected as landslides, as shown in false positives. A few landslide areas are detected as non-landslides, as demonstrated by false negatives.

Flow direction is utilized to separate amalgamated landslides in the detection result. Figure 3.6a-c show the histogram of landslide area in the landslide detection result for the analysis with the top 10 features for the 2015 Lefkada earthquake event, the top 30 features for the 2016 Kaikōura earthquake event, and the top 20 features for the 2021 Nippes earthquake event compared to the landslide inventory. As shown in Figure 3.6a-c, the model detects small landslides with a landslide area <2500 m^2 the most but overestimates the number of small landslides for the 2015 Lefkada earthquake and 2016 Kaikōura earthquake.

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

48
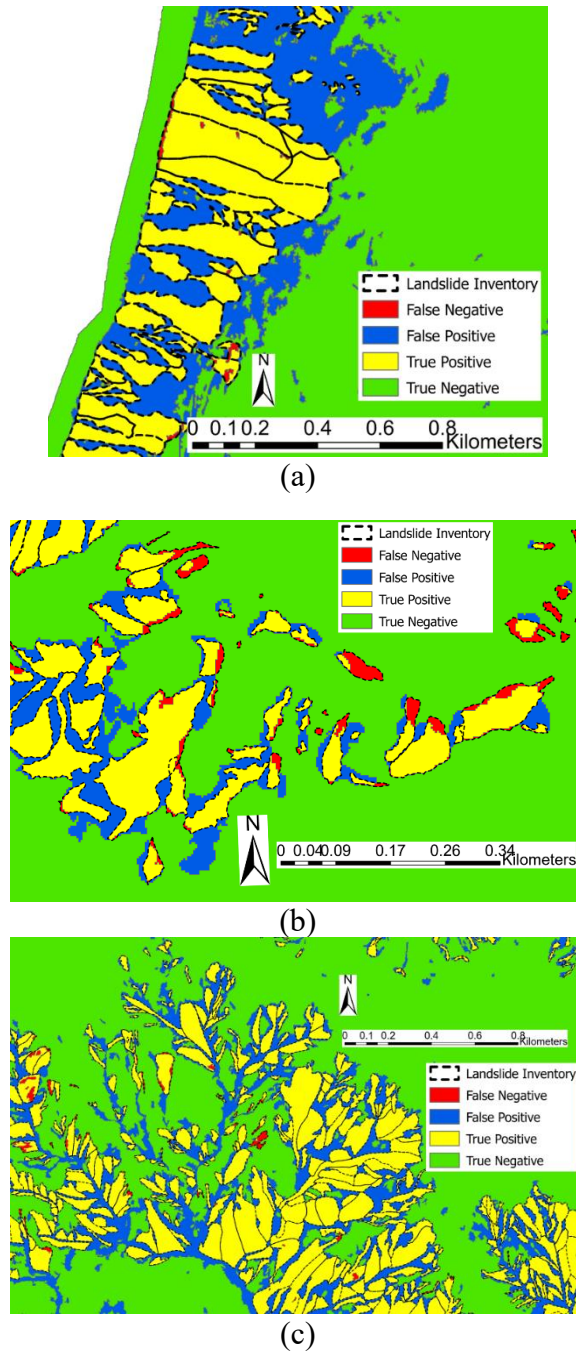
Figure 3.5: Landslide detection result (a) from the top 10 features in the 2015 Lefkada
earthquake event (b) from the top 30 features in the 2016 Kaikoura earthquake event (c)
from the top 20 features in the 2021 Nippes earthquake event

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

49
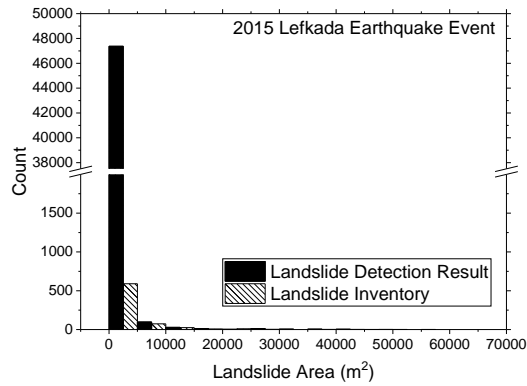
(a)
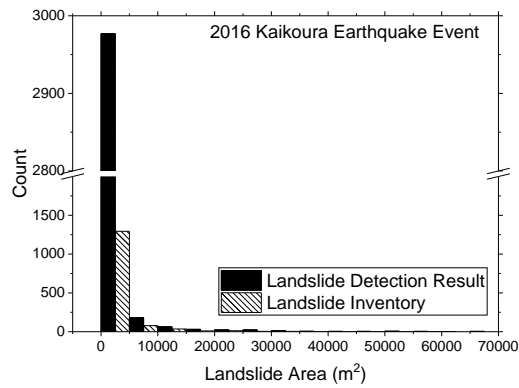
(b)
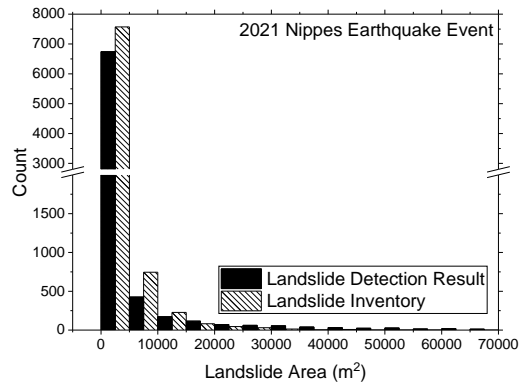
(c)

Figure 3.6: Histogram of landslide area in the landslide detection result and landslide inventory (a) from the top 10 features in the Lefkada earthquake (b) from the top 30 features in the Kaikoura earthquake (c) from the top 20 features in the Nippes earthquake

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

50

3.7.2 Effect of Sample Size

The effect of training sample size on the landslide detection model's performance is investigated, and the results are shown in Figure 3.7. This investigation is essential for practical cases. Although large training samples are known to be better for detection, in a post-disaster assessment stage, we do not know how big the entire inventory is yet, and we do not have the opportunity to have very large samples for model training. The training sample is geospatially distributed over the entire event area for the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake event. For the 2015 Lefkada earthquake, the OA is 59.9 % with a very small training sample size that is equal to 0.5 % of the inventory and increases to 85.6 % with a training sample size of 20 %. The model's performance becomes stable after the training sample size achieves 20 %, with an OA near 85 %. As the training sample size reaches 20 % and beyond, the larger size of training samples does not improve the model's performance. For the 2016 Kaikōura earthquake, the OA is 69.0 % with a training sample size equal to 3.3 % of the inventory and increases to 78.5 % for a training sample size of 50 %, remaining around 78 % for larger training sample sizes. For the 2021 Nippes earthquake event, the OA is 84.1 % with the training sample size as 2.5 % of the inventory and modestly increases to 88.6 % with the training sample size as 5 %. The model's performance becomes stable after the training sample size achieves 5 % and beyond. Overall, for the training sample size of about 2-4 % of the inventory (specifically 3.02 %, 3.26 %, and 2.5 % for the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake, respectively), OA>70 % can be achieved, which is valuable for practical purposes. The model's performance is stable as the training sample size reaches 20 %, 50 %, and 5 % of the inventory for the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake, respectively.

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

51
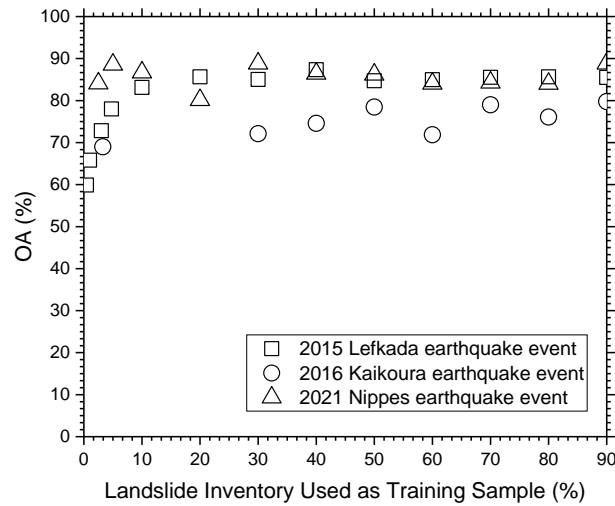
Figure 3.7: The influence of training sample size on the landslide detection model

3.7.3 Effect of Machine Learning Algorithm

Figure 3.8 shows the effect of machine learning algorithms on landslide detection results. For the 2015 Lefkada earthquake, RF has an OA of 82.9 %, SVM has an OA of 71.2 %, and MLE has an OA of 66.2 %. For the 2016 Kaikōura earthquake, RF has an OA of 79.6 %, SVM has an OA of 77.2 %, and MLE has an OA of 50.4 %. For the 2021 Nippes earthquake, RF has an OA of 85.8 %, SVM has an OA of 82.0 %, and MLE has an OA of 50 %. For all three study areas, RF outperforms SVM, and MLE performs worst for all three events. RF and SVM have more robust and better landslide detection results than MLE.

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

52
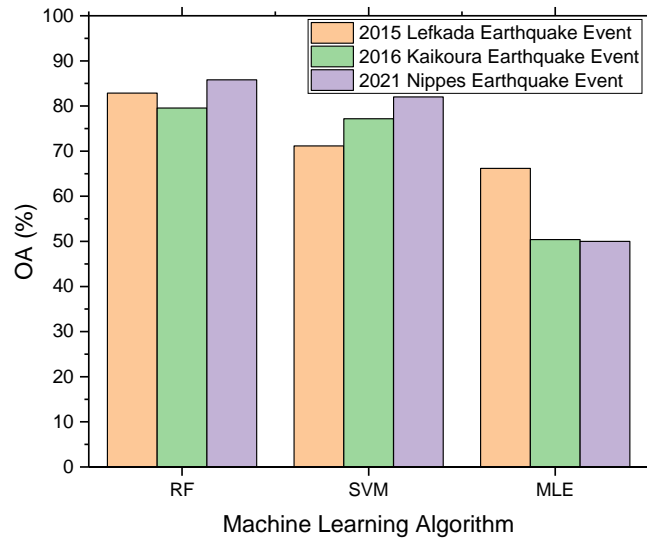
Figure 3.8: The influence of machine learning algorithms on the landslide detection model

3.7.4 Effect of Data Resolution

The top 10 features for the 2015 Lefkada earthquake event, the top 30 features for the 2016 Kaikōura earthquake event, and the top 20 features for the 2021 Nippes earthquake event are composited as the input data with the original data resolution of 2 m for each event. They are re-run using the same data resampled to different resolutions in 5 m, 10 m, 30 m, and 50 m. The RF model is re-trained with different data resolutions to access the data resolution effect. The OA of all three models is systematically reduced as the data resolution is reduced. As the data resolution changes from 2 m to 50 m, OA decreases from 84.8 % to 63.9 % for the 2015 Lefkada earthquake event, OA decreases from 75.6 % to 48.8 % for the 2016 Kaikōura earthquake event, and OA decreases from 85.8 % to 66.9 % for the 2021 Nippes earthquake event as shown in Figure 3.9.

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

53
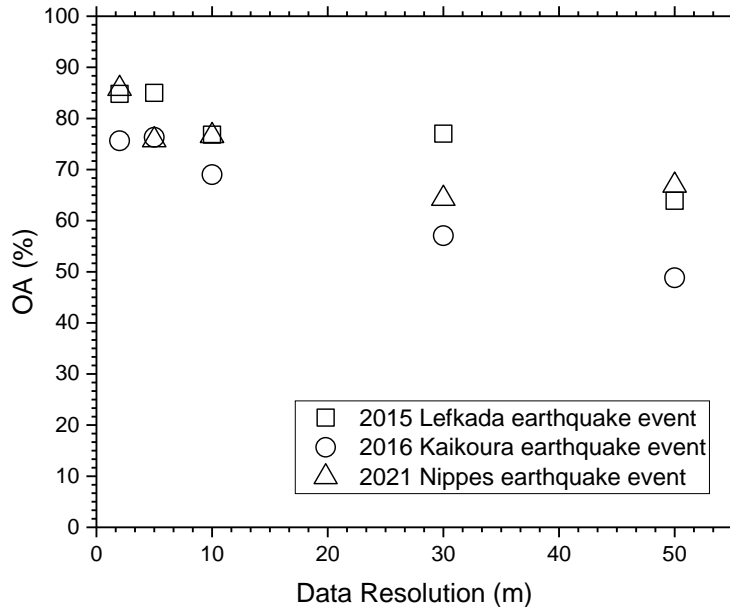
Figure 3.9: The effect of data resolution on the landslide detection model

3.7.5   Effect of Segmentation

The manner by which the objects are defined in the OBIA method is an important consideration. For segmentation, three parameters, spectral detail, spatial detail, and minimum segment size in pixels, are vital in grouping neighboring pixels into objects. The spectral detail and spatial detail range between 1 and 20. The spectral detail and spatial detail values of 1, 10, and 20 are tested to see the influence of spectral detail and spatial detail on the model. The impact of segmentation on the landslide model is shown in Figure 3.10, Figure 3.11, and Figure 3.12.

As shown in Figure 3.10a-c, Figure 3.11a-c, and Figure 3.12a-c, as the spectral detail increases, there are more segments, and the value of OA increases. For the 2015 Lefkada earthquake event, the model with the spectral detail of 1, 10, and 20 has an OA of 87.4 %, 89.6 %, and 92.2 %, respectively. For the 2016 Kaikōura earthquake event, the model with the spectral detail of 1, 10, and 20 has an OA of 41.1 %, 53.5 %, and 77.8 %, respectively. For the 2021 Nippes earthquake event, the model with the spectral detail of 1, 10, and 20 has an OA of 66.8 %, 78.2 %, and 88.3 %, respectively.

Changes in spatial detail do not affect the results as systematically as spectral detail. For the 2015 Lefkada earthquake event, the model with the spatial detail of 1, 10, and 20 has an OA of 91.6 %, 89.6 %, and 90.4 %, respectively. As shown in Figure 3.10b, d, and e, the size and shape of segments of the landslide area do not change

significantly as the spatial detail increases. For the 2016 Kaikōura earthquake event, the model with the spatial detail of 1, 10, and 20 has an OA of 43.7 %, 53.5 %, and 74.8 %, respectively. When the value of spatial detail increases, the segments become smaller, as shown in Figure 3.11b, d, and e. For the 2021 Nippes earthquake event, the model with the spatial detail of 1, 10, and 20 has an OA of 70.2 %, 78.2 %, and 45.1 %, respectively.

The minimum segment size also has a lesser effect on the results of the analyses. For the 2015 Lefkada earthquake event, the values of 1, 10, and 20 of minimum segment size in pixels are used, and the OA is 89.2 %, 89.6 %, and 89.6 %, respectively. For the 2016 Kaikōura earthquake event, the values of 1, 10, and 20 of minimum segment size in pixels are used, and the OA is 73.9 %, 53.5 %, and 56.3 %, respectively. As shown in Figure 3.11b, f, and g, the segment size increases as the value of minimum segment size in pixels increases. For the 2021 Nippes earthquake event, the values of 1, 10, and 20 of minimum segment size in pixels are used, and the OA is 73.7 %, 78.2 %, and 79.6 %, respectively.

For all three events, higher spectral details have a higher OA and can produce a better landslide detection model. The variation in the value of spatial detail in the 2015 Lefkada earthquake event does not change the OA significantly. However, the higher spatial detail for the 2016 Kaikōura earthquake event has a higher OA. The minimum segment size in pixels does not affect the model's performance in the 2015 Lefkada earthquake event. However, the OA decreases with increasing the minimum segment size in pixels for the 2016 Kaikōura earthquake event. The OA increases with the increase of the minimum segment size in pixels for the 2021 Nippes earthquake event.

Low spectral detail, low spatial detail, and large minimum segment size in pixels result in under-segmentation. Higher spectral detail, high spatial detail, and small minimum segment size in pixels cause over-segmentation. Figure 3.6 shows that in the inventory, most landslides have a landslide area below 2500 m^2 for all three cases. In the case of under-segmentation, segments are too coarse to capture the shape of small landslides, so under-segmentation has a lower OA. Although over-segmentation might divide the landslides into small segments, over-segmentation has a higher OA because small segments can capture the landslide characteristics for small landslides better than large segments.

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

55
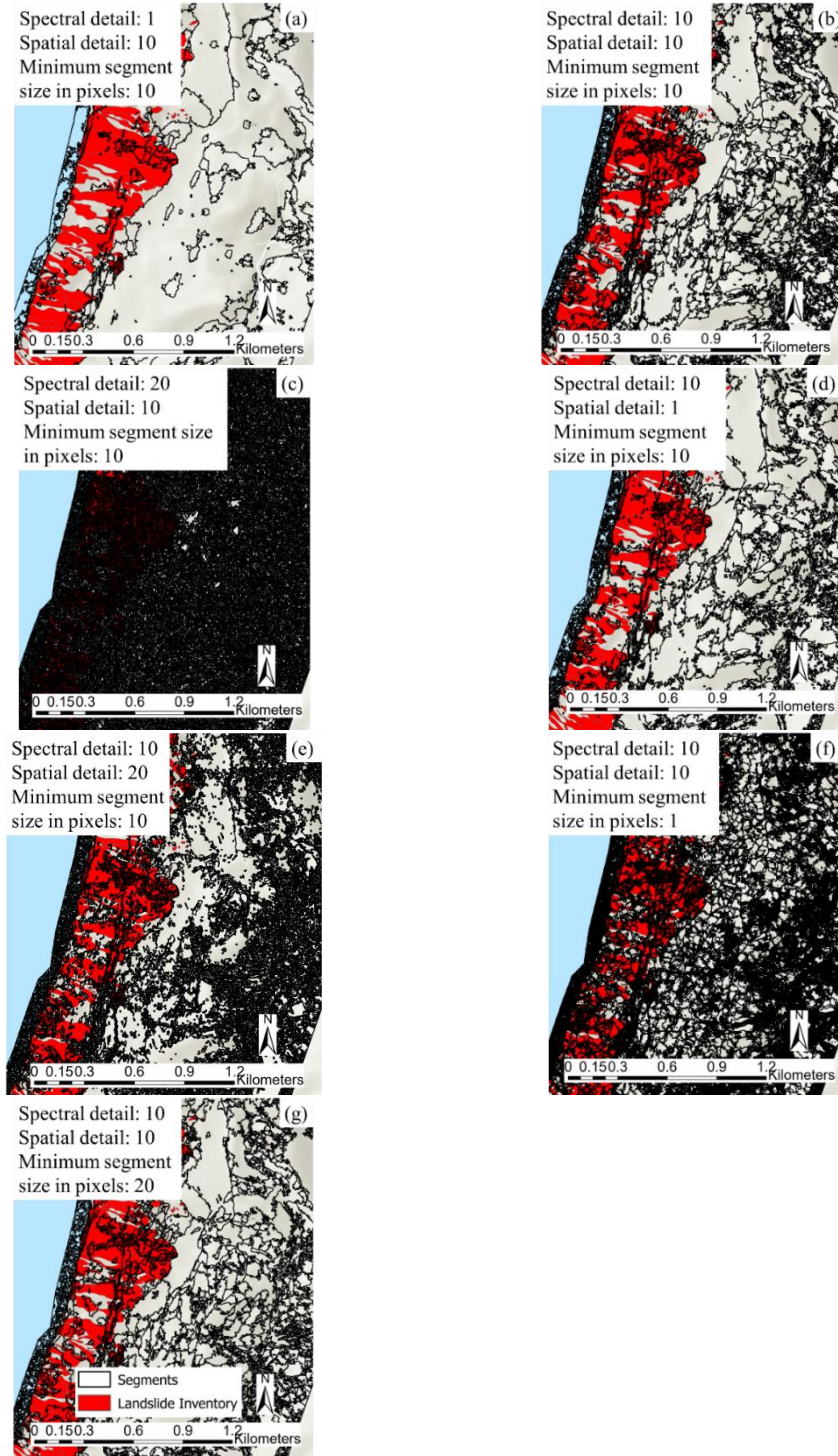
Figure 3.10: The effect of segmentation for the 2015 Lefkada earthquake event

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

56
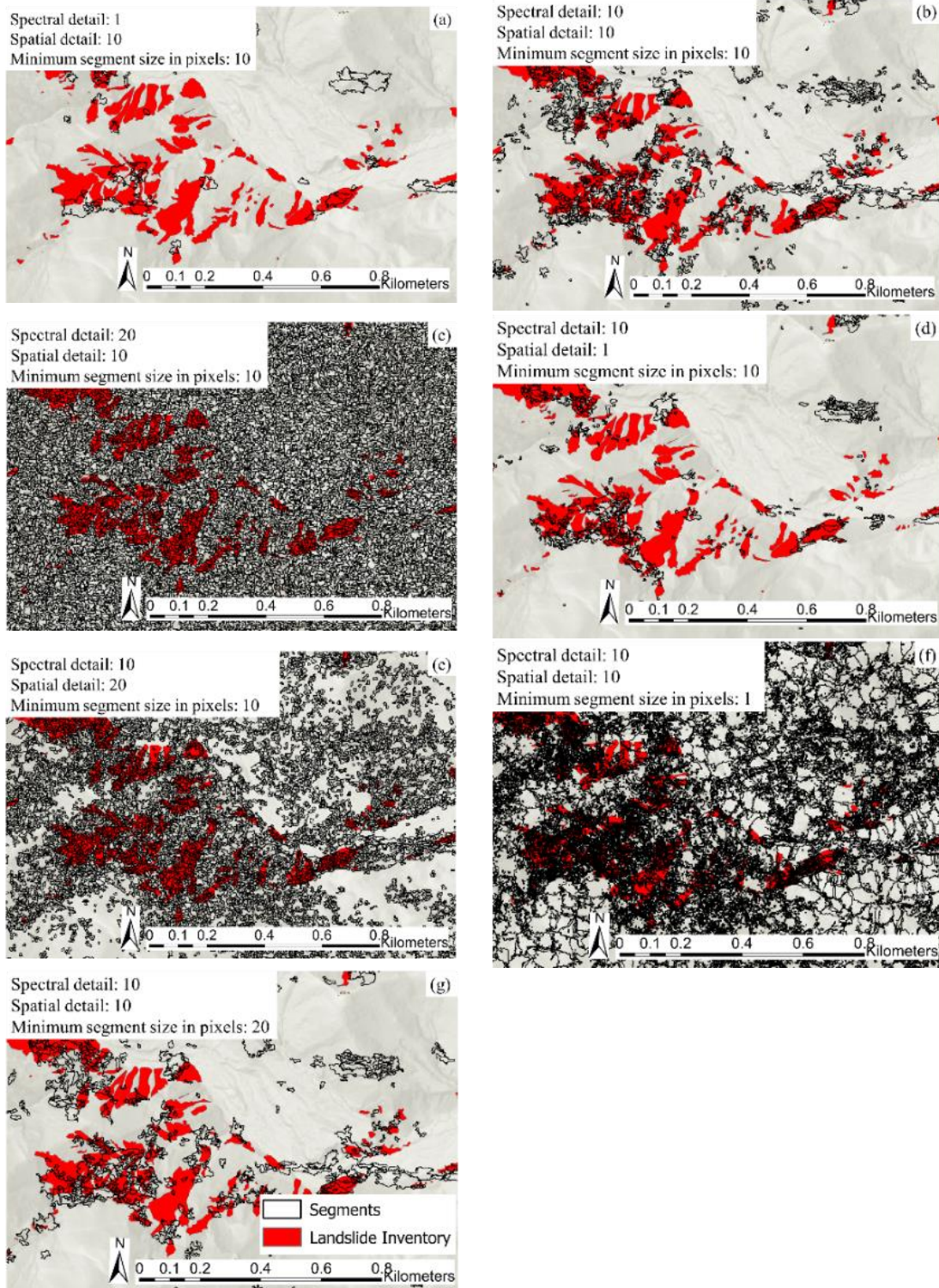
Figure 3.11: The effect of segmentation for the 2016 Kaikōura earthquake event

CHAPTER 3.   LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

57
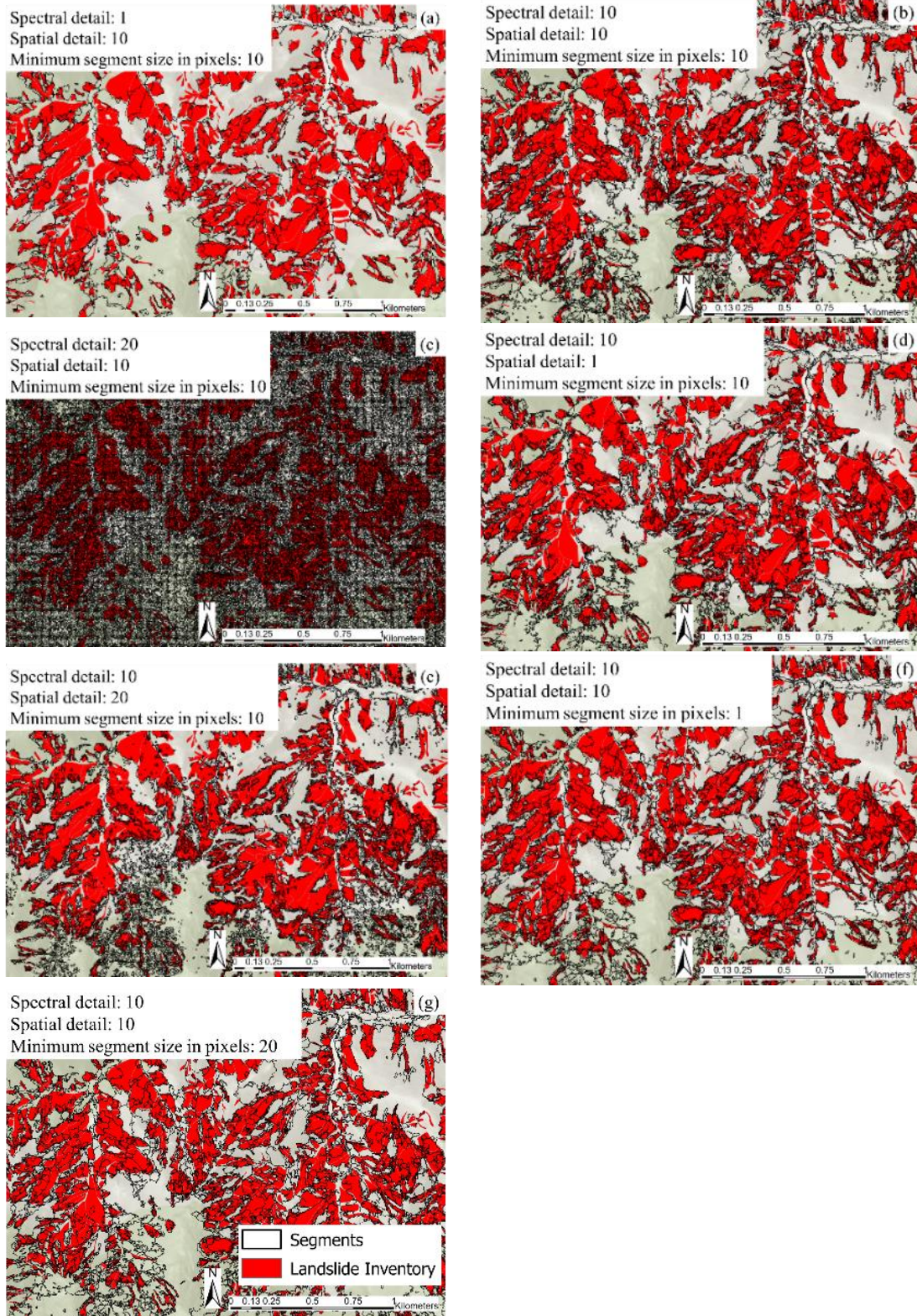
Figure 3.12: The effect of segmentation for the 2021 Nippes earthquake event

3.7.6 Importance of Different Factors in Landslide Detection

An effort is made to quantify the relative importance of the factors that affect the landslide detection model in the three study areas. This may be done by quantifying the variation in OA due to changes in a particular factor. The results are summarized in Table 3.4, which shows that significant variation in OA (>40 %) can be observed due to changes in the factors for all three events. The post-slope is the most critical feature, and the GEMI difference is the feature with the lowest OA for the 2015 Lefkada earthquake event. The OA difference between the best and worst features of the 2015 Lefkada earthquake event is 67 %. The NDVI difference has the highest OA, and the post-DSM has the lowest OA for the 2016 Kaikōura earthquake event. The post-CIg has the highest OA, and the pre-EVI has the lowest OA for the 2021 Nippes earthquake event. The OA difference of the feature selection is 47.2 % for the 2016 Kaikōura earthquake and 46.1 % for the 2021 Nippes earthquake. Overall, the selection of the features is overwhelmingly the most important factor, but as discussed earlier, it is not the larger number of features that improve the results but the type of features. Among the three cases, the Nippes is the most vegetated case, and Lefkada is in the least vegetated area, with overall low vegetation along the slope. The vegetation index is more important in the highly vegetated case. However, the terrain features are more critical in the less vegetated case.

Data resolution is also important. Data resolution results in an OA difference of >20 % for all three events, as shown in Table 3.4. The various ML algorithms result in the OA difference of 16.7 %, 29.2 %, and 35.8 % for the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake event, respectively. RF and SVM have more stable landslide detection results in all three cases. However, the performance of MLE varies and can be worse than the other two, with OA ~50 % for the 2016 Kaikōura earthquake and 2021 Nippes earthquake event. Therefore, the OA difference caused by the ML algorithms for the 2016 Kaikōura earthquake and 2021 Nippes earthquake event is higher than the 2015 Lefkada earthquake event. The influence of segmentation on the landslide detection model's performance varies with cases. The 2015 Lefkada earthquake event has an OA difference of 6.5 %, the 2016 Kaikōura earthquake has an OA difference of 36.7 %, and the 2021 Nippes earthquake has an OA difference of 43.3 %. High spectral detail, high spatial detail, and small minimum segment size in pixels lead to over-segmentation, and the small segments in over-segmentation can distinguish the irregular landslide shape and sizes better. The event with a smaller landslide size is more sensitive to segmentation. As shown in Figure 3.6, the 2021 Nippes earthquake has the smallest landslides with a landslide area of <2500 m^2, and the 2015 Lefkada earthquake event has the least landslides with a landslide area of <2500 m^2 in the landslide inventory. Therefore, the segmentation leads to the most significant OA difference in the 2021 Nippes earthquake and the least OA difference in the 2015 Lefkada earthquake event. When the training sample is distributed over the entire event area, the size of the training sample can generate an

CHAPTER 3. LESSONS LEARNED ON THE EFFICACY OF MACHINE LEARNING-BASED
LANDSLIDE DETECTION FOLLOWING THREE RECENT EARTHQUAKE EVENTS

59

OA difference of around 10 % for all three events.

Table 3.4: Ranking of importance for different factors in landslide detection

	2015 Lefkada earthquake OA Difference % (Maximum OA - Minimum OA)	2016 Kaikōura earthquake OA Difference % (Maximum OA - Minimum OA)	2021 Nippes earthquake OA Difference % (Maximum OA - Minimum OA)
Feature	67.0	47.2	46.1
Resolution	21.1	27.5	28.2
ML	16.7	29.2	35.8
Segmentation	6.5	36.7	43.3
Sample size	12.7	10.7	8.7

3.8 Conclusions

This study evaluates the importance of different factors, the selection of features, data resolution, machine learning algorithms, segmentation, and training sample size in the landslide detection model for the landslides triggered by the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake event. The selection of features affects the landslide detection model the most. The vegetation index plays a more critical role in the highly vegetated case, but the terrain features are more important in the less vegetated case. The importance of data resolution is similar in all three cases. The effect of machine learning algorithms and segmentation on landslide detection models varies with cases. The impact of training sample size is similar for all three cases.

