# Earthquake-Induced Hazards Modeling and Damage Mapping Using Computer Vision, Machine Learning, and Geospatial Data Analytics

A dissertation submitted by

Adel Asadi

In partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

In

Civil and Environmental Engineering

TUFTS UNIVERSITY

May 2023



Academic Advisor: Professor Laurie Gaskins Baise

# Abstract

Earthquakes can result in catastrophes in terms of loss of human lives and property damage. Post-disaster rapid response, rescue operations, and reconstruction needs adequate and timely information about the damaged areas. Thus, development of such post-event damage maps is essential. In addition, preventative disaster management and mitigation require knowledge of the susceptibility of earthquake-induced hazards which should be provided to policy makers and urban planners. Thus, efforts in modeling and predicting earthquake-induced hazards are consequential to save lives and properties. This dissertation used advanced machine learning, computer vision, geographical information systems, remote sensing, and geospatial data analytics to provide methods related to earthquake modeling in these two themes: 1) Post-earthquake damage mapping; and 2) Earthquake hazard modeling. Chapters 2, 3 and 4 fall within the first theme (damage mapping), and the scope of Chapter 5 is within the second theme (hazard modeling).

The focus of this dissertation is on modeling and mapping earthquake-induced landslides and soil liquefaction as two highly destructive surface damages happening as a result of ground shaking. When an earthquake occurs, the transmission of seismic waves can cause shaking and vibration of ground surface. This often trigger the collapse of potential landslide areas, based on some influential factors including geographical, hydrological and meteorological conditions of the landslide location. Soil liquefaction is a type of ground failure in which a granular soil is transformed from a solid state to a liquefied state because of increased pore pressure and reduced effective stress. Liquefaction occurrence also physically depends on some contributing factors including soil density, saturation, and earthquake load.

In summary: 1) Chapter 2, as a geospatial and image data fusion approach, proposes a post-event landslide mapping method based on optical remotely sensed imagery, and investigates the performance of the supervised learning-based landslide detection model by the addition of information including data of bi-temporal change and geospatial parameters to the image data; 2) Chapter 3, as an image processing approach, proposes a semi-supervised liquefaction detection method, and investigates the classification performance by the addition of image-derived information including color transformation, texture analysis, statistical indices, and dimensionality reduction outputs, to the post-event optical imagery; 3) Chapter 4, as a computer vision approach, proposes a deep transfer learning framework to develop a regional landslide mapping model, and investigates the cross-event transferability of the model; and 4) Chapter 5, and a geospatial model, proposes a system of voting machine learning classifiers to develop a regionally informed global liquefaction model to predict liquefaction hazards using a set of explanatory geospatial parameters by the aid of advanced machine learning algorithms.

The results of the dissertation suggest that: 1) Based on Chapter 2, Adding selected bi-temporal change detection and effective geospatial parameters to the color channels of the post-event imagery can help in improving the accuracy of classification model for landslide mapping; 2) Based on Chapter 3, Adding selected statistical indices and dimensionality reduction bands derived from the color imagery to the color channels of the post-event imagery can improve the accuracy of semi-supervised classification model for liquefaction mapping, and the semi-supervised approach works better than supervised learning due to the iterative self-training process leading to learn from unlabeled data as well as labeled data; 3) Based on Chapter 4, a regionally transferable landslide mapping model can be

developed by exploiting the power of deep transfer learning methods, and following some recommendations including the use of data augmentation techniques, and down-sampling of the post-event imagery if the resolution is considerably lower than the imagery the model has been trained on; and 4) Based on Chapter 5, The accuracy of global geospatial liquefaction methods can be increased by exploiting the power of different advanced machine learning classification algorithms, a higher number of explanatory geospatial parameters, and different models learning from data in global, coastal/non-coastal, and regional datasets as subsets of a geospatial liquefaction inventory of 53 earthquakes.

In the last chapter (Chapter 6), a short review of the slope monitoring techniques used in the open pit mining industry is provided. In summary, different commonly used slope monitoring techniques in the open pit mining industry was presented, in addition to the geomechanical, hydrological and geological considerations of pit slope design briefly. The instrumentations vary from surface to subsurface and airborne. Visual investigations are also a crucial part of the monitoring which is complementary to the remote sensing data.

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Civil and Environmental Engineering with specialization in Geosystems Engineering.

Department of Civil and Environmental Engineering, Tufts University

| | |
|---|---|
| Dissertation Advisor: | *Dr. Laurie G. Baise, Tufts University* |
| Committee Member: | *Dr. Babak Moaveni, Tufts University* |
| Committee Member: | *Dr. Abani Patra, Tufts University* |
| Committee Member: | *Dr. Magaly Koch, Boston University* |
| Committee Member: | *Dr. Snehamoy Chatterjee, Michigan Technological University* |
| Department Chair: | *Dr. Laurie G. Baise* |

## Acknowledgements

I would like to appreciate the great supervision, help and support by my kind advisor, Professor Laurie Basie, during the past three years of my doctoral studies at Tufts University. I can never appreciate enough for her trust and patience while I was going through this amazing learning process, building my skills, and developing what I have accomplished during my PhD study. It was a great honor for me to study under her supervision. I am also thankful to the whole team at the Tufts Geohazards Research Lab for all the help and support.

I would also like to appreciate my co-advisors - Prof. Magaly Koch, Prof. Babak Moaveni, and Prof. Snehamoy Chatterjee - who had a great impact on the output of my research by providing feedback and sharing their knowledge and experiences through numerous discussions and communications. Their expertise in their fields of research has significantly improved the methods presented in this dissertation. I would also Like to thank Prof. Abani Patra for serving as a committee member, and for providing valuable feedback to improve the content of the dissertation.

In addition, I would like to express my gratitude to the Department of Civil and Environmental Engineering at Tufts University, and the Graduate School of Engineering for supporting me financially through the PhD program. I would also appreciate the US Geological Survey (US Department of the Interior), and the National Geospatial Intelligence Agency (US Department of Defense), for funding my research studies. In addition, I would like to thank Freeport-McMoRan Inc. for the providing me with the amazing internship opportunities in the mining industry, which helped me understand the

practical aspects of what I was learning at Tufts and put me in a position to solve real-world problems and learn from highly experienced engineers.

A very important aspect of my doctoral study at Tufts has been my involvement in coursework, both as a Teaching Assistant, and as a student passing the courses. I can't emphasize enough on how great the learning experience has been. I would like to especially thank Prof. Magaly Koch for the amazing courses in the area of Remote Sensing, Prof. Juan Pestana for the excellent teaching of the Geotechnical Earthquake Engineering course, and Profs. Jonathan Lamontagne and Shafiqul Islam for the courses building my knowledge in Statistics and Data Analysis.

I also want to thank all individuals who helped me achieving this goal. My unlimited gratitude goes to my parents, family and friends for their encouragements, and for supporting me to pursue higher education, especially my smart brother and my lovely wife who motivated me the most to achieve my goals. I must especially appreciate my great uncle, Dr. Manouchehr Bagheri, for helping me pass extremely difficult situations.

# Table of Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Landslide Mapping by Fusion of Geospatial and Image Data

A series of earthquakes hit Kumamoto, Japan, over a period of two days in April 2016, which caused numerous landslides. In this study, high-resolution pre-event and post-event optical imagery, plus bi-temporal Synthetic Aperture Radar (SAR) data are paired with geospatial data to train a pixel-based machine learning classification algorithm using logistic regression to identify landslides occurred because of the Kumamoto earthquakes. The geospatial data used include a categorical variable (surficial geology), and continuous variables including elevation, slope, aspect, curvature, annual precipitation, and landslide probability derived by the USGS preferred geospatial model which incorporates ground shaking in the input parameters. Grayscale index change and vegetation index change are also calculated from the optical imagery and used as input variables, in addition to temporal differences in HH and HV amplitudes of SAR data. A detailed human-drawn landslide occurrence inventory was used as ground-truth for model development and testing. The selection of optimal features was done using a supervised feature ranking method based on the Receiver Operating Characteristic (ROC) curve. To weigh the benefit of combining different types of imagery, temporal change information and geospatial environmental indicators for landslide mapping after earthquakes, five different combinations of features were tested, and the results showed that adding data of selected geospatial parameters (landslide probability, slope, curvature, precipitation, and geology) plus selected change indices (grayscale index change, vegetation index change, and HV amplitude difference of

SAR data) to the imagery (post event optical) lead to the highest classification accuracy of 86.5% on class-balanced independent testing data.

## 1.2 Liquefaction Mapping by Semi-Supervised Learning and Image Processing

Soil liquefaction often occurs as a secondary hazard during earthquakes and can lead to significant structural and infrastructure damage. Liquefaction is most often documented through field reconnaissance and recorded as point locations. Complete liquefaction inventories across the impacted area are rare but valuable for the development of empirical liquefaction prediction models. Remote sensing analysis can be used to rapidly produce the full spatial extent of liquefaction ejecta after an event to inform and supplement the field investigations. Visually labeling liquefaction ejecta from remotely sensed imagery is a time-consuming task and prone to human error and inconsistency. This study uses a partially labeled liquefaction inventory and proposes a pixel-based approach to detect unlabeled liquefaction using advanced machine learning and image processing techniques. The research goal is to generate an augmented inventory of liquefaction ejecta with high spatial completeness using a set of partial labels created from visual annotations by experts. The proposed methodology is applied to aerial imagery taken from the 2011 Christchurch earthquake and considers the available partial liquefaction labels as high certainty liquefaction features. This study consists of two specific comparative analyses: 1) To tackle the limited availability of labeled data and their spatial incompleteness, a semi-supervised self-training classification via linear discriminant analysis is presented, and the performance of the semi-supervised learning approach is compared with supervised learning classification; 2) Post-event aerial image is used to extract color transformation

bands, statistical indices and texture components from the RGB (red-green-blue) channels, plus dimensionality reduction outputs, and performance of the proposed model with different combinations of selected features from these four groups are compared to find the optimal set of features. Building footprints are also used as the only non-imagery geospatial information to improve classification accuracy by masking out building roofs from the classification process. To develop the multi-class labeled data used for classification, regions of interest (ROIs) were drawn to collect samples of seven landcover and land use classes. The labeled samples of liquefaction were also clustered into two groups (dark and light) using Fuzzy c-means clustering algorithm to split the liquefaction pixels into two classes. Comparing the generated maps with fully and manually labeled liquefaction data, showed that the proposed semi-supervised method outperforms supervised learning, and can better augment the liquefaction labels across the image in terms of spatial completeness, when selected high-ranked features of the two groups of statistical indices (gradient weight and sum of the band squares) and dimensionality reduction outputs (first and second principal components of PCA analysis) are used.

## 1.3 Regional Landslide Mapping Model by Deep Transfer Learning

Landslides are a significant natural disaster typical of mountainous terrains that can be triggered by earthquakes and extreme rainfall events. Manual delineation of landslide features to create inventories by visual interpretations of optical remotely sensed imagery is a tedious and time-consuming process; therefore, post-disaster automatic detection of landslides is desirable. Deep convolutional neural networks are a state-of-the-art computer vision methods that can be used for landslide detection. A few recent studies use both pre-

event and post-event RGB channels in a change detection framework for detection of landslides or they include different sources of data, including digital elevation model (DEM) products. Since the performance of deep learning depends on the quality and quantity of training data, it is challenging to use such methods in a timely manner after the natural hazard events. Previous studies have not well investigated and reported the deep learning landslide mapping models' transferability from a training event to a testing event. In this study, the transferability of a trained deep convolutional network using the landslide inventory from the 2016 Kumamoto Earthquakes 6.6 Mw only based on post-event true-color optical imagery is analysed and a regional landslide mapping model is introduced. The landslide affected areas after the 2018 Hokkaido earthquake in Japan are used to test cross-event transferability. The landslide affected areas after the 2017 Asakura Rainfall event in Japan are used for a second cross-event transferability test with a different landslide triggering factor. The rationale for selecting these three case studies is the regional similarity of the events, all occurring in Japan, containing vegetated and forested mountainous areas, including urban settings, with comparable geology, topography and landcover. The presented deep transfer learning model uses a ResNet50-based DeepLabV3+ semantic segmentation method as an automated way to extract landslide features from post-event imagery instantly given the availability of the imagery without the need to fine-tune the model or to develop event-specific training data. High accuracy was achieved on the testing data from the 2016 Kumamoto Earthquakes (88.77 %), and on the case studies (unseen events) of the 2018 Hokkaido earthquake (85.75 %) and the 2017 Asakura Rainfall (86.84 %).

## 1.4 Global Liquefaction Hazard Modeling by a Voting Classification System

Data-driven geospatial liquefaction models are useful tools for real-time post-event impact and regional seismic hazard assessments. Geospatial liquefaction models are based on liquefaction occurrence inventories, widely available geospatial variables, and earthquake-specific parameters. This research uses an updated inventory with geospatial data from non-liquefaction and liquefaction occurrence locations in 53 earthquakes around the world, including categorical and continuous variables representing proxies for soil saturation, soil density, and earthquake loading. In this study, the performance of advanced machine learning (ML) algorithms in learning complex nonlinear patterns in the large dataset to predict liquefaction susceptibility, as an alternative approach to previously published logistic regression classifiers is evaluated. The proposed methodology starts with an exploratory data analysis and feature selection scheme on the sampled data across both the liquefaction and non-liquefaction classes to remove redundant and less relevant features, run data transformations, and perform statistical analysis to define some classification thresholds. The class and event imbalance issue is treated in an innovative manner by under-sampling large events and distributing the datasets over several balanced subsets. Instead of training a single classifier, a system of voting machine learning classifiers is designed to achieve higher accuracy and certainty, using different classification techniques to balance the performance, as not a single classification technique can be considered the best for all earthquake events. The voting classifiers use different portions of the data, representing global, coastal/non-coastal and regional information, to optimize the predictive power of the proposed model by taking the majority votes of six different

classifiers. The classifiers include three global classifiers, two coastal/non-coastal classifiers, and a regional classifier. For each data subset, the ML-based binary classification model is trained and validated via a K-fold cross-validation approach, and the voting system is tested based on the leave-one-out approach, excluding individual earthquake events one at a time for a fair accuracy, reliability, and applicability assessment. The final class assignment is generated by majority vote of the system of 6 voting classifiers trained by the individual data subsets. The results of this study are compared with a model developed using logistic regression to investigate the benefits and limitations of the proposed voting approach.

## 1.5 Slope Stability and Monitoring in Open Pit Mining Industry

In the last chapter (Chapter 6), a short review of the slope monitoring techniques used in the open pit mining industry is provided. In summary, different commonly used slope monitoring techniques in the open pit mining industry was presented, in addition to the geomechanical, hydrological and geological considerations of pit slope design briefly. The instrumentations vary from surface to subsurface and airborne. Visual investigations are also a crucial part of the monitoring which is complementary to the remote sensing data.

# 2 Pixel-based Classification Method for Earthquake-Induced Landslide Mapping Using Imagery, Geospatial Data and Temporal Change Information

Adel Asadi[1], Laurie G. Baise[1], Magaly Koch[2], Babak Moaveni[1], Snehamoy Chatterjee[3], Yusupujiang Aimaiti[2]

[1] Geohazards Research Lab, Department of Civil and Environmental Engineering, School of Engineering, Tufts University, Medford, MA 02155, USA

[2] Center for Remote Sensing, Department of Earth and Environment, Boston University, Boston, MA 02215, USA

[3] Geological and Mining Engineering and Sciences Department, Michigan Technological University, Houghton, MI 49931, USA

(The material contained in this chapter is submitted for review and possible publication in Natural Hazards Journal – Springer Publications)

## 2.1 Abstract

A series of earthquakes hit Kumamoto, Japan, over a period of two days in April 2016, which caused numerous landslides. In this study, high-resolution pre-event and post-event optical imagery, plus bi-temporal Synthetic Aperture Radar (SAR) data are paired with geospatial data to train a pixel-based machine learning classification algorithm using logistic regression to identify landslides occurred because of the Kumamoto earthquakes. The geospatial data used include a categorical variable (surficial geology), and continuous variables including elevation, slope, aspect, curvature, annual precipitation, and landslide probability derived by the USGS preferred geospatial model which incorporates ground shaking in the input parameters. Grayscale index change and vegetation index change are also calculated from the optical imagery and used as input variables, in addition to temporal differences in HH and HV amplitudes of SAR data. A detailed human-drawn landslide occurrence inventory was used as ground-truth for model development and testing. The

selection of optimal features was done using a supervised feature ranking method based on the Receiver Operating Characteristic (ROC) curve. To weigh the benefit of combining different types of imagery, temporal change information and geospatial environmental indicators for landslide mapping after earthquakes, five different combinations of features were tested, and the results showed that adding data of selected geospatial parameters (landslide probability, slope, curvature, precipitation, and geology) plus selected change indices (grayscale index change, vegetation index change, and HV amplitude difference of SAR data) to the imagery (post event optical) lead to the highest classification accuracy of 86.5% on class-balanced independent testing data.

## 2.2 Introduction

Landslides are a significant natural disaster that can be triggered by earthquakes and extreme rainfall events (Guzzetti 2006). Landslides tend to occur in mountainous areas and can be influenced by the local terrain, geology, and geomorphology of the area, as well as land use, and land cover (Cruden and Varnes 1996). As a result, there are numerous empirical models for landslides that use geospatial inputs such as slope, elevation, aspect, curvature, land cover, geology, and precipitation to predict the location of landslides (Godt et al. 2008; Nowicki et al. 2014, 2018). In addition, landslides' spatial extent (area) range in size from less than a few square meters, for shallow soil slides, to thousands of square kilometers, for large submarine failures (Cruden and Varnes 1996). Due to this scale, landslides are often easily visible in aerial and optical remotely sensed imagery. This has led to numerous landslide inventories developed using visual inspection of imagery or automated machine learning and computer vision approaches (Dou et al. 2015; Zhong et al. 2020). In this work, since an integrated approach can lead to improved performance,

these two related efforts are combined to determine relative accuracies and best practices for generating landslide maps using automated classification methods that take advantage of geospatial data as well as pre- and post-event imagery.

The spatial distribution of seismically induced landslides depends on geospatial variables related to earthquake shaking (epicenters, ground motion, and focal depth of the main shock), and environmental variables related to the geology, land use, hydrology, and topography (Sidle et al. 2005; Dai et al. 2011). The occurrence, extent, and volume of landslides are also influenced by land cover conditions (Reichenbach et al. 2014), including vegetation and surficial geology. In shallow soils, tree roots can penetrate below the potential slip surface, providing reinforcement and anchoring against sliding as well as removing soil moisture via transpiration at depth; whereas grass roots are typically shallow and weak, providing negligible reinforcement against landslides (Sidle and Bogaard 2016). Geological settings of the study area are influential in landslide susceptibility mapping (Segoni et al. 2020). The susceptibility ranking analysis by Parise and Jibson (2000) showed a clear distinction of susceptibilities of various geologic units to failure during seismic shaking. Volcanic deposits, specifically, are made of weak materials and contain landslide-prone soils and rocks (Reid et al. 2001; Frolova et al. 2015).

The U.S. Geological Survey (USGS) preferred landslide probability model, introduced by Nowicki Jessee et al. (2018), uses lithology and land cover as two of its input parameters alongside peak ground velocity, slope, and compound topographic index (wetness proxy). Several alternative global landslide probability models are published and used to map the landslide probability (Nadim et al. 2006; Hong et al. 2007; Stanley and

Kirschbaum 2017; Parker et al. 2017; Tanyas et al. 2019), including the ones recommended by USGS as alternative models (Godt et al. 2008; Nowicki et al. 2014), each using a combination of geospatial inputs. Most of the global-scale models were carried out via logistic regression as one of the earliest machine learning algorithms (Parker et al. 2017; Nowicki Jessee et al. 2018; Tanyas et al. 2019). Logistic regression is the most used, preferred, and recommended method for landslide mapping as well as susceptibility modeling (Fan et al. 2019; He et al. 2021; Sreelakshmi et al. 2022; Shao and Xu 2022). Recently, more advanced machine learning algorithms have also been used for global landslide susceptibility mapping (Stanley et al. 2021; He et al. 2021).

In addition to empirical models based on geospatial inputs, landslides can be mapped using image processing techniques that extract surface changes from optical images or synthetic aperture radar (SAR) intensity images (Bai et al. 2017; Ji et al. 2018). Remote sensing data has been widely exploited to create landslide inventories, to map landslides, and to perform landslide susceptibility analysis (Zhong et al. 2020; Casagli et al. 2023). High resolution imagery within multispectral (Martha et al. 2011; Hölbling et al. 2017; Mohan et al. 2021) and microwave (Solari et al. 2020; Mondini et al. 2021) domains of the electromagnetic spectrum are used to create landslide inventories. Current research studies focus on identifying landslides using available images by visual interpretation, supervised classification and change detection based on the temporal changes of those images (Savvaidis 2003; Wang et al. 2021). Change detection has been exploited for landslide mapping by analyzing the change in multi-temporal optical imagery (Ding et al. 2016; Yu et al. 2017; Huang et al. 2019) or high-resolution DEM data from LiDAR point clouds (Jaboyedoff et al. 2012; Liu et al. 2019).

Advanced machine learning-based classification and computer vision techniques have also been used in a wide range of geotechnical applications, including landslide identification (mapping) via remotely sensed data (Moosavi et al. 2014; Zhang and Goh 2016; Fanos et al. 2018; Tehrani et al. 2021; Mohan et al. 2021), and landslide susceptibility modeling (Kumar et al. 2017; Dou et al. 2020; Fang et al. 2020; Liu et al. 2021). However, complex machine learning classifiers are computationally expensive, time-intensive, more difficult to learn and interpret by the users, and also require a large number of high-quality training data (Zhong et al. 2020). In addition, these methods are prone to overfitting which means the trained model might not perform equally well when applied to independent testing data (Ma et al. 2021). Applying deep learning-based computer vision techniques for landslide mapping is also subject to additional unique limitations. They are not yet well-utilized for real-time post-event mapping, and they are supposed to perform the best when applied on RGB (Red, Green and Blue) channels, while satellite imagery used for landslide mapping generally has multiple additional bands which might be beneficial for the purpose of landslide mapping (Thirugnanam, 2023). The transferability of the trained model to other events might also be difficult and might need additional model parameters tuning.