Chapter 4 Derivation of Voellmy Model Parameters for Landslide Runout based on Co-seismic Rock Avalanches Inventory

4.1 Introduction

Landslides can result in severe damage to infrastructure and loss of life. Thus, understanding the mechanisms of landslide triggering and mobility is crucial in reducing the risk of damage and loss (Zekkos et al., 2019). Landslide mobility is particularly critical for landslide risk assessment (Guo et al., 2014), as landslides' runout (or travel distance) increases risk to communities. Models for landslide mobility prediction are primarily empirical or mechanistic. Empirical landslide mobility models are often statistical models derived from case histories and past observations (Liu et al., 2021). Table 4.1 lists existing empirical models used in estimating landslide travel distance. Most studies (Geo et al., 2021; Zhang et al., 2022a; Legros, 2002; Rickenmann, 1999; Zou et al., 2017; Corominas, 1996; Hunter and Fell, 2003; Devoli et al., 2009; Hattanji and Moriwaki, 2009; Whittall et al., 2017; Moncayo and Ávila, 2023; Finlay et al., 1999; Scheidegger, 1973; Wan and Xue, 2023) are based on a limited number of landslides caused by a specific triggering event, and thus landslide data from different locales and historical events is aggregated in studies (Qiu et al., 2018; Li et al., 2011; Geo et al., 2014; Qiu et al., 2017). Area-specific landslide runout data are preferable to reduce uncertainty associated with runout analyses since the geologic setting influences the types of landslides that occur and their runout. However, even within a specific area, landslides may be triggered by various factors, such as rainfall or earthquakes, and the landslide material may vary depending on the subsurface profile.

Triggering factors, geology, topography, landslide volume, slope angle, material, and landslide size are known to affect landslide mobility (Zou et al., 2017; Wan and Xue, 2023; Legros, 2002; Qiu et al., 2018; Corominas, 1996; Li et al., 2011;

Geo et al., 2014; Crosta et al., 2018; Zhang et al., 2022a; Tsunetaka et al., 2022; Roback et al., 2018; Hattanji and Moriwaki, 2009; Qiu et al., 2017; Zhao et al., 2021; Moncayo and Ávila, 2023; Whittall et al., 2017; Johnson and Campbell, 2017). Zou et al. (2017) pointed out that landslide volume is the controlling predictor in estimating landslide runout for loess landslides as well as non-seismic and seismic rockslides. Also, the slope angle of the landslide source affects the runout distance for non-seismic rockslides. The 55 historical landslides in their dataset are in colluvium, loess, and rock; three different material types. They showed that landslide material significantly impacts landslide runout. The triggering factor is also vital in landslide runout; rainfall-induced rockslides have higher mobility than gravity-induced rockslides, followed by earthquake-induced rockslides. Wan and Xue (2023) concluded that the landslide source volume and slope angle significantly affect landslide travel distance for rainfall-induced and earthquake-induced landslides. Legros (2002) proposed that the runout distance mainly depends on the volume of landslides for non-volcanic, volcanic, Martian, and submarine landslides. Qiu et al. (2018) showed that loess slide volume, slope height, and slope angle affect the landslide travel distance, with travel distance increasing with the increase in slope height and slide volume. Corominas (1996) stated that the angle of reach, the ratio between the vertical drop of landslide and the horizontal projection of landslide runout distance, is a suitable indicator of landslide mobility, and the angle of reach reduces with volume increase for rockfall, debris flow, earthflow, and translational slide. Li et al. (2011) concluded that volume and slope angle influence the landslide travel distance, and the logarithmic value of landslide volume positively correlated with landslide travel distance in seismic landslides. Geo et al. (2014) reported that rock type, landslide source volume, and slope transition angle were predominant factors in travel distance for landslides in the 2008 Wenchuan earthquake. Crosta et al. (2018) investigated landslide mobility on Mars and concluded that the landslide location, landslide typology, the presence of ice, and how ice melts affect landslide mobility. Landslide mobility increased with latitude because of the presence of ice. Zhang et al. (2022a) observed a linear relationship between landslide height and travel distance and concluded that larger landslides have greater mobility. Tsunetaka et al. (2022) observed that the tree height around the runout area affects the landslide runout for rainfall-induced landslides. Hattanji and Moriwaki (2009) stated that the ratio of landslide height and length strongly correlates with the slope angle of the landslide source for relict landslides. Qiu et al. (2017) revealed that for the loess slides in China, the landslide-affected area increases with the increase in landslide volume, and the landslide travel distance increases with landslide height. Zhao et al. (2021) showed that the landslide area, slope angle, and peak ground acceleration affect landslide mobility for seismic earth flows. Moncayo and Ávila (2023) revealed that the landslide volume, height, and slope angle control the landslide travel distance in Colombia. Whittall et al. (2017) reported that landslide mobility is strongly sensitive to the fall height, slope angle, and material properties in open pit slope failures, but is only modestly sensitive to volume. Weathered and saturated rock mass material led to the high mobility of the open pit slope. Johnson and Campbell (2017) pointed out that the drop height and volume control the landslide mobility on both Earth and Mars.

CHAPTER 4.   DERIVATION OF VOELLMY MODEL PARAMETERS FOR LANDSLIDE RUNOUT
BASED ON CO-SEISMIC ROCK AVALANCHES INVENTORY

62

Roback et al. (2018) showed that the landslide size influences the landslide runout distance and topography controls landslide mobility and found that in the steep and high mountainous terrain of Nepal, co-seismic landslides from the 2015 Gorkha earthquake traveled an area that was equal to 10 times the source area.

Most often, landslide mobility empirical models provide regressions of the landslide travel distance using landslide characteristics of the entire landslide (i.e., source+runout), such as landslide volume, landslide area, and landslide height (Legros, 2002; Scheidegger, 1973; Corominas, 1996; Rickenmann, 1999; Qiu et al., 2018; Li et al., 2011; Tsunetaka et al., 2022; Zhang et al., 2022a; Devoli et al., 2009; Qiu et al., 2017; Whittall et al., 2017; Zhao et al., 2021; Moncayo and Ávila, 2023; Finlay et al., 1999). Table 4.1 is a compilation of many such empirical models. However, from a practical perspective, estimating the landslide travel distance from the landslide source's characteristics is more valuable than the entire landslide characteristics. Tsunetaka et al. (2022) and Zhao et al. (2021) documented more than one hundred landslides from a single event, but the source and runout areas of the landslides are not separated in the inventories. A few studies (Wan and Xue, 2023; Zou et al., 2017; Geo et al., 2014; Hattanji and Moriwaki, 2009) considered landslide characteristics of the landslide source in the estimation of landslide travel distance, but considered a limited number of landslides from a specific event. Wan and Xue (2023) considered the slope angle of the landslide source and landslide source volume for 111 landslides from different studies and events. Zou et al. (2017) took account of landslide source height and travel distance for 55 historical landslides in China. Geo et al. (2014) examined the landslide source volume and area for 54 landslides triggered by the 2008 Wenchuan earthquake. Hattanji and Moriwaki (2009) included landslide source area and slope angle of landslide source for 338 relic landslides from four different locations: Hachimantai, Nagano, Ichinomiya, and Gojo, in Japan. In Table 4.1, the landslide travel distance, L (L$_{2D}$), is defined as the maximum horizontal distance that parallels the slide direction between the crown and toe, as shown in Figure 4.1.
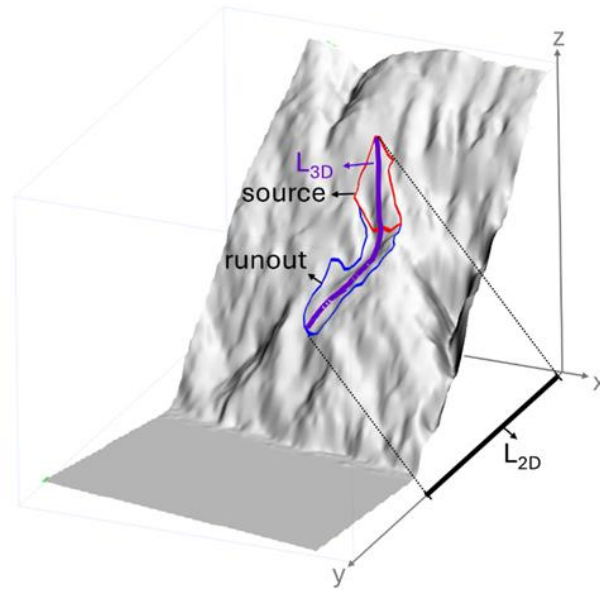
Figure 4.1: 3D oblique view of landslide travel distance

Although some studies considered the average slope angle, none of the empirical landslide mobility models in Table 4.1 considered the 3-dimensional (3D) topography effect in estimating landslide travel distance. The 2D definition allows for a simplified equation and provides a reasonable runout estimate, especially given the common situation where high-quality, reliable field data is lacking to validate the landslide mobility. In reality, the landslide travel path is not a straight line but a more complex geometry because of topographic changes along the path, as shown in L_{3D} in Figure 4.1.

Several mechanistic models are also available for landslide runout simulation, such as RAMMS-DF (RAMMS User Manual, 2022), FLO-2D (Wu et al., 2013), and D-Claw (Barnhart et al., 2021). They have been used for a wide range of runout analyses, including debris flow (Mikoš and Bezak, 2021), snow avalanches (Christen et al., 2013), and rock avalanches (De Pedrini et al., 2022).

Barnhart et al. (2021) compared three different 3D continuum models, RAMMS-DF, FLO-2D, and D-Claw, and concluded that all three models did well at simulating the inundation pattern and peak depths observed in the debris flow runout for the 2018 Montecito California, post-wildfire event if the debris flow volume is constrained well. The Rapid Mass Movement Simulation Debris Flow (RAMMS-DF) model is one of the most used models. RAMMS-DF is a single-phase model in which the material is modeled as a bulk flow and does not distinguish between the fluid and solid phases (RAMMS User Manual, 2022). RAMMS is a 3D continuum model using a depth-averaged 2D solution to motion law over a 3D topography and depth-integrated equilibrium equations and makes some assumptions on the flow vertical structure (Liu et al., 2021). D-Claw represents fluid and solid fractions as quasi-two-phase flow and needs a time-dependent solid volume fraction input, hydraulic

permeability, basal friction angle, debris' elastic compressibility, and effective shear viscosity of the pore fluid. FLO-2D uses five inputs, including three quadratic empirical relationship coefficients introduced by O'Brien and Julien (1985), Manning's n value, and the flow volume. RAMMS-DF is simpler than D-Claw and FLO-2D and requires only three inputs in RAMMS-DF, two Voellmy friction parameters, and the landslide source volume.

RAMMS-DF implements the Voellmy friction law to simulate mass movements (RAMMS User Manual, 2022). The biggest challenge in the runout simulation is the wide variety of landslide conditions, such as material, triggering mechanism, moisture content, runout magnitude, simulation resolution, study location, and environment that affect the choice of Voellmy friction parameters (RAMMS User Manual, 2022). Since the material is modeled as a bulk flow in RAMMS-DF, the Voellmy friction parameters should be calibrated to represent the material for the field condition.

The RAMMS-DF model was initially developed for debris flow runout simulation. Debris flow is a rapid mass movement of solid-fluid mixtures (Iverson, 2005). However, the RAMMS-DF runout simulation is not limited to debris flow. It has also been applied to other runout simulations, such as rockslides, rock avalanches (De Pedrini et al., 2022), and glacier avalanches (Zhang et al., 2022b; Gilany and Iqbal, 2019). A rockslide involves a relatively intact rock mass that moves along a distinct weak zone, separating the slide material from the more stable underlying material (Hungr et al., 2014). Rock avalanche also starts with a relatively initially intact rock mass, but moves as masses of fragments in a flow-like manner. It is often associated with large-volume (> 1 Mm^3) bedrock mass movement (Hungr et al., 2001). Rock-ice avalanches are rock and debris avalanches that contain ice or snow from steep glaciers, rock-ice walls' failure, or ice entrainment along the propagation path (Schneider et al., 2011). A glacier avalanche is a vast glacier slope failure leading to a massive portion of breaking glacier (with million cubic meters or more) moving through the landscape with a velocity of up to 100 m/s (Voight and Pariseau, 1978).

Table 4.2 lists Voellmy friction parameters for the RAMMS simulation available in the literature. Debris flows caused by rainfall have been primarily investigated (Schraml et al., 2017; Gan et al., 2019; Zhuang et al., 2021; Peethambaran et al., 2023; Kumar et al., 2019). Mikoš and Bezak (2021) explored 35 debris flow cases worldwide in alpine environments. Schraml et al. (2015) studied two debris flow cases in Austria. Gardezi et al. (2021) analyzed one debris flow in Pakistan. Roy et al. (2022) evaluated two debris flows triggered by rainfall in India. De Pedrini et al. (2022) performed one rock avalanche simulation for the 1513 Monte Crenone rock avalanche, and Zhang et al. (2022b) utilized RAMMS-DF to simulate two rock avalanches in China triggered by rainfall.

As shown in Table 4.2, the range of best-fit μ is between 0.001 and 0.7 (i.e. a wide range) and the range of best-fit ξ is between 10 m/s^2 and 2000 m/s^2 for the debris flow simulation in RAMMS-DF (RAMMS User Manual, 2022; Roy et al., 2022; Schraml et al., 2017; Gan and Zhang, 2019; Zhuang et al., 2021; Peethambaran et al., 2023; Kumar et al., 2019; Gardezi et al., 2021; Mikoš and Bezak, 2021; Schraml et al.,

2015). For the rock avalanches simulation in RAMMS-DF, the range of best-fit μ is between 0.12 and 0.7, and the range of best-fit ξ is between 250 m/s^2 and 1200 m/s^2 (De Pedrini et al., 2022; Zhang et al., 2022b). Schraml et al. (2015) and Luna et al. (2013) collected Voellmy friction parameters from published values using different software. They summarized that rock avalanches have similar μ with debris flows, but have higher ξ. Rock-ice avalanches have higher ξ and lower μ than debris flows, possibly due to the ice and the water generated from the melting ice. For the debris flows and rock-ice avalanches, the μ decreases and ξ increases with an increase in landslide volume. However, there is no clear trend between Voellmy friction parameters and the landslide volume for rock avalanches.

Mikoš and Bezak (2021) reported that Voellmy friction parameters vary widely. No clear pattern and correlation were observed for the frequency of the most representative Voellmy friction parameters based on 35 worldwide debris flow cases in an alpine environment. Thus, the most suited Voellmy friction parameters depend significantly on the local area characteristics, such as topography, geology, and hydrology, and model calibration is the most appropriate method for deriving the best-fit Voellmy friction parameters (RAMMS User Manual, 2022). Pre-event and post-event field data are needed to correctly calibrate Voellmy friction parameters, and according to Mikoš and Bezak (2021), further research is necessary to explore the correlation between the Voellmy friction parameters and the landslide characteristics and properties.

In general, Voellmy friction parameters have been calibrated for a single or a limited few debris flow or rock avalanche cases in a single event (Roy et al., 2022; Schraml et al., 2017; Gan and Zhang, 2019; Zhuang et al., 2021; Peethambaran et al., 2023; Kumar et al., 2019; Gardezi et al., 2021; Mikoš and Bezak, 2021; Schraml et al., 2015; De Pedrini et al. 2022; Zhang et al., 2022b). However, the statistical significance of these back-analyses is not well established due to the limited number of cases. Calibration of Voellmy friction parameters against a large population of debris flows, rockslides, or rock avalanches from a specific event may provide statistically significant trends and an assessment of variance as well as the bias of the model against field observations. To date, some studies have provided Voellmy friction parameters for several debris flows. However, similar calibrations for rockslides or rock avalanches, especially co-seismic ones, are missing.

This study's contribution lies in calibrating Voellmy friction parameters based on thousands of simulations by leveraging a high-resolution inventory of hundreds of landslides. Specifically, the 716 landslides that occurred during the 2015 Lefkada earthquake event are compared against thousands of simulations to derive Voellmy friction parameters that match each landslide runout. Because the Voellmy friction model simulates the runout material as a bulk flow and does not distinguish between the fluid and solid phases, it is a simple and efficient model choice for the simulation of co-seismic landslides, which do not have a fluidized mass. In doing so, these parameters (average and variation) can provide a basis for simulations on these (or similar) geologic and topographic environments. In addition, based on a large number of simulations, the influence of various factors, such as landslide size and topography,

on the calculated Voellmy parameters can be quantified.

4.2 Methodology

4.2.1 Numerical Model

This study considers the Rapid Mass Movement Simulation Debris Flow (RAMMS-DF) model, which is formulated based on the Voellmy friction law (Salm, 1993). The Voellmy friction law considers the resistance of the solid phase, dry-Coulomb friction (μ), and the resistance of the turbulent fluid phase, viscous-turbulent friction (ξ), as shown in Equation 4.1 (RAMMS User Manual, 2022):

$$S = \mu N + \frac{\rho g u^2}{\xi} \text{ with } N = \rho h \cos(\phi) \tag{4.1}$$

where S (Pa) is frictional resistance, ρ is density, g is the gravitational acceleration, ϕ is the slope angle, h is the flow height, u is the flow velocity, and N is the normal stress.

The two Voellmy friction parameters define the flow behavior: μ governs when the flow is close to the deposition zone, and ξ controls as the flow moves quickly (RAMMS User Manual, 2022). The combination of μ and ξ controls the landslide runout geometry (Hungr, 2010). Calibrating the Voellmy friction parameters is essential in getting a realistic simulation. In this study, the block release simulation is used where the release area and release volume are defined based on the landslide inventory, and the sliding depth of the release area is calculated accordingly. For each landslide, hundreds of analyses are conducted, with μ varying between 0.01 and 1 and ξ varying between 50 m/s^2 and 10000 m/s^2, to identify the best-fit Voellmy friction parameters. The best match between the model runout area and the mapped runout area is quantified using the F1 score shown in Equation 4.2 (Powers 2015).

$$\text{F1 score} = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \tag{4.2}$$

where TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative.
The comparison is made on a pixel-by-pixel basis. TP occurs when a mapped landslide pixel is modeled as a landslide. Otherwise, it is FP. TN occurs when a pixel that is not a landslide (i.e., outside the landslide footprint) is modeled to be not a landslide. Otherwise, if it is modeled to be a landslide, that is FN.

4.2.2 Study Area and Landslide Inventory

A co-seismic landslide inventory from the M$_w$ 6.5 earthquake that occurred on

November 17, 2015, on the island of Lefkada in Greece is used in this study. The inventory was developed by Zekkos and Clark (2020) and consists of 716 landslides on the highly tectonized and brecciated limestone along the west coast of Lefkada island (Huang and Zekkos, 2023). Most landslides triggered by the 2015 Lefkada earthquake were along Lefkada island's west coast. Once triggered, the landslide debris moved at a high velocity, and thus, these landslides would be classified as rock avalanches per Hungr et al. (2014). The inventory consists of fully three-dimensional landslide polygons derived based on mapping using imagery and digital surface models developed using satellite imagery and imagery produced by Unmanned Aerial Vehicles (UAVs). Each landslide in the inventory consists of a polygon with a source area, runout area, and volume.

Earlier work by Huang et al. (2024) leveraged the landslide inventory (Zekkos and Clark, 2020) and a 2-m digital elevation model (DEM) to derive empirical landslide mobility models based on multivariate linear and machine-learning regression models. The pre-earthquake 2-m resolution digital elevation model (DEM) is used to estimate the landslide travel distance considering the 3D topography. Figure 4.2 compares landslide travel distances in 3D and 2D. The landslide travel distance in 3D considers all deviations from a linear (2D) travel path. The 3D path is, as expected, higher than the 2D travel distance projected as a straight line and simplified as the maximum horizontal distance that parallels the slide direction between the landslide crown and toe. The landslide travel distance in 2D, as used in the literature in Table 4.1, is lower than the landslide travel path influenced by the topography. The influence of several parameters on the landslide travel distance was considered, and following the analyses, the equations with the most important parameters are presented, as shown in Equation 4.3 and Equation 4.4. The height, slope angle, and area of the entire landslide or the source-only are the three most significant parameters in estimating landslide travel distance, as shown in Equation 4.3 and Equation 4.4, respectively:

$$D = 0.981(H_e) - 2.979(S_e) + 0.003(A_e) + 134.93 \qquad (4.3)$$

$$D = 2.368(H_s) - 4.074(S_s) + 0.0003(A_s) - 0.002(V_s) + 216.13 \qquad (4.4)$$

where D = 3D landslide travel distance (m); H_e = Height, entire (m);
S_e = Slope angle, entire (degree); A_e = Area, entire (m^2); H_s = Height, source (m);
S_s = Slope angle, source (degree); A_s = Area, source (m^2); V_s = Volume loss, source (m^3).

Despite the ease of use of such empirical models, due to the range of factors affecting landslide mobility discussed earlier, it is practically difficult to assess the conditions for which such a regression model is valid. Thus, the empirical landslide mobility models are often used in regions with similar geological and geomorphological environments (Geo et al., 2014). On the other hand, mechanistic models provide a mathematical function based on the underlying physics, such as

landslide size, velocity, and material properties that provide a basis for making decisions on input variables. Compared to the empirical model, such as the one derived by Huang et al. (2024) for this event, the calibrated mechanistic model presented in this study has the advantage that it can potentially be more broadly applicable in changing topographies and can also provide helpful intuition on how the model parameters may change due to changing conditions.
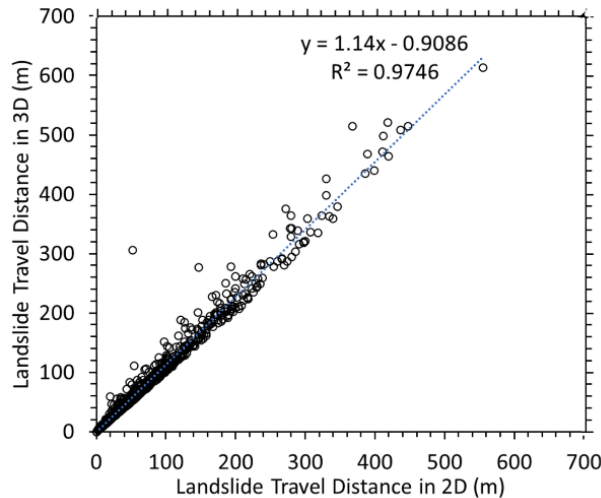
Figure 4.2: Comparison of landslide travel distance in 2D and 3D

4.3   Results and Discussion

Subsurface conditions vary geospatially and temporally due to variations in moisture content. Variations in subsurface conditions should affect the landslide characteristics and the input parameters of the runout model. Model calibration against numerous landslides can provide representative Voellmy friction parameters that are statistically significant and represent the conditions at the time of landsliding (RAMMS User Manual, 2022). Figure 4.3 shows an example of a landslide mobility simulation for one of the landslides triggered by the 2015 Lefkada earthquake, compared to the mapped landslide. The source area and the volume of the landslide source are assigned to the source polygon based on the volume calculated from the inventory, and the sliding depth of the landslide source is calculated accordingly. For the simulation example of Figure 4.3a, the landslide source volume of 151,543 m^3 has an average depth of 3.45 m. Figure 4.3b displays the maximum flow height in the simulation result. The colored area represents the modeled runout area. The purple color means the flow height is near 0 m, dark blue is around 4 m, green is ~13 m, yellow is ~22 m, and red is ~26 m. Figure 4.3b reveals that the landslide travels along the topography, and the landslide arrests when the landslide material reaches the flat coastline. By comparing the entire landslide area and landslide travel path between the inventory and simulation results using the F1 score, the pair of landslide runout properties, i.e., dry-Coulomb friction

(μ) and viscous-turbulent friction (ξ) that best match the field observations, is
identified.
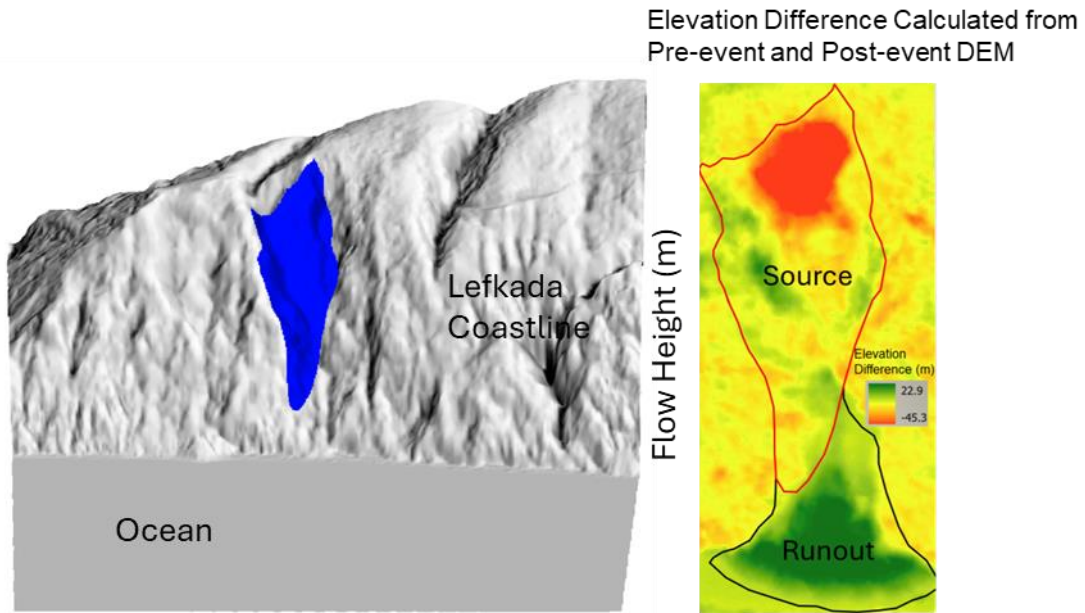
(a)
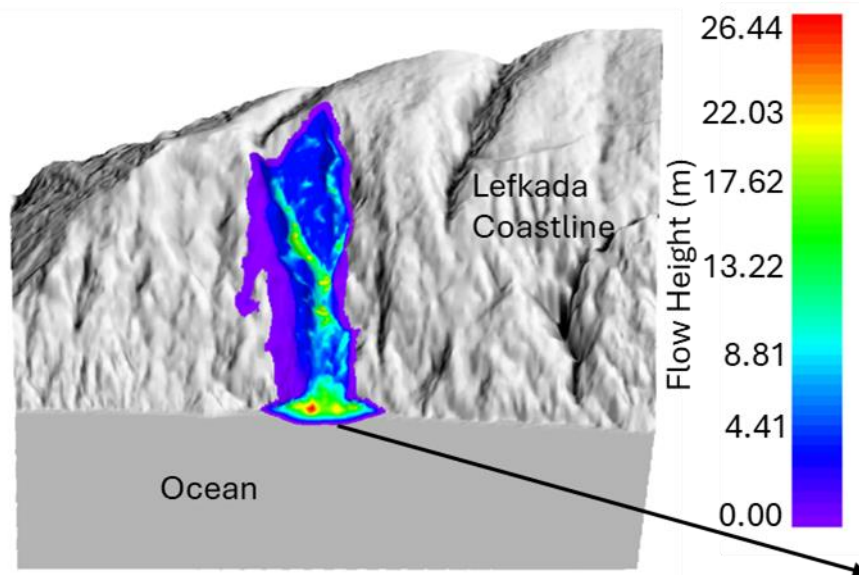
(b)

Figure 4.3: Landslide simulation example in RAMMS-DF (a) landslide source specified
based on inventory (b) simulated landslide travel path based on topography

Table 4.1: Empirical landslide mobility model in the literature

Empirical Equation	Data Size, number of landslides	Type of Event	Reference	Note
$\log L = 0.0277 + 0.0963 \log V + 0.9468 \log H$	11,622	Rainfall-induced landslides in Hong Kong during the period 1984-2013	Geo et al. (2021)	Historical events
$H/L = 0.81 - 0.11 \log A$	12,110	Historical landslides in Lvliang City, Shanxi Province in China	Zhang et al. (2022a)	Historical events
$H/L = 0.16 V^{-0.15}$ $H/L = 0.11 V^{-0.19}$ $H/L = 0.42 V^{-0.19}$ $H/L = 0.03 V^{-0.09}$	203	Non-volcanic landslides Volcanic landslides Martian landslides Submarine landslides	Legros (2002)	Several events in the literature
$L = 1.9 V^{0.16} H^{0.83}$	232	Debris flow from different events	Rickenmann (1999)	Several events in the literature
$H/L = 3.5033 - 0.623 \log V$ $H/L = 0.7233 - 0.0686 \log V + 0.0725 H_0/L_0$ $H/L = 1.2241 - 0.1226 \log V$	55	Loess landslide Non-seismic rockslide Seismic rockslide	Zou et al. (2017)	Different events in China
$\log (H/L) = -0.109 \log V + 0.210$ $\log (H/L) = -0.105 \log V - 0.012$ $\log (H/L) = -0.070 \log V - 0.214$ $\log (H/L) = -0.068 \log V - 0.159$	204	Rockfall Debris flow Earth flow Translational slide	Corominas (1996)	Several events in the literature
$H/L = 0.54 \tan\alpha + 0.16$	350	Debris flow	Hunter and Fell (2003)	Several events in the literature
$L = 0.8653 \log H + 0.5085$ $L = 1.1172 \log H + 0.228$ $L = -3000 \ln(1 - H/1600)$ $\log (H/L) = -0.0916 \log V - 0.041$ $H/L = -0.1047 \log V - 0.0253$	13 81 256 527 33	Rockfall Slide Debris flow International data Central American data	Devoli et al. (2009)	Several events in the literature
$H/L = 0.77 \tan\theta_r + 0.040$ $H/L = 0.82 \tan\theta_r + 0.034$ $H/L = 0.91 \tan\theta_r + 0.035$ $H/L = 0.83 \tan\theta_r + 0.055$	73 92 93 80	Relict landslides-Hachimantai Relict landslides- Nagano Relict landslides- Gojo Relict landslides-Ichinomiya	Hattanji and Moriwaki (2009)	338 relict landslides in the four mountainous areas in Japan
$H/L = 0.488 \tan\theta + 0.117$	105	Open pit slope failure	Whittall et al. (2017)	Cases from different mines
$\log L = 0.286 + 0.159 \log V + 0.563 \log H - 0.618 \log \tan\theta$	123	-	Moncayo and Ávila (2023)	Landslides that occurred from the beginning of the twentieth century to the present in the Andean region of Colombia
$\log L = 0.109 + 1.010 \log H - 0.506 \log \tan\theta$ $\log DT = 0.678 + 0.695 D + 0.0537 H$ $\log L = 0.178 + 0.587 \log H + 0.309 \log V/W$ $\log L = 0.253 + 0.703 \log H - 0.417 + \log \tan\theta$	515 68 50 61	Cut slope Fill slope Retaining wall Boulder fall	Finlay et al. (1999)	Landslides in man-modified slopes in Hong Kong for the period 1984-1993
$\log (H/L) = -0.15666 \log V + 0.62419$	33	-	Scheidegger (1973)	Several events in the literature
$L = 0.6350 V^{0.2466} H^{0.5355} (\tan\theta)^{-0.2154}$	84	Loess slide on the Loess Plateau in China	Qiu et al. (2018)	
$\log L = -0.0785 \log V + 1.2347 \log(H/\tan\theta)$	46	Seismic landslides triggered by the 2008 Wenchuan earthquake	Li et al. (2011)	
$\log L = 0.136 RT + 0.159 \log V + 0.529 \sin\sixi + 1.497$	54	Seismic landslides triggered by the 2008 Wenchuan earthquake	Geo et al. (2014)	
$H/L = 2.2061 + V^{-0.0926}$ $H/L = 7.0109 + V^{-0.2245}$ $H/L = 3.1442 + V^{-0.1448}$	69	Landslide in Loess Red clay contact landslide Bedrock contact landslide	Qiu et al. (2017)	Loess slides in China
$L = 1.57 A^{0.75}; L = 8.11 V^{0.43}$	366	Rainfall-induced landslides triggered by Typhoon Prapiroon in Japan in 2018	Tsunetaka et al. (2022)	
$H/L = 1.46 - 0.24 \log A; H/L = 0.013\theta + 0.03$	7058	Seismic earthflow induced by the 2018 Eastern Iburi earthquake in Japan	Zhao et al. (2021)	

where L (m) is the landslide travel distance, H (m) is the entire landslide height, V (m³) is the entire landslide volume, A (m²) is the entire landslide area, H_0 is the height of the landslide source, L_0 is the landslide travel distance of landslide source, α is the travel distance angle, $\tan\theta_r$ is the slope angle of the landslide source, θ is the slope angle of the entire landslide, DT is the depth of debris at the toe of the slope, D is the maximum landslide depths for the entire landslide, W is the landslide width for the entire landslide, RT is rock type, and $\sixi$ is the slope transition angle, the angle change between the landslide source which is the upper slope and the lower slope.