Fang et al. (2020) assessed different machine learning algorithms and reported that overfitting has been an issue in their study. The large gap between the training and testing accuracy in the study by Xie et al. (2021) can also be indicator of overfitting as discussed in the study by Ado et al. (2022). Kalantar et al. (2018) have evaluated the performance of logistic regression in landslide susceptibility mapping and compared the accuracy with Support Vector Machines (SVM) and Artificial Neural Networks (ANN). The results of

their study showed higher accuracy derived from the logistic regression model. Xie et al. (2018) have also compared the performance of logistic regression and ANN in earthquake-induced landslides susceptibility mapping, and they have reported slight superior performance by the logistic regression model.

A few studies have been conducted to assess earthquake-induced damage using optical satellite imagery (Kim and Kim 2018; Xu et al. 2018), multi-temporal airborne LiDAR (Yamazaki et al. 2018; Konishi and Suga 2018; Liu et al. 2019) and SAR data (Nava et al. 2022) for landslide mapping after Kumamoto 2016 earthquake. UAV (Unmanned Aerial Vehicles) imagery-derived three-dimensional (3D) models for damage mapping (Yamazaki et al. 2017), and a combination of non-optical remote sensing data for surface displacement and landslide mapping (He et al. 2019; Uemoto et al. 2019) are also performed in the aftermath of the 2016 Kumamoto Earthquakes. Building damage detection using LiDAR data (Hajeb et al. 2020; Yamazaki et al. 2022), post-event aerial imagery (Zhan et al. 2022) and SAR imagery (Bai et al. 2017) have also been done for this event. Tamkuan and Nagai (2017) have also fused interferometric coherence data and optical imagery for building damage detection after Kumamoto 2016 earthquake.

This research is an attempt to showcase that logistic regression has the capability to classify images efficiently with a significant reduction in computational time and can also waive the tedious task of building and tuning complex models for image segmentation. The proposed pixel-based classification model is simple, fast to train, less prone to overfitting, and easily applicable for future events. Pixel-based analysis of the images for landslide detection performs slightly better and has been reported to provide better results

than object-based techniques (Sreelakshmi et al. 2022). Calculating the probability of each pixel belonging to the landslide makes the output of the model more reliable (Chen et al. 2017; Yu et al. 2017; Ullo et al. 2021). It is also difficult to apply an object-based method with the same set of parameters to different images, because the characteristics of different landslides and their surrounding backgrounds are mostly different (Yu et al. 2017).

In this work, empirical geospatial landslide modeling methods and image-based classification methods are merged to demonstrate how data fusion can improve landslide inventory development. The datasets used in this research are collected in the region impacted by landslides resulting from the 2016 Kumamoto earthquake series, including high resolution optical aerial and satellite imagery, SAR imagery and additional environmental geospatial parameters. Logistic regression is used as the machine learning classification algorithm for detecting and mapping seismically induced landslides on post-event aerial imagery. The proposed model's accuracy is quantified across different input features to determine which input features are the most effective and useful in developing pixel-based landslide classification for inventory development. The optimization of features was done via supervised feature ranking method based on the Receiver Operating Characteristic (ROC) curve. The hypothesis of the study is that visual identification of the landslides can be improved using data fusion with relevant temporal change indices, environmental geospatial characteristics and empirical geospatial models.

## 2.3 Datasets

The 2016 Kumamoto earthquakes in Kyushu Island, Japan, began with a magnitude 6.2 M earthquake on the Hinagu Fault on April 14, followed by another foreshock of magnitude

6.0 M on the Hinagu Fault on April 15, and a larger magnitude 7.0 M earthquake on the Futagawa Fault on April 16, 2016, at an epicentral depth of about 11 kilometers beneath Kumamoto City, Kumamoto Prefecture on Kyushu, Japan (GEER Report 2016). The mainshock had a focal depth of 10 km, and there were several co-seismic surface ruptures along the fault zone, so most of the landslides by the mainshock roughly tracked the area of the triggering fault (Shirahama et al. 2016).

The 2016 Kumamoto earthquakes were the strongest earthquakes ever recorded in the region, with large mass movements damaging properties, including houses, roads, and bridges, and even cultural heritage sites (GEER Report 2016; Japanese Water and Disaster Management Bureau 2016). The damages and loss of lives were largely due to the collapse of wooden buildings in Mashiki town and landslides in Minami-Aso village. Thousands of buildings collapsed, and lifelines, such as gas, electric power, and water, were also disrupted by the earthquake. Landslides and surface ruptures disrupted road and railway networks, making it impossible to reach the afflicted area quickly after the earthquake (Cabinet Office of Japan 2016; Liu et al. 2019).

Data collected for this study include a landslide inventory, high resolution optical imagery, SAR imagery, environmental geospatial characteristics, and imagery-derived change information. Each of the inputs and their derived products are discussed in the following subsections. Figure 1 shows the Kumamoto region with the location of the earthquake epicenters.

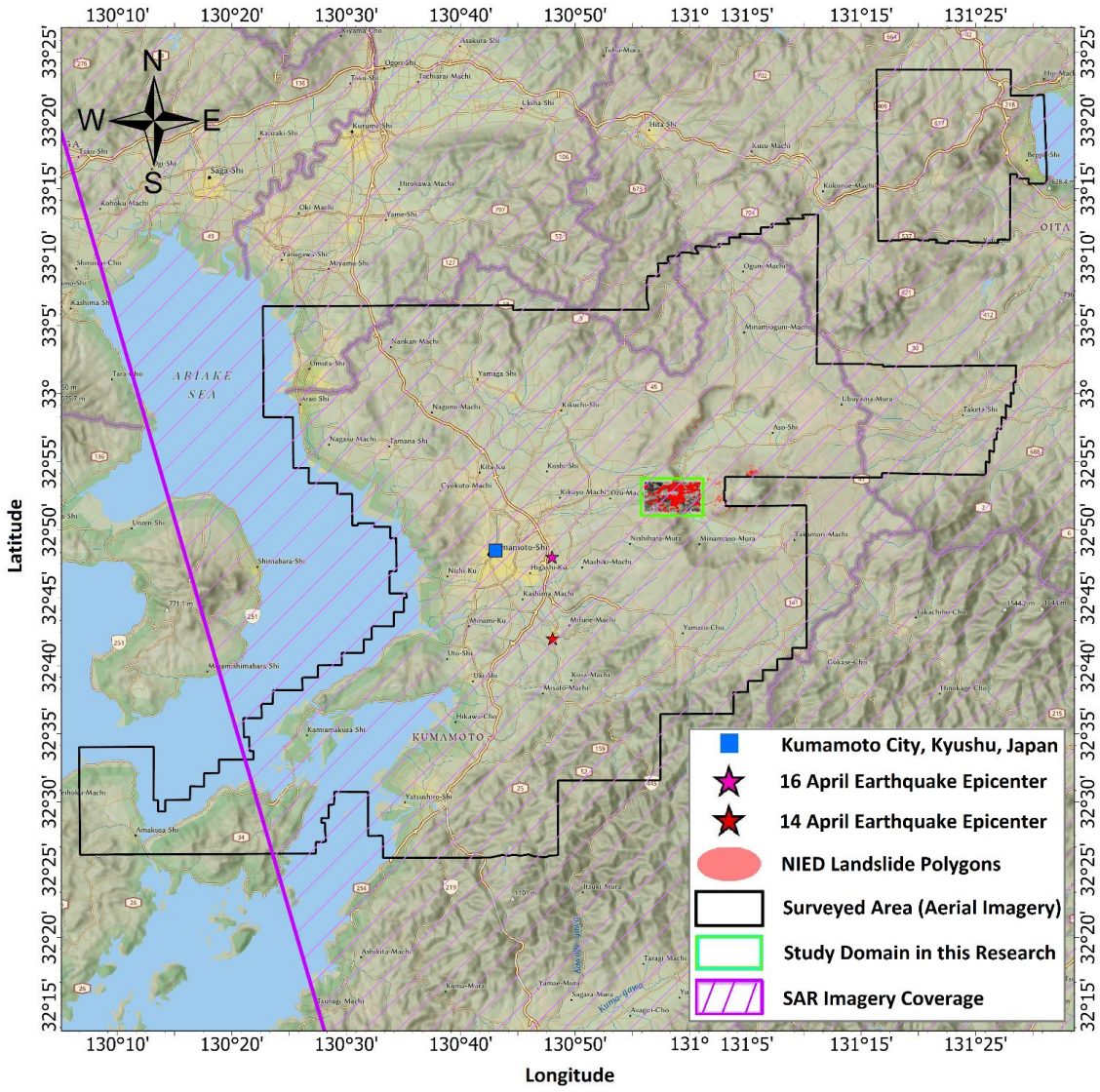



Figure 2-1. Area of study for Kumamoto 2016 earthquake's landslide mapping in this study (rectangular green area). The map shows the coverage of aerial imagery provided by GSI (Geospatial Authority of Japan), SAR data coverage by JAXA (Japanese Aerospace Exploration Agency), surveyed areas by NIED (National Research Institute for Earth Science and Disaster Resilience) for the landslide inventory used in this study, epicenters of the earthquakes sequence, and landslide features delineated by NIED (2016).

### 2.3.1 Landslide Inventory

The National Research Institute for Earth Science and Disaster Resilience (NIED) of Japan has released a few versions of landslide inventory for the 2016 Kumamoto events (NIED, 2016). NIED compared aerial photographs, satellite images, and airborne SAR images

before and after the earthquake centering on Kumamoto to identify the locations of sediment movement and create a sediment movement distribution map. In this study, the version published on April 21, 2016, is used which was created by comparing the aircraft SAR image on the 17$^{th}$ with the satellite image posted on Google Earth before the disaster, and by inspecting aerial imagery published by the Geospatial Authority of Japan (GSI) on April 16, 19 and 20, 2016 (GSI, 2016).

The 2016 Kumamoto earthquake caused mainly small to medium-sized landslides, with 47% of the landslide feature having area of less than $10^3$ m$^2$ (Chen et al. 2017b), consistent with many other earthquakes (Koyanagi et al. 2020; Xu et al. 2018). The distribution of the 264 landslide features used in this study is highly positively skewed. The area ranges from 23.4 m$^2$ to over 0.23 km$^2$ with a mean of 5373 m$^2$, standard deviation of 18726, and median value of 1049 m$^2$. The total area of the landslides is ~1.42 km$^2$. Figure 2 shows the area-frequency distributions of the landslide features used in this study in a logarithmic scale, plus the box plot of the area of the features. It shows that the landslide feature size is highly skewed towards zero, which means that most of the landslide features in the inventory are small, and large features are significantly less frequent.

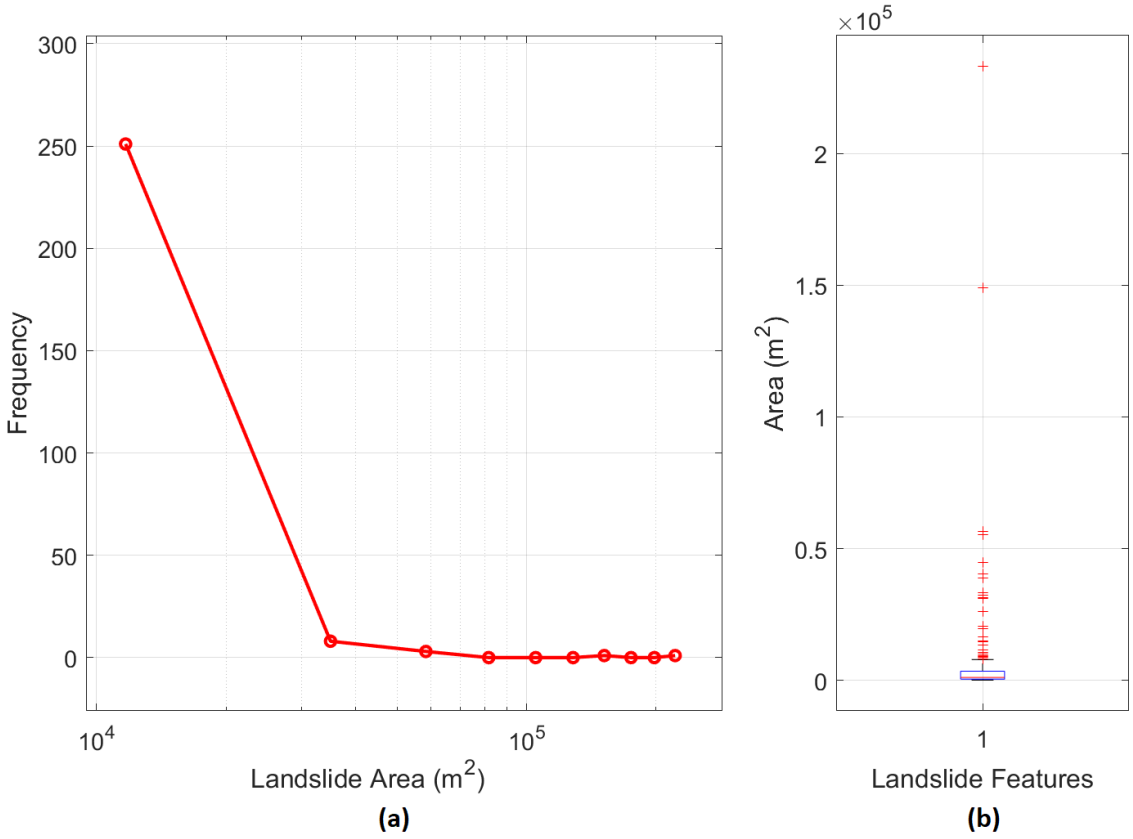



Figure 2-2. (a) Log-binned area-frequency plot of the distribution of the landslide features' area, generated considering 10 bins. (b) The box plot of the feature area data also shows a significant positive skewness. In the box plot, a box (in blue) is drawn from the first quartile to the third quartile of the area distribution, and a line (in red) goes through the box at the median value.

The NIED inventory maps were developed through visual inspection of the post-earthquake aerial imagery; however, the polygons are not assumed to match perfectly with the landslide areas, since the delineation has been a part of the rapid response process. The inventory does not cover the whole earthquake-affected area, and is human-drawn so has limited accuracy. Interpretation of the aerial photos is the most common method to identify and map landslides (Guzzetti et al. 2012). The results of the study by Carrara et al. (1992) showed that the mapping error decreases with the increasing resolution of imagery and the experience of the investigators. However, the identification and mapping of landslides

through visual investigation is still empirical and uncertain (Guzzetti et al. 2012). The NIED inventory is created based on the definition of the boundary of the landslide features as the area covering displacement source areas, travel paths, and accumulation areas (Shinoda et al. 2019). Due to reasons such as insufficient field surveys (NIED 2016), the inventory should not necessarily be considered as completely correct indicator of landslide features. NIED set a criterion of having width of higher than 100 m (moderate-size relic landslides) to identify and map landslide features (Chen et al. 2017b), thus, some moderate- or small-scale landslides were probably omitted (Xu et al. 2018).

Measuring the accuracy of landslide inventories is difficult as there are no standards on the procedure (Galli et al. 2008). The accuracy of landslide inventories depends on their completeness, and their geographical and thematic correctness (Guzzetti et al. 2012). Assessing the completeness depends on knowing the real complete map of landslides which is unknown most of the times. Specially, smallest landslides have the highest potential to be missed in inventories. A few studies have stated that their inventories are nearly complete, but with a threshold on the minimum area of covered landslide features (Harp and Jibson 1995; Malamud et al. 2004). Geographical correctness means the similarity of the mapped landslide features with the reality in the field in terms of shape, size, and area (Santangelo et al. 2010). Thematic correctness depends on the accuracy of the additional information provided with the inventory, including, but not limited to type, age, and depth of landslide features, for which the certainty information is rarely provided in literature (Guzzetti et al. 2012).

In order to match the resolution of post-event imagery, nearest neighbor interpolation approach was used to downscale the rasterized polygons in the inventory to generate the final raster file of the landslide labels used as the ground-truth label for classification modeling.

## 2.3.2 Imagery

The imagery used in this research include pre-event optical satellite imagery, post-event aerial color imagery, and pre- and post-event synthetic aperture radar imagery. The specifications of the datasets are provided in the following subsections. Then, the change information parameters derived by the temporal differences of the imagery are presented.

### *2.3.2.1 Optical Imagery*

Aerial images of the study area, taken on April 20, 2016, were collected from GSI repository. The post-event orthographic images taken near Nishihara 2 district have 0.5 m resolution. The pre-event multi-spectral optical images, taken on March 17, 2016, were used in this study were collected via GeoEye-1 satellite from DigitalGlobe (MAXAR) with 1.64 meters multispectral and 0.41 m panchromatic resolution. The spectral range of the RGB channels are: Blue: 450-510 nm; Green: 510-580 nm; and Red: 655-690 nm. The 1.64 m resolution pre-event image was pansharpened using the panchromatic band, and the resulting 0.41 m resolution image was resampled via bilinear method to convert to 0.5 m resolution to match the post-event image.

The cropped post-event image of the study area with 0.5 m resolution, which was considered as the reference for interpolation of the other parameters, covers an area of almost 42.71 $km^2$, consisting approximately 16.61 million image pixels., Figure 3 shows

the pre-event natural-color satellite image used for this study. Figure 4 shows the post-event aerial image area of our study, with landslide polygons overlayed.

The post-event aerial imagery has the highest resolution among the parameters used in this research, all the other parameters have been resampled via bilinear interpolation (for continuous variables) and nearest neighbor (for categorical variables) methods. Thus, it should be noted that all the parameters introduced and visualized in section 2, were interpolated to match the resolution of the post-event imagery before further processing. In addition, to co-register all the layers used in this study, the imagery and geospatial layers where all projected to the same coordinate system used for the post-event aerial imagery.

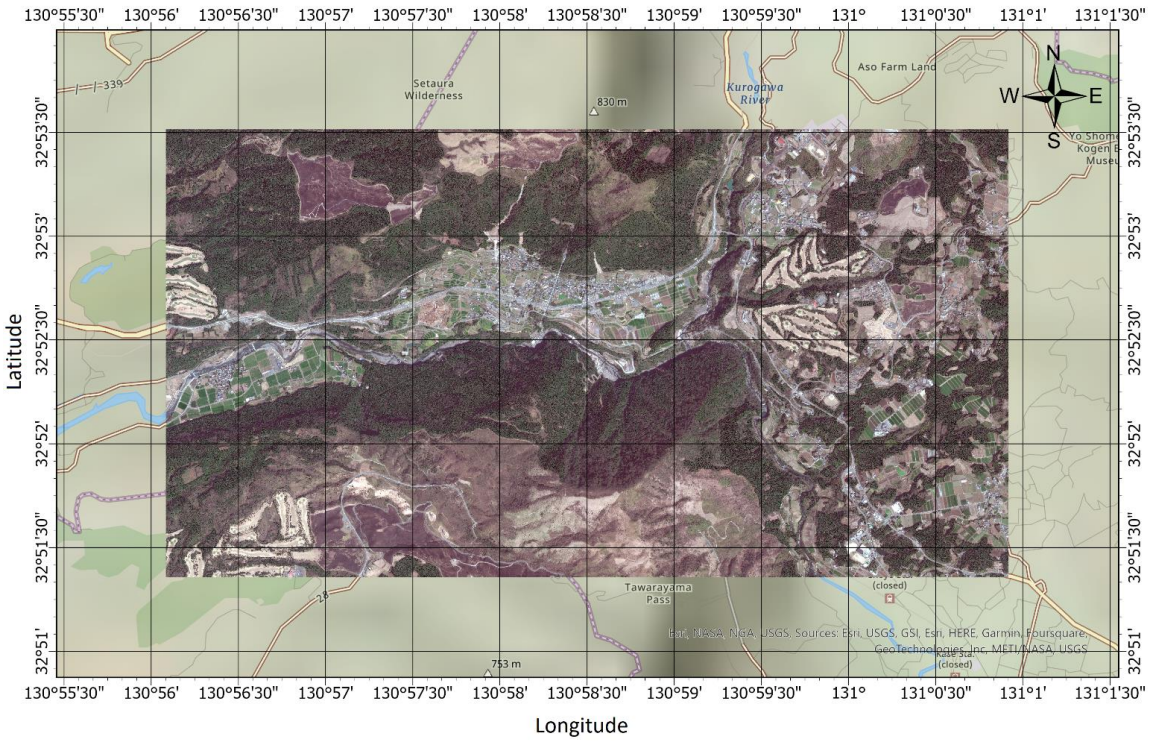



Figure 2-3. Pre-event natural-color GeoEye-1 satellite image (1.64 m resolution) provided by DigitalGlobe (2016). The image was further processed by pan-sharpening and resampling to match the resolution of the post-event aerial image (0.5 m).

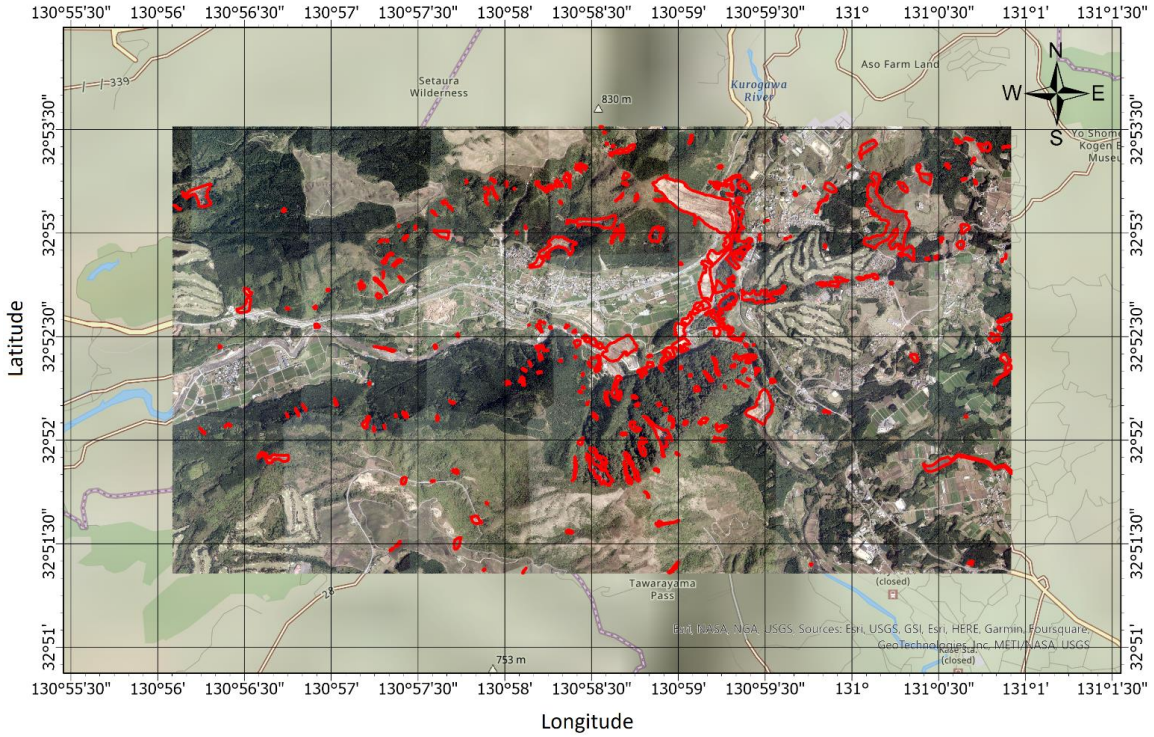



Figure 2-4. Post-event RGB aerial image with 0.5 m resolution, provided by GSI (2016), with NIED landslide polygons (red polygons) overlaying the image.