Table 4.2: RAMMS-DF Voellmy friction parameters available in the literature

Dry-Coulomb friction		Viscous-turbulent friction (m/s²)		Numerical model	Number of Cases	Location	Triggering factor	Type, failure mode	Reference
min	max	min	max						
0.001	0.7	10	2000		Cases from 35 papers	Worldwide Alpine Environment			Mikoš and Bezak (2021)
0.07	0.11	200	300		2	Austria			Schraml et al. (2015)
0.1		200			1	Pakistan			Gardezi et al. (2021)
0.21	0.29	100	200		2	India	Heavy rainfall		Roy et al. (2022)
0.1		200			1	Nepalese Himalayas	Earthquake + monsoon	Debris flow	Schraml et al. (2017)
0.07		1500		RAMMS-DF	1	Luzhuang Gully in China	Rainfall		Gan et al. (2019)
0.2	0.32	200	850		1	China	Heavy rainfall-induced		Zhuang et al. (2021)
0.06		1450			1	India	Rainfall		Peethambaran et al. (2023)
0.1		500			1	India	Rainfall		Kumar et al. (2019)
0.4	0.7	250	600		1	Switzerland			De Pedrini et al. (2022)
0.12		1200			2	China	Gravity, rainfall, weather changes	Rock avalanche	Zhang et al. (2022b)

Overall, 248 landslides have an F1 score >60 %, as shown in Figure 4.4. Figure 4.5 presents the relationship between the F1 score and the landslide source area, landslide source height, and landslide 3D travel distance. Figure 4.7 illustrates the back-calculated values of μ and ξ for landslides with F1 score >60 %. In the 2015 Lefkada earthquake event, 248 landslides have F1 scores >60 %, with the remaining having lower F1 scores. The mean, mode, and standard deviation of ξ are 4264 m/s², 3000 m/s², and 3314 m/s², respectively, as shown in Figure 4.7a. The mean, mode, and standard deviation of μ are 0.8, 0.9, and 0.28, respectively, as shown in Figure 4.7b. Figure 4.7c shows that the mode of density plot has ξ =3000 m/s² and μ =0.9.

Figure 4.8 displays an example of the influence of μ on the simulation results with the same ξ value of 1689 m/s². Figure 4.8a shows that the F1 score increases significantly from 19 % to 76.7 % when the μ increases from 0.05 to 0.81. Figure 4.8b shows that as μ increases from 0.05 to 0.81, the landslide area in the simulation result becomes similar to the landslide area from the inventory. Figure 4.9 displays the influence ξ on the simulation results with the same μ value of 0.81. Figure 4.9a shows that the F1 score increases modestly from 61.83 % to 73.12 % when the ξ increases from 50 m/s² to 378 m/s². After the ξ reaches around 1000 m/s², the F1 score becomes stable around 76 %. Figure 4.9b shows that as the ξ increases from 50 m/s² to 378 m/s², the landslide area in the simulation result becomes similar to the landslide area from the inventory. Interestingly, after ξ reaches 1000 m/s², the F1-score becomes stable, and the change of ξ does not significantly impact the simulation result. Although the mode of ξ for rock avalanches triggered by the 2015 Lefkada earthquake is around 3000 m/s², the best-fit value of ξ can be between 1000 m/s² and 3000 m/s². Figure 4.8 and Figure 4.9 reveal that the change of μ can lead to a significant increase in the F1 score by 60 %, and the change of ξ can cause an increase in the F1 score by 11 %. In general, the F1 score (i.e., the metric used to quantify the similarity of the simulation result to the inventory) is more sensitive to μ than ξ. The change of μ governs the accuracy of RAMMS-DF simulation. The F1 score positively correlates with the landslide source area, as shown in Figure 4.5a. Landslides with small source areas tend to have an F1 score <40 %, and landslides with source areas >10,000 m² achieve an F1 score >50 %. Figure 4.5b shows the same trend expressed in terms of source height, which is commonly used in empirical methods); landslides with larger source heights are more accurately simulated. Similarly, Figure 4.5c reveals that landslides with longer 3D travel distances have a higher F1 score. Overall, landslides with larger source areas, larger landslide source height, and longer 3D landslide travel distances are more accurately simulated and have a higher F1 score.

Although the simulation becomes less accurate for smaller landslides, the model bias is always towards overpredicting the runout. Recognizing model bias is important, and this bias may appeal to a number of scenarios. For example, in risk assessment studies, it may be preferable to have a model that over-predicts runout and thus has an inherent conservatism. Figure 4.6 illustrates that when simulation results are inaccurate (lower F1 scores), they are conservative as they predict a larger landslide area than the inventory. When the F1 score drops below 20 %, the modeled landslide area may be significantly larger (5-30 times) than the inventory area. As the F1 score

increases, the ratio between the landslide area in the simulation and the inventory decreases and is closer to one. When the F1 score reaches 60 %, the modeled landslide area is approximately equal to 1-3 times the landslide area in the inventory. Landslides with an F1 score >60 % have approximately the same mapped and modeled area.

Luna et al. (2013) summarized 270 past events in different geologic settings and environments. They concluded that the best-fit μ for the debris flow is between 0 and 0.5, and for the rock avalanche is between 0 and 0.6. Also, the best-fit ξ for the debris flow is between 0 m/s^2 and 1000 m/s^2, and for the rock avalanche is between 0 m/s^2 and 3000 m/s^2. Schaub and Cochachin (2016) reported that μ affects the runout distance, and μ can represent the tangent of the average slope angle at the deposition zone. Rock avalanches triggered by the 2015 Lefkada earthquake are on steep slopes along the coastline and were very dry when the earthquake occurred. This is possibly the reason that the mode of the best-fit value of μ of 0.9 for the co-seismic rock avalanches triggered by the 2015 Lefkada earthquake is higher than the best-fit μ debris flow and rock avalanche database in Luna et al. (2013). ξ influences the flow velocity (Schaub and Cochachin, 2016). De Pedrini et al. (2022) mentioned that the higher value of ξ suggests a quick-running flow of rock avalanches caused by the high interaction of fragmented rocks. The mode of the best-fit value of ξ is equal to 3000 m/s^2 for the co-seismic rock avalanches triggered by the 2015 Lefkada earthquake, which is higher than the best-fit ξ debris flow database and similar to the values for rock avalanches that represent the quick movement of rock avalanches triggered by the earthquake.
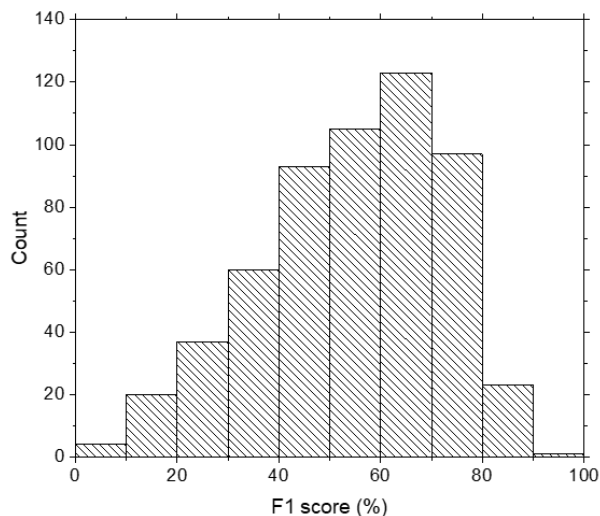
Figure 4.4: F1 score in the simulation result

(a) F1 score versus landslide source area

(b) F1 score versus landslide source height

(c) F1 score versus landslide 3D travel distance

Figure 4.5: The influence of landslide source area, height, and 3D travel distance on F1 score
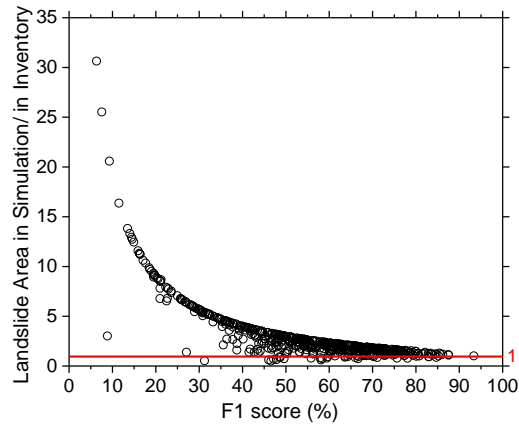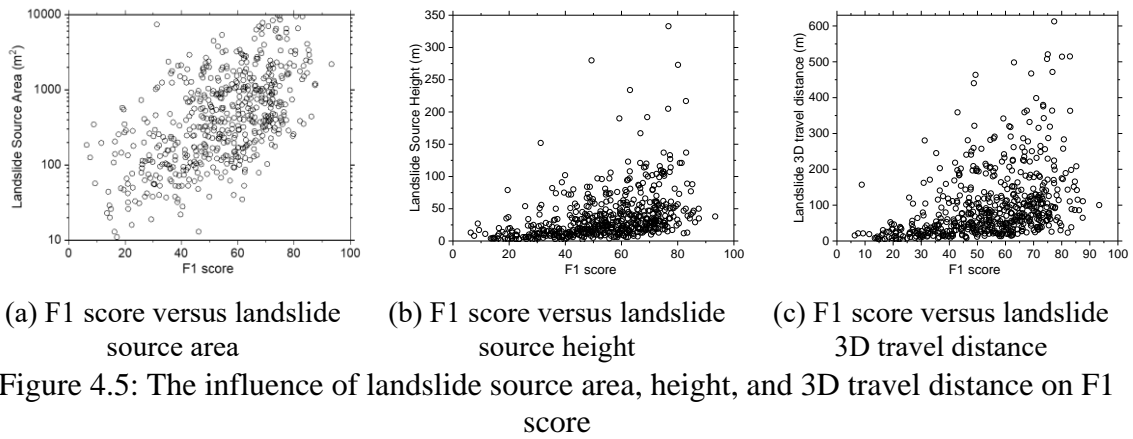
Figure 4.6: F1 score versus the ratio of landslide area in RAMMS-DF simulation result and the inventory
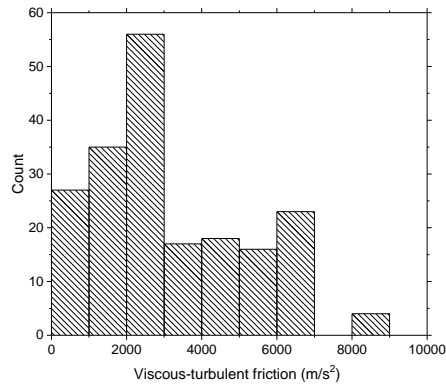
(a) The histogram of viscous-turbulent friction, ξ
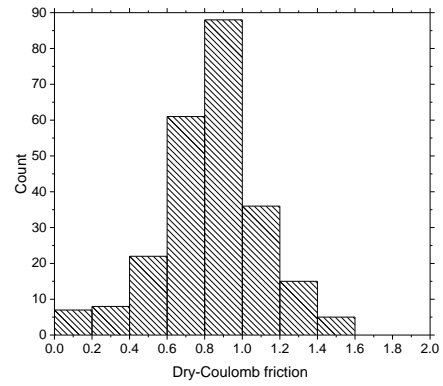
(b) The histogram of dry-Coulomb friction, μ
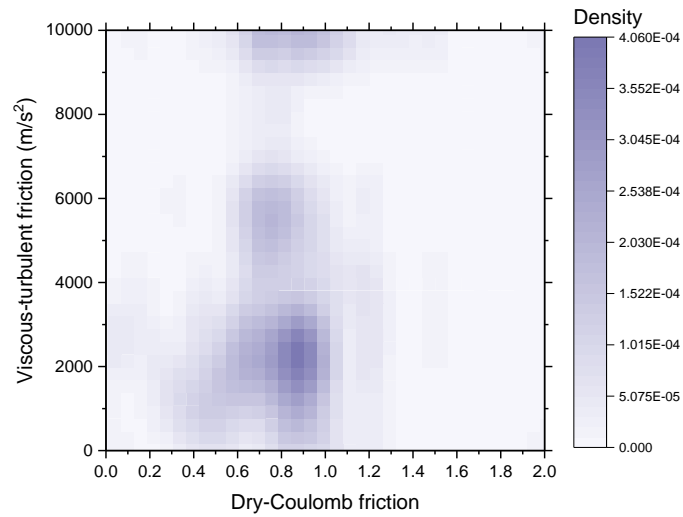
(c) Density plot for dry-Coulomb friction, μ, and viscous-turbulent friction, ξ

Figure 4.7: The histogram of viscous-turbulent friction (ξ) and dry-Coulomb friction (μ) for landslides with F1 score >60 %
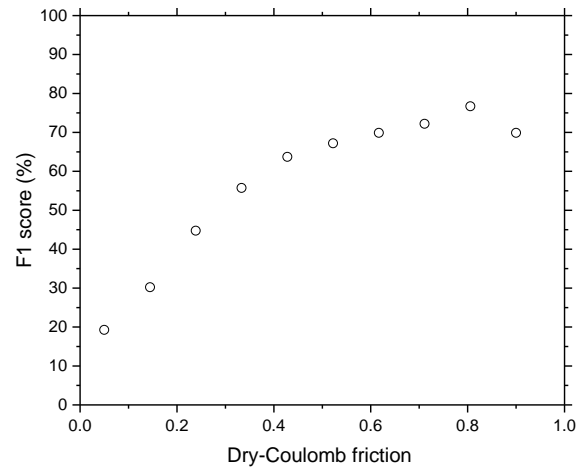
(a) F1 score versus μ
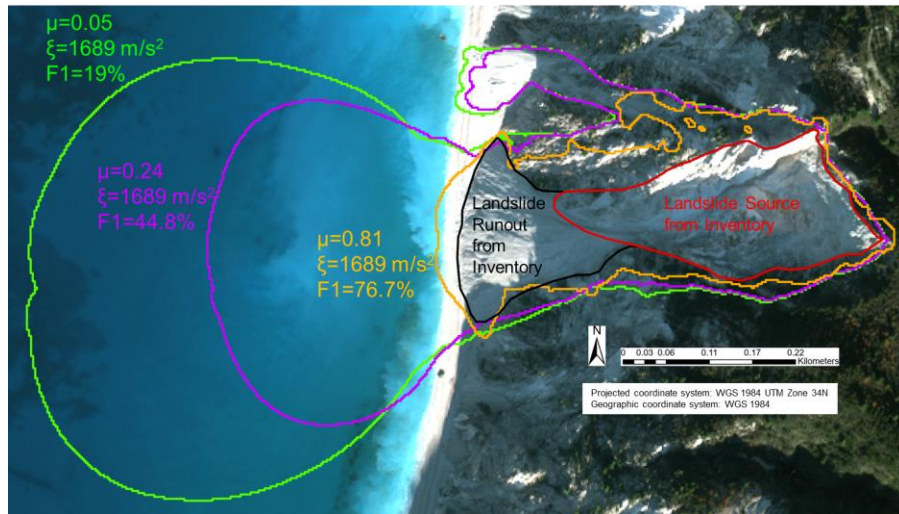
(b) Landslide area with different μ

Figure 4.8: The influence of μ on the RAMMS simulation
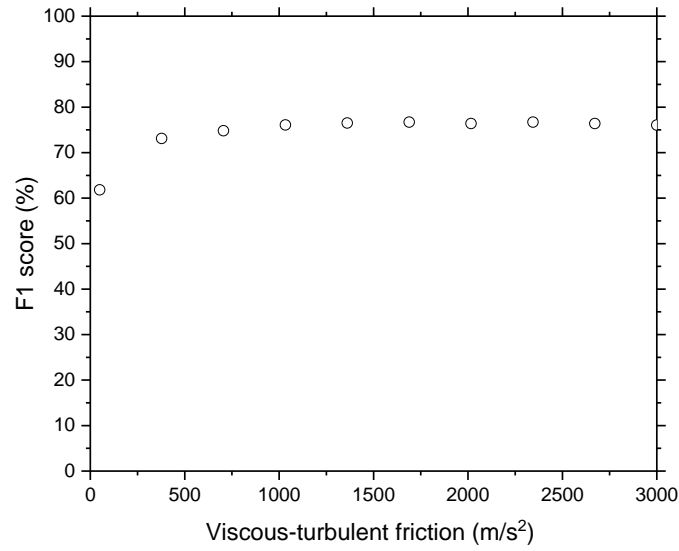
(a) F1 score versus ξ
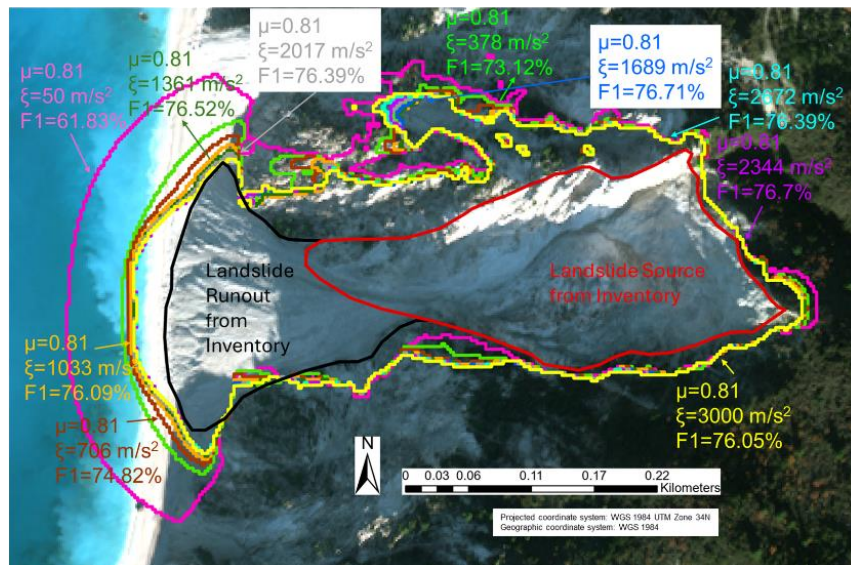
(b) Landslide area with different ξ

Figure 4.9: The influence of ξ on the RAMMS simulation

Figure 4.10a-d show the relationship of the landslide source volume, landslide source height, landslide source area, and landslide 3D travel distance with ξ. Landslides with ξ >3000 m/s^2 (i.e., 4000 m/s^2 and 10000 m/s^2) are all small landslides with small source volume, source height, small landslide source area, and short landslide 3D travel distance. As shown in the example of Figure 4.9, ξ values that are greater than 1000 m/s^2 produce similar landslide areas and travel paths in the simulation result. Thus, large ξ values do not represent the landslide population, but better approximate only small landslides, resulting in longer runout distances than those observed in the inventory.

Figure 4.11a-d show the relationship between the best-fit μ and the landslide source volume, landslide source height, landslide source area, and landslide 3D travel distance, respectively. The mean, mode, and standard deviation of the best-fit value of μ are 0.8, 0.9, and 0.28, respectively. Landslides with lower values of μ (<0.5) and higher values μ (>1.5) are small landslides with small source volume, source height, small landslide source area, and short landslide 3D travel distances. As shown in Figure 4.5, the simulation for the small landslide tends to have low F1 scores, so the best-fit value of μ for small landslides is not representative.
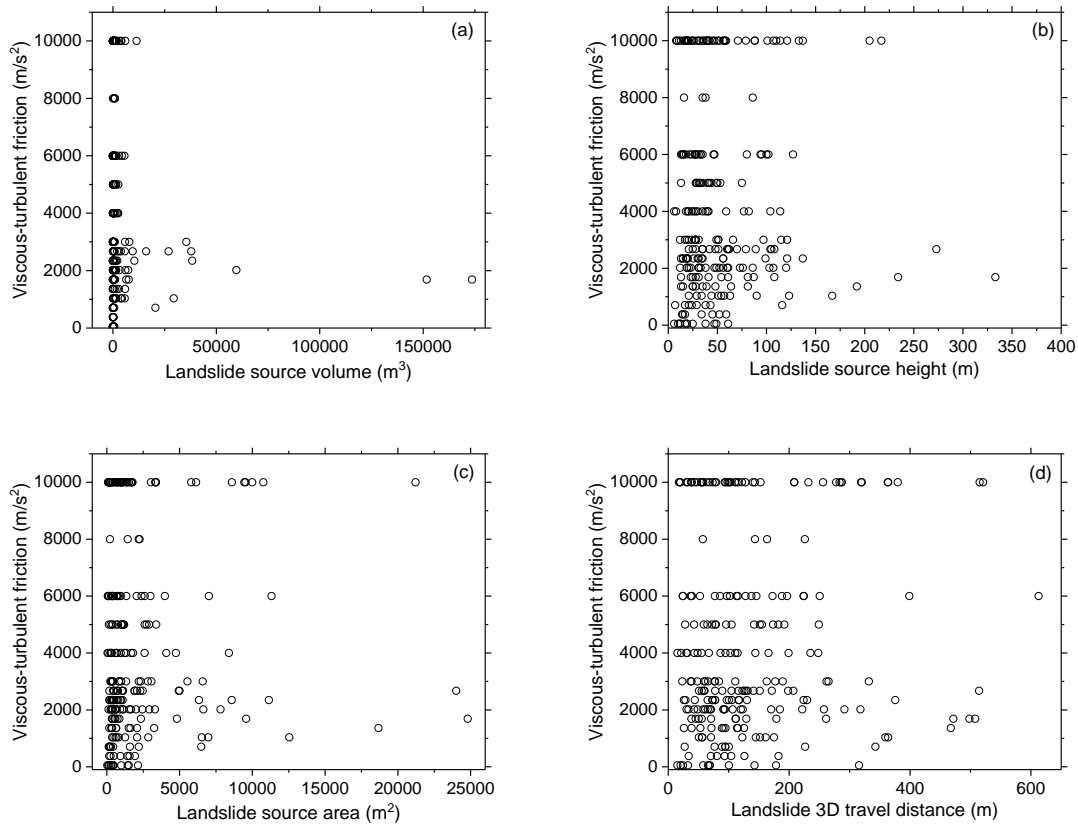
Figure 4.10: Relationship between back-calculated viscous-turbulent friction (ξ) and (a) landslide source volume, (b) landslide source height, (c) landslide source area, and (d) landslide 3D travel distance

For the landslides of the 2015 Lefkada earthquake event, a value of μ between 0.6 and 1 is observed. Smaller landslides have a wider range of μ and ξ but still do not match the inventory data well. This may be associated with modeling limitations. Corominas (1996) demonstrated that the scale effect should be considered in the landslide mobility assessment since most small landslides (less than 0.5×10^6 m^3) do not have any change in the mechanism of progression while in motion, and they have travel distances similar to large landslides. In the 2015 Lefkada earthquake event, the RAMMS-DF, a fluidized model, tends to overpredict the area and travel path for the small rock avalanches. The trend that small landslides have low F1 scores is consistent with this observation. It indicates that the actual release energy from the small rock avalanches might not be suitable and might not be large enough to be simulated as the fluidized model.
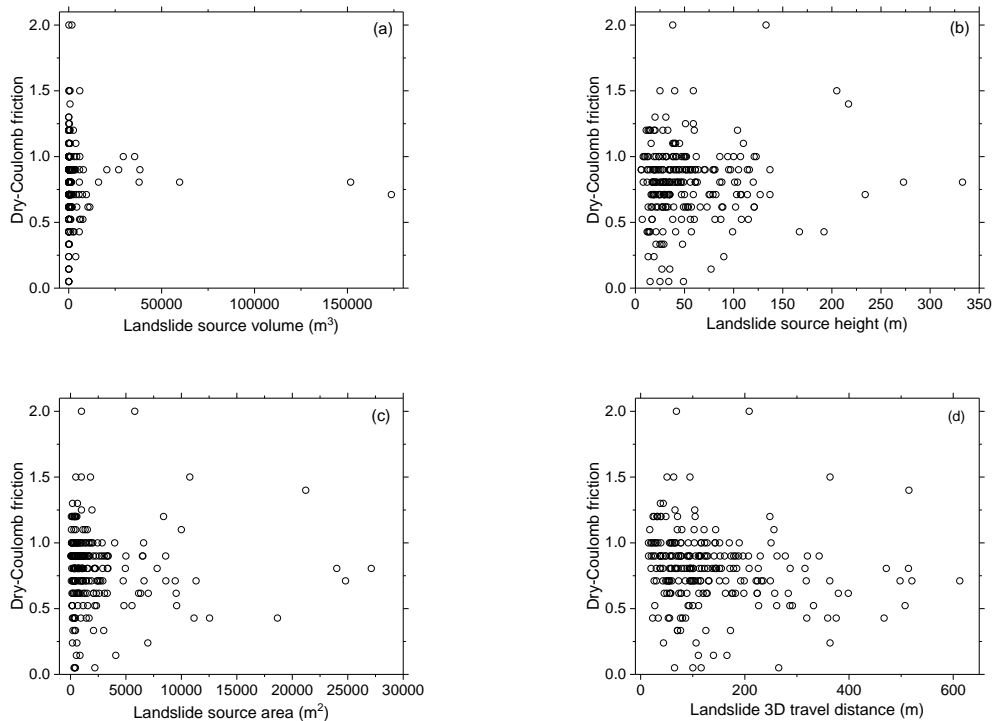
Figure 4.11: Relationship between back-calculated dry-Coulomb friction (μ) and (a) landslide source volume, (b) landslide source height, (c) landslide source area, and (d) landslide 3D travel distance

4.4 Conclusions

Understanding landslide mobility is essential to address landslide risk. Although primarily empirical landslide mobility models have been developed to estimate landslide travel distance, such models are difficult to provide insights into the reasons for the observed travel distance and, by extension, have significant uncertainties if implemented in different geologic environments. Instead, mechanistic models that are properly calibrated against field observations may provide better insights into the observed behavior. In this study, a simple mechanical model that is based on Voellmy friction law in the Rapid Mass Movement Simulation Debris Flow (RAMMS) is calibrated against an inventory of hundreds of landslides triggered by the 2015 Lefkada earthquake event to derive statistically significant model parameters. The mode and standard deviation of the dry-Coulomb friction (μ) are 0.9 and 0.28, and the mode and standard deviation of the viscous-turbulent friction (ξ) are 3000 m/s^2 and 3314 m/s^2. It is found that dry-Coulomb friction (μ) governs the accuracy of the Voellmy friction law simulation. Voellmy friction law simulation is not sensitive to viscous-turbulent friction (ξ), especially for $\xi >1000$ m/s^2. The performance of Voellmy friction law simulation is correlated with the size of the rock avalanches, such as the source area, source height, and 3D travel distance. The Voellmy friction law, a fluidized model, tends to overpredict the small rock avalanches' release energy and travel path compared to the inventory.

Chapter 5 Remote Sensing-based System-level Monitoring of Highway 1 Following the 2021 Atmospheric River Event

5.1 Introduction

Extreme events such as storms and hurricanes can trigger ground mass instability such as debris flows and landslides that can lead to loss of life or damage to the infrastructure, disrupting the road network and causing economic loss (Hilker et al., 2009, Trigg et al. 2023, Loli et al. 2022). Storms and hurricanes have damaged several highways and bridges worldwide in recent decades (Chen et al., 2009). Damage to the highway network could isolate the affected area and block access for rescuers and supplies (Han et al., 2009). Infrastructure retrofit and restoration before and after natural disasters and regular monitoring can reduce infrastructure damage. Guevara et al. (2017) noted that highway systems do not often achieve sustained improvements because authorities have to choose short-term reactive actions rather than long-term proactive efforts due to financial or other constraints. Also, transportation authorities do not always have sufficient resources to inspect the entire breadth of infrastructure systematically and sustainably.

California's coastline is an active geologic environment with a long history of landslides and debris flows associated with seismic activity, storms, complex geology, and steep topography. Highway 1 is a scenic highway along the California coastline, but landslides, debris flow, and sediment mobility have repeatedly damaged the road and resulted in the closure of Highway 1. Highway 1 has been closed more than 55 times between 1937 and 2001 due to landslides. As a result, the California Department of Transportation (Caltrans) systematically expends efforts for repairs (JRP Historical Consulting Services, 2001). In 1952, an 11-km section of Highway 1, north of San Simeon to Big Sur, was closed because of several large landslides (Provost, 2017). In

1955, Highway 1 was closed in Big Sur due to rainfall-induced landslides (Caltrans, n.d.). In 1957, earthquake-induced landslides damaged a segment of Highway 1 between Pacifica and Daly City (Woods, 2019). In 1983, a 14-month closure of Highway 1 was caused by massive landslides triggered by several storms between January and April. The largest landslide during that event mobilized 3.1 million cubic meters of material that buried Highway 1 near Julia Pfeiffer Burns State Park, and the debris from that event reached the ocean and created a beach in McWay Cove (JRP Historical Consulting Services, 2001 and Palmer, 2002). In 1998, El Niño storms damaged 40 locations of Highway 1, and a landslide caused a 3-month Highway 1 closure in the south of Gorda (National Oceanic and Atmospheric Administration 2016). The Devil's Slide area had two Highway 1 closure records: one was the 5-month closure in 1995, and another one was the 4-month closure in 2006 (Cabanatuan, 2010). In 2011, a 12-m section of Highway 1 collapsed south of the Rocky Creek Bridge, leading to road closure for several months (KSBW 2011). In 2016, a mudslide resulted in a Highway 1 closure near Julia Pfeiffer Burns State Park, and the Soberanes Fire brought about road closures in several locations in the summer (Magallon, 2016). In the winter of 2017, the Pfeiffer Big Sur State Park received the most rainfall since 1915 (Johnson, 2017). By early February of 2017, mudslides blocked more than twelve locations of Highway 1, resulting in a closure for more than 14 months because of Paul's slide, the failure of Pfeiffer Canyon Bridge, and the Mud Creek slide (Caltrans n.d.). Paul's slide contained around 1.5 million cubic meters of debris, leading to the road closure north of the Nacimiento-Ferguson Road (Caltrans n.d.). Pfeiffer Canyon Bridge was damaged beyond repair, so the road was closed for 8 months to construct a bridge that cost $24 million (Wright, 2017). The 2017 Mud Creek slide was the largest landslide in Highway 1 history that caused a 48-km Highway 1 closure (Caltrans n.d.). The Mud Creek slide brought around 5 million cubic meters of debris blocking Highway 1, and the repair took over one year and cost $ 54 million (Reyes, 2018).

The California coast received more than 300 mm of rainfall from an atmospheric river from January 26 to 28, 2021. This event came shortly after the Dolan wildfire, impacting a portion of Big Sur. The Dolan wildfire burned for four months, from August to December 2020, affecting more than 500 square kilometers in the Big Sur region (USDA Forest Service 2020). More than 90% of the area within the fire perimeter was considered burned, and 55% of the total area within the fire perimeter had moderate to high burn severity (Burned Area Emergency Response 2020). It is well known that a burned area is vulnerable to post-fire debris flows within two to three years after the fire because the affected area has a lower storage capacity, higher flood potential, rising surface water flow, and higher erosion rate due to vegetation removal during the wildfire (Rengers et al., 2020; USGS, 2021).

The 2021 precipitation washed rock, mud, and trees downslope from the 2020 Dolan wildfire burned the area, and triggered debris flows in the Big Sur region that had experienced several wildfire and drought events the previous summer (Canon, 2021). During the January 2021 storm event, Mill Creek, Big Creek, and Rat Creek were identified as debris flow sites by Li et al. (2022). The debris flow did not result

in the damage and closure of Highway 1 at the Mill Creek and Big Creek drainage crossing because both Mill Creek and Big Creek drainage crossing of Highway 1 are engineered with bridges that allow the evacuation of debris flows with minimum impact on the system. Mill Creek Bridge was built in 1993 to carry Highway 1 over water (BRIDGEREPORTS.COM, access 2023). Big Creek Bridge is a 180-m concrete arch bridge opened for traffic in 1938 (BRIDGEREPORTS.COM, access 2023). The Rat Creek drainage crossing of Highway 1 is one of the largest tributary areas that does not have a bridge, and thus, it is susceptible to significant surface water flow and debris flow (Zekkos and Stark, 2023). During the January 2021 storm, surface water transported burned tree trunks and debris from the 2020 Dolan wildfire area and breached a natural debris dam, resulting in a high-velocity debris flow scouring Rat Creek channel. The debris flow damaged the riser pipe connected to the main culvert passing under Highway 1. As a result, the additional water and sediment overtopped the roadway and caused the erosion-induced failure that thoroughly washed out the Highway 1 section, consisting of two lanes. The failure of Highway 1 at Rat Creek was investigated extensively by the Embankments, Dams, and Slopes Technical Committee of the Technical Coordination Council of the Geo-Institute of ASCE (Zekkos and Stark, 2023). The highway was closed for rebuilding for three months, and the embankment was reconstructed with off-site fill material, costing $11.5 million (Reynolds, 2021).