#### *2.3.2.2 Synthetic Aperture Radar*

Landslides can be mapped using image processing techniques that extract surface changes from synthetic aperture radar intensity images (Yamaguchi, 2012). In this research, L-band imagery from the PALSAR-2 sensor of the Advanced Land Observing Satellite 2 (ALOS-2) was acquired from the Japan Aerospace Exploitation Agency (JAXA, 2016) in ScanSAR Wide Beam Dual (WBD) mode with pixel spacing of 25 meters. The images are level 1.5 data which are multi-looked and geocoded. We collected the pre-event (01/26/2016) and the post-event (04/19/2016) ALSO-2 PALSAR-2 imagery to calculate the changes in HH and HV polarization bands. The datasets used in this study are made publicly available by JAXA (JAXA, 2016).

The geocoded pre- and post-event amplitude images were first filtered in ENVI using Enhanced Lee Filter (Lopes et al. 1990) with 3×3 pixels window size to reduce speckle in radar imagery while simultaneously preserving texture information. Filtered data was then calibrated based on the instructions provided by JAXA (Motohka et al. 2018) to convert digital numbers ($DN$) to the normalized backscattering coefficients (sigma naught) using Equation 1 (Shimada et al. 2009).

$$\sigma_0 = 10 \times \log_{10}(DN^2) - CF_1 \quad (1)$$

where $\sigma_0$ is the backscattering coefficient, DN is digital number of SAR amplitude images, and $CF_1$ is a calibration factor equal to -83 for Level 1.5 data. The unit of both $\sigma_0$ and $CF_1$ is dB.

### 2.3.3 Change Detection

Besides using raw post-event imagery, the temporal change from pre-event and post-event raw imagery can also be exploited in landslide mapping. In addition, processing the imagery to derive additional change indices and geospatial products can be useful in mapping the impacted areas. In this work, three change detection products are used, including grayscale index change and vegetation index change derived from the RGB optical imagery, and the temporal change products derived from the pre- and post-event SAR data which includes two amplitude difference bands of HH (horizontally transmitted and horizontally received polarization) and HV (horizontally transmitted and vertically received polarization). The change detection results of the two polarization bands of SAR data are visualized via Figure 5 in addition to the other two derived change indices which are presented through the following subsections.

#### 2.3.3.1 Vegetation Change

Vegetation indices are commonly used to highlight vegetation and are useful for the landslide identification as landslides often result in the removal of vegetation along the path of the landslide. When pre- and post-imagery is available, change detection applied to these indices can be used. There are several RGB-derived vegetation indices introduced in the literature (Sánchez-Sastre et al. 2020). In this study, vegetation indices of pre- and post-event imagery are derived using Equation 2 which has been used to distinguish vegetation and living plants from soil or residue background (Meyer et al. 2013).

$$Vegetation\ Index\ =\ (1.4\ \times\ Red)\ -\ Green \tag{2}$$

The vegetation index is calculated using the pre-event and post-event optical images. Vegetation index change is calculated by subtracting the pre-event vegetation index form the post-event vegetation index and the resulted interpolated difference is shown in Figure 5.

#### 2.3.3.2 Grayscale Change

The grayscale index allows the information across three bands to be combined into a single image input which may reduce redundant information (Kazmi et al. 2015) and can provide valuable compact information about the occurred change. The grayscale index is derived using Equation 3 (Burger and Burge 2010).

$$Grayscale\ =\ (0.2898 \times Red)\ +\ (0.5870 \times Green)\ +\ (0.1140 \times Blue) \tag{3}$$

The grayscale change is calculated by subtracting the pre-event from the post-event grayscale index, and the resulted interpolated difference is shown in Figure 5.

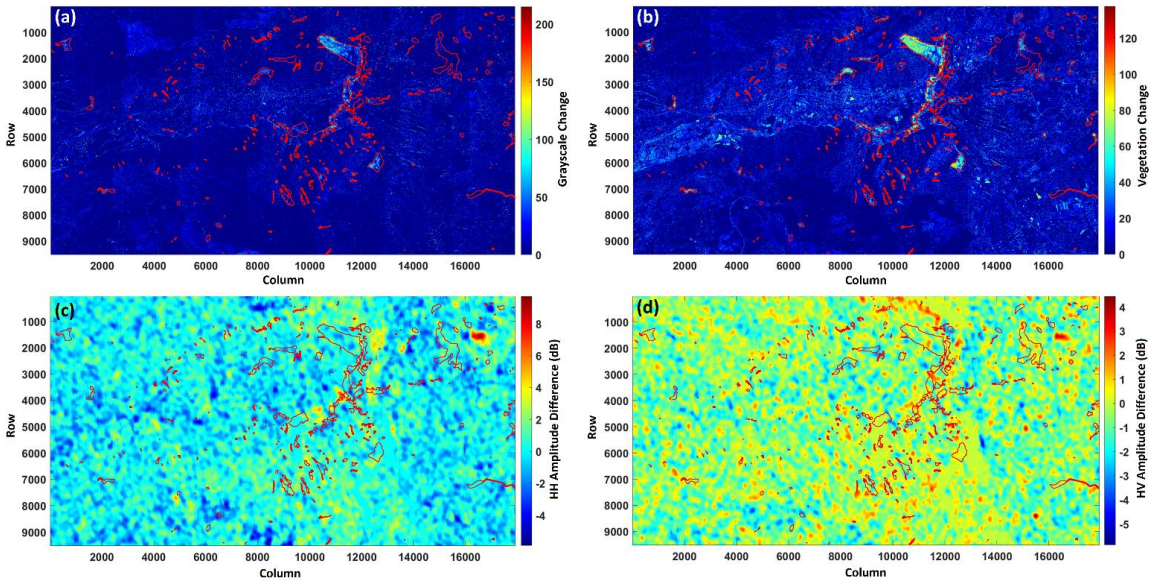



Figure 2-5. (a) Grayscale index change, (b) Vegetation index change, (c) HH amplitude difference, and (d) HV amplitude deference of SAR data. SAR data were interpolated via bilinear method to match the resolution of the optical post-event image. The grayscale and vegetation change indices have no units, and they are derived from RGB digital numbers in 0-255 range. All layers are shown with the landslide features overlayed. Rows and columns as axes titles are indicators of pixel number (pixel size is 0.5 m).

### *2.3.4* Environmental Geospatial Inputs

Seismically induced landslides are controlled by seismic, topographic, and geological factors (Xu et al. 2018). Environmental geospatial parameters that are commonly used in empirical landslide models are selected for use in the pixel-based landslide classification algorithm derived herein, as shown via Table 1. The geospatial inputs include categorical and continuous variables which are introduced in the following subsections. The interpolated version of the six continuous environmental geospatial variables used in the study are plotted in Figure 6. They include landslide probability, DEM, slope, aspect, curvature, and mean annual precipitation. The Box plot statistics of these variables are also shown via Figure 7.

Table 2-1. Parameters used in this study for landslide mapping modeling.

| Resource | No. | Variable | Type | Unit | Min | Max | Mean | Median | St. Dev. |
|---|---|---|---|---|---|---|---|---|---|
| Topographic | 1 | Digital Elevation | Geospatial Continuous | m | 141.00 | 806.00 | 423.03 | 422.00 | 125.34 |
| | 2 | Slope | Geospatial Continuous | deg | 0.00 | 55.18 | 13.77 | 10.98 | 9.55 |
| | 3 | Aspect | Geospatial Continuous | deg | 0.00 | 359.65 | 224.90 | 228.68 | 71.58 |
| | 4 | Curvature | Geospatial Continuous | $m^{-1}$ | -4.91 | 3.93 | 0.00 | -0.01 | 0.41 |
| Seismic | 5 | Landslide Probability | Geospatial Continuous | - | 0.00 | 0.22 | 0.09 | 0.08 | 0.05 |
| Geological | 6 | Precipitation | Geospatial Continuous | mm | 2033.00 | 2654.00 | 2315.50 | 2287.57 | 118.73 |
| | 7 | Geology | Geospatial Categorical | - | Number of Categories = 9 | | | | |

#### *2.3.4.1 Topographic Inputs*

The topographic inputs are derived from the Digital Elevation Model (DEM) for the globe, provided by The Ministry of Economy, Trade, and Industry (METI) of Japan and the United States National Aeronautics and Space Administration (NASA), released as the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 3 (GDEM-003, 2018). The resolution of the digital elevation data used is ~30 meters. The users are advised by the data providers that the DEM

data contains anomalies and artifacts which might reduce or eliminate effectiveness for certain applications.

Stefanescu et al. (2015) analyzed uncertainties in DEM datasets as inputs of probabilistic hazard models, and they concluded that the errors due to DEM data should be incorporated into the model error, and generating multiple realizations of the outputs, leading to ensemble maps, was suggested in their study to deal with such uncertainty. In this study, since the area of study is very small, compared to the global-scale 30 m resolution of the DEM data, the possible uncertainty due to DEM error is neglected.

Slope (derivative of the surface), Aspect (orientation of the slope) and Curvature layers were then generated using ArcGIS software deriving from the downloaded DEM layer. Curvature is indicator of the shape of the slope (concave versus convex) calculated by taking the second derivative of the surface.

According to the research done by Koyanagi et al. (2020), 59 percent of the landslides in the 2016 Kumamoto incident in Japan happened on slopes less than 30 degrees. Previous research in volcanic regions, relatively soft soils such as volcanic ash, and places underlain by poorly cemented rocks have found that more earthquake-initiated landslides typically occur on gentler slopes (Song et al. 2019; Koyanagi et al. 2020). The 2016 Kumamoto Earthquakes generated greater landslides than recent rainfall-induced landslides in the region, perhaps due to the distinct topographic trigger points as well as the strong ground motion (Koyanagi et al. 2020).

#### *2.3.4.2 Seismic Input*

Although they can be effective in landslide mapping, peak ground acceleration (PGA) and peak ground velocity (PGV) from USGS ShakeMap were not used because of the relatively small area of study compared to the ESRI raster's resolution (USGS ShakeMap). Instead, the USGS preferred landslide model with output resolution of ~250 m is used which is published through USGS Ground Failure products to be used after earthquakes (Nowicki Jessee et al. 2018; Allstadt et al. 2022).

The Nowicki Jessee et al. (2018) landslide model includes PGV, lithology, land cover, slope, and compound topographic index (CTI) as its input parameters. The model was developed by relating 23 inventories of landslides triggered by past earthquakes with different combinations of predictor variables using logistic regression (Nowicki et al. 2018; USGS ShakeMap; Allstadt et al. 2022). It is decided in this study to evaluate the USGS preferred landslide model as a potential input since it is a readily available product after earthquakes.

#### *2.3.4.3 Precipitation*

The mean annual precipitation data was downloaded from the WorldClim database version 2, which has monthly averages for precipitation from 1970-2000. The monthly averages were summed to get the mean annual precipitation data at a resolution of 30 arc-seconds (Fick and Hijmans 2017; WorldClim Database), and resampled to match the post-event aerial image resolution.

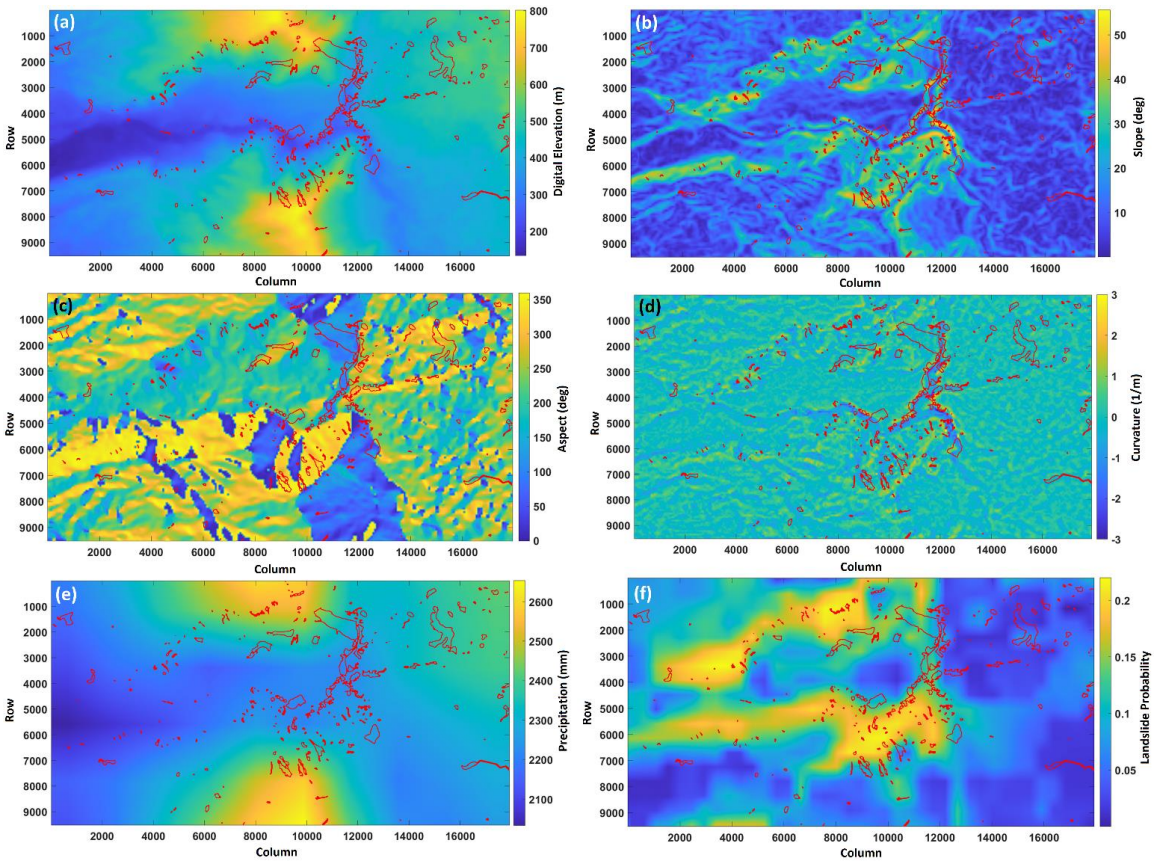



Figure 2-6. Plots of six of the interpolated continuous geospatial explanatory variables used in this study, including (a) digital elevation (b) slope, (c) aspect, (d) curvature (e) landslide probability, and (f) mean annual precipitation. All layers are shown with the landslide features overlayed. Rows and columns as axes titles are indicators of pixel number (pixel size is 0.5 m).

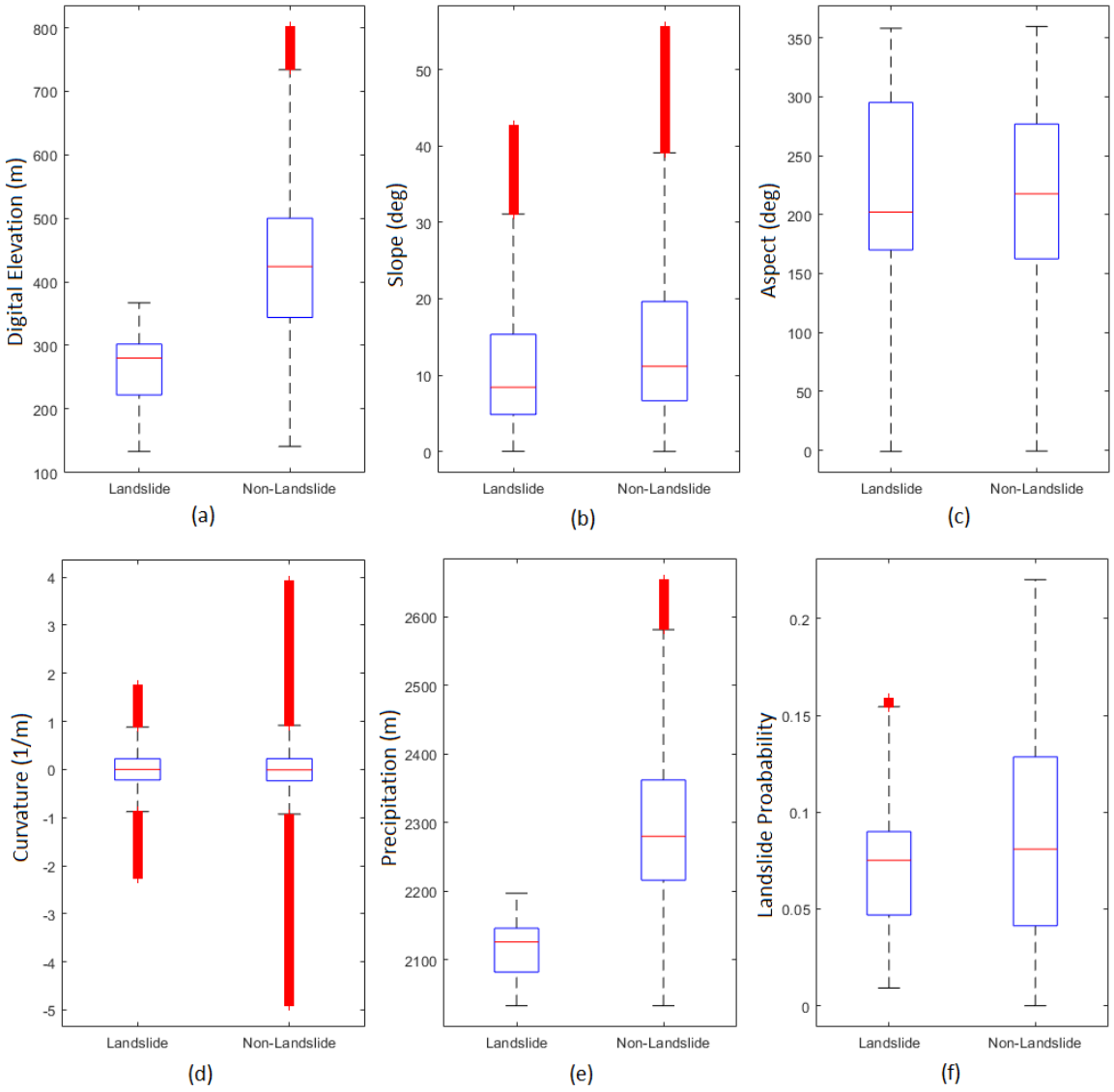



Figure 2-7. Box plots of the continuous environmental geospatial variables used, including (a) Digital elevation, (b) Slope, (c) Aspect, (d) Curvature, (e) Precipitation, and (f) Landslide probability.

#### *2.3.4.4 Surficial Geology*

The Seamless Digital Geological Map of Japan (1:200,000) provided by the Geological Survey of Japan and National Institute of Advanced Industrial Science and Technology, was used to generate the geology layer in this study (Geological Survey of Japan, 2015). Figure 8 shows the map outlining 9 different geological units in the study area. Although

“water” is not considered a geological category, it is considered as a geology class in this study, since the geology map used in this study has a class of waterbodies, and the intention is to use data with full coverage in the area of study, instead of removing the water class.

Many natural slopes and road embankments failed as moderate to large-scale landslides in the Aso Caldera area. Volcanic ash, pumice and andosol, a highly porous dark-colored substance made up of volcanic ash combined with organic materials, made up most landslide masses in that area (Kiyota et al. 2017).

To match the resolution of post-event imagery, nearest neighbor interpolation approach was used to resample the categorical variable (geology). Figure 8 shows the interpolated version of the categorical variable used in this study. Figure 9 also shows the bar plot of the variable to compare the histograms of the categories per the binary classes of landslide and non-landslide.

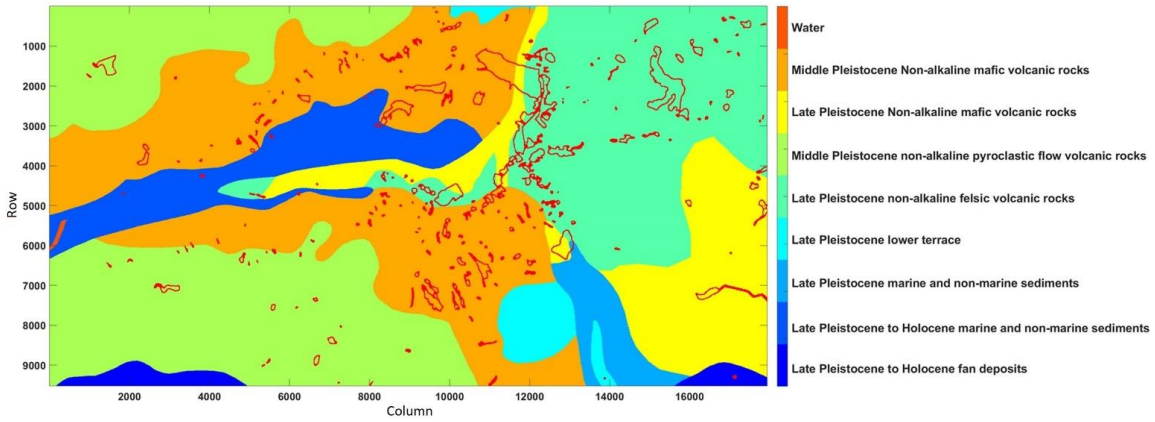



Figure 2-8. Seamless Digital Geological Map (1: 200,000) by the Geological Survey of Japan. The layer is shown with the landslide features overlayed the categories. Rows and columns as axes titles are indicators of pixel number (pixel size is 0.5 m).