Although Rat Creek was indeed the most impacted location along the highway and resulted in the complete closure, there were several locations where damage was observed and repairs were needed. Thus, system-level response and assessment are essential, not just for the highway section with the most severe damage but for the entire roadway system. In this study, information on the performance of several locations along the entire highway system is leveraged. Regular infrastructure condition assessment is necessary to enable the maintenance and repair work at the early damage stage with reduced loss. Conventional visual inspection has been the most common method in structural damage detection (Teughels et al., 2004). However, visual inspection and non-destructive structure testing are time-consuming and labor-intensive, especially on linear distributed infrastructure extending for multiple kilometers. Such approaches can only be applied locally and have accessibility limitations.

Ground instability, such as debris flow and landslide monitoring, has gained more attention in developing prevention measures. Heavy rainfall, the rapid rise of pore water pressure, and increasing water levels are significant factors triggering debris flow. Although rain gauges, soil moisture, and pore-water pressure sensors can be used to monitor, identify, and predict the occurrence of debris flow (Angeli et al., 2020), these field sensors are resource-intensive and can only provide data on a small and local scale. Field observations are the most common monitoring method for distributed highway systems (Chae et al., 2017) and require inspection crews to drive along the highway and make frequent stops to assess the highway condition. This is inefficient and prone to missing out on ground instability that may not be immediately visible from the highway corridor but may have short-term or long-term impacts on

the highway. Remote sensing data such as satellite imagery can provide geospatial information over a large area at a short time interval that can assist in regular infrastructure monitoring and provide early warning of ongoing problematic regions. There are numerous examples of remote sensing data and techniques implemented to inspect highways and transportation infrastructure. Feng et al. (2021) quantified the environmental impacts of highway construction using remote sensing data, including vegetation index, moisture index, and land surface temperature. Duffell and Rudrum (2005) tried to identify highway earthwork and detect slope instability utilizing satellite, aerial photography, airborne laser scanning, thermal scanning, and videography. Vaghefi et al. (2012) evaluated potential remote sensing techniques such as LiDAR and InSAR for assessing the bridge deck and superstructure condition. Omar and Nehdi (2017) utilized unmanned aerial vehicle (UAV) infrared thermography to detect subsurface delaminations in concrete bridge decks. Shaghlil and Khalafallah (2018) investigated the use of UAV in assessing highway maintenance needs. Ozden et al. (2016) assessed pavement deformation and deformation velocities using synthetic aperture radar (SAR) satellite data. Hoppe et al. (2016) used interferometric synthetic aperture radar (InSAR) data to detect sinkholes and monitor slope stability, bridges, and pavement. However, limited studies have used remote sensing data, especially satellite imagery, to monitor and assess the damage severity of highways caused by debris flow, landslides, and sediment mobility.

Highway 1 crosses through several drainage basins with different characteristics, such as size, steepness, topography, geology, and landslide or debris flow susceptibility. The characteristics of these drainage basins can reflect Highway 1's performance during natural hazard events and can be monitored using remote sensing. This study leverages the data on Highway 1 performance during the January 2021 atmospheric rivers to develop a methodology to detect damages along the highway. The methodology involves using satellite-based multispectral imagery and the thresholding of a remote sensing index to indicate damage along Highway 1. The methodology can be used to monitor distributed infrastructure such as Highway 1 and provide expedient damage assessment along the highway that can help inspection crews prioritize inspection efforts along the most affected areas. The quantified failure threshold value can be used to monitor the condition of Highway 1 autonomously and support maintenance so that the authority can focus on the most critical drainage basins compared to the less susceptible ones. The system can issue early warning of infrastructure damage when the thresholding of a remote sensing index is reached. Furthermore, the monitoring system can immediately reveal the critical damaged area whenever a land surface hazard happens. So, authorities can use the system as a preliminary assessment, confirm on the ground, and prioritize their mitigation plan.

5.2 Methodology

5.2.1 Basin Analysis

Performance analyses using remote sensing are conducted along a stretch of 30 km of Highway 1 from Half Moon Bay to San Luis Obispo. In addition, forty-four main drainage basins were inspected by CalTrans along Highway 1 immediately after the January 2021 storm event. Additional drainage basins exist along that stretch but were too small or had no observable damage after the January 2021 storm. For the analyses, a 3.4-meter (1/9 arc-second) digital elevation model (DEM) from the United States Geological Survey (USGS) is used to measure the size and characteristics of the basins that Highway 1 intersects using Arc GIS Pro 2.9.0.

5.2.2 Observed Damage and Ranking
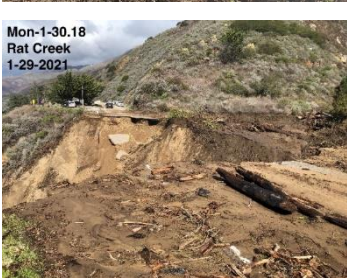
The observed damage along the Highway varied from minimal damage and sediment deposition to landslides, rockfalls, and, worst case, complete loss of the Highway embankment at Rat Creek. A damage ranking system that includes five classes is developed based on storm impact descriptions and photos taken by CalTrans. Table 5.1 displays example photographs along Highway 1 and their corresponding class. Class 0 represents no impact on the inspected drainage crossing and the surface water drainage system, typically consisting of one or more culverts. Class 1 is indicative of a minor impact on the drainage crossing. In this case, debris or sediment was observed, and the culvert may have been partially plugged, requiring modest cleanup work to restore full functionality. Class 2 represents moderate damage where portions of the roadway were obstructed by debris, requiring cleanup to eliminate any obstacles that may result in a slowdown of traffic flow. Class 3 represents moderate to severe damage, where post-storm damage partially closes the highway. This may include excessive sediment on roadways, rocks, flooding, or other obstacles that do not allow for any traffic along one of the highway lanes. Class 4 represents severe damage where the highway is entirely closed and requires repair to restore operation.

5.2.3 Remote Sensing-based Monitoring

In this study, Planet© satellite imagery is used for the remote sensing-based assessment of Highway 1 condition. Planet© provides daily high-resolution satellite imagery with four bands and a resolution of 3 m. The daily frequency of imagery, which is more frequent than other sources, facilitates the continuous monitoring of the highway and is also more likely to provide <24 hours of imagery following a major land surface hazard. Worldview© imagery is also used. Although WorldView© satellites offer higher resolution imagery than Planet© sensors, there is no good quality WorldView imagery near the January 2021 storm event, and the first images of the damage that are available from WorldView© satellites after the January 2021 event are 1.84 m resolution WorldView-2 imagery with eight multispectral bands in April 2021 and 1.24 m resolution WorldView-3 imagery with eight multispectral bands in May 2021.

Vegetation indices estimated from remote sensing data can be quite effective for quantitative and qualitative evaluations of vegetation cover and growth dynamics (Xue and Su, 2017). In this study, one of the most common indices, the normalized difference vegetation index (NDVI), is calculated using the red and near-infrared (NIR) band from Planet©, as shown in Equation 5.1. The value of NDVI ranges between -1.0 and +1.0. Low NDVI values (from -1.0 to 0.1) represent barren rock. Sparse vegetation has moderate NDVI values, around 0.2 to 0.5. Dense vegetation has high NDVI values, approximately 0.6 to 0.9 (USGS Remote Sensing Phenology, n.d.).

Table 5.1: Damage ranking class description and example photos

Class	Description	Photos taken by CalTrans after the January 2021 storm
0	No damage	
1	Minor. Minor cleanup work is needed along the sides; the highway remains operational.	
2	Moderate. Some cleanup work is needed; the road is partially obstructed but still operational.	
3	Moderate to severe. The road is partially closed or expensive to repair due to significant damages to related infrastructure, such as the surface water drainage system.	
4	Severe. Complete damage and road closure; significant work and time needed to repair and operate the highway	

$$\text{Normalized Difference Vegetation Index (NDVI)} = \frac{\text{NIR}-\text{Red}}{\text{NIR}+\text{Red}} \qquad (5.1)$$

Calculating the NDVI difference (dNDVI), as shown in Equation 5.2, allows for detecting change, most often vegetation change. dNDVI is calculated by subtracting pre-event NDVI from post-event NDVI to quantify the influence of an event, in this case, the January 2021 storm on drainage crossings along Highway 1. A negative dNDVI represents vegetation loss or exposure of rock, as well as sediment mobilized on the highway sides. Positive dNDVI represents vegetation gain. Remote sensing data can reveal the evidence of significant change on a large scale. This study focuses on Highway 1 damage caused by the January 2021 storm. So, dNDVI within a 30-m buffering distance from Highway 1 is used to capture major changes that can eventually damage Highway 1.

$$\text{dNDVI} = \text{NDVI}_{\text{post-event}} - \text{NDVI}_{\text{pre-event}} \qquad (5.2)$$

5.3   Result and Discussion

5.3.1   Relationship between Damage Severity and Basin Characteristic

Figure 5.1 illustrates the relationship between the 2D drainage basin area and the damage ranking caused by the January 2021 event. Overall, larger basin areas are found to have a higher damage ranking. The exceptions are the Mill Creek (Basin 35) and McWay (Basin 1) Creek locations, which are discussed subsequently.

For damage, Class 0 (no damage), 1 (minor damage), or 2 (moderate damage), the area of the drainage basins ranges between 0 and 1 km^2. Drainage basins assigned to Class 3 (moderate to severe damage) have drainage basin areas around 0.01 to 0.75 km^2. Drainage crossings assigned to Class 4 (severe damage) have drainage basin areas near or above 1 km^2, as shown in Figure 5.1.

There are three basins assigned to Class 4. Rat Creek (Basin 11) is the third largest drainage basin among all forty-four drainage basins, as displayed in Figure 5.2. It is also the largest drainage basin in Class 4, with a basin area of 2.17 km^2. Although WorldView© provides multispectral satellite imagery with a resolution of 1.84 m and 1.24 m, the available good quality pre-event and post-event WorldView© imagery for the January 2021 storm is only on May 2020, April 2021 and May 2021, as shown in Figure 5.3. Figure 5.3 shows the new embankment at Rat Creek crossing with off-site fill four months after the storm. The other two Class 4 basins are Basin 31 and Basin 32, with drainage areas of 1.37 and 0.81 km^2, respectively. The trace of debris flow in Basin 32 caused by the January 2021 storm can still be observed four months after the rainfall, as shown in Figure 5.4.

Although Mill Creek (Basin 35) has a basin area of 7.06 km^2, about three times the area of Rat Creek, Mill Creek had only Class 1 damage during the January 2021 event (Figure 5.1). The reason is that Mill Creek (Basin 35) has a bridge carrying Highway 1 over water, so the drainage crossing of Highway 1 at Mill Creek does not rely on culverts to convey surface water below the embankment. Thus, although debris flow happened at Mill Creek in the January 2021 event, the water and debris can pass through Highway 1 under the bridge deck without interfering with the roadway. Figure 5.5 is the WorldView© Imagery at Mill Creek in May 2020 and April 2021 that presents the vegetation loss along the channel at Mill Creek caused by the debris flow during the January 2021 storm.
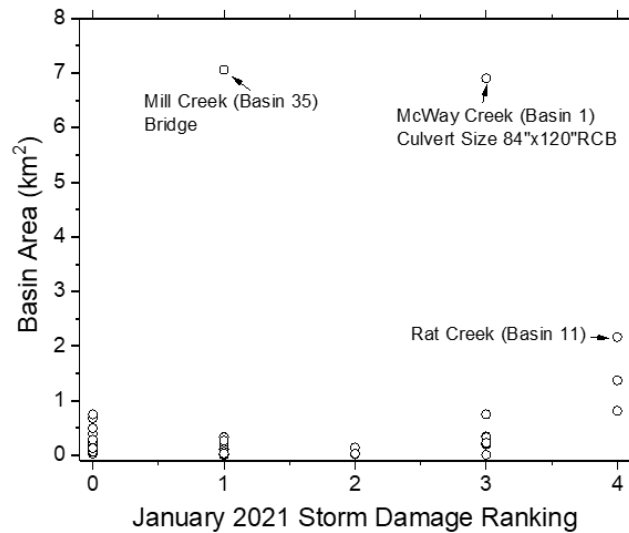
Figure 5.1: Basin area versus damage ranking of the basin during the January 2021 storm
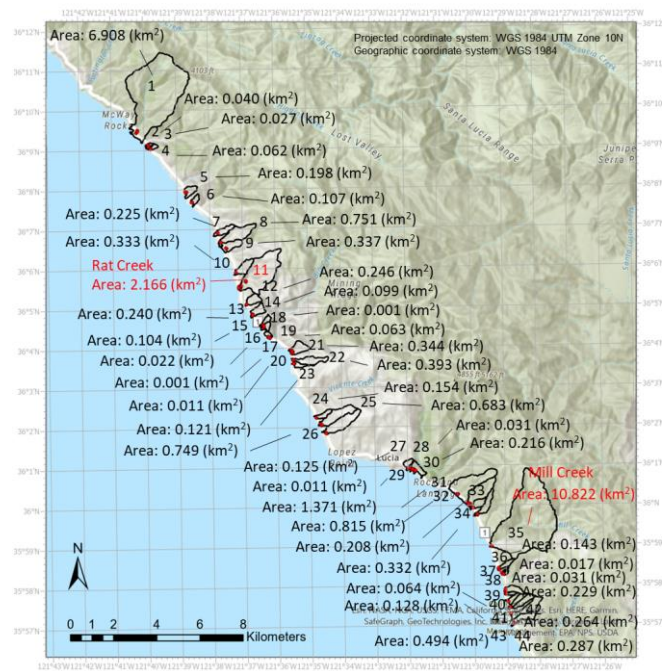
Figure 5.2: Drainage basins inspected by CalTrans along Highway 1 after the January 2021 storm

Figure 5.3: WorldView© Imagery at Rat Creek on May 2020 and May 2021 (WorldView© 2021, DigitalGlobe, Inc., a Maxar company). The imagery from May 2021 illustrates the new embankment that was placed following the washout of the previous event. The debris flow in that location is not easily discernible along its length as it is obstructed by forest vegetation
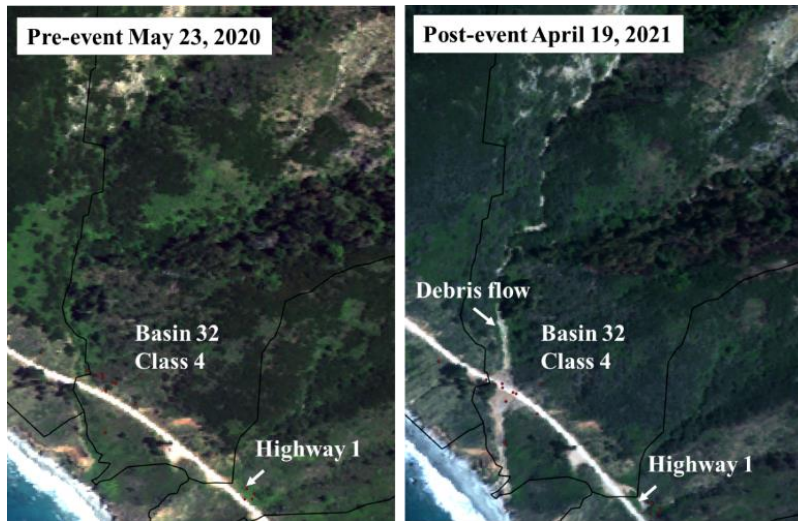
Figure 5.4: WorldView© Imagery (WorldView© 2021, DigitalGlobe, Inc., a Maxar company) at Basin 32 on May 2020 and April 2021 illustrates the occurrence of debris flow

Figure 5.5: WorldView© Imagery (WorldView© 2021, DigitalGlobe, Inc., a Maxar company) at Mill Creek on May 2020 and April 2021 illustrating the occurrence of a major debris flow event that passed below the bridge and thus did not affect the highway performance

Figure 5.6 correlates damage to the ratio of total basin area to the culvert nominal cross-sectional area for the forty-four drainage basins inspected by Caltrans. Most basins have a culvert size between 60.96 cm (24 inches) and 152.4 cm (60 inches). Locations with lower ratios (i.e., high flow capacity culverts compared to the area that they are draining) have lower observed damage. The most significant damage

is at Rat Creek, which also had the largest ratio of basin area to culvert cross-sectional area. Figure 5.6 shows the pre-1985 and post-1985 geometric configurations at Rat Creek. The Rat Creek drainage crossing was built in 1930 with a 167.6 cm (66 inch) corrugated metal pipe (CMP) culvert that passed through the Highway 1 embankment and transported water into the Pacific Ocean (Zekkos and Stark, 2023). In 1985, following the Gorda-Rat fire, a post-wildfire debris flow buried the 167.6 cm (66 inches) diameter CMP culvert. Hence, Caltrans installed an 88.9 cm (35 inch) diameter vertical riser pipe to carry water from the ground surface to the buried 167 cm diameter CMP culvert, reducing the total water flow capacity of the drainage system. The ratio of drainage basin area and culvert cross-sectional area in Figure 5.6 shows that the Rat Creek drainage crossing had become significantly more vulnerable to storm and debris flow due to the smaller riser pipe and lower culvert flow capacity after the 1985 Gorda-Rat fire wildfire and represented the weakest point along the entire highway stretch.

Among forty-four basins, McWay Creek (Basin 1) has the largest culvert, a 213.36 cm x 304.8 cm (84 inch x 120 inch) reinforced concrete box culvert (RCB). Although McWay Creek (Basin 1), as shown in Figure 5.2, has a basin area of 6.91 km^2, around three times the basin area of Rat Creek, the culvert size in the McWay Creek drainage crossing is around ten times the vertical riser pipe in Rat Creek. Figure 5.6 shows that Rat Creek's ratio of drainage basin area and culvert cross-sectional area (Basin 11) is 3.5 times that of McWay Creek (Basin 1). As described earlier, a debris flow in the January 2021 storm event damaged the 88.9 cm (35 inches) diameter riser pipe in Rat Creek, causing water accumulation and overtopping of Highway 1. In contrast, the culvert in McWay Creek was only partially plugged during January 2021, so the drainage crossing at McWay Creek suffered moderate damage rather than severe damage in the January 2021 storm event, as shown in Figure 5.6.

Figure 5.7 shows the repair cost estimate of the inspected drainage crossing as reported by Caltrans as a function of damage ranking. Increasing damage ranking is correlated to increasing repair costs. The Rat Creek crossing has Class 4 damage and the highest repair cost, around $11.5 million, among all damaged crossings of Highway 1 in the January 2021 storm (Reynolds, 2021). Basin 21 has the second highest repair cost due to partial embankment failure following culvert plugging, as shown in the photo of Class 3 in Table 5.1. The damage at Rat Creek and Basin 21 highlights how vital it is to maintain the flow capacity of culverts. Regular cleaning and maintenance are required to maintain adequate drainage capacity and highway functionality.

Figure 5.6: Basin area/culvert cross-sectional area versus damage ranking of the basin during the January 2021 storm

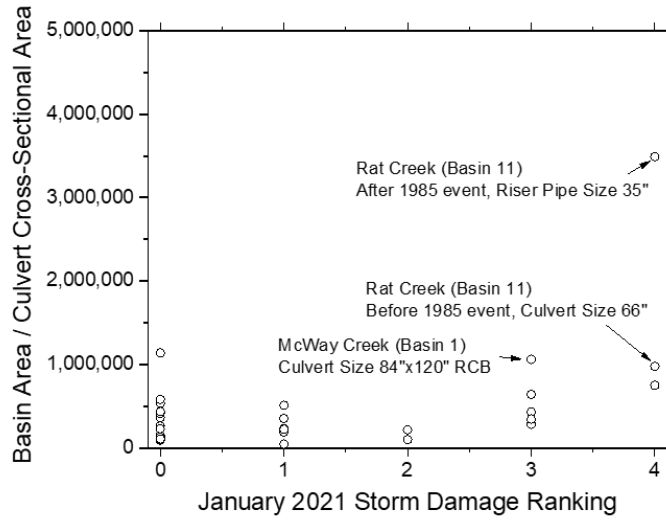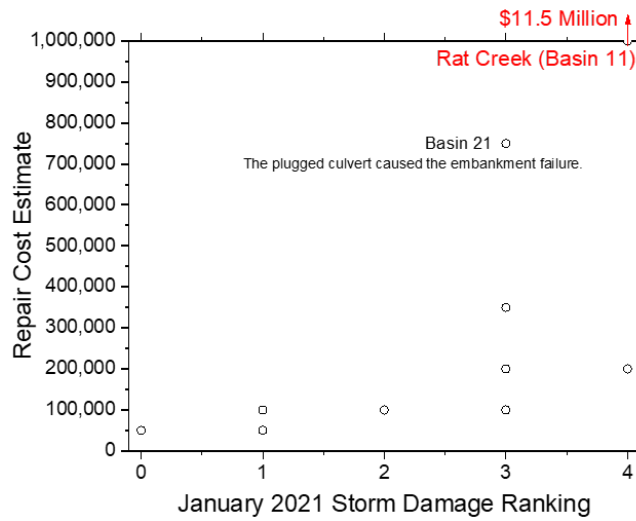Figure 5.7: Repair cost estimate versus damage ranking of the basin during the January 2021 storm

5.3.2 Remote Sensing-based Indicator of Highway Damage

The January 26-28, 2021, storm brought a large amount of precipitation to the California coastline. During the January 2021 storm, debris flow removed vegetation, scoured the channel, plugged the culvert, and blocked Highway 1. Considering the cloud coverage and imagery quality, the Planet© Imagery on January 20, 2021, is used as the pre-event imagery, and the imagery on March 28, 2021, is used as the post-event imagery. Figure 5.8 shows the dNDVI along Highway 1 between January 20, 2021 and March 28, 2021.

 Among forty-four drainage basins, drainage crossings at Rat Creek, Basin 32, and Mill Creek (Basin 35) have the most extensive vegetation loss, negative dNDVI area, along Highway 1, as shown in Figure 5.8a-c. The vegetation loss reveals the traveling path of debris flow along the channel and some sediment mobility along Highway 1 at Rat Creek, Basin 32, and Mill Creek. Basin 7, 2, and 13 are assigned a Ranking 1, 0, and 0, with only minor or no damage. Figure 5.8d-f show little vegetation loss, indicating no sediment and debris, near Highway 1 at Basin 7, 2, and 13.

 Water and debris flow during storm events can wash out vegetation, and sediment can cover vegetation. Hence, vegetation loss is able to reflect the occurrence of debris flow, sediment mobility, and the effect of storms. The forty-four drainage basins inspected by CalTrans during the January 2021 storm have different sizes of area, and each basin crossing intersects with Highway 1 at a different length. A distribution curve can display the occurrence frequency of each value in a dataset. Figure 5.9 is the vegetation loss, negative dNDVI, distribution curve for the forty-four drainage basins. The y-axis in Figure 5.9 shows the value that normalizes the vegetation loss area by the 30-m buffering area of Highway 1 at each inspected drainage basin. The distribution is only demonstrated for negative dNDVI, which is the focus and the expected observation when vegetation is removed or debris is deposited.

 Among forty-four drainage basins, three basins are assigned to Class 4, eight basins are Class 3, three basins are Class 2, thirteen basins are Class 1, and seventeen basins are Class 0. The result shows that Class 4 locations have a minimum dNDVI of -0.33, a maximum dNDVI of -0.01, and a mode dNDVI of -0.05. Class 3 locations have a minimum dNDVI of -0.03 and a mode of dNDVI of -0.01. Class 1 locations have a mode dNDVI of around 0. Class 4 has the lowest minimum dNDVI among the five classes, indicating severe change associated with vegetation loss or debris accumulation. Class 3 and Class 2 have less negative dNDVI. Class 1 has even less area (pixels) with negative dNDVI and higher near zero dNDVI pixels that are indicative of no vegetation change. Among all classes, Class 0 has the least negative dNDVI values.

 The magnitude of vegetation loss, or negative dNDVI, area along Highway 1 can reflect the damage severity of Highway 1. Figure 5.10 is the conceptual negative dNDVI distribution curve, and the dashed lines represent the shape of the distribution curve for each class. Class 4 has the most vegetation loss area and the broadest

distribution of negative dNDVI, shown in Figure 5.10a-f. Class 1 has a mode of dNDVI at zero, displayed in blue in Figure 5.10d. Class 0 indicates that the drainage basin of Highway 1 is not damaged and does not have vegetation loss, which is presented in yellow in Figure 5.10e. As the damage severity decreases, the distribution curve narrows, and the minimum dNDVI value increases, as shown in Figure 5.10a-e. The mode of dNDVI shifts from negative dNDVI to zero when the damage level reduces in Figure 5.10f. As shown in Figure 5.10f, Class 4 has the widest distribution and the smallest minimum negative dNDVI, so it is the easiest to detect severe damage compared to Class 0-3. The result suggests that the shape and minimum value of negative dNDVI distribution curves can reflect and capture the damage severity of Highway 1 and form the basis for an autonomous framework for damage detection along Highway 1, as described subsequently.
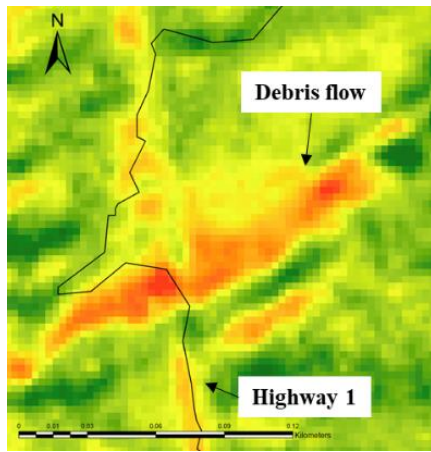
(a) Rat Creek, Basin 11, Ranking 4
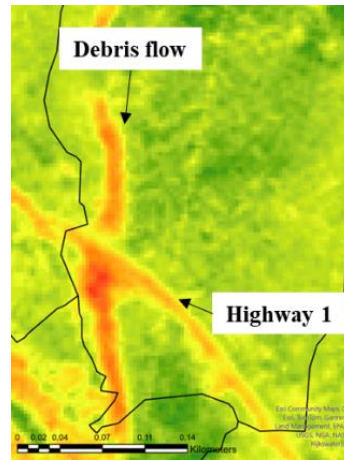
(b) Basin 32, Ranking 4
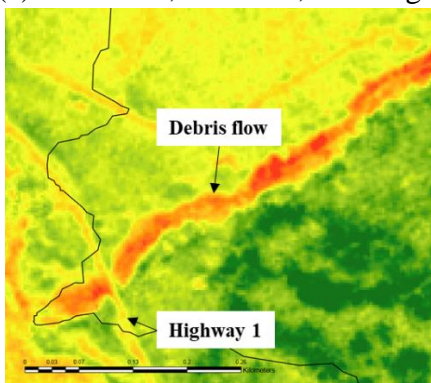
(c) Mill Creek- bridge, Basin 35, Ranking 1
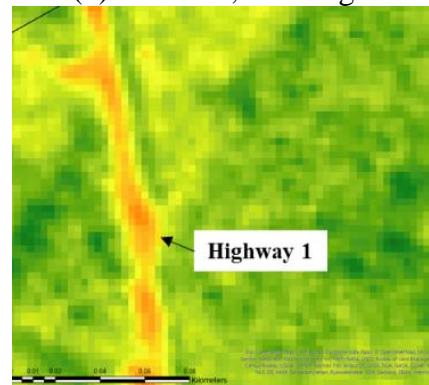
(d) Basin 7, Ranking 1
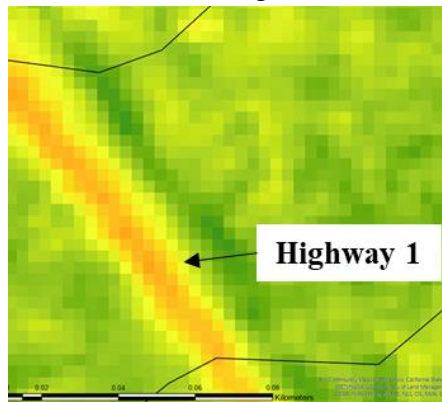
(e) Basin 2, Ranking 0
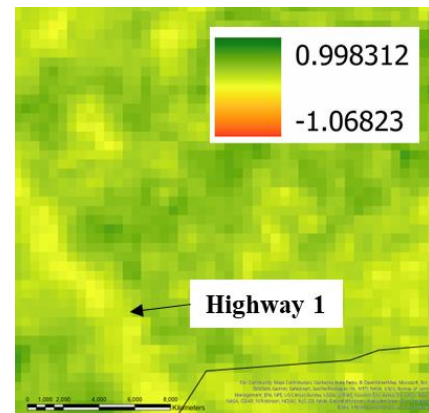
(f) Basin 13, Ranking 0

Figure 5.8: Vegetation change, dNDVI, caused by the January 2021 storm

5.3.3 Proposed Framework of Autonomous Remote Sensing Monitoring System

The change of remote sensing indexes such as NDVI in this study can provide insight into the severity of infrastructure damage. With the availability of daily high-resolution satellite imagery, remote sensing indexes can be developed as an autonomous monitoring system to track infrastructure conditions regularly and remotely. During regular highway operations and following natural hazard events (such as storms and earthquake-induced landslides), the proposed monitoring system can continuously provide a first-order screening of infrastructure conditions on a large scale.

Satellite imagery can provide information over a large area with high temporal frequency. The resolution of imagery governs how detailed infrastructure damage can be detected, and the temporal frequency of imagery affects whether the infrastructure damage caused by a specific event can be captured. Even with a high temporal frequency, the quality of satellite imagery might be affected by cloud coverage and other climate factors. Therefore, the first step in developing an autonomous remote sensing monitoring system is collecting good quality satellite imagery before and after historical damage events.

Secondly, in this study, threshold values for dNDVI are shown based on the experience with the 2021 atmospheric river damage. An algorithm should be trained using imagery from historical damage events to get the thresholding of a remote sensing index to indicate infrastructure damage. When the monitoring system considers more and more damage events, the damage threshold will become more accurate, and the system can be more accurate in detecting critical areas. The incorporation of other data, such as precipitation, ground water level, and soil type, can enhance the accuracy and reliability of the monitoring system. Training (or, in this case, thresholding) using data at regular times or smaller events can enhance the monitoring system's performance since the data at regular times or smaller events are more extensive. Data collection during regular times can help more accurately detect "anomalies" instead of changes due to seasonal variations in the environment. Moreover, while the monitoring system is applied system-wide, regular monitoring can reveal potential spatial differences among different locations.