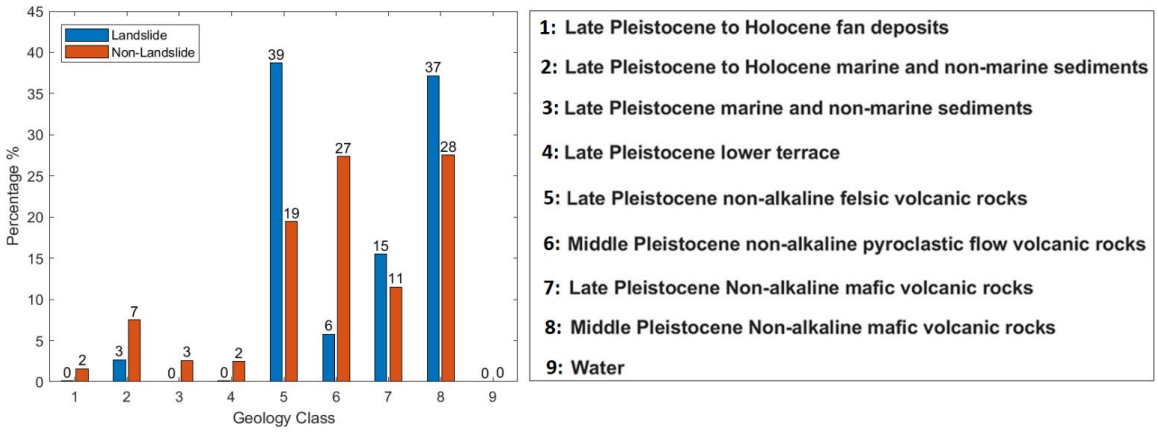



Figure 2-9. Comparative normalized bar charts of the geology parameter. Percentage labels on the graph are rounded to the nearest integer value. The percentage values are rounded to the closest integer, so the class of water (9) has very low number of samples in the map, but not absolutely zero.

## 2.4 Methodology

Given the available data described in Section 2, a methodology for feature selection followed by a pixel-based binary classification method is presented in this study. The feature selection step helps to determine which data inputs help the model to distinguish better between the two classes (landslide and non-landslide) without overfitting, since one way to reduce the overfitting is to reduce the dimensionality of the data. Feature analysis can, therefore, help us understand variable importance in the classification modeling.

In this study, landslide mapping is a binary classification problem, which means that the goal of developing a model is to predict each pixel as either positive (landslide) or negative (non-landslide) for landslide mapping. The flowchart of the proposed methodology is shown in Figure 10. In this section, data processing and modeling is presented via subsections starting with the data normalization as a pre-processing step and ending with model accuracy evaluation and post-processing steps.

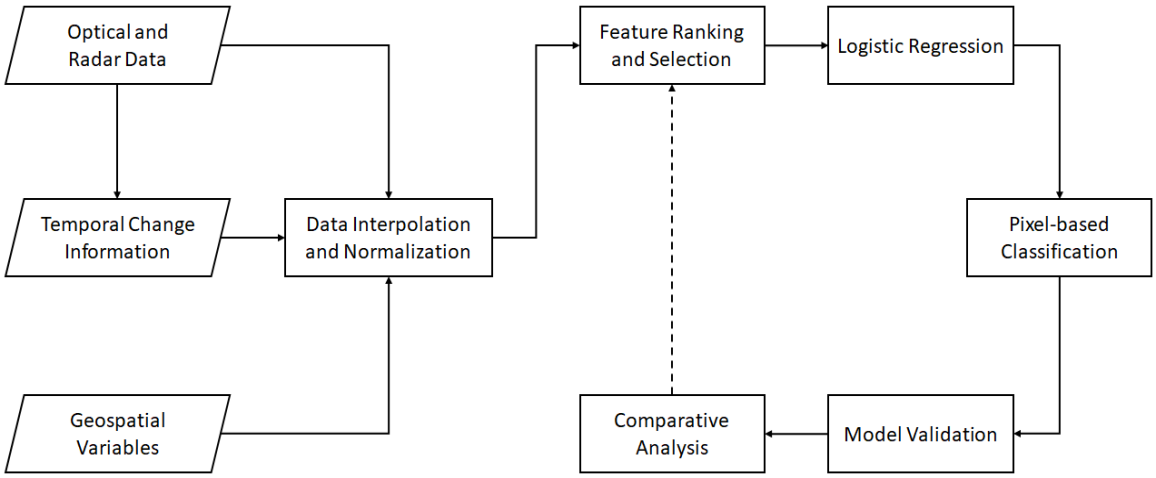



Figure 2-10. Flowchart of the proposed landslide mapping algorithm.

### 2.4.1 Data Normalization

After the resampling and stacking of the datasets, all the data layers were normalized as a pre-processing step. Since the geology variable is originally categorical, but categorical data is not desirable for the classification method used in this research, the geological categories are sorted by the percentage of landslide occurrence in the categories within the training data, and the percentage of landslide pixels is substituted with the category numbers. This leads to a meaningful set of numbers for the geology variable, where the higher number indicates a higher probability of landslide occurrences. Then, the geology feature is considered same as other continuous variables, and the variable is normalized as well.

Both the feature selection and classification models are developed based on the normalized data of the explanatory variables or data inputs, $\boldsymbol{X}$, which significantly reduces the impact of data range on the final landslide prediction model and increases the processing speed with fast convergence (Rojas 1996). Using the mean ($\mu$) and standard

deviation (σ) of individual explanatory variables, data of all variables were normalized using Equation 4 to have a mean of 0 and standard deviation of 1.

$$z_i = \frac{x_i - \mu}{\sigma} \quad (4)$$

### 2.4.2 Feature Ranking

Different combinations of features are used in the model validation step based on the feature ranking results, to perform a comparative analysis and to choose the best performing model based on the accuracy of the results for each combination.

The implemented algorithm ranks the features in data using an independent evaluation criterion for binary classification. Assume that *Z* is a matrix where every row is an observed vector ($Z_i$), and the number of columns correspond to the original number of features. The criterion for feature ranking is the area under the empirical receiver operating characteristic (ROC) or the Area Under the Curve (AUC) (Theodoridis 1999). The receiver operating characteristic (ROC) curve is a graphical plot that illustrates the diagnostic ability of the binary classification system. It is generated by plotting the true positive rate for a given classifier against the false positive rate for various thresholds. For a perfect classifier, AUC = 1. For a classifier that randomly assigns observations to classes, AUC = 0.5 (Fawcett 2004).

ROC curves and AUC scores also allow us to compare the performance of different classifiers for the same problem. The ROC curve is produced by calculating and plotting the true positive rate against the false positive rate for a single classifier at a variety of thresholds between 0 and 1. The optimal threshold leads to the highest AUC and highest

classification accuracy. At the feature ranking step in this research, each variable $Z_i$ is tested through various thresholds within its range (min-max) to calculate TPR and FPR for those thresholds, which are needed to generate the ROC curve for the feature.

### 2.4.3 Binary Classification Modeling

In this subsection, different aspects of the classification modeling are explained in detail.

#### *2.4.3.1 Logistic Regression*

The pixel-based binary classification step uses logistic regression algorithm to predict landslide and non-landslide pixels. Logistic Regression is a supervised machine learning algorithms which is commonly used for binary class classification problems. Logistic regression was borrowed by machine learning from the field of statistical models and is currently one of the most well-known and widely used machine learning algorithms (James et al. 2013). Logistic regression operates using a sigmoidal function for values ranging between 0 and 1. In the context of image processing, this could mean identifying whether a given pixel belongs to a particular class.

A logistic regression algorithm takes as its input a feature vector $Z_i$ and outputs a probability, $\hat{y}$=P(y=1|z), where the feature vector represents an object belonging to the class *y*. For image processing, the feature vector might be just the values of the red, green, and blue (RGB) channels for each pixel in the image, a one-dimensional array of real numbers formed by flattening the three-dimensional array of RGB pixel values. A logistic regression model is described by the logistic (sigmoid) function shown via Equation 5.

$$\sigma(w) = \frac{1}{1+e^{-w}} \tag{5}$$

and *w* contains the weights (coefficients) and the bias (intercept) of the linear equation. To calculate the coefficients in a way that the error is minimized, a loss error function is defined as shown via Equation 6, which describes how far the estimates are from actual observations.

$$L(\hat{y}, y) = \frac{(\hat{y}-y)^2}{2} \quad (6)$$

#### *2.4.3.2 Data Partitioning*

To train the classifier, it is recommended to separate data samples into distinct groups for model development and model testing, to avoid overfitting or underfitting. The difference in accuracies between the training and test sets can indicate whether a model is overfitting. The model is overfitting when the difference is large and the accuracy on the training set is very high. Underfitting, on the other hand, occurs when a model is not sophisticated enough to precisely capture relationships between predictors and responses, resulting in low accuracy on both training and test sets.

In this study, a special sampling process is designed to create model development and testing data. The area of study was split into 10-by-20 grid regions (100 model development and 100 model testing regions). The resulting dataset was significantly unbalanced with only 3.32% of pixels belonging to the landslide class. Class imbalance can cause overfitting in the classification algorithm as the model is rewarded by only predicting the majority class. To balance the datasets, only 10% of landslide points in the model development grid regions were sampled and added to only 0.4% of non-landslide pixels in those grid regions to form the sampled and balanced model development dataset.

Based on the sampling strategy, the model development data of the landslide class contains only 5.67% of the total pixels assigned as landslide pixels in the study area, which means that the final model has not ever seen 94.33% of ground-truth landslide samples. The reason for the sampling process is to highlight the fact that no significant amount of data is needed to train the algorithms. This will also be helpful in a trustworthy validation step by considering most pixels in the final predicted landslide map of the study area as new independent predictions. In addition, results of the study by Rashidian et al. (2020) on pixel-based Earthquake-induced liquefaction modeling showed that although increasing the number of training pixels increases the accuracy and reliability of the classifier, the improvement will not be meaningful after reaching a specific number of pixels (2000 pixels recommended in their study).

Figure 11 shows the ground-truth label based on the NIED landslide inventory, and the regional distinction for sampling model development and testing data, which has been implemented to avoid model bias. As shown in the Figure, model development data points have been sampled from blue boxes, whereas the testing points were sampled from the gray boxes. The testing dataset was also randomly sampled to create a balanced testing dataset consisting of all landslide pixels plus the same number of pixels from the non-landslide class (3.92% of available non-landslide pixels in the testing grid regions).