5.4 Conclusions

Remote sensing methodologies can detect the impact of land surface hazards, such as storms and earthquake-induced landslides, on highways and roadway infrastructure that is inherently distributed. The atmospheric river and associated precipitation event of January 28-30, 2021, highlights that drainage crossings at larger drainage basins are more likely to experience severe damage during storm events, especially if the culvert sizes are smaller or poorly maintained at full flow capacity. Measuring the negative differential NDVI, typically indicative of vegetation loss or debris accumulation, in the vicinity of the highway system can reveal the severity of highway damage. A lower minimum value and a broader distribution of vegetation loss indicate severe damage. Thresholds of dNVDI are identified based on the January 2021 event for different damage classes. An autonomous remote sensing monitoring system for distributed infrastructure can be developed based on the workflow in this study, including data collection and analyses that have quantified failure threshold values calibrated from historical events. The monitoring system can track the distributed infrastructure condition and assess expedient damage on a large scale in regular operation times and immediately after natural hazard events.
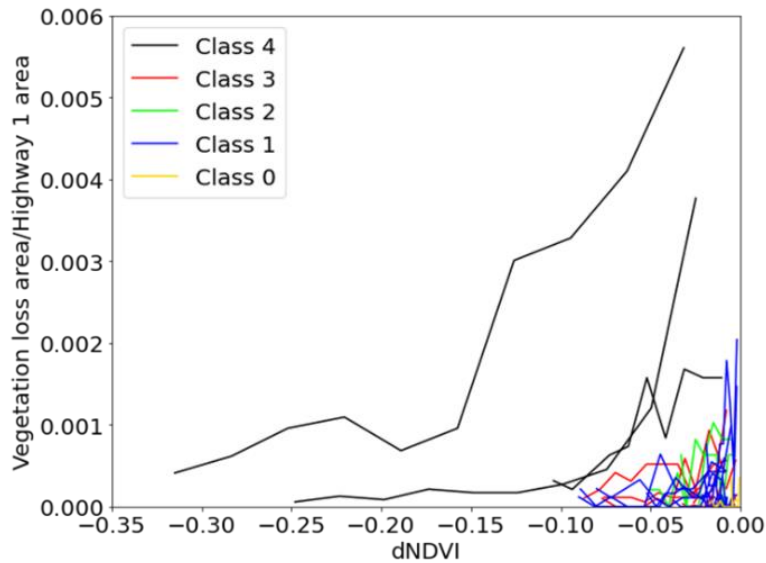
Figure 5.9: dNDVI distribution curves along Highway 1 caused by January 2021 storm

(a) Class 4

(b) Class 3

(c) Class 2

(d) Class 1

(e) Class 0

(f) All Classes
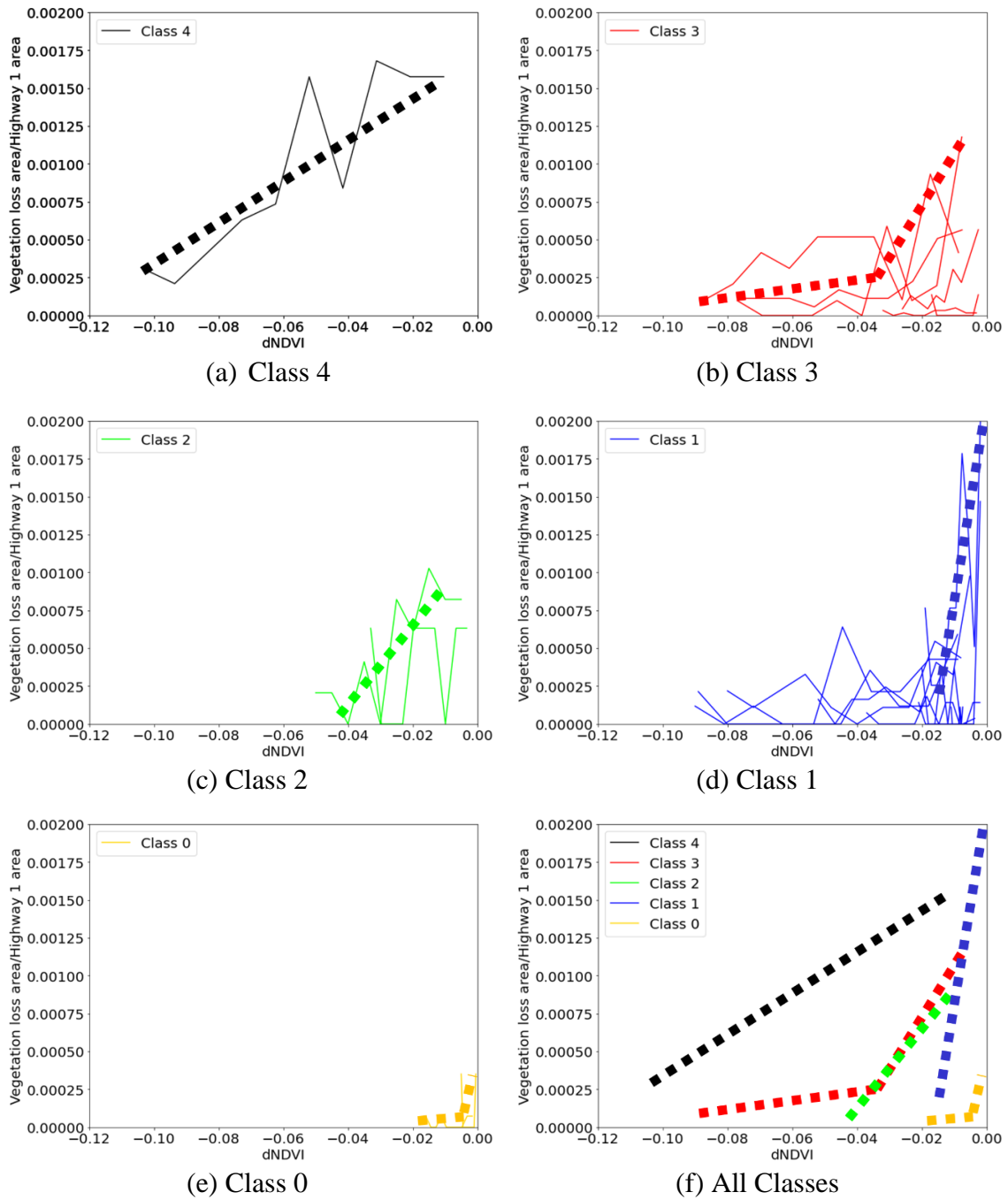
Figure 5.10: Conceptual dNDVI distribution curves

Chapter 6 An Autonomous Methodology for Monitoring Mining Instability Using Remote Sensing

6.1 Introduction

Mining activities produce significant amounts of waste by-products, such as crushed rock, processed water, and chemicals (Kossoff et al., 2014). The waste by-products are usually disposed of and retained by engineered embankments known as tailings dams (Lyu et al., 2019). Billions of tons of mine tailings have been generated by the mining industry globally (Owen et al., 2020), and the decreasing ore grades and increasing consumption of raw materials show that vast amounts of tailings will need to be stored in the future (Bowker and Chambers, 2017).

Tailings dam facilities are vulnerable to failure and require high maintenance costs during the operation and after mine closure (Rico et al., 2008). Tailings dam failures can affect areas several kilometers downstream from the dam and adversely impact the environment. The rate of tailings dam failure was estimated to be approximately 1.2% in the last 100 years, which is a higher failure rate than water retention dams (Azam and Li, 2010; Bowker and Chambers, 2015). Islam and Murakami (2021) pointed out that the location of mining failure cases has shifted from developed to developing countries. The number of tailings dam failure events is rising, and many smaller-size failure cases are under-reported (Villarroel et al., 2006). Even though tailings dam failures are preventable, failures have also reoccurred at the same site due to several factors, including regulation requirements and lack of oversight (Yangon, 2020).

Recent advances in remote sensing techniques and data availability have increased the value of remote sensing technologies in disaster prevention and response (Metternicht et al., 2005). For example, Mura et al. (2018) used DInSAR to monitor

the vulnerability of dams and dikes after the collapse of the Fundao tailings dam. Rudorff et al. (2018) analyzed the collapse of an iron ore mining wastewater dam using Landsat imagery. Rauhala et al. (2017) conducted unmanned aerial vehicle surveillance of a mine tailings impoundment in sub-arctic conditions. Remote sensing techniques have been applied to study mining activities but have yet to become standard practice. Existing research has focused on using remote sensing data to analyze a specific failure event rather than developing a scalable monitoring system for mining failure detection.

This study proposes a fully autonomous approach for monitoring mining failures. Environmental protection agencies, mining boards, and other regulators typically do not have the resources to independently monitor all active and closed facilities. A fully autonomous approach for monitoring mining failures can provide a completely independent approach to detect potential instabilities in operational, closed, and abandoned mines. As shown in Figure 6.1, the multispectral satellite imagery can be collected regularly on a daily, weekly, or monthly basis. Secondly, the remote sensing indexes can be calculated from the multispectral satellite imagery. Thirdly, change detection can be performed on historical failure cases, and the failure threshold value can be derived. Lastly, the monitoring system can monitor the mining sites autonomously. If there is a significant change in remote indexes and the failure threshold value is exceeded, the failure warning will be issued to the authority.

The key to the process is identifying a detection method that can automatically indicate a failure. Thus this study aims to identify the detection indexes that are more likely to successfully detect a failure and quantify the threshold values that can be used to do so. To that end, this study leverages eight recent mining instability failures and tests different satellite-based change detection indexes to assess their performance. Specifically, the following events are considered: the 2022 Jagersfontein tailings dam failure (Motsau and Van Wyk, 2022), the 2022 Pau Branco iron ore mine landslide (Petley, 2022), the 2020 Carmen copper mine landslide (Petley, 2020), the 2020 Singrauli fly ash dam breach (South Asia Network on Dams, Rivers and People, 2020), the 2019 Córrego De Feijão tailings dam failure (Robertson et al., 2019), the 2018 Cadia gold mine tailings dam failure (Petley, 2018), the 2014 Mount Polley mine tailing dam failure (WISE Uranium Project, 2023), and the 2013 Bingham Canyon copper mine landslide (Pankow et al. 2014).
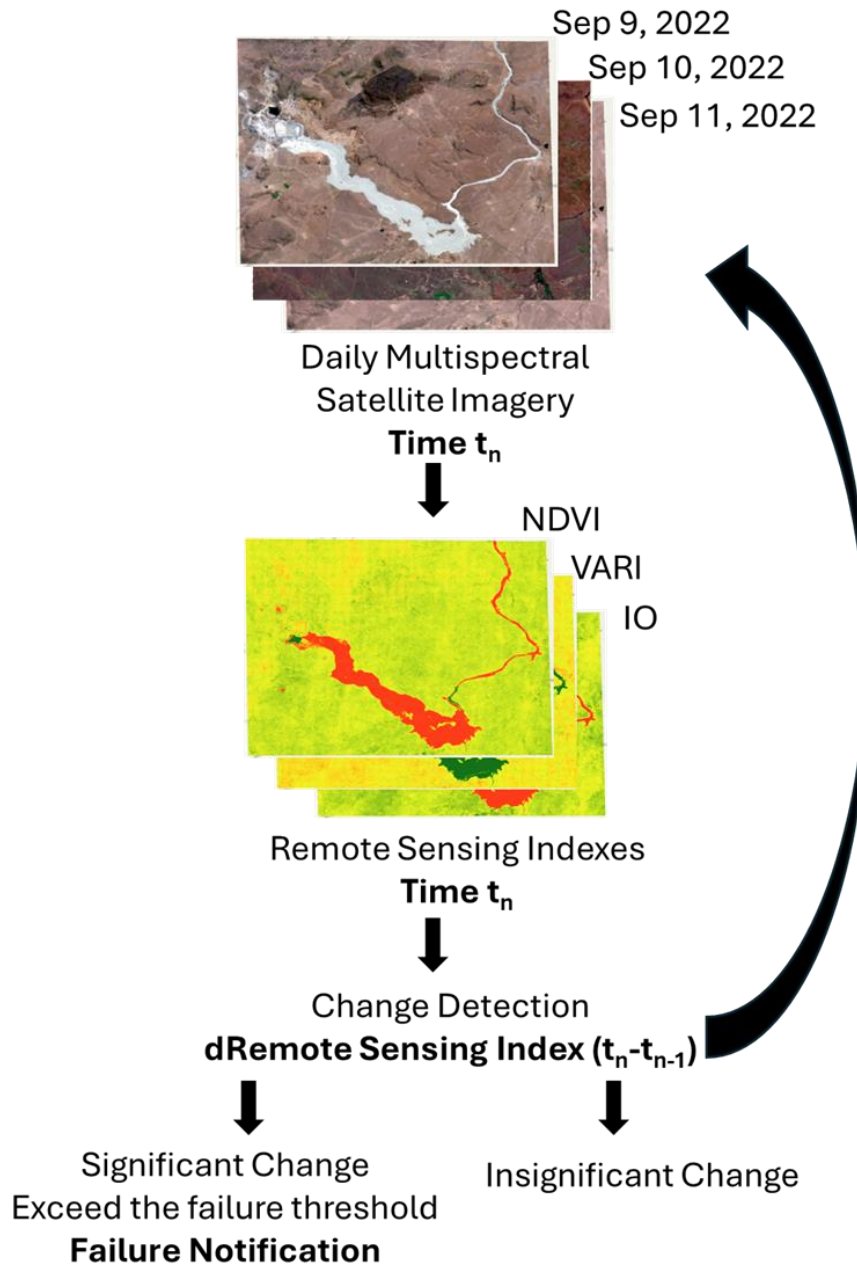
Figure 6.1: Workflow for the mining instability monitoring

6.2 Failure Cases

Table 6.1 lists the failure case histories considered in the analyses. Each case is located in a different place and has various climates. Pau Branco iron ore mine (PB) is in the tropical highland, Jagersfontein tailings dam (JF) is in semi-desert, Carmen Copper mine (CCM) is in tropical wet and dry, Singrauli fly ash dam (SFA) is in subtropical,

Córrego de Feijão tailings dam (CDF) is in tropical, Cadia gold mine tailings dam (CGM) is in temperate oceanic, Mount Polley mine tailing dam (MPM) is in humid continental, and Bingham Canyon Copper mine (BCC) is in dry-summer continental climate. CDF, JF, SFA, CGM, and MPM are tailings dam failures. PB, CCM, and BCC are landslides.

Córrego de Feijão (CDF) mine near Brumadinho, Região in Brazil, had a catastrophic rotational failure and released around 12 million cubic meters of ore tailings into the river on January 25, 2019 (Grebby et al., 2021). Figure 6.2a is a satellite imagery of the Córrego de Feijão mine before the failure. Figure 6.2b shows that the fluidized mass traveled approximately 10 km downstream and reached the river. The fluidized mass led to 260 deaths, destroyed a bridge, contaminated the drinking water, and damaged the habitat. Robertson et al. (2019) mentioned that liquefaction and vertical subsidence occurred before the dam collapsed. The lack of internal drainage, the steep upstream constructed slope, and rainfall contributed to the instability.

The Jagersfontein mine (JF) is the deepest hand-excavated mine in the world (Streeter, 1892). On September 11, 2022, the Jagersfontein tailings dam collapsed. Figure 6.3a shows the Jagersfontein tailings dam before the 2022 failure. Figure 6.3b shows that the slide had a significant runout and affected a large downstream area. The tailings runout was about 1.5 km wide and extended 8.5 km toward the southeast (NASA Earth Observatory, 2022). The Jagersfontein tailings dam had a steep tailings gradient and was built on top of a shallow aquifer with steep retaining walls on a slope, all contributing to the failure (Motsau and Van Wyk, 2022).

The Pau Branco iron ore mine landslide (PB) occurred in Brazil on January 8, 2022. Petley (2022) mentioned that the failure was a rotational landslide that occurred following heavy rainfall. The failure of the uphill slope gave rise to the overtopping of the downstream retention structure (Riskslope, 2022). Figure 6.4a displays the Pau Branco iron ore mine before the 2022 failure event. Figure 6.4b indicates that the landslide affected the road downslope. The landslide is approximately 300 m wide and 870 m long.

On December 21, 2020, a landslide occurred after a tropical storm at the Carmen Copper mine (CCM) in the Philippines. The primary failure was located on the upper part of the slope and collapsed into the lake downslope (Petley, 2020). Figure 6.5a is a satellite image before the 2020 Carmen Copper mine failure event. The landslide is 330 m wide and 800 m long, as shown in Figure 6.5b.

The Singrauli fly ash dam (SFA) was breached in India on April 10, 2020. The massive liquid fly ash affected villages, destroyed crops, and contaminated water reservoirs (South Asia Network on Dams, Rivers, and People, 2020). Figure 6.6a presents the Singrauli fly ash dam before the failure. Figure 6.6b illustrates the long travel distance of the fly ash that exceeded 5 km from the dam after the failure.

On March 9, 2018, the Cadia gold mine tailings dam (CGM) failed. Cadia is one of Australia's largest open copper mines (Petley, 2018). Figure 6.7a is the satellite imagery one day before the failure event. Figure 6.7b shows that the failed section of the tailings dam is 270 m wide and 430 m long. The failure of the Cadia gold mine tailings dam released around 1.3 cubic meters of tailings (WISE Uranium Project,

2023).

The Mount Polley mine (MPM) in Canada failed on August 4, 2014, releasing 7.3 million cubic meters of tailings (WISE Uranium Project, 2023). Figure 6.8a shows the imagery two weeks before the failure event. Figure 6.8b shows that the collapse mass from the Mount Polley mine reached the adjacent lakes.

On April 10, 2013, a massive landslide occurred at the Bingham Canyon copper mine (BCC) in the United States. Around 65 million cubic meters of material was deposited, making the landslide the largest non-volcanic landslide in North America in modern times (Pankow et al. 2014). Figure 6.9a shows the imagery before the failure event. The landslide is 450 m wide and 2.3 km long, as shown in Figure 6.9b.

Maus et al. (2020) generated a global-scale mining site polygons dataset considering all mining activities features derived from visual interpretation of satellite imagery. The size of each mining site in Table 6.1 is estimated according to the mining activities polygons from Maus et al. (2020). The size of the failure source for each site is estimated based on Figure 6.2 to Figure 6.9, and the percentage of failure sources compared to the entire mining site is approximated accordingly to estimate the size of the failure compared to the facility. In addition, the Normalized difference vegetation index (NDVI) is calculated for each site for the whole imagery in Figure 6.10 and within the mining site. NDVI is a vegetation index with values ranging between -1 and 1. The value of NDVI is descriptive of the density of vegetation. Densely vegetated areas tend to have higher NDVI (closer to 1). In contrast, barren and non-vegetated areas have a lower value of NDVI. A negative NDVI indicates water, a positive value near zero indicates bare soil, a positive NDVI around 0.1 and 0.5 indicates sparse vegetation, and an NDVI of 0.6 and above indicates dense vegetation (Pettorelli, 2013).

Figure 6.10 displays the NDVI density curves of eight mining sites from the entire imagery before the failure. BCC, JF, SFA, and CGM are the mining sites with lower NDVI and have a statistical mode of NDVI below 0.5. MPM, CDF, PB, and CCM have a mode of NDVI higher than 0.5, which indicates more vegetated sites.

(a) January 24, 2019 (pre-failure)   (b) February 1, 2019 (post-failure)
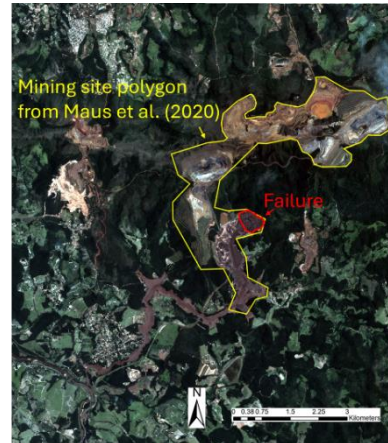
Figure 6.2: 2019 Córrego de Feijão tailings dam failure, Planet© Imagery

(a) September 9, 2022 (pre-failure)   (b) September 12, 2022 (post-failure)
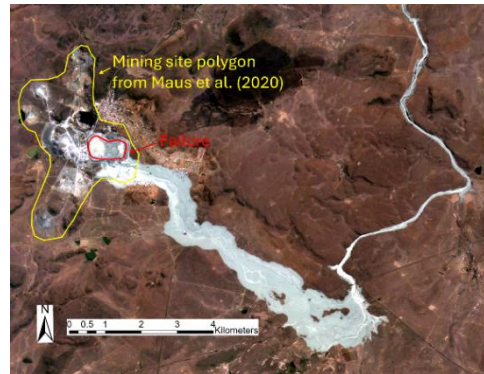
Figure 6.3: 2022 Jagersfontein tailings dam failure, Planet© Imagery
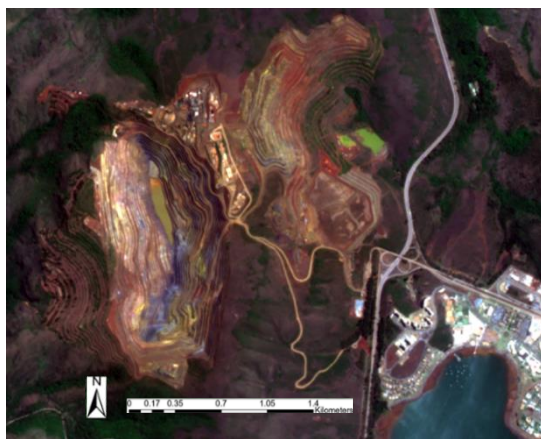
(a) October 03, 2021 (pre-failure)   (b) January 23, 2022 (post-failure)
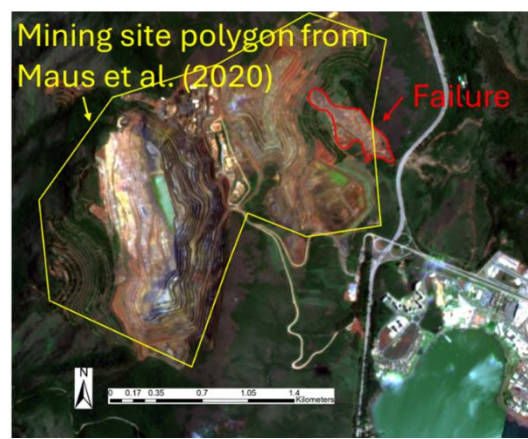
Figure 6.4: 2022 Pau Branco iron ore mine landslide, Planet© Imagery
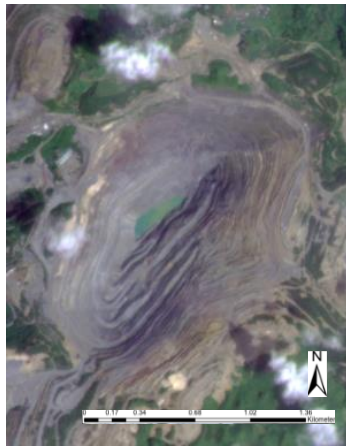
(a) October 14, 2020 (pre-failure)
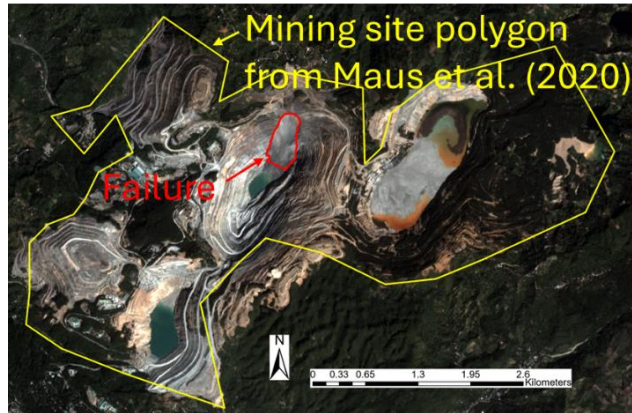
(b) January 26, 2021 (post-failure)

Figure 6.5: 2020 Carmen Copper mine landslide, Planet© Imagery

(a) April 09, 2020 (pre-failure)
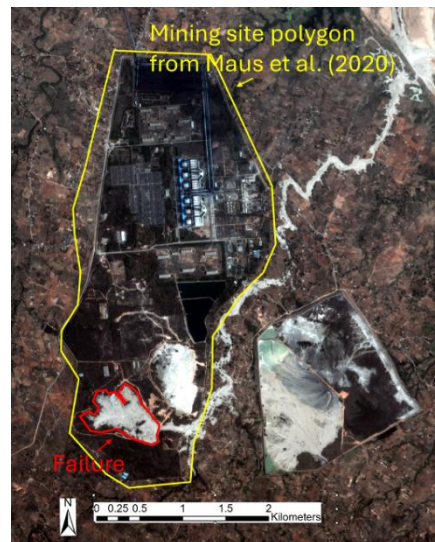
(b) April 13, 2020 (post-failure)

Figure 6.6: 2020 Singrauli fly ash dam breach, Planet© Imagery

(a) March 08, 2018 (pre-failure)   (b) March 10, 2018 (post-failure)
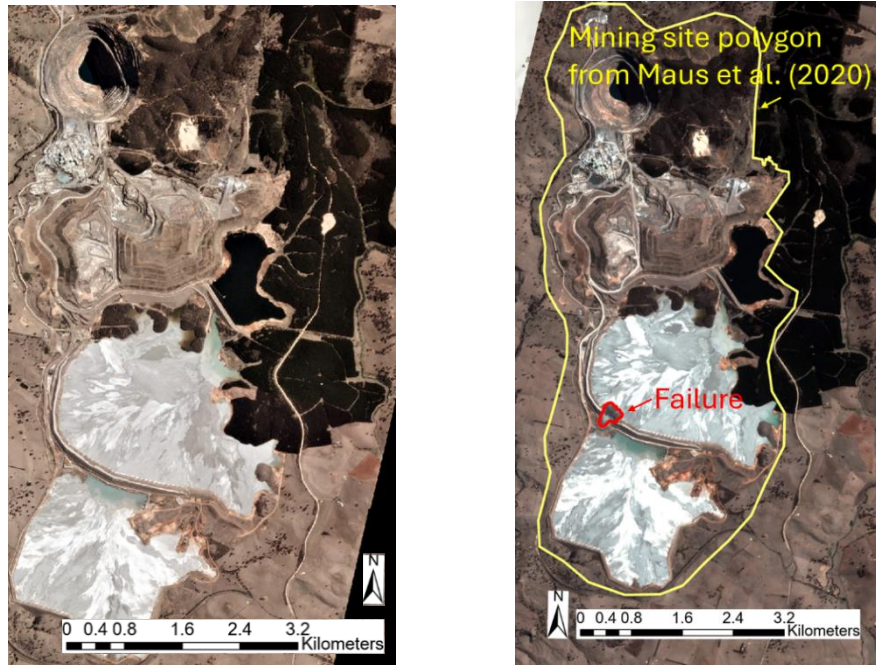
Figure 6.7: The 2018 failure of the Cadia gold mine tailings dam, Planet© Imagery

(a) July 12, 2014 (pre-failure)   (b) August 14, 2014 (post-failure)

Figure 6.8: Mount Polley mine tailings dam failure, Planet© Imagery

(a) October 05, 2012 (pre-failure)   (b) May 23, 2013 (post-failure)
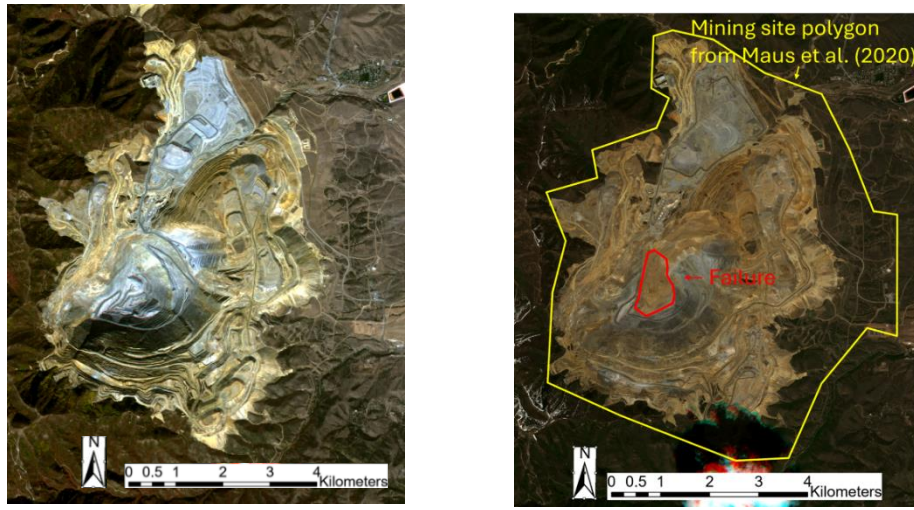
Figure 6.9: 2013 Bingham Canyon copper mine landslide, Planet© Imagery
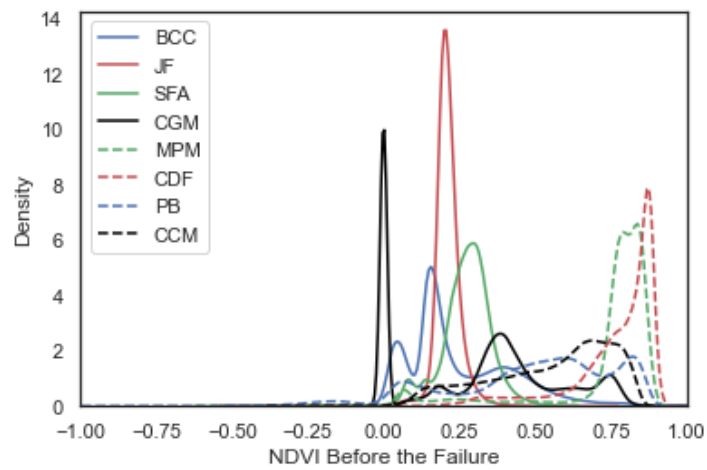
Figure 6.10: NDVI from the entire imagery before the failure event

Table 6.1: The background of mining sites

Mining Site	Year of the Failure Event	Country	Climate	Size of Mining Site (km^2)	Size of Failure Source (km^2)	Failure Source/ Mining Site (%)
Pau Branco iron ore mine (PB)	2022	Brazil	Tropical highland	4.13	0.13	3.05
Jagersfontein tailings dam (JF)	2022	South Africa	Semi-desert	9.15	0.53	5.78
Carmen Copper mine (CCM)	2020	Philippines	Tropical wet and dry	17.67	0.19	1.05
Singrauli fly ash dam (SFA)	2020	India	Subtropical	8.76	0.36	4.07
Córrego de Feijão tailings dam (CDF)	2019	Brazil	Tropical	11.51	0.25	2.20
Cadia gold mine tailings dam (CGM)	2018	Australia	Temperate oceanic	29.75	0.08	0.27
Mount Polley mine tailing dam (MPM)	2014	Canada	Humid continental	18.56	2.58	13.92
Bingham Canyon Copper mine (BCC)	2013	United States	Dry-summer continental climate	42.91	0.68	1.59

6.3 Remote Sensing Data

Satellite imagery is used in the detection analyses. Specifically, Planet© (Planet Team, 2017) imagery is selected to provide daily satellite imagery with global coverage from different sensors. This study uses Planet Scope Scene with a resolution of 3 m, Planet Psorthotile with a resolution of 3.125 m, and Planet Reorthotile with a resolution of 5 m with four bands of blue, green, red, and near-infrared to detect the mining instability.

As shown in Table 6.2, satellite imagery at different time instances is collected considering cloud coverage and imagery quality for the 2022 Jagersfontein tailings dam failure, 2022 Pau Branco iron ore mine landslide, 2020 Carmen copper mine landslide, 2020 Singrauli fly ash dam breach, 2019 Córrego de Feijão tailings dam failure, 2018 Cadia gold mine tailings dam failure, 2014 Mount Polley mine tailing dam failure, and 2013 Bingham Canyon copper mine landslide.

Two images are used before the event to assess variations in the indexes under "normal" operations (as shown in the second column of Table 6.2), whereas an image before and an image after the event is used to compare the changes in indexes associated with the failure (as shown in the third column of Table 6.2).

Table 6.2: Planet© satellite imagery is collected for change detection analysis

	Images before the failure (duration between two images)	Images for the failure event (duration between two images)
2019 Córrego de Feijão tailings dam failure	12/20/2018 and 01/24/2019 (1 month)	01/24/2019 and 02/01/2019 (9 days)
2022 Jagersfontein tailings dam failure	09/10/2021 and 09/12/2021 (2 days)	09/09/2022 and 09/12/2022 (3 days)
2022 Pau Branco iron ore mine landslide	09/30/2020 and 10/03/2021 (12 months)	10/03/2021 and 01/23/2022 (4 months)
2020 Carmen Copper Mine landslide	10/04/2019 and 05/28/2020 (8 months)	10/14/2020 and 01/26/2021 (3 months)
2020 Singrauli fly ash dam breach	04/09/2019 and 04/12/2019 (4 days)	04/09/2020 and 04/12/2020 (4 days)
2018 Cadia gold mine tailings dam failure	03/05/2017 and 03/26/2017 (21 days)	03/07/2018 and 03/10/2018 (3 days)
2014 Mount Polley mine tailing dam failure	07/21/2013 and 09/11/2013 (1.5 months)	07/12/2014 and 08/14/2014 (1 month)
2013 Bingham Canyon copper mine landslide	10/18/2011 and 05/10/2012 (8 months)	10/05/2012 and 05/23/2013 (8 months)

6.4 Change Detection

The change detection model uses eight remote sensing indexes that can be derived from satellite imagery. Indexes often related to vegetation that are used in this study are normalized difference vegetation index (NDVI), green normalized difference vegetation index (Green NDVI), modified soil adjusted vegetation index (MSAVI), visible atmospherically resistant index (VARI), simple ratio (SR), modified triangular vegetation index (MTVI2) and chlorophyll index - green (CIg) are described subsequently and are shown in Equation 6.1-6.7.

NDVI is a very common vegetation index that uses the red and near-infrared (NIR) band, as shown in Equation 6.1. The NIR band can distinguish the high reflectivity of vegetation material, and the red band can show chlorophyll pigment absorption (Pettorelli, 2013). NDVI ranges between -1 and +1, and the value of NDVI is proportional to the amount of vegetation. Green NDVI (Moges et al., 2005) can estimate photosynthetic activity in Equation 6.2 and ranges from -1 to 1. MSAVI (Qi et al. 1994) is a vegetation index that minimizes the effect of bare soil using NIR and red band in Equation 6.3 and ranges between -1 and 1. VARI (Stow et al., 2005) is a vegetation index that utilizes three visible bands to reduce the atmospheric effect, as shown in Equation 6.4. SR (Chen 1996) is a vegetation index that uses the NIR and red band ratio to minimize the atmosphere effect in Equation 6.5. MTVI2 (Haboudane et al., 2004) detects the leaf chlorophyll content in Equation 6.6. CIg (Wu et al., 2012) is a vegetation index to estimate leaf chlorophyll content utilizing NIR and green band in Equation 6.7. Lastly, iron oxide (IO) in Equation 6.8, a geological index, can detect rock experiencing oxidation of iron-bearing sulfides (Dogan, 2009). Change detection is performed in ArcGIS Pro 2.9.0 by calculating the absolute value of the remote sensing index difference between two rasters at different times, as shown in Equation 6.9.

$$\text{Normalized Difference Vegetation Index (NDVI)} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}} \tag{6.1}$$

$$\text{Green Normalized Difference Vegetation Index (Green NDVI)} = \frac{\text{NIR} - \text{Green}}{\text{NIR} + \text{Green}} \tag{6.2}$$

$$\text{Modified Soil Adjusted Vegetation Index (MSAVI)} = \frac{1}{2}\left(2(\text{NIR} + 1) - \sqrt{(2\text{NIR} + 1)^2 - 8(\text{NIR} - \text{Red})}\right) \tag{6.3}$$

$$\text{Visible Atmospherically Resistant Index (VARI)} = \frac{\text{Green} - \text{Red}}{\text{Green} + \text{Red} - \text{Blue}} \tag{6.4}$$

$$\text{Simple Ratio (SR)} = \frac{\text{NIR}}{\text{Red}} \tag{6.5}$$

$$\text{Modified Triangular Vegetation Index (MTVI2)} = \frac{1.5(1.2(\text{NIR} - \text{Green}) - 2.5(\text{Red} - \text{Green}))}{\sqrt{(2\,\text{NIR} + 1)^2 - (6\,\text{NIR} - 5\sqrt{\text{Red}}) - 0.5}} \tag{6.6}$$

$$\text{Chlorophyll Index} - \text{ Green (CIg)} = \frac{\text{NIR}}{\text{Green}} - 1 \tag{6.7}$$

$$\text{Iron Oxide (IO)} = \frac{\text{Red}}{\text{Blue}} \tag{6.8}$$

$$\text{Remote Sensing Index Difference} = |\text{Remote Sensing Index}_{\text{time2}} - \text{Remote Sensing Index}_{\text{time1}}| \tag{6.9}$$

where NIR is the near-infrared band, red is the red band, green is the green band, blue is the blue band, and time 1 is earlier than time 2.

6.5 Result and Discussion

Change detection of remote sensing indexes is conducted to detect mining instability. The change detection of remote sensing indexes can reveal the tailings runout and landslide traveling path, as shown in the example of Figure 6.11 for the 2019 Córrego de Feijão tailings dam failure and in Figure 6.12-6.16 for the other cases.

The Córrego de Feijão tailings dam failed on January 25, 2019. Satellite imagery on December 20, 2018, and January 24, 2019, is used to look for the change detection before the failure. Imageries on January 24, 2019, and February 01, 2019, are utilized to explore the detection of the Córrego de Feijão tailings dam failure incident, as shown in Figure 6.11a-h. Figure 6.11 demonstrates that the change in practically all remote sensing indexes can detect the Córrego de Feijão tailings dam failure. The traveling path of the failure material downstream is particularly visible. Similarly, Figure 6.12, Figure 6.14, and Figure 6.16 illustrate that the change of NDVI and Green NDVI can reveal the traveling path of the fluidized mass caused by the Jagersfontein tailings dam failure, disclose the traveling path of fly ash released from the collapse of Singrauli fly ash dam, and display the path of tailings runout for the Mount Polley mine tailing dam failure. In Figure 6.13 and Figure 6.15, the change of remote sensing indexes can reveal the Pau Branco iron ore mine landslide and disclose the Cadia gold mine tailings dam failure.

(a) Failure event- NDVI

(b) Failure event- Green NDVI

(c) Failure event- MSAVI

(d) Failure event- VARI

(e) Failure event- SR

(f) Failure event- MTVI2

(g) Failure event- CIg
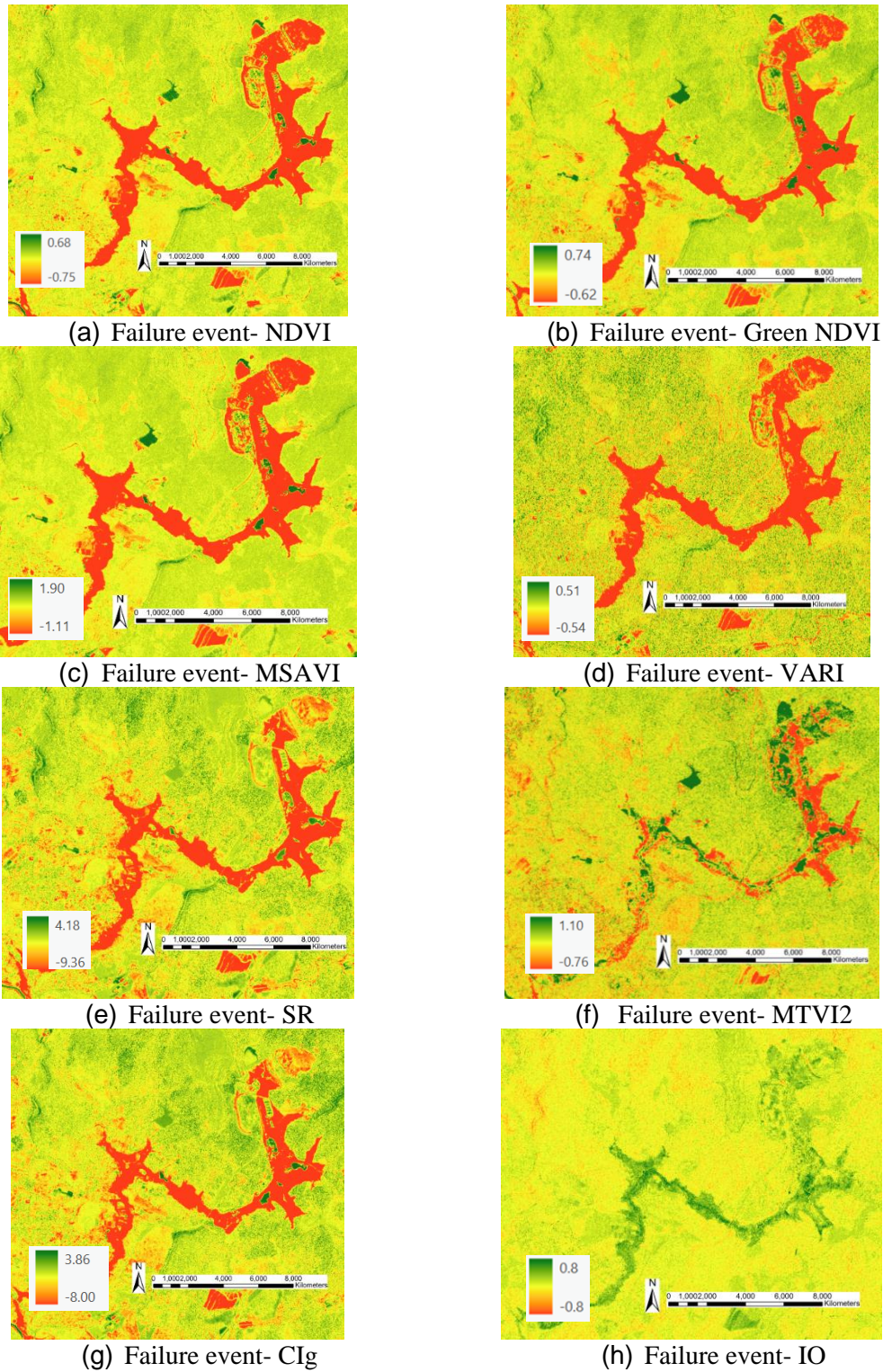
(h) Failure event- IO

Figure 6.11: Change detection of remote sensing indexes- 2019 Córrego de Feijão
tailings dam failure
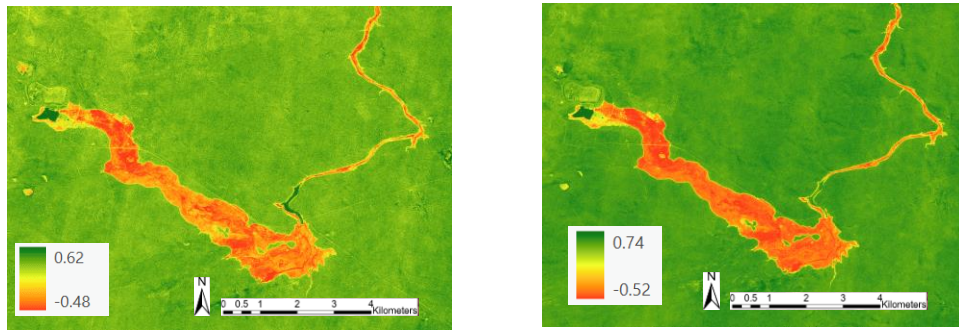
(a) Failure event- NDVI   (b) Failure event- Green NDVI

Figure 6.12: Change detection of remote sensing indexes using imageries on September 9, 2022, and September 12, 2022- 2022 Jagersfontein tailings dam failure
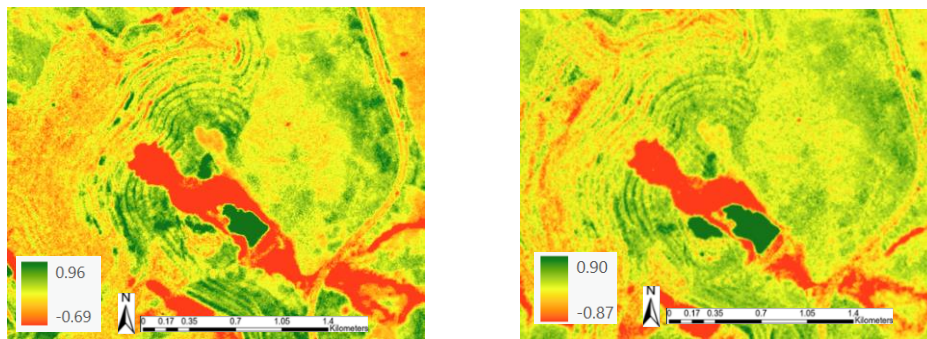
(a) Failure event- NDVI   (b) Failure event- Green NDVI

Figure 6.13: Change detection of remote sensing indexes using imageries October 3, 2021, and January 23, 2022- 2022 Pau Branco iron ore mine landslide
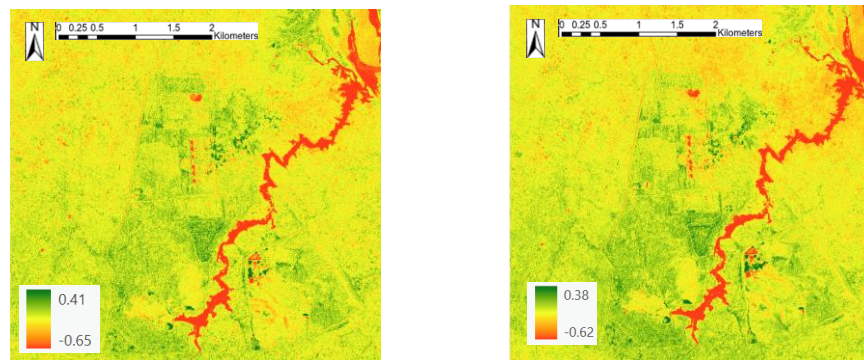
(a) Failure event- NDVI   (b) Failure event- Green NDVI

Figure 6.14: Change detection of remote sensing indexes using imageries on April 9, 2020, and April 14, 2020- 2020 Singrauli fly ash dam breach

(a) Failure event- NDVI

(b) Failure event- Green NDVI
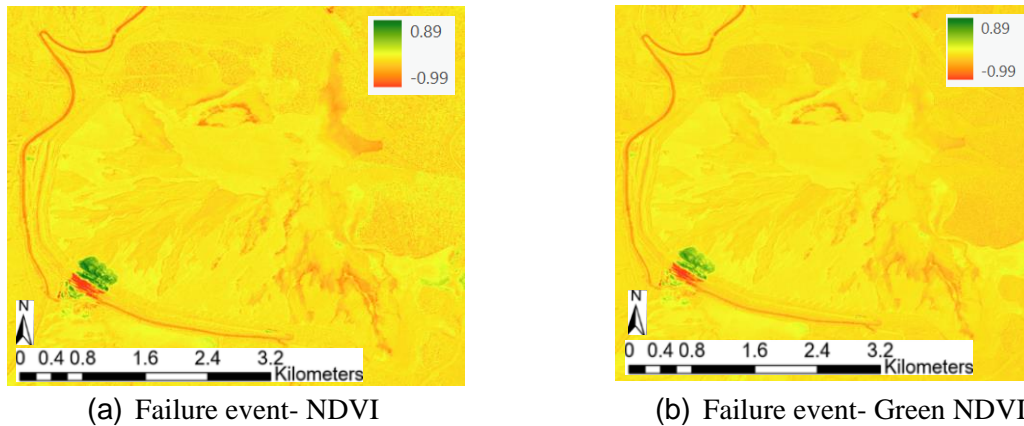
Figure 6.15: Change detection of remote sensing indexes using imageries on March 7, 2018, and March 10, 2018- 2018 Cadia gold mine tailings dam failure

(a) Failure event- NDVI

(b) Failure event- Green NDVI
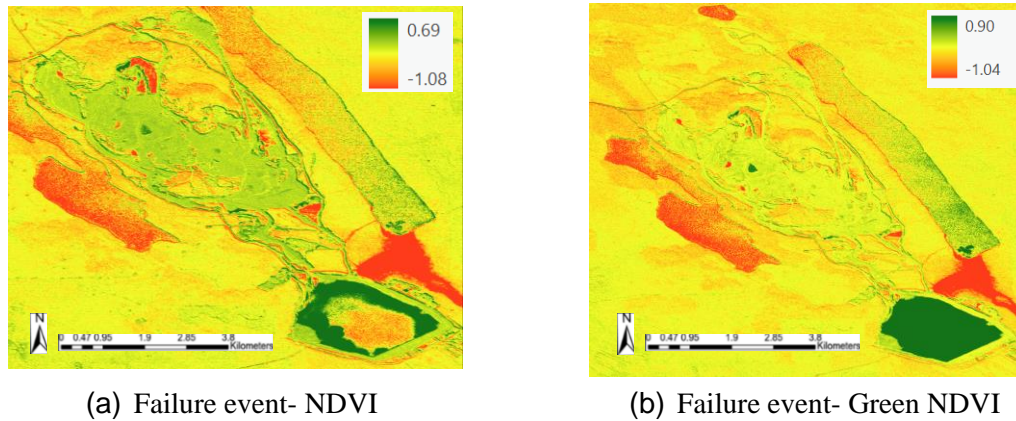
Figure 6.16: Change detection of remote sensing indexes using Imageries on July 12, 2014, and August 14, 2014- 2014 Mount Polley mine tailing dam failure

Figure 6.17 illustrates the remote sensing index density curves for the Córrego de Feijão tailings dam failure event from the entire imagery with the area of 202.2 km^2 as shown in Figure 6.2b, mining site polygon with the area of 11.51 km^2 from Maus et al. (2020), and the failure source with the area of 0.25 km^2 is annotated with a red polygon in Figure 6.2b. Different density curves are shown for different portions of imagery. The solid line is the density curve for the failure source only (as annotated in Figure 6.2), the dashed line is the density curve when the image portion that shows the mining facility is selected, and the dotted line is the density curve of the entire imagery as shown in Figure 6.2b. The density curve before the failure is shown in blue, and after the failure is shown in red.

As shown in Figure 6.17a, in all cases (failure source only, mining facility, and the entire imagery), the density curves translate significantly after the event, indicating that for all scales, a discernible change has occurred. As expected, the differences are highest for the case where change detection is taking place for the failure source only, but the change is also very discernible for the mining facility scale and the entire

imagery scale. The statistical mode of the NDVI difference density curve within the failure source changes from 0.068 before the failure to -0.195 for the failure event, i.e., the NDVI difference has changed by 0.263. On the other hand, the mode of NDVI difference density curve at the mining site scale and entire imagery has a value change of 0.06. Correspondingly, the statistical mean of NDVI difference within the failure source changes from 0.07 before the failure to -0.15 for the failure event and has a value change of 0.22. For the mining site, the mean NDVI difference has a value change of 0.08. Similarly, the statistical median of NDVI difference on the failure source has a change of 0.238 and has a value change of 0.052 on the entire mining site. In general, the density curve from the failure source has a major change between before the failure and failure event compared to the curve from the mining site and entire imagery, especially NDVI, Green NDVI, MSAVI, SR, and CIg as shown in Figure 6.17a-h. The density curves tend to have a normal-type shape with a pronounced mode. Differences in the density curves before and after the failure are observed, and the difference in the remote sensing indexes has a similar change in the statistical mean, median, and mode caused by the failure event. This confirms that at all scales, these indexes could be automatically calculated and used for detection. However, in the detection result, it is important to get a sense of what a significant change in the index is.

(a) NDVI

(b) Green NDVI

(c)MSAVI

(d)VARI

(e) SR

(f) MTVI2

(g) CIg

(h) IO
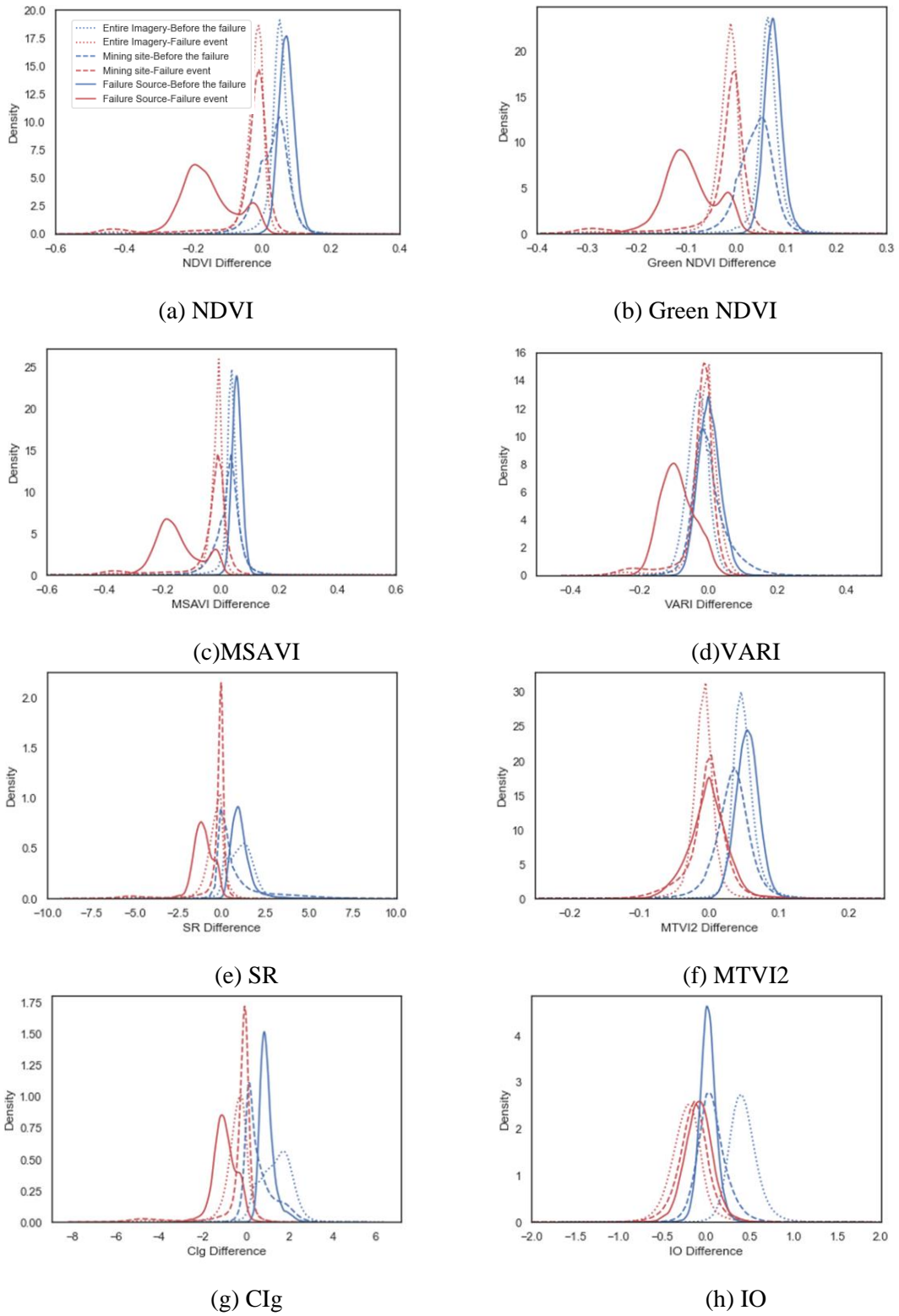
Figure 6.17: Remote sensing indexes difference- 2019 Córrego de Feijão tailings dam failure

Each remote sensing index has a different range of values; thus, normalization is necessary to quantify its performance. The change of statistical mean, median, and mode in remote sensing indexes difference is normalized by the standard deviation of the density curve before the failure to provide a more quantitative assessment of the range of each remote sensing index and the significance of the change that has occurred. To clarify, if the ratio of the change in the mean or median or mode divided by the standard deviation of the curve before failure, as shown in the y-axis in Figure 6.18, is high, that would be indicative of significant change. For example, for the Córrego de Feijão tailings dam failure incident, NDVI, Green NDVI, and MSAVI are the three indexes that have the highest normalized change in mean, median, or mode of the remote sensing index difference and are better predictors of change than the other indexes.
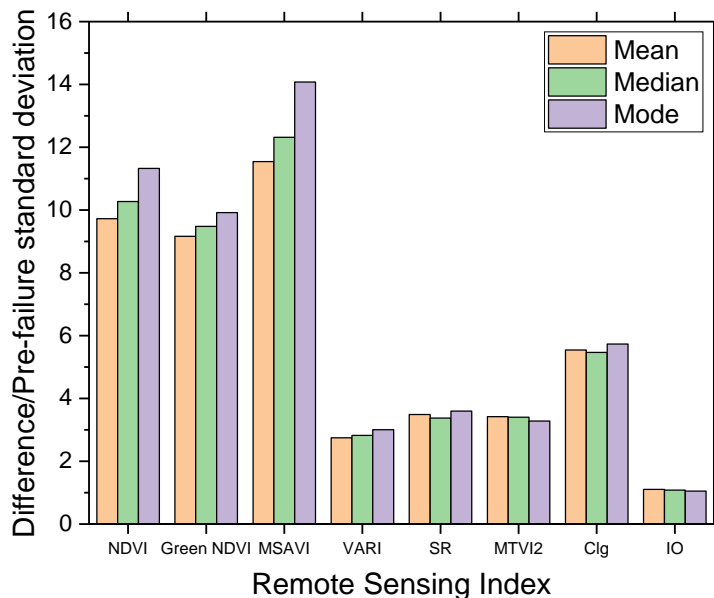
Figure 6.18: Histogram of statistical mode differences/ pre-failure standard deviation of remote sensing indexes within failure source- 2019 Córrego de Feijão tailings dam failure

This study aims to quantify the remote sensing index for autonomous monitoring of mining instability. As described earlier and shown in Figure 6.17, when one considers the failure source only, the most notable remote sensing index change is observed compared to considering the mining site and the entire imagery. However, the location of the failure source is unknown before the failure incident happens, so the analysis within the mining site is of particular interest and focus in the autonomous mining instability monitoring system. Despite the clear differences in the entire density curves, it is easier to focus on just the mode of the density curve. Table 6.3 summarizes the change in the statistical mode of remote sensing index difference on the mining site for each failure event. The top three remote sensing indexes with the highest

difference in the normalized statistical mode for each failure event are highlighted in bold format. It is clear that a number of indexes are sensitive to the observed changes, and some of them are quite common. For example, the MSAVI and VARI appear to be among the top three performing indicators in five out of eight case histories, the IO in four cases, the MTVI2 in three cases, the NDVI, Green NDVI, and CIg in two cases, while SR appears only once as the top performing indicator.

For the 2019 Córrego de Feijão tailings dam failure, the NDVI, Green NDVI, and MSAVI are the top three performing indicators, with the value of normalized change in remote sensing mode as 1.267, 1.952, and 1.410, respectively. MSAVI, CIg, and IO are the best three failure indicators for the 2022 Jagersfontein tailings dam failure and have the value of normalized change in remote sensing mode as 0.592, 0.985, and 1.600. The top three failure indicators for the 2022 Pau Branco iron ore mine landslide are NDVI, MSAVI, and VARI, with the normalized change in remote sensing mode as 1.326, 1.067, and 1.226, respectively. Green NDVI, MTVI2, and IO are the top three remote sensing indicators for the 2020 Carmen copper mine landslide and have the normalized change in remote sensing mode as 1.513, 1.215, and 1.734, respectively. MSAVI, VARI, and IO are the best three failure indicators for the 2020 Singrauli fly ash dam breach and have the normalized change in remote sensing mode as 0.926, 0.961, and 1.855, respectively. For the 2018 Cadia gold mine tailings dam failure, the VARI, MTVI2, and IO are the top three performing indicators, with the value of normalized change in remote sensing mode as 4.383, 1.737, and 6.563, respectively. The top three failure indicators for the 2014 Mount Polley mine tailing dam failure are VARI, SR, and MTVI2, with the normalized change in remote sensing mode as 0.814, 1.680, and 0.687, respectively. MSAVI, VARI, and CIg are the best three failure indicators for the 2013 Bingham Canyon copper mine landslide and have the value of normalized change in remote sensing mode as 1.450, 20.290, and 2.051.

The mean value of NDVI within each of the mining sites before the failure event can quantify the vegetation condition of mining sites. CDF, CGM, CCM, MPM, SFA, PB, JF, and BCC have the mean values of NDVI of 0.52, 0.37, 0.36, 0.36, 0.24, 0.22, 0.15, and 0.11, respectively. Among eight cases, CDF, CGM, CCM, and MPM are the more vegetated sites and have a mean NDVI between 0.36 and 0.52, whereas SFA, PB, JF, and BCC are the less vegetated sites with a mean NDVI between 0.11 and 0.24. MTVI2 and Green NDVI are the top two remote sensing indexes for the four vegetated sites. MSAVI and CIg are the best remote sensing indexes for the four non-vegetated sites. Generally, VARI and IO detect mining instability well for both vegetated and non-vegetated sites. MTVI2 can point out the mining instability of vegetated mining sites the best. MSAVI performs the best in indicating the mining instability in non-vegated sites.

As mentioned, the change in the remote sensing index within the failure source can reveal the mining instability more distinctly than the index value when the entire mining site or entire imagery is considered. Although the size of the failure source for the 2019 Córrego de Feijão tailings dam failure event is only 2.2 percent of the mining site, the normalized remote sensing index mode difference on the failure source is around ten times, nine times, and five times as the value from the mining site for

MSAVI, NDVI, and Green NDVI respectively as shown in Table 6.4. The normalized remote sensing mode difference within the failure source can be more significant than the mining site. However, in some cases, the mining site has a more notable normalized remote sensing mode difference than the failure source, probably due to the construction and other factors, so satellite imagery with higher temporal frequency is required to access the difference caused by the failure event only.

6.6 Conclusions

Mining instability and failures can happen anytime during the operation or several years after mining activities have stopped and may reoccur at the same location. Therefore, mining sites need regular maintenance and monitoring to ensure stability. Daily high-resolution satellite data allows for the efficient and economical monitoring of mining failures in short intervals. Change detection analyses using vegetation and geological indexes for eight recent study cases: the 2022 Jagersfontein tailings dam failure, the 2022 Pau Branco iron ore mine landslide, the 2020 Carmen copper mine landslide, the 2020 Singrauli fly ash dam breach, the 2019 Córrego De Feijão tailings dam failure, the 2018 Cadia gold mine tailings dam failure, the 2014 Mount Polley mine tailing dam failure, and the 2013 Bingham Canyon copper mine landslide are conducted in this study. The result shows that remote sensing indexes can be promising mining instability and failure indicators. Both visible atmospherically resistant index (VARI) and Iron oxide (IO) effectively reveal mining instability in vegetated and non-vegetated sites. The modified triangular vegetation index (MTVI2) is particularly useful in identifying the failure source in vegetated mining sites, while the modified soil-adjusted vegetation index (MSAVI) performs best in indicating mining instability in non-vegetated sites. The change in remote sensing indexes can be applied as a reference for failure and assist the mining industry and supervisory boards in enhancing monitoring, safety management, and planning during and after mining operations.