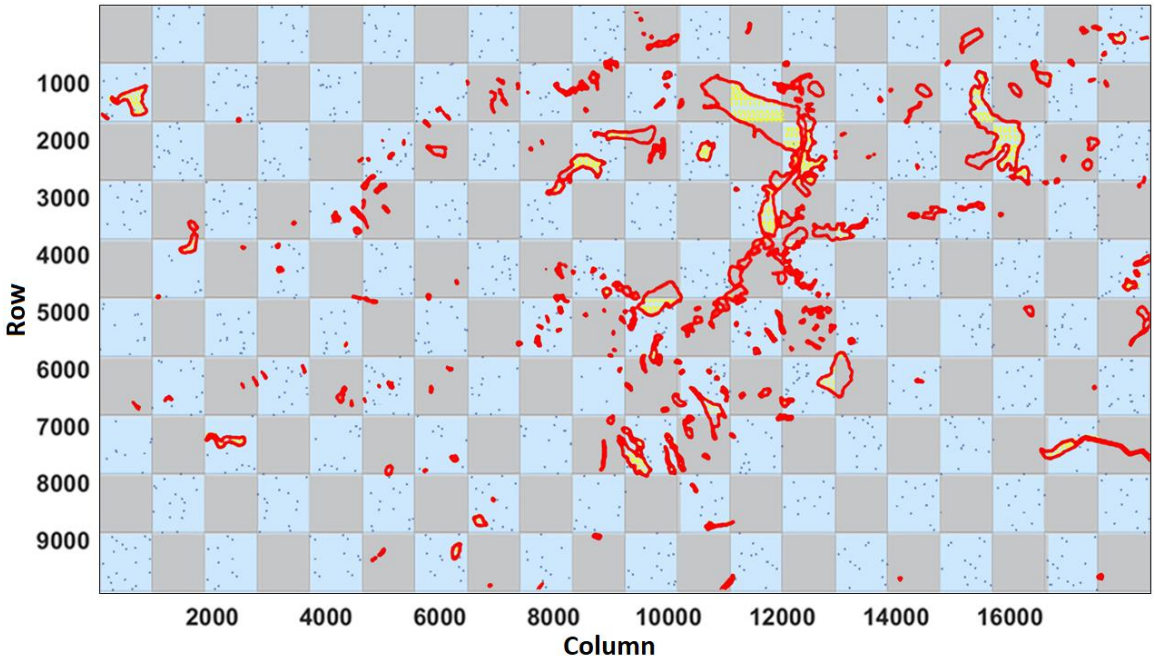



Figure 2-11. Ground-truth binary landslide label, with distinct model development and testing regions. Model development regions are shown via light blue, and testing regions are shown via gray color. Yellow dots are the sampled training data for the landslide class, and blue dots are the sampled training data for the non-landslide class. Rows and columns as axes titles are indicators of pixel number (pixel size is 0.5 m). The sample points are schematical, and are not real indicator of the sampled data, since the points are very small, and needed to be enlarged for visualization. The number of sampled data is higher that visualized points in this figure.

#### *2.4.3.3 Model Evaluation*

After random sampling of the model development data, the collected samples are divided into two groups of training and validation. The validation process was implemented by hold-out validation of 25% of the sampled and balanced model development data to help the training process. The validation accuracy is reported for each of the evaluated case models. Then, the distinct testing data samples from testing grid regions are used to test the predictive ability of the classification model on an independent unseen dataset, and to calculate the accuracy indices explained below.

There are several classification accuracy indices which were used in this study to make sure the model performs well. The proportion of correctly classified observations per positive class is referred to as the Sensitivity or True Positive Rate (TPR). Specificity or True Negative Rate (TNR) is the rate of correct prediction in the negative class. The sensitivity and specificity formulas are provided via Equations 7 and 8. The proportion of incorrectly classified observations per true class is referred to as the False Negative Rates (FNR). The Area Under Curve (AUC) curve depicts the true positive versus false positive rate. The AUC number is a measure of the overall quality of the classifier. Higher AUC (maximum is 1) values indicate better classification performance.

$$Sensitivity = TPR = \frac{TP}{TP+FN} \tag{7}$$

$$Specificity = TNR = \frac{TN}{TN+FP} \tag{8}$$

The proportion of correctly classified observations in the positive (landslide) class is referred to as the Positive Predictive Values (PPV) or precision. A high precision indicates that the model has a high probability of correctly classifying positive samples. The recall or Negative Predictive Value (NPV) is proportion of correctly classified observations in the negative (non-landslide) class. While the recall explains how sensitive the model is to identifying actual positive samples, it also quantifies the probability of detecting actual positive samples. Because precision and recall assess different aspects of the model, an index that combines the two is also used. The F1 score is the harmonic mean of the precision and recall, where an F1 score reaches its best value at 1 and worst at 0. The precision, recall and $F_1$ score formulas are provided via Equations 9 to 11.

$$Precision = PPV = \frac{\mathrm{TP}}{\mathrm{TP+FP}} \tag{9}$$

$$Recall = NPV = \frac{\mathrm{TN}}{\mathrm{TN+FN}} \quad (10)$$

#### *2.4.3.4 Landslide Map Post-Processing*

After generating the final landslide maps, to remove the noise and smooth the maps, two-dimensional median filtering was performed on the generated maps, which substitutes isolated pixels with the dominant class around them. Five rounds of smoothing are performed to avoid pixelated landslide maps, and to smooth the boundaries of the features. The window size selected for the post-processing step is 21×21 pixels in this study as a moderate size window as a trade-off to isolate and filter noise pixels while not losing very small-size landslide features. To analyze the impact of post-processing on the landslide map in terms of accuracy statistics, the generated map is compared with the ground-truth label, and the map overall accuracies before and after the smoothing are reported.

#### *2.4.3.5 Feature Importance Analysis*

Feature importance is also assessed by multiplying the coefficients of the logistic regression model by the standard deviation of the corresponding continuous parameters. Since all variables were normalized in the pre-processing step, having standard deviation of 1, the coefficients of the logistic regression model themselves are also the indicators for the feature importance in this study.

## 2.5 Analysis Results

In this section, the results of feature ranking step and the binary classification modeling are provided. Model validation is performed, and the visualizations are provided plus a discussion of the methods and outputs.

### 2.5.1 Feature Selection

In this study, it was decided to keep all RGB color channels of the post-event imagery active in the training process for the classification models. Feature ranking of the geospatial features and the temporal change features was performed to evaluate different combinations. The goal of performing this step is to select high-weight features per the categories of change information and geospatial information. The supervised feature ranking results via ROC calculation are shown in Table 2. The weights of features in the landslide and non-landslide classes show that the two change bands of vegetation difference and grayscale change have the highest impact on the classification as compared to the other variables. The third high-weight band in the change information category is HV amplitude difference. Geology, slope, precipitation, curvature, and landslide probability are the other high-weight geospatial features. The criteria to select the high-weight features was to have a weight of at least 0.55 in the feature ranking step (Table 2). It should be noted that having a weight of 0.5 is indicator of a random classifier.

Table 2-2. Feature ranking results by ROC scores for the continuous variables.

| Rank | Feature Description | Feature Type | Weight |
|---|---|---|---|
| 1 | Grayscale Change | Change | 0.792 |
| 2 | Vegetation Change | Change | 0.791 |
| 3 | Geology | Geospatial | 0.783 |
| 4 | Slope | Geospatial | 0.669 |
| 5 | Precipitation | Geospatial | 0.582 |
| 6 | Curvature | Geospatial | 0.580 |
| 7 | HV Amplitude Difference | Change | 0.565 |
| 8 | Landslide Probability | Geospatial | 0.562 |
| 9 | Digital Elevation | Geospatial | 0.542 |
| 10 | HH Amplitude Difference | Change | 0.529 |
| 11 | Aspect | Geospatial | 0.524 |

### 2.5.2 Classification and Model Evaluation

To evaluate how data fusion can aid in the development of landslide inventories from post-event imagery, different subsets of features were created with increasing complexity to test the performance of the logistic regression classification algorithm. All models used the post-event RGB imagery, including Model 1 developed solely via RGB channels. Next, a model that adds the optimal temporal change indices was tested (Model 2 = RGB + Change). Next, a model that adds the optimal geospatial features was tested (Model 3 = RGB + Geospatial). Next, a model that adds both optimal change indices and optimal geospatial features was tested (Model 4 = RGB + Change + Geospatial). Then, Model 5 was tested via all available data (14 bands).

Table 3 provides the classification accuracy results of different base models. Based on the overall classification accuracy of the model applied on the testing data, as shown in the table, the recommended model is Model 4 which uses RGB, selected change indices (grayscale change, vegetation change, and HV amplitude difference), plus selected geospatial features (geology, slope, precipitation, curvature, and landslide probability). The proposed model 4 outperforms other models for most of the accuracy indices, especially the overall testing accuracy (86.5%). The color imagery alone (Model 1 = RGB), proved to have the worst performance among the models, relatively. Although, it is observed that all models achieved a high level of learning and predictive performance, with some variations related to the different inputs fed to the algorithm.

Table 2-3. Classification performance results in different case models. (* is indicator of selected high-ranked feature)

| Case Model | | | Overall Classification Accuracy % | | Area Under Curve (AUC) | | Non-Landslide Testing Accuracy % | | Landslide Testing Accuracy % | |
|---|---|---|---|---|---|---|---|---|---|---|
| No. | Variable Categories | Channel Number | Validation | Testing | Validation | Testing | TNR | NPV | TPR | PPV |
| 1 | RGB | 3 | 85.00 | 82.48 | 0.92 | 0.89 | 82.85 | 82.04 | 82.11 | 82.92 |
| 2 | RGB<br>Change* | 6 | 86.39 | 83.90 | 0.93 | 0.91 | 85.08 | 82.33 | 82.79 | 85.47 |
| 3 | RGB<br>Geospatial* | 8 | 87.95 | 86.13 | 0.94 | 0.93 | 87.06 | 84.99 | 85.25 | 87.29 |
| 4 | RGB<br>Change*<br>Geospatial* | 11 | 88.63 | 86.50 | 0.95 | 0.93 | 87.99 | 84.63 | 85.11 | 88.38 |
| 5 | All Channels | 14 | 88.64 | 86.44 | 0.95 | 0.93 | 87.91 | 84.59 | 85.06 | 88.30 |

Table 4 shows the specifications of the logistic regression models developed for each of the five case models with acceptable performances. The information includes the intercept constant value and the coefficients of the different variables used within each model. It should be noted again that the models were trained and developed using normalized data. Model 4 (the preferred Model), as well as Models 3 and 5, have higher prediction accuracy, compared to Models 1 and 2, which can be the result of using geospatial parameters, especially the geology variable, which showed very high weights in the feature ranking step. Among the Models 3, 4 and 5, the Model 3 has the lowest accuracy, probably because of not exploiting the change information. Among Models 4 and 5, the preferred Model 4 has the superior performance, with slightly better accuracy results, compared to Model 5,

which can be attributed to optimal feature selection in Model 4, compared to redundant features in Model 5, leading to overfitting. The model parameters show small estimation uncertainty based on the standard errors (SE) reported via Table 4, and all selected variables are statistically significant based on the t-statistic for a two-sided test with the null hypothesis that the coefficient is zero, and the near-zero associated p-values of the hypothesis test.

Table 2-4. Binomial Logistic Regression model specifications. “SE” in the table stands for standard error, and “Coeff.” stands for model coefficient.

| Variables | | Model 1 (Base RGB) | | Model 2 | | Model 3 | | Model 4 (Preferred) | | Model 5 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| No. | Name | Coeff. | SE | Coeff. | SE | Coeff. | SE | Coeff. | SE | Coeff. | SE |
| 1 | *Intercept* | -1.756 | 0.006 | -1.909 | 0.006 | -2.249 | 0.007 | -2.330 | 0.007 | -2.337 | 0.007 |
| 2 | Red | 9.313 | 0.023 | 8.701 | 0.025 | 9.737 | 0.026 | 9.202 | 0.029 | 9.225 | 0.030 |
| 3 | Green | -12.583 | 0.031 | -12.156 | 0.033 | -11.956 | 0.033 | -11.635 | 0.035 | -11.569 | 0.035 |
| 4 | Blue | 3.468 | 0.018 | 3.160 | 0.019 | 2.558 | 0.020 | 2.356 | 0.021 | 2.278 | 0.022 |
| 5 | Grayscale Change | - | - | 0.354 | 0.006 | - | - | 