Table 6.3: Summary of the difference in the mode of remote sensing indexes in the mining site

		NDVI	Green NDVI	MSAVI	VARI	SR	MTVI2	CIg	IO
2019 Córrego de Feijão tailings dam failure	Before the Failure	0.050	0.060	0.035	0.000	0.335	0.035	0.200	0.050
	For the failure event	-0.010	-0.010	-0.030	-0.020	0.050	0.010	-0.200	-0.150
	Difference	0.060	0.070	0.065	0.020	0.285	0.025	0.400	0.200
	Pre-failure standard deviation	0.047	0.036	0.046	0.033	1.528	0.031	0.675	0.173
	Difference/Pre-failure standard deviation	**1.267**	**1.952**	**1.410**	0.601	0.187	0.804	0.593	1.159
2022 Jagersfontein tailings dam failure	Before the Failure	0.000	-0.005	-0.025	0.000	-0.015	-0.015	-0.150	-0.150
	For the failure event	0.010	0.000	0.005	0.010	0.010	-0.010	0.000	-0.015
	Difference	0.010	0.005	0.030	0.010	0.025	0.005	0.150	0.135
	Pre-failure standard deviation	0.028	0.038	0.051	0.039	0.084	0.081	0.152	0.084
	Difference/Pre-failure standard deviation	0.364	0.132	**0.592**	0.253	0.296	0.062	**0.985**	**1.600**
2022 Pau Branco iron ore mine landslide	Before the Failure	-0.020	-0.020	0.000	0.015	0.050	-0.050	-0.100	-0.150
	For the failure event	0.090	0.040	0.200	0.090	-0.100	-0.100	0.150	-0.100
	Difference	0.110	0.060	0.200	0.075	0.150	0.050	0.250	0.050
	Pre-failure standard deviation	0.083	0.077	0.187	0.061	0.382	0.127	0.416	0.342
	Difference/Pre-failure standard deviation	**1.326**	0.775	**1.067**	**1.226**	0.392	0.393	0.601	0.146
2020 Carmen Copper mine landslide	Before the Failure	0.025	0.055	0.005	-0.060	0.100	0.085	0.050	0.135
	For the failure event	-0.040	-0.030	-0.045	-0.025	-0.300	-0.010	-0.050	0.000
	Difference	0.065	0.085	0.050	0.035	0.400	0.095	0.100	0.135
	Pre-failure standard deviation	0.065	0.056	0.090	0.046	0.344	0.078	0.255	0.078
	Difference/Pre-failure standard deviation	1.000	**1.513**	0.553	0.757	1.163	**1.215**	0.392	**1.734**
2020 Singrauli fly ash dam breach	Before the Failure	0.015	0.000	-0.005	0.005	0.050	-0.020	0.005	0.015
	For the failure event	0.020	-0.005	0.050	0.030	0.050	-0.005	0.010	-0.075
	Difference	0.005	0.005	0.055	0.025	0.000	0.015	0.005	0.090
	Pre-failure standard deviation	0.036	0.039	0.059	0.026	0.106	0.068	0.120	0.049
	Difference/Pre-failure standard deviation	0.138	0.128	**0.926**	**0.961**	0.000	0.222	0.042	**1.855**
2018 Cadia gold mine tailings dam failure	Before the Failure	0.080	0.055	0.080	-0.250	-4.850	0.030	-5.150	2.615
	For the failure event	0.000	-0.020	-0.035	0.030	-0.150	-0.065	-0.250	-0.040
	Difference	0.080	0.075	0.115	0.280	4.700	0.095	4.900	2.655
	Pre-failure standard deviation	0.088	0.071	0.095	0.064	5.176	0.055	4.865	0.405
	Difference/Pre-failure standard deviation	0.910	1.051	1.215	**4.383**	0.908	**1.737**	1.007	**6.563**
2014 Mount Polley mine tailing dam failure	Before the Failure	-0.010	0.010	-0.200	-0.150	-15.450	1.100	-3.400	6.600
	For the failure event	0.000	-0.020	0.050	0.050	0.300	-0.150	0.050	-1.380
	Difference	0.010	0.030	0.250	0.200	15.750	1.250	3.450	7.980
	Pre-failure standard deviation	0.139	0.114	0.394	0.246	9.376	1.820	16.141	54.588
	Difference/Pre-failure standard deviation	0.072	0.262	0.634	**0.814**	**1.680**	**0.687**	0.214	0.146
2013 Bingham Canyon copper mine landslide	Before the Failure	-0.005	-0.010	-0.050	0.050	-0.100	-0.050	-8.850	-13.750
	For the failure event	-0.025	-0.040	0.100	2.460	-0.005	0.000	-1.350	-51.450
	Difference	0.020	0.030	0.150	2.410	0.095	0.050	7.500	37.700
	Pre-failure standard deviation	0.069	0.113	0.103	0.119	0.540	0.238	3.656	42.465
	Difference/Pre-failure standard deviation	0.289	0.265	**1.450**	**20.290**	0.176	0.210	**2.051**	0.888

Table 6.4: Summary of the difference in the mode of remote sensing indexes in the failure source

		NDVI	Green NDVI	MSAVI	VARI	SR	MTVI2	CIg	IO
2019 Córrego de Feijão tailings dam failure	Before the Failure	0.068	0.066	0.053	-0.005	1.050	0.054	0.850	0.020
	For the failure event	-0.195	-0.115	-0.195	-0.105	-1.150	0.000	-1.150	-0.080
	Difference	-0.263	-0.181	-0.248	-0.100	-2.200	-0.054	-2.000	-0.100
	Pre-failure standard deviation	0.023	0.018	0.018	0.033	0.612	0.016	0.349	0.095
	Difference/Pre-failure standard deviation	**11.328**	**9.918**	**14.077**	3.005	3.596	3.280	5.732	1.048
2022 Jagersfontein tailings dam failure	Before the Failure	-0.017	-0.022	-0.040	0.001	-0.060	-0.025	-0.065	-0.060
	For the failure event	-0.016	-0.025	-0.030	0.015	-0.050	-0.070	-0.075	-0.080
	Difference	0.001	-0.003	0.010	0.014	0.010	-0.045	-0.010	-0.020
	Pre-failure standard deviation	0.023	0.023	0.038	0.024	0.057	0.041	0.064	0.052
	Difference/Pre-failure standard deviation	0.044	0.133	0.265	**0.582**	0.175	**1.090**	0.155	**0.387**
2022 Pau Branco iron ore mine landslide	Before the Failure	0.025	0.010	0.050	0.050	0.000	-0.075	-0.200	-0.050
	For the failure event	-0.225	-0.220	-0.350	-0.005	0.150	-0.065	0.400	0.350
	Difference	-0.250	-0.230	-0.400	-0.055	0.150	0.010	0.600	0.400
	Pre-failure standard deviation	0.178	0.218	0.475	0.106	0.641	0.501	0.701	0.295
	Difference/Pre-failure standard deviation	**1.402**	**1.055**	0.842	0.519	0.234	0.020	0.856	**1.355**
2020 Carmen Copper mine landslide	Before the Failure	0.010	0.060	0.025	-0.062	0.080	0.135	0.140	0.115
	For the failure event	-0.050	-0.055	-0.075	0.020	-0.120	-0.110	-0.155	-0.005
	Difference	-0.060	-0.115	-0.100	0.082	-0.200	-0.245	-0.295	-0.120
	Pre-failure standard deviation	0.051	0.047	0.086	0.029	0.130	0.074	0.118	0.043
	Difference/Pre-failure standard deviation	1.166	2.441	1.162	**2.781**	1.535	**3.320**	2.503	**2.820**
2020 Singrauli fly ash dam breach	Before the Failure	-0.007	-0.011	-0.025	0.001	-0.040	-0.014	-0.040	0.025
	For the failure event	0.005	-0.005	0.015	0.026	0.040	-0.015	-0.010	-0.030
	Difference	0.012	0.006	0.040	0.025	0.080	-0.002	0.030	-0.055
	Pre-failure standard deviation	0.021	0.018	0.028	0.017	0.074	0.021	0.074	0.039
	Difference/Pre-failure standard deviation	0.548	0.337	**1.440**	**1.482**	1.076	0.070	0.404	**1.395**
2018 Cadia gold mine tailings dam failure	Before the Failure	0.080	0.070	0.105	0.021	0.500	0.062	0.250	-0.065
	For the failure event	0.155	0.125	0.215	0.001	0.450	0.155	0.450	0.015
	Difference	0.075	0.055	0.110	-0.020	-0.050	0.094	0.200	0.080
	Pre-failure standard deviation	0.060	0.052	0.052	0.031	1.274	0.039	1.187	0.062
	Difference/Pre-failure standard deviation	1.242	1.049	**2.129**	0.632	0.039	**2.368**	0.169	**1.292**
2014 Mount Polley mine tailing dam failure	Before the Failure	0.020	-0.020	-0.200	-0.100	0.100	1.100	0.050	2.700
	For the failure event	0.000	0.705	0.050	0.045	0.000	0.050	1.000	1.085
	Difference	-0.020	0.725	0.250	0.145	-0.100	-1.050	0.950	-1.615
	Pre-failure standard deviation	0.244	0.177	0.999	0.452	0.425	4.503	0.299	0.686
	Difference/Pre-failure standard deviation	0.082	**4.085**	0.250	0.321	0.235	0.233	**3.175**	**2.353**
2013 Bingham Canyon copper mine landslide	Before the Failure	-0.005	-0.030	-0.050	0.015	-0.010	0.000	-0.050	0.020
	For the failure event	0.020	0.080	0.025	-0.085	0.035	0.230	0.255	0.450
	Difference	0.025	0.110	0.075	-0.100	0.045	0.230	0.305	0.430
	Pre-failure standard deviation	0.089	0.135	0.305	0.154	0.145	0.567	0.345	2.540
	Difference/Pre-failure standard deviation	0.280	**0.816**	0.246	**0.649**	0.310	0.405	**0.884**	0.169

Chapter 7 Conclusions

In this dissertation remote sensing data is analyzed using various Artificial Intelligence (AI) and big data analysis techniques with the goal to detect ground failure, assess landslide mobility, and monitor ground instability of infrastructure. Detailed conclusions are provided in each of the chapters, and a brief summary and broad conclusions, as well as recommendations for future work, are provided as follows.

7.1   Summary and Conclusions

- **Factors affecting machine learning-based landslide detection for the 2015 Lefkada earthquake (chapter 2)** The influence of different inputs and computing steps of a machine learning-based landslide detection model for the 2015 Lefkada earthquake event on the detection results is evaluated. It is found that for successful landslide detection, the selection of features is the most critical, with the post-event slope and pre-event slope being the two most essential features in the Lefkada event. Incorporating ten features results in a robust landslide detection model that is similar to the model performance using ninety-two features. The size and geospatial distribution of training samples are also critical. Secondary factors affecting detection accuracy are the resolution of input data and machine learning algorithms. Other factors, such as segmentation parameters, the spatial distance between non-landslide and landslide training samples, and the geometry of the training sample are less critical. The geospatial distribution influences the training sample size required to generate a reliable landslide detection model. Wider geospatial distribution of training samples results in better landslide detection ability.

- **Lessons learned on the efficacy of machine learning-based landslide detection following three recent earthquake events (chapter 3)** Machine learning models can be valuable for the expedited detection of large numbers of landslides following major earthquakes. The performance of the landslide detection model is affected by the selection of features, resolution, machine learning algorithms, segmentation, and training sample size. Since landslides occur in different

geologic and climatic environments, the importance of each factor on the landslide detection may vary. The influence of various factors on the landslide detection model has been investigated for three recent earthquakes, namely the 2015 Lefkada earthquake, the 2016 Kaikōura earthquake, and the 2021 Nippes earthquake event, to assess the commonalities and the differences in detection success. The results reveal that feature selection is the most critical factor in landslide detection. However, the features that may yield the best results vary. Topographic terrain features play an essential role in events in less vegetated areas, whereas indexes that consider textural characteristics, such as vegetation index, can improve the model's performance significantly in highly vegetated areas. The importance of machine learning algorithms and segmentation varies with the event. However, the data resolution and training sample size similarly affect the landslide detection model for the 2015 Lefkada earthquake, 2016 Kaikōura earthquake, and 2021 Nippes earthquake event.

- **Derivation of Voellmy model parameters for landslide runout based on co-seismic rock avalanches inventory (chapter 4)** Landslide mobility is essential in addressing landslide risk. A simple mechanistic model that is based on the Voellmy friction law as incorporated in Rapid Mass Movement Simulation Debris Flow (RAMMS-DF) is regressed against hundreds of mapped rock avalanches that were triggered by the Mw 6.5 Lefkada earthquake on November 17, 2015, with the goal to calibrate the two model parameters against statistically significant observations of landslide runout. The result reveals that the dry-Coulomb friction (μ) governs the performance of the simulation, whereas the results are less sensitive to viscous-turbulent friction (ξ), especially for large values of ξ. The RAMMS-DF simulation's accuracy positively correlates with landslide source area, height, and 3D travel distance. The model does not match well landslides with small source areas ($<$4,000 m^2), but in these cases, it systematically overestimates the runout compared to field observations.

- **Remote sensing-based system-level monitoring of Highway 1 following the 2021 atmospheric river event (chapter 5)** The damage observed along Highway 1 from the January 2021 atmospheric river is classified, and a remote-sensing-based methodology is developed for system-level monitoring and assessment following natural disasters. Forty-four drainage basins inspected by Caltrans along Highway 1 are classified into five damage classes based on the severity of damage along the highway. As expected, surface water management systems need to be properly sized to accommodate the size of the drainage basins, but larger drainage basins are usually prone to debris flow after heavy rainfall. Remotely sensed indicators of vegetation loss can reflect sediment mobility and the occurrence of landslides or debris flows. Planet© satellite imagery was used to detect changes in the normalized difference vegetation index (NDVI) before and after the January 2021 atmospheric river. The result reveals that the vegetation loss distribution curve can represent the severity of highway damage. A lower minimum value and

broader distribution of the vegetation loss curve are indicative of severe impact. Sections of the highway not impacted by the storm usually have a higher minimum value and a narrower distribution of vegetation loss. The minimum value and shape of the vegetation loss distribution curve can serve as an indicator of highway damage to detect the most critical and susceptible drainage crossings in an expedited manner.

- **An autonomous methodology for monitoring mining instability using remote sensing (chapter 6)** The detection of mining instability using high-resolution satellite imagery by evaluating the performance of different remote sensing indexes when failures occur is investigated. Eight recent failure cases are examined: the 2022 Jagersfontein tailings dam failure, the 2022 Pau Branco iron ore mine landslide, the 2020 Carmen copper mine landslide, the 2020 Singrauli fly ash dam breach, the 2019 Córrego De Feijão tailings dam failure, the 2018 Cadia gold mine tailings dam failure, the 2014 Mount Polley mine tailing dam failure, and the 2013 Bingham Canyon copper mine landslide. Vegetation indexes (normalized difference vegetation index, NDVI; green normalized difference vegetation index, Green NDVI; modified soil adjusted vegetation index, MSAVI; visible atmospherically resistant index, VARI; simple ratio, SR; modified triangular vegetation index, MTVI2; chlorophyll index – green, CIg) and geological indexes (iron oxide, IO) derived from satellite imagery are used to track spatial and temporal change at mining facilities. The result shows that remote sensing indexes can be good indicators of mining failures. Change in VARI and IO detects the mining instability well for both vegetated and non-vegetated sites. Change in MTVI2 detects the failure well in vegetated sites, while MSAVI performs best in identifying mining instability in non-vegetated sites.

7.2 Recommendations for Future Work

- **Leveraging big data in the machine learning-based landslide detection model and incorporating global landslide cases to develop a landslide detection guideline** Landslides are catastrophic, and various characteristics, such as ground displacement and fluctuation of soil water content, can indicate the occurrence of landslides. Therefore, integrating remote sensing data and in-situ monitoring field data can enhance the capability of the landslide detection model. Data from ground-based remote sensing sensors, including lidar and radar, can benefit the model's accuracy and acquisition frequency. In general, leveraging big data and incorporating global landslide cases can improve the efficiency of landslide risk assessment and management.

- **Developing a mechanistic runout model parameters database** Although empirical landslide mobility models have been developed to estimate the landslide travel distance, these models have limitations when implemented in different

geologic environments. This dissertation uses the 2015 Lefkada earthquake event as a study case to derive parameters in a simple mechanistic model calibrated against field observations. Extending the derivation of the mechanistic model parameters to global landslide cases can provide an improved understanding of landslide behavior and enhance the applicability of existing empirical landslide mobility models.

- **Calibrating the autonomous remote sensing-based monitoring model using historical and future infrastructure failure events** This dissertation uses the 2021 atmospheric river event to develop the remote sensing-based monitoring system of Highway 1 and eight mining failure cases for the mining instability monitoring system. Leveraging weather and in-situ monitoring data and adding historical and future failure events can improve the model's accuracy and reduce uncertainty while eliminating bias associated with model calibrations against specific events. This study demonstrated that such a remote sensing-based monitoring system can be fully autonomous and continuously track roadway infrastructure conditions and mining facilities on a large scale and after natural hazards. However, additional model optimizations against a wider set of events/data are expected to be needed.

Bibliography

Aaron, J., McDougall, S., & Nolde, N. (2019). Two methodologies to calibrate landslide runout models. Landslides, 16, 907-920.

Acharya AS, Prakash A, Saxena P, et al (2013) Sampling: Why and how of it. researchgate.net. https://doi.org/10.7713/ijms.2013.0032

Alimohammadlou, Y., Najafi, A., & Yalcin, A. (2013). Landslide process and impacts: A proposed classification method. Catena, 104, 219-232.

Alipouri, E., Nami, M. H., & Naderi, M. (2024). An Overview of the Application of Remote Sensing Technologies in Disaster Management (With an Emphasis on Natural Hazards). Disaster Prevention and Management Knowledge (quarterly), 13(4), 490-507.

Amatya, P., Kirschbaum, D., Stanley, T., & Tanyas, H. (2021). Landslide mapping using object-based image analysis and open source tools. Engineering Geology, 282, 106000.

Amatya, P., Scheip, C., Déprez, A., Malet, J. P., Slaughter, S. L., Handwerger, A. L., ... & Boissier, E. (2023). Learnings from rapid response efforts to remotely detect landslides triggered by the August 2021 Nippes earthquake and Tropical Storm Grace in Haiti. Natural Hazards, 118(3), 2337-2375.

Angeli, M. G., Pasuto, A., & Silvano, S. (2000). A critical review of landslide monitoring experiences. Engineering Geology, 55(3), 133-147.

Argentin, A. L., Hauthaler, T., Liebl, M., Robl, J., Hergarten, S., Prasicek, G., ... & Dabiri, Z. (2022). Influence of rheology on landslide-dammed lake impoundment and sediment trapping: Back-analysis of the Hintersee landslide dam. Geomorphology, 414, 108363.

Azam, S., & Li, Q. (2010). Tailings dam failures: a review of the last one hundred years. Geotechnical news, 28(4), 50-54.

Barlow, J., Franklin, S., & Martin, Y. (2006). High spatial resolution satellite imagery, DEM derivatives, and image segmentation for the detection of mass wasting processes. Photogrammetric Engineering & Remote Sensing, 72(6), 687-692.

Barnhart, K. R., Jones, R. P., George, D. L., McArdell, B. W., Rengers, F. K., Staley, D. M., & Kean, J. W. (2021). Multi-model comparison of computed debris flow runout for the January 9 2018 Montecito, California post-wildfire event. Journal of Geophysical Research: Earth Surface, 126(12), e2021JF006245.

Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation measures for models assessment over imbalanced data sets. J Inf Eng Appl, 3(10).

Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24-31.

Blaschke, T., Hay, G. J., Kelly, M., Lang, S., Hofmann, P., Addink, E., ... & Tiede, D. (2014). Geographic object-based image analysis–towards a new paradigm. ISPRS journal of photogrammetry and remote sensing, 87, 180-191.

Bowker, L. N., & Chambers, D. M. (2015). The risk, public liability, & economics of tailings storage facility failures. Earthwork Act, 24, 1-56.

Bowker, L. N., & Chambers, D. M. (2017). In the dark shadow of the supercycle, tailings failure risk & public liability reach all time highs. Environments, 4(4), 75.

Breiman, L. (1996). Bagging predictors. Machine learning, 24, 123-140.

Breiman, L. (2001). Random forests Mach Learn 45 (1): 5–32.

Bridgehunter.com "Big Creek Bridge, Monterey County, California" historic and notable bridges of the u.s. https://bridgehunter.com/ca/monterey/440056/. (Accessed January 05, 2023).

BridgeReports.com "STATE ROUTE 1 over MILL CREEK, Mendocino County, California" National Bridge Inventory data. https://bridgereports.com/1040415. (Accessed January 05, 2023).

Burned Area Emergency Response: Dolan Postfire BAER Soil Burn Severity Map Release (2020), U.S. Forest Service, https://inciweb.nwcg.gov/incident/7231/ (Accessed January 5, 2023).

Cabanatuan, Michael (June 19, 2010). "Major Progress on Devil's Slide Tunnel". San Francisco Chronicle. (Accessed January 3, 2023).

Cai, H., Chen, T., Niu, R., & Plaza, A. (2021). Landslide detection using densely connected convolutional networks and environmental conditions. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 5235-5247.

Caltrans (California Deptartment of Transportation). n.d. "Big Sur inventory reports and history road closures.". http://www.dot.ca.gov/dist05/projects/bigsur/inventory_reports/history_road_closures.pdf.

Canon, Gabrielle (2021). "California's famed Highway 1 collapsed last week. It's

sure to happen again". The Guardian. ISSN 0261-3077. (Accessed January 2023).

Casagli, N., Frodella, W., Morelli, S., Tofani, V., Ciampalini, A., Intrieri, E., ... & Lu, P. (2017). Spaceborne, UAV and ground-based remote sensing techniques for landslide mapping, monitoring and early warning. Geoenvironmental Disasters, 4, 1-23.

Casagli, N., Intrieri, E., Tofani, V., Gigli, G., & Raspini, F. (2023). Landslide detection, monitoring and prediction with remote-sensing techniques. Nature Reviews Earth & Environment, 4(1), 51-64.

Chae, B. G., Park, H. J., Catani, F., Simoni, A., & Berti, M. (2017). Landslide prediction, monitoring and early warning: a concise review of state-of-the-art. Geosciences Journal, 21, 1033-1070.

Chen, W., Peng, J., Hong, H., Shahabi, H., Pradhan, B., Liu, J., ... & Duan, Z. (2018). Landslide susceptibility modelling using GIS-based machine learning techniques for Chongren County, Jiangxi Province, China. Science of the total environment, 626, 1121-1135.

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. Canadian Journal of Remote Sensing, 22(3), 229-242.

Chen, Q., Wang, L., & Zhao, H. (2009). Hydrodynamic investigation of coastal bridge collapse during Hurricane Katrina. Journal of Hydraulic Engineering, 135(3), 175-186.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21, 1-13.

Christen, M., Kowalski, J., & Bartelt, P. (2010). RAMMS: Numerical simulation of dense snow avalanches in three-dimensional terrain. Cold Regions Science and Technology, 63(1-2), 1-14.

Clevers, J. G., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3. International Journal of Applied Earth Observation and Geoinformation, 23, 344-351.

Constantin, M., Bednarik, M., Jurchescu, M. C., & Vlaicu, M. (2011). Landslide susceptibility assessment using the bivariate statistical analysis and the index of entropy in the Sibiciu Basin (Romania). Environmental earth sciences, 63, 397-406.

Corominas, J. (1996). The angle of reach as a mobility index for small and large landslides. Canadian Geotechnical Journal, 33(2), 260-271.

Crosetto, M., Monserrat, O., Cuevas-González, M., Devanthéry, N., & Crippa, B. (2016). Persistent scatterer interferometry: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 78-89.

Crosta, G. B., De Blasio, F. V., & Frattini, P. (2018). Global scale analysis of Martian landslide mobility and paleoenvironmental clues. Journal of Geophysical Research: Planets, 123(4), 872-891.

Dai, F. C., Lee, C. F., & Ngai, Y. Y. (2002). Landslide risk assessment and management: an overview. Engineering geology, 64(1), 65-87.

Das, S., Sharma, P., Pain, A., Kanungo, D. P., & Sarkar, S. (2023). Deep learning based landslide detection using open-source resources: Opportunities and challenges. Earth Science Informatics, 16(4), 4035-4052.

De Pedrini, A., Ambrosi, C., & Scapozza, C. (2022). The 1513 Monte Crenone rock avalanche: numerical model and geomorphological analysis. Geographica Helvetica, 77(1), 21-37.

Dehnavi, A., Aghdam, I. N., Pradhan, B., & Varzandeh, M. H. M. (2015). A new hybrid model using step-wise weight assessment ratio analysis (SWARA) technique and adaptive neuro-fuzzy inference system (ANFIS) for regional landslide hazard assessment in Iran. Catena, 135, 122-148.

Devoli, G., De Blasio, F. V., Elverhøi, A., & Høeg, K. (2009). Statistical analysis of landslide events in Central America and their runout distance. Geotechnical and Geological Engineering, 27, 23-42.

Dhakal, A. S., Amada, T., & Aniya, M. (2000). Landslide hazard mapping and its evaluation using GIS: an investigation of sampling schemes for a grid-cell based quantitative method. Photogrammetric engineering and remote sensing, 66(8), 981-989.

Dogan, H. M. (2009). Mineral composite assessment of Kelkit River Basin in Turkey by means of remote sensing. Journal of Earth System Science, 118, 701-710.

Dong, S., Gao, X., Mostafavi, A., & Gao, J. (2022). Modest flooding can trigger catastrophic road network collapse due to compound failure. Communications Earth & Environment, 3(1), 38.

Dong, T., Liu, J., Shang, J., Qian, B., Ma, B., Kovacs, J. M., ... & Shi, Y. (2019). Assessment of red-edge vegetation indices for crop leaf area index estimation. Remote Sensing of Environment, 222, 133-143.

Dou J, Yunus AP, Bui DT, et al (2020) Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 17:641–658

Draves, J. F., & Fox, M. G. (1998). Effects of a mine tailings spill on feeding and metal concentrations in yellow perch (Perca flavescens). Environmental Toxicology and Chemistry: An International Journal, 17(8), 1626-1632.

Duffell, C. G., & Rudrum, D. M. (2005). Remote sensing techniques for highway earthworks assessment. In Site Characterization and Modeling (pp. 1-13).

Eeckhaut, M., Poesen, J., Verstraeten, G., Vanacker, V., Moeyersons, J., Nyssen, J., & Van Beek, L. P. H. (2005). The effectiveness of hillshade maps and expert knowledge in mapping old deep-seated landslides. Geomorphology, 67(3-4), 351-363.

Esri Inc. (2021). ArcGIS Pro (Version 2.9). Esri Inc. https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview.

Fathani, T. F., Legono, D., & Alfath, M. A. (2017). Sensitivity analysis of depth-integrated numerical models for estimating landslide movement. Journal of disaster research, 12(3), 607-616.

Feizizadeh, B., Roodposhti, M. S., Jankowski, P., & Blaschke, T. (2014). A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping. Computers & geosciences, 73, 208-221.

Fell R (1994) Landslide risk assessment and acceptable risk. Can Geotech J 31:261–272. https://doi.org/10.1139/T94-031

Feng, S., Liu, S., Jing, L., Zhu, Y., Yan, W., Jiang, B., ... & Jia, J. (2021). Quantification of the environmental impacts of highway construction using remote sensing approach. Remote Sensing, 13(7), 1340.

Feroz, S., & Abu Dabous, S. (2021). Uav-based remote sensing applications for bridge condition assessment. Remote Sensing, 13(9), 1809.

Finlay, P. J., Mostyn, G. R., & Fell, R. (1999). Landslide risk assessment: prediction of travel distance. Canadian Geotechnical Journal, 36(3), 556-562.

Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of online learning and an application to boosting. Journal of computer and system sciences, 55(1), 119-139.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, 1189-1232.

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.

Gagliardi, V., Tosti, F., Bianchini Ciampoli, L., Battagliere, M. L., D'Amato, L., Alani, A. M., & Benedetto, A. (2023). Satellite remote sensing and non-destructive testing methods for transport infrastructure monitoring: Advances, challenges and perspectives. Remote Sensing, 15(2), 418.

Galli, M., Ardizzone, F., Cardinali, M., Guzzetti, F., & Reichenbach, P. (2008). Comparing landslide inventory maps. Geomorphology, 94(3-4), 268-289.

Gan, J., & Zhang, Y. S. (2019). Numerical simulation of debris flow runout using Ramms: a case study of Luzhuang Gully in China. Computer Modeling in Engineering and Sciences, 981-1009.

Gao, L., Zhang, L. M., Chen, H. X., Fei, K., & Hong, Y. (2021). Topography and geology effects on travel distances of natural terrain landslides: Evidence from

a large multi-temporal landslide inventory in Hong Kong. Engineering Geology, 292, 106266.

Gardezi, H., Bilal, M., Cheng, Q., Xing, A., Zhuang, Y., & Masood, T. (2021). A comparative analysis of attabad landslide on january 4, 2010, using two numerical models. Natural Hazards, 107, 519-538.

Ghorbanzadeh, O., Blaschke, T., Gholamnia, K., Meena, S. R., Tiede, D., & Aryal, J. (2019). Evaluation of different machine learning methods and deep-learning convolutional neural networks for landslide detection. Remote Sensing, 11(2), 196.

Ghorbanzadeh O, Shahabi H, Crivellari A, et al (2022) Landslide detection using deep learning and object-based image analysis. Landslides 19:929–939.

Ghorbanzadeh, O., Shahabi, H., Crivellari, A., Homayouni, S., Blaschke, T., & Ghamisi, P. (2022a) b. Landslide detection using deep learning and object-based image analysis. Landslides, 19(4), 929-939.

Ghorbanzadeh, O., Xu, Y., Ghamisi, P., Kopp, M., & Kreil, D. (2022b). Landslide4sense: Reference benchmark data and deep learning models for landslide detection. arXiv preprint arXiv:2206.00515.

Gilany, N., & Iqbal, J. (2019). Simulation of glacial avalanche hazards in Shyok Basin of upper Indus. Scientific reports, 9(1), 20077.

Gómez, D., García, E. F., & Aristizábal, E. (2023). Spatial and temporal landslide distributions using global and open landslide databases. Natural Hazards, 117(1), 25-55.

Grebby, S., Sowter, A., Gluyas, J., Toll, D., Gee, D., Athab, A., & Girindran, R. (2021). Advanced analysis of satellite data reveals ground deformation precursors to the Brumadinho Tailings Dam collapse. Communications Earth & Environment, 2(1), 1-9.

Guevara, J., Garvin, M. J., & Ghaffarzadegan, N. (2017). Capability trap of the US highway system: Policy and management implications. Journal of Management in Engineering, (4), 04017004.

Guikema, S. (2020). Artificial intelligence for natural hazards risk analysis: Potential, challenges, and research needs. Risk Analysis, 40(6), 1117-1123.

Guo, D., Hamada, M., He, C., Wang, Y., & Zou, Y. (2014). An empirical model for landslide travel distance prediction in Wenchuan earthquake area. Landslides, 11, 281-291.

Guzzetti, F., Mondini, A. C., Cardinali, M., Fiorucci, F., Santangelo, M., & Chang, K. T. (2012). Landslide inventory maps: New tools for an old problem. Earth-Science Reviews, 112(1-2), 42-66.

Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. B. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI

of crop canopies: Modeling and validation in the context of precision agriculture. Remote sensing of environment, 90(3), 337-352.

Hagner, O., & Reese, H. (2007). A method for calibrated maximum likelihood classification of forest types. Remote sensing of environment, 110(4), 438-444.

Han, Q., Du, X., Liu, J., Li, Z., Li, L., & Zhao, J. (2009). Seismic damage of highway bridges during the 2008 Wenchuan earthquake. Earthquake Engineering and Engineering Vibration, 8(2), 263-273.

Hattanji, T., & Moriwaki, H. (2009). Morphometric analysis of relic landslides using detailed landslide distribution maps: Implications for forecasting travel distance of future landslides. Geomorphology, 103(3), 447-454.

Heckmann, T., Gegg, K., Gegg, A., & Becht, M. (2014). Sample size matters: investigating the effect of sample size on a logistic regression susceptibility model for debris flows. Natural Hazards and Earth System Sciences, 14(2), 259-278.

Herath, S., & Wang, Y. (2009). Case studies and national experiences. Landslides– disaster risk reduction, 475-497.

Highland, L. M., Godt, J. W., Howell, D. G., & Savage, W. Z. (1998). El Nino 1997-98; damaging landslides in the San Francisco Bay area (No. 089-98). US Dept. of the Interior, US Geological Survey, National Landslide Information Center.

Hilker, N., Badoux, A., & Hegg, C. (2009). The Swiss flood and landslide damage database 1972–2007. Natural Hazards and Earth System Sciences, 9(3), 913-925.

Hong, H., Pradhan, B., Xu, C., & Bui, D. T. (2015). Spatial prediction of landslide hazard at the Yihuang area (China) using two-class kernel logistic regression, alternating decision tree and support vector machines. Catena, 133, 266-281.

Hoppe, E., Bruckno, B., Campbell, E., Acton, S., Vaccari, A., Stuecheli, M., ... & Morgan, J. (2016). Transportation infrastructure monitoring using satellite remote sensing. Materials and Infrastructures 1, 5, 185-198.

Huang F, Zhang J, Zhou C, et al (2020) A deep learning algorithm using a fully connected sparse autoencoder neural network for landslide susceptibility prediction. Landslides 17:217–229.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., & Shao, G. (2021). A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of Forestry Research, 32(1), 1-6.

Huang, J. R., & Zekkos, D. (2023). Effect of Machine Learning Algorithms on Detection of Landslides Caused by the 2015 Lefkada Earthquake. In Geo-Congress 2023 (pp. 347-356).

Huang, J. R., & Zekkos, D. (2023a). Effect of segmentation and machine learning models on co-seismic landslide detection. 4th International Symposium on

Machine Learning and Big Data in Geoscience.

Huang, J. R., & Zekkos, D. (2023b). Satellite-based Change Detection for Mining Instability.

Huang, J. R., Zekkos, D., & Clark, M. (2024). Co-Seismic Landslide Mobility Assessment Using Machine Learning Models. In Geo-Congress 2024 (pp. 475-484).

Hungr, O., Leroueil, S., & Picarelli, L. (2014). The Varnes classification of landslide types, an update. Landslides, 11, 167-194.

Hungr, O. (2010). Prospects for prediction of landslide dam geometry using empirical and dynamic models. In Natural and Artificial Rockslide Dams (pp. 463-477). Berlin, Heidelberg: Springer Berlin Heidelberg.

Hungr, O., & Evans, S. G. (1996). Rock avalanche runout prediction using a dynamic model. In Landslides (pp. 233-238).

Hunter, G., & Fell, R. (2003). Travel distance angle for" rapid" landslides in constructed and natural soil slopes. Canadian Geotechnical Journal, 40(6), 1123-1141.

Iannacone, J. P., Quan Luna, B., & Corsini, A. (2013). Forward simulation and sensitivity analysis of runout scenarios using MassMov2D at the Trafoi rockslide (South Tyrol, Italy). Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards, 7(4), 240-249.

Islam, K., & Murakami, S. (2021). Global-scale impact analysis of mine tailings dam failures: 1915–2020. Global environmental change, 70, 102361.

Iverson, R. M. (2005). Debris-flow mechanics. Debris-flow hazards and related phenomena, 8, 105-134.

Johnson, B. C., & Campbell, C. S. (2017). Drop height and volume control the mobility of long-runout landslides on the Earth and Mars. Geophysical Research Letters, 44(24), 12-091.

Johnson, Lizzie (2017). "Bridge failure severs Big Sur's ties to outside world". San Francisco Chronicle. (Accessed January 2023).

JRP Historical Consulting Services. November (2001). "A History of Road Closures Along Highway 1, Big Sur Monterey and San Luis Obispo Counties, California" Big Sur Coast Highway Management Plan.

Juang, C. S., Stanley, T. A., & Kirschbaum, D. B. (2019). Using citizen science to expand the global map of landslides: Introducing the Cooperative Open Online Landslide Repository (COOLR). PloS one, 14(7), e0218657.

Karantanellis, S., Clark, M. K., Zekkos, D., & Willis, M. J. (2023). Landslide volume estimates for the November 2016 M7. 8 Kaikoura Earthquake region using satellite derived Digital Surface Models (DSM). AGU23.

Kavzoglu T, Sahin EK, Colkesen I (2014) Landslide susceptibility mapping using

GIS-based multi-criteria decision analysis, support vector machines, and logistic regression. Landslides 11:425–439.

Keefer, D. K. (2002). Investigating landslides caused by earthquakes–a historical review. Surveys in geophysics, 23, 473-510.

Kelly, S. (2015). Estimating economic loss from cascading infrastructure failure: a perspective on modelling interdependency. Infrastructure Complexity, 2, 1-13.

Kjekstad, O., & Highland, L. (2009). Economic and social impacts of landslides. Landslides–disaster risk reduction, 573-587.

Klose, M., Maurischat, P., & Damm, B. (2016). Landslide impacts in Germany: A historical and socioeconomic perspective. Landslides, 13, 183-199.

Kossoff, D., Dubbin, W. E., Alfredsson, M., Edwards, S. J., Macklin, M. G., & Hudson-Edwards, K. A. (2014). Mine tailings dams: Characteristics, failure, environmental impacts, and remediation. Applied Geochemistry, 51, 229-245.

Krichen, M., Abdalzaher, M. S., Elwekeil, M., & Fouda, M. M. (2023). Managing natural disasters: An analysis of technological advancements, opportunities, and challenges. Internet of Things and Cyber-Physical Systems.

Krishnamoorthi, N. (2016). Role of remote sensing and GIS in natural-disaster management cycle. Imp J Inter Res, 2(3), 144-154.

KSBW (2011) "Highway 1 to Be Closed at Least 1 Month". Salinas, California. https://web.archive.org/web/20111004040123/http://www.ksbw.com/news/27 237744/detail.html (Accessed January 3, 2023).

Kucharczyk, M., & Hugenholtz, C. H. (2021). Remote sensing of natural hazard-related disasters with small drones: Global trends, biases, and research opportunities. Remote Sensing of Environment, 264, 112577.

Kumar, A., Misra, R., Singh, T. N., & Dhiman, G. (2024). APO-AN feature selection based Glorot Init Optimal TransCNN landslide detection from multi source satellite imagery. Multimedia Tools and Applications, 83(14), 40451-40488.

Kumar, V., Gupta, V., Jamir, I., & Chattoraj, S. L. (2019). Evaluation of potential landslide damming: Case study of Urni landslide, Kinnaur, Satluj valley, India. Geoscience Frontiers, 10(2), 753-767.

Legros, F. (2002). The mobility of long-runout landslides. Engineering geology, 63(3-4), 301-331.

Levner, I. (2005). Feature selection and nearest centroid classification for protein mass spectrometry. BMC bioinformatics, 6, 1-14.

Li, C., Handwerger, A. L., Wang, J., Yu, W., Li, X., Finnegan, N. J., ... & Horton, D. E. (2022). Augmentation of WRF-Hydro to simulate overland-flow-and streamflow-generated debris flow susceptibility in burn scars. Natural Hazards and Earth System Sciences, 22(7), 2317-2345.

Li, X. Z., Kong, J. M., & Li, S. W. (2011). Travel distance prediction of landslides

triggered by the M8. 0 Wenchuan earthquake. Applied Mechanics and Materials, 71, 1736-1740.

Lin Q, Wang L, Buda S, et al (2022) The influences of the spatial extent selection for non-landslide samples on statistical-based landslide susceptibility modelling: A case study of Anhui Province in China. Springer.

Lindley, S. M., Pindilli, E. J., & Huber, C. (2023). Societal Benefits of Science: A study of the economic value of US Geological Survey landslide susceptibility maps in Puerto Rico.

Little, R. G. (2002). Controlling cascading failure: Understanding the vulnerabilities of interconnected infrastructures. Journal of Urban Technology, 9(1), 109-123.

Liu, K., Shi, W., & Zhang, H. (2011). A fuzzy topology-based maximum likelihood classification. ISPRS Journal of Photogrammetry and Remote Sensing, 66(1), 103-114.

Liu, P., Wei, Y., Wang, Q., Chen, Y., & Xie, J. (2020). Research on post-earthquake landslide extraction algorithm based on improved U-Net model. Remote Sensing, 12(5), 894.

Liu, Z., L'heureux, J. S., Glimsdal, S., & Lacasse, S. (2021). Modelling of mobility of Rissa landslide and following tsunami. Computers and Geotechnics, 140, 104388.

Loli, M., Manousakis, J., Mitoulis, S. A., & Zekkos, D. (2022). UAV-enabled flood damage assessment and recovery monitoring of bridges following Medicane Ianos. In Bridge Safety, Maintenance, Management, Life-Cycle, Resilience and Sustainability (pp. 828-836). CRC Press.

Lu, D., Hetrick, S., & Moran, E. (2011). Impervious surface mapping with Quickbird imagery. International journal of remote sensing, 32(9), 2519-2533.

Lu, W., Hu, Y., Zhang, Z., & Cao, W. (2023). A dual-encoder U-Net for landslide detection using Sentinel-2 and DEM data. Landslides, 20(9), 1975-1987.

Luna, B. Q., Cepeda, J., Stumpf, A., Van Westen, C. J., Remaître, A., Malet, J. P., & Van Asch, T. W. (2013). Analysis and uncertainty quantification of dynamic runout model parameters for landslides. Landslide Science and Practice: Volume 3: Spatial Analysis and Modelling, 315-318.

Lyu, Z., Chai, J., Xu, Z., Qin, Y., & Cao, J. (2019). A comprehensive review on reasons for tailings dam failures based on case history. Advances in Civil Engineering, 2019(1), 4159306.

Magallon, Stephanie. (2016) "UPDATE: Hwy 1 at Big Sur back open after rock slide". KION.

Maltinti, F., Melis, D., & Annunziata, F. (2012). Road network vulnerability: A review of the literature. ICSDC 2011: Integrating Sustainability Practices in the Construction Industry, 677-685.

Mantovani, F., Soeters, R., & Van Westen, C. J. (1996). Remote sensing techniques for landslide studies and hazard zonation in Europe. Geomorphology, 15(3-4), 213-225.

Martha TR, Kerle N, Jetten V, et al (2010) Characterising spectral, spatial and morphometric properties of landslides for semi-automatic detection using object-oriented methods. Geomorphology 116:24–36.

Massey, C. I., Townsend, D. T., Lukovic, B., Morgenstern, R., Jones, K., Rosser, B., & de Vilder, S. (2020). Landslides triggered by the MW 7.8 November 14 2016 Kaikōura earthquake: an update. Landslides, 17, 2401-2408.

Matsushita, B., Yang, W., Chen, J., Onda, Y., & Qiu, G. (2007). Sensitivity of the enhanced vegetation index (EVI) and normalized difference vegetation index (NDVI) to topographic effects: a case study in high-density cypress forest. Sensors, 7(11), 2636-2651.

Maus, V., Giljum, S., Gutschlhofer, J., da Silva, D. M., Probst, M., Gass, S. L., ... & McCallum, I. (2020). A global-scale data set of mining areas. Scientific data, 7(1), 289.

McFeeters SK (1996) The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int J Remote Sens 17:1425–1432.

McKinnon, M., Hungr, O., & McDougall, S. (2008, May). Dynamic analyses of Canadian landslides. In Proceedings of the Fourth Canadian Conference on GeoHazards: From Causes to Management (pp. 20-24). Canadian Geotechnical Society, Engineering Geology Division, Presse de l'Université de Laval CA.

Meena SR, Soares LP, Grohmann CH, et al (2022) Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides 19:1209–1229.

Metternicht, G., Hurni, L., & Gogu, R. (2005). Remote sensing of landslides: An analysis of the potential contribution to geo-spatial systems for hazard assessment in mountainous environments. Remote sensing of Environment, 98(2-3), 284-303.

Meyrat, G., McArdell, B., Ivanova, K., Müller, C., & Bartelt, P. (2022). A dilatant, two-layer debris flow model validated by flow density measurements at the Swiss illgraben test site. Landslides, 19(2), 265-276.

Mikoš, M., & Bezak, N. (2021). Debris Flow Modelling Using RAMMS Model in the Alpine Environment With Focus on the Model Parameters and Main Characteristics. Frontiers in Earth Science, 8, 605061.

Moges, S. M., Raun, W. R., Mullen, R. W., Freeman, K. W., Johnson, G. V., & Solie, J. B. (2005). Evaluation of green, red, and near infrared bands for predicting winter wheat biomass, nitrogen uptake, and final grain yield. Journal of plant nutrition, 27(8), 1431-1441.

Mohan A, Singh A, … BK-T on, 2021 undefined (2020) Review on remote sensing

methods for landslide detection using machine and deep learning. Wiley Online Libr 32:.

Moncayo, S., & Ávila, G. (2023). Landslide travel distances in Colombia from national landslide database analysis. In Progress in Landslide Research and Technology, Volume 1 Issue 1, 2022 (pp. 315-325). Cham: Springer International Publishing.

Mondini, A. C., Guzzetti, F., Chang, K. T., Monserrat, O., Martha, T. R., & Manconi, A. (2021). Landslide failures detection and mapping using Synthetic Aperture Radar: Past, present and future. Earth-Science Reviews, 216, 103574.

Motsau, B., & Van Wyk, D. (2022). Report on the Jagersfontein tailings disaster.

Mountrakis, G., Im, J., & Ogole, C. (2011). Support vector machines in remote sensing: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 247-259.

Mühlhofer, E., Koks, E. E., Kropf, C. M., Sansavini, G., & Bresch, D. N. (2023). A generalized natural hazard risk modelling framework for infrastructure failure cascades. Reliability Engineering & System Safety, 234, 109194.

Mura, J. C., Gama, F. F., Paradella, W. R., Negrão, P., Carneiro, S., De Oliveira, C. G., & Brandão, W. S. (2018). Monitoring the vulnerability of the dam and dikes in Germano iron mining area after the collapse of the tailings dam of Fundão (Mariana-MG, Brazil) using DInSAR techniques with TerraSAR-X data. Remote Sensing, 10(10), 1507.

Murphy, K. P. (2006). Naive bayes classifiers. University of British Columbia, 18(60), 1-8.

NASA Earth Observatory (2022). Jagersfontein Covered in Mining Waste. https://earthobservatory.nasa.gov/images/150497/jagersfontein-covered-in-mining-waste (Accessed February 8, 2023).

National Oceanic and Atmospheric Administration (2016) "Resource Issues: Landslide and Debris" Monterey Bay National Marine Sanctuary.

Nava, L., Monserrat, O., & Catani, F. (2021). Improving landslide detection on SAR data through deep learning. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.

Nguy-Robertson, A. L. (2013). The mathematical identity of two vegetation indices: MCARI2 and MTVI2. International journal of remote sensing, 34(21), 7504-7507.

Omar, T., & Nehdi, M. L. (2017). Remote sensing of concrete bridge decks using unmanned aerial vehicle infrared thermography. Automation in Construction, 83, 360-371.

Owen, J. R., Kemp, D., Lébre, É., Svobodova, K., & Murillo, G. P. (2020). Catastrophic tailings dam failures and disaster risk disclosure. International

journal of disaster risk reduction, 42, 101361.

Ozden, A., Faghri, A., Li, M., & Tabrizi, K. (2016). Evaluation of Synthetic Aperture Radar satellite remote sensing for pavement and infrastructure monitoring. Procedia Engineering, 145, 752-759.

Palmer, Tim (2002). Pacific High: Adventures in the Coast Ranges from Baja to Alaska. Washington: Island Press. p. 157.

Pankow, K. L., Moore, J. R., Hale, J. M., Koper, K. D., Kubacki, T., Whidden, K. M., & McCarter, M. K. (2014). Massive landslide at Utah copper mine generates wealth of geophysical data. Gsa Today, 24(1), 4-9.

Paola, J. D., & Schowengerdt, R. A. (1995). A detailed comparison of backpropagation neural network and maximum-likelihood classifiers for urban land use classification. IEEE Transactions on Geoscience and remote sensing, 33(4), 981-996.

Pardeshi, S. D., Autade, S. E., & Pardeshi, S. S. (2013). Landslide hazard assessment: recent trends and techniques. SpringerPlus, 2, 1-11.

Pawluszek-Filipiak, K., & Borkowski, A. (2020). On the importance of train–test split ratio of datasets in automatic landslide detection by supervised classification. Remote Sensing, 12(18), 3054.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. the Journal of machine Learning research, 12, 2825-2830.

Peethambaran, B., Nandakumar, V., & Sweta, K. (2023). Engineering geological investigation and runout modelling of the disastrous Taliye landslide, Maharashtra, India of July 22 2021. Natural Hazards, 1-16.

Petley, Dave (2018). Cadia Gold Mine – another tailings dam failure. American Geophysical Union. https://blogs.agu.org/landslideblog/2018/03/12/cadia-1/ (Accessed February 27, 2023).

Petley, Dave (2020). The Carmen Copper Mine landslide in the Philippines – the slope evolution before failure. American Geophysical Union. https://blogs.agu.org/landslideblog/2020/12/26/the-carmen-copper-mine-landslide-in-the-philippines-the-slope-evolution-before-failure/ (Accessed February 25, 2023).

Petley, Dave (2022). Pau Branco: another significant mining-related landslide in Brazil. American Geophysical Union. https://blogs.agu.org/landslideblog/2022/01/11/pau-branco-1/ (Accessed February 25, 2023).

Petrucci, O. (2022). Landslide fatality occurrence: a systematic review of research published between January 2010 and March 2022. Sustainability, 14(15), 9346.

Pettorelli, N. (2013). The normalized difference vegetation index. Oxford University

Press.

Pinty, B., & Verstraete, M. M. (1992). GEMI: a nonlinear index to monitor global vegetation from satellites. Vegetation, 101, 15-20.

Piralilou ST, Shahabi H, Jarihani B, et al (2019) Landslide detection using multi-scale image segmentation and different machine learning models in the higher himalayas. Remote Sens 11:.

Planet Team (2017). Planet Application Program Interface: In Space for Life on Earth. San Francisco, CA.

Pourghasemi, H. R., Kariminejad, N., Amiri, M., Edalat, M., Zarafshar, M., Blaschke, T., & Cerda, A. (2020). Assessing and mapping multi-hazard risk susceptibility using a machine learning technique. Scientific reports, 10(1), 3203.

Powers, D. M. (2015). What the F-measure doesn't measure: Features, Flaws, Fallacies and Fixes. arXiv preprint arXiv:1503.06410.

Provost, Stephen H. (2017) "What price beauty? On Highway 1, Mother Nature decides". The Tribune. (Accessed January 3, 2023).

Qi W, Wei M, Yang W, et al (2020) Automatic mapping of landslides by the ResU-Net. Remote Sens 12:1–14.

Qi, J., Chehbouni, A., Huete, A. R., Kerr, Y. H., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. Remote sensing of environment, 48(2), 119-126.

Qin, H., Wang, J., Mao, X., Zhao, Z. A., Gao, X., & Lu, W. (2024). An Improved Faster R-CNN Method for Landslide Detection in Remote Sensing Images. Journal of Geovisualization and Spatial Analysis, 8(1), 2.

Qiu, H., Cui, P., Hu, S., Regmi, A. D., Wang, X., & Yang, D. (2018). Developing empirical relationships to predict loess slide travel distances: a case study on the Loess Plateau in China. Bulletin of Engineering Geology and the Environment, 77, 1299-1309.

Qiu, H., Cui, P., Regmi, A. D., Hu, S., Wang, X., Zhang, Y., & He, Y. (2017). Influence of topography and volume on mobility of loess slides within different slip surfaces. Catena, 157, 180-188.

Ramli, M. F., Yusof, N., Yusoff, M. K., Juahir, H., & Shafri, H. Z. M. (2010). Lineament mapping and its application in landslide hazard assessment: a review. Bulletin of engineering Geology and the Environment, 69, 215-233.

RAMMS User Manual (2022) RAMMS rapid mass movement simulation A numerical model for debris flows in research and practice User Manual v1.8.0 Debris Flow.

Rauhala, A., Tuomela, A., Davids, C., & Rossi, P. M. (2017). UAV remote sensing surveillance of a mine tailings impoundment in sub-arctic conditions. Remote sensing, 9(12), 1318.

Rengers, F. K., McGuire, L. A., Oakley, N. S., Kean, J. W., Staley, D. M., & Tang, H. (2020). Landslides after wildfire: Initiation, magnitude, and mobility. Landslides, 17(11), 2631-2641.

Reyes, J. (2018). CalTrans sets new target date to reopen Highway 1 in Big Sur at Mud Creek, <http://www.montereyherald.com/general-news/20180430/caltrans-sets-new-target-date-to-reopen-highway-1-in-big-sur-at-mud-creek> (Accessed January 3, 2023).

Reynolds, C.: Highway 1 washout near Big Sur expected to be fixed by summer, Los Angeles Times (2021), https://www.latimes.com/travel/story/2021-02-25/highway-1-to-big-sur-will-reopen-by-summer-caltrans-says. (Accessed January 05, 2023).

Rickenmann, D. (1999). Empirical relationships for debris flows. Natural hazards, 19, 47-77.

Rico, M., Benito, G., Salgueiro, A. R., Díez-Herrero, A., & Pereira, H. G. (2008). Reported tailings dam failures: a review of the European incidents in the worldwide context. Journal of hazardous materials, 152(2), 846-852.

Riskslope (2022). Recent failure at Pau Branco Mine, MG, Brazil. https://www.riskope.com/2022/01/13/recent-failure-at-pau-branco-mine-mg-brazil/. (Accessed February 25, 2023).

Roback, K., Clark, M. K., West, A. J., Zekkos, D., Li, G., Gallen, S. F., ... & Godt, J. W. (2018). The size, distribution, and mobility of landslides caused by the 2015 Mw7. 8 Gorkha earthquake, Nepal. Geomorphology, 301, 121-138.

Robertson, P. K., De Melo, L., Williams, D. J., & Wilson, G. W. (2019). Report of the expert panel on the technical causes of the failure of Feijão Dam I. Commissioned by Vale, 81.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. ISPRS journal of photogrammetry and remote sensing, 67, 93-104.

Rogan, J., Franklin, J., Stow, D., Miller, J., Woodcock, C., & Roberts, D. (2008). Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. Remote Sensing of Environment, 112(5), 2272-2283.

Roy, P., Martha, T. R., Khanna, K., Jain, N., & Kumar, K. V. (2022). Time and path prediction of landslides using InSAR and flow model. Remote Sensing of Environment, 271, 112899.

Rudorff, N., Rudorff, C. M., Kampel, M., & Ortiz, G. (2018). Remote sensing monitoring of the impact of a major mining wastewater disaster on the turbidity of the Doce River plume off the eastern Brazilian coast. ISPRS Journal of Photogrammetry and Remote Sensing, 145, 349-361.

Salm, B. (1993). Flow, flow transition and runout distances of flowing avalanches.

Annals of Glaciology, 18, 221-226.

Sameen, M. I., & Pradhan, B. (2019). Landslide detection using residual networks and the fusion of spectral and topographic information. IEEE Access, 7, 114363-114373.

Scheidegger, A. E. (1973). On the prediction of the reach and velocity of catastrophic landslides. Rock mechanics, 5(4), 231-236.

Schraml, K., Thomschitz, B., McArdell, B. W., Graf, C., & Kaitna, R. (2015). Modeling debris-flow runout patterns on two alpine fans with different dynamic simulation models. Natural Hazards and Earth System Science, 15(7), 1483-1492.

Schraml, K., Uhlir, C., & Hübl, J. (2017). Natural Hazard Analysis for a Small Alpine Catchment in the Nepalese Himalayas. In Advancing Culture of Living with Landslides: Volume 4 Diversity of Landslide Forms (pp. 459-465). Springer International Publishing.

Segal, D. (1982, December). Theoretical basis for differentiation of ferric-iron bearing minerals, using Landsat MSS data. In Proceedings of symposium for remote sensing of environment, 2nd Thematic Conference on Remote Sensing for Exploratory Geology, Fort Worth, TX (Vol. 949, p. 951).

Shaghlil, N., & Khalafallah, A. (2018). Automating highway infrastructure maintenance using unmanned aerial vehicles. In Construction Research Congress 2018 (pp. 486-495).

Shirzadi, A., Solaimani, K., Roshan, M. H., Kavian, A., Chapi, K., Shahabi, H., ... & Bui, D. T. (2019). Uncertainties of prediction accuracy in shallow landslide modeling: Sample size and raster resolution. Catena, 178, 172-188.

Sim, K. B., Lee, M. L., & Wong, S. Y. (2022). A review of landslide acceptable risk and tolerable risk. Geoenvironmental Disasters, 9(1), 3.

Sohouenou, P. Y., Christidis, P., Christodoulou, A., Neves, L. A., & Presti, D. L. (2020). Using a random road graph model to understand road networks robustness to link failures. International Journal of Critical Infrastructure Protection, 29, 100353.

South Asia Network on Dams, Rivers and People (2020) Singrauli Fly Ash Dam Breach: Who regulates these dams in India? https://sandrp.in/2020/04/12/singruali-fly-ash-dam-breach-who-regulates-these-dams-in-india/ (Accessed February 27, 2023).

Stow, D., Niphadkar, M., & Kaiser, J. (2005). MODIS-derived visible atmospherically resistant index for monitoring chaparral moisture content. International journal of remote sensing, 26(17), 3867-3873.

Streeter, E. W. (1892). Precious stones and gems: Their history, sources and characteristics. G. Bell & Sons.

Stumpf A, Kerle N (2011) Object-oriented mapping of landslides using Random Forests. Remote Sens Environ 115:2564–2577.

Sun, J., Yang, J., Zhang, C., Yun, W., & Qu, J. (2013). Automatic remotely sensed image classification in a grid environment based on the maximum likelihood method. Mathematical and Computer Modelling, 58(3-4), 573-581.

Tarboton, D. G. (1997). A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water resources research, 33(2), 309-319.

Tehrani FS, Santinelli G, Meylin &, et al (2021) Multi-Regional landslide detection using combined unsupervised and supervised machine learning. Taylor Fr 12:1015–1038.

Tehrani, F. S., Calvello, M., Liu, Z., Zhang, L., & Lacasse, S. (2022). Machine learning and landslide studies: recent advances and applications. Natural Hazards, 114(2), 1197-1245.

Teillet, P. M., Guindon, B., & Goodenough, D. G. (1982). On the slope-aspect correction of multispectral scanner data. Canadian Journal of Remote Sensing, 8(2), 84-106.

Teughels, A., & De Roeck, G. (2004). Structural damage identification of the highway bridge Z24 by FE model updating. Journal of Sound and Vibration, 278(3), 589-610.

The World Bank (2021). Haiti Digital Terrain Model 2014 - 2016. Distributed by OpenTopography. https://doi.org/10.5069/G9GX48R8. (Accessed June 10, 2024).

Tien Bui D, Ho TC, Pradhan B, et al (2016) GIS-based modeling of rainfall-induced landslides using data mining-based functional trees classifier with AdaBoost, Bagging, and MultiBoost ensemble frameworks. Environ Earth Sci 75:.

Tien Bui, D., Shahabi, H., Shirzadi, A., Chapi, K., Alizadeh, M., Chen, W., ... & Tian, Y. (2018). Landslide detection and susceptibility mapping by airsar data using support vector machine and index of entropy models in cameron highlands, malaysia. Remote Sensing, 10(10), 1527.

Trigg, M. A., Dehghani, M. S., Kesete, Y. Y., Carr, A. B., Trigg, S. G., Zekkos, D., ... & Ogden, F. L. (2023). Realities of bridge resilience in Small Island Developing States. Mitigation and Adaptation Strategies for Global Change, 28(1), 1.

Tsunetaka, H., Asano, S., & Murakami, W. (2022). Do standing trees affect landslide mobility on forested hillslopes in Japan?. Earth Surface Processes and Landforms, 47(14), 3332-3347.

Twigg, J. (2004). Disaster risk reduction: mitigation and preparedness in development and emergency programming. Overseas Development Institute (ODI).

U.S. Geological Survey (USGS) (2021): Postfire Debris flow Hazards: Dolan Fire (Los Padres National Forest, CA), Landslide Hazards Program, USGS, https://landslides.usgs.gov/hazards/postfire_debrisflow/detail.php?objectid=3 12 (Accessed January 5, 2023).

Ullo, S. L., Mohan, A., Sebastianelli, A., Ahamed, S. E., Kumar, B., Dwivedi, R., & Sinha, G. R. (2021). A new mask R-CNN-based method for improved landslide detection. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 3799-3810.

USDA Forest Service. (2020). Dolan Fire Burned-Area Report, USDA Forest Service, Los Padres National Forest, California.

USGS Remote Sensing Phenology. (n.d.). NDVI, the Foundation for Remote Sensing Phenology, <https://www.usgs.gov/core-science-systems/eros/phenology/science/ndvi-foundation-remote-sensing-phenology?qt-science_center_objects=0#qt-science_center_objects> (Accessed January 5, 2023).

Vaghefi, K., Oats, R. C., Harris, D. K., Ahlborn, T. T. M., Brooks, C. N., Endsley, K. A., ... & Dobson, R. (2012). Evaluation of commercially available remote sensors for highway bridge condition assessment. Journal of Bridge Engineering, 17(6), 886-895.

Vapnik V (1999) The nature of statistical learning theory.

Villarroel, L. F., Miller, J. R., Lechler, P. J., & Germanoski, D. (2006). Lead, zinc, and antimony contamination of the Rio Chilco-Rio Tupiza drainage system, Southern Bolivia. Environmental Geology, 51(2), 283-299.

Wan, J., & Xue, X. (2023). Application of the group method of data handling (GMDH) approach for travel distance prediction of landslides. Landslides, 20(3), 645-661.

Wang, C., Yu, Q., Law, K. H., McKenna, F., Stella, X. Y., Taciroglu, E., ... & Cetiner, B. (2021a). Machine learning-based regional scale intelligent modeling of building information for natural hazard risk management. Automation in Construction, 122, 103474.

Wang H, Zhang L, Yin K, et al (2021b) Landslide identification using machine learning. Geosci Front 12:351–364.

Wang, Fu-Min, Jing-feng Huang, Yan-lin Tang, and Xiu-zhen Wang. "New vegetation index and its application in estimating leaf area index of rice." Rice Science 14, no. 3 (2007a): 195-203.

Wang, L., Chu, F., & Xie, W. (2007b). Accurate cancer classification using expressions of very few genes. IEEE/ACM Transactions on computational biology and bioinformatics, 4(1), 40-53.

Wang, H., Zhang, L., Wang, L., Fan, R., Zhou, S., Qiang, Y., & Peng, M. (2023). Machine learning powered high-resolution co-seismic landslide detection.

Gondwana Research, 123, 217-237.

Ward, P. J., Blauhut, V., Bloemendaal, N., Daniell, J. E., de Ruiter, M. C., Duncan, M. J., ... & Winsemius, H. C. (2020). Natural hazard risk assessments at the global scale. Natural Hazards and Earth System Sciences, 20(4), 1069-1096.

Whitehead, K., & Hugenholtz, C. H. (2014). Remote sensing of the environment with small unmanned aircraft systems (UASs), part 1: A review of progress and challenges. Journal of Unmanned Vehicle Systems, 2(3), 69-85.

Whittall, J., Eberhardt, E., & McDougall, S. (2017). Runout analysis and mobility observations for large open pit slope failures. Canadian Geotechnical Journal, 54(3), 373-391.

WISE Uranium Project (2023) Chronology of major tailings dam failures. https://www.wise-uranium.org/mdaf.html (Accessed February 27, 2023).

Woods Arnold (2019) "The 1957 Earthquake: A Closer Look". OpenSFHistory. (Accessed January 3, 2023).

Wright, Tom (2017) "Caltrans: Pfeiffer Canyon Bridge on Highway 1 in Big Sur beyond repair". Monterey Herald. (Accessed January 3, 2023).

Wu, C., Niu, Z., & Gao, S. (2012). The potential of the satellite derived green chlorophyll index for estimating midday light use efficiency in maize, coniferous forest and grassland. Ecological Indicators, 14(1), 66-73.

Wu, Y. H., Liu, K. F., & Chen, Y. C. (2013). Comparison between FLO-2D and Debris-2D on the application of assessment of granular debris flow hazards with case study. Journal of Mountain Science, 10, 293-304.

Xu, Y., Ouyang, C., Xu, Q., Wang, D., Zhao, B., & Luo, Y. (2024). CAS Landslide Dataset: A Large-Scale and Multisensor Dataset for Deep Learning-Based Landslide Detection. Scientific Data, 11(1), 12.

Xue, J., & Su, B. (2017). Significant remote sensing vegetation indices: A review of developments and applications. Journal of sensors, 2017.

Yang, L., Wei, Y., Wang, W., & Zhu, S. (2019). Numerical runout modeling analysis of the loess landslide at Yining, Xinjiang, China. Water, 11(7), 1324.

Yangon, C. D. (2020). Mining disasters continue in Myanmar despite regulations [WWW Document]. URL https://www.dw.com/en/mining-disasters-continue-in-myanmar-despite-regulations/a-54042106 (Accessed December 26, 2022).

Ye, C., Li, Y., Cui, P., Liang, L., Pirasteh, S., Marcato, J., ... & Li, J. (2019). Landslide detection of hyperspectral remote sensing data based on deep learning with constrains. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(12), 5047-5060.

Yilmaz C, Topal T, Süzen ML (2012) GIS-based landslide susceptibility mapping using bivariate statistical analysis in Devrek (Zonguldak-Turkey). Environ Earth Sci 65:2161–2178.

Yousefi, S., Pourghasemi, H. R., Emami, S. N., Pouyan, S., Eskandari, S., & Tiefenbacher, J. P. (2020). A machine learning framework for multi-hazards modeling and mapping in a mountainous area. Scientific Reports, 10(1), 12144.

Youssef AM, Pourghasemi HR, Pourtaghi ZS, Al-Katheeri MM (2016) Landslide susceptibility mapping using random forest, boosted regression tree, classification and regression tree, and general linear models and comparison of their performance at Wadi Tayyah Basin, Asir Region, Saudi Arabia. Landslides 13:839–856.

Yu B, Chen F, Xu C (2020) Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015. Comput Geosci 135:1–8.

Zekkos, D., & Clark, M. (2020). Characterization of landslides and rock mass strength leveraging the 2015 Mw 6.5 Lefkada Earthquake in Greece (No. G17AP00088). Final Technical Report. US Geology Survey. National Earthquake Hazards Reduction Program, Award.

Zekkos, D., Clark, M., Cowell, K., Medwedeff, W., Manousakis, J., Saroglou, H., & Tsiambaos, G. (2017, September). Satellite and UAV-enabled mapping of landslides caused by the November 17 2015 Mw 6.5 Lefkada earthquake. In Proc. 19th Int. Conference on soil mechanics and geotechnical engineering (pp. 17-22).

Zekkos, D., Clark, M., Gong, W., Manousakis, J., & Greenwood, W. (2019). Landslides Driven by Extreme Events: Can We Learn More from More of Them?. GeoStrata Magazine Archive, 23(3), 44-51.

Zekkos, D., Stark (2023) (editors). Highway 1 Rat Creek Embankment Failure: 2021 Reconnaissance and Analysis. Reston, VA: American Society of Civil Engineers.

Zhang, X., Li, L., & Xu, C. (2022a). Large-scale landslide inventory and their mobility in Lvliang City, Shanxi Province, China. Natural Hazards Research, 2(2), 111-120.

Zhang, T., Yin, Y., Li, B., Gao, Y., & Wang, M. (2022b). Characteristics of two complex rock and glacier avalanches and RAMMS simulation of a long-runout disaster in the Sedongpu Basin of the Yarlung Zangbo River downstream in October 2018, Tibet, China.

Zhao, B., Su, L., Wang, Y., Ji, F., Li, W., & Tang, C. (2021). Insights into the mobility characteristics of seismic earthflows related to the Palu and Eastern Iburi earthquakes. Geomorphology, 391, 107886.

Zhuang, Y., Xing, A., Leng, Y., Bilal, M., Zhang, Y., Jin, K., & He, J. (2021). Investigation of characteristics of long runout landslides based on the multi-source data collaboration: a case study of the Shuicheng Basalt Landslide in Guizhou, China. Rock Mechanics and Rock Engineering, 54(8), 3783-3798.

Zou, Z., Xiong, C., Tang, H., Criss, R. E., Su, A., & Liu, X. (2017). Prediction of landslide runout based on influencing factor analysis. Environmental Earth Sciences, 76, 1-17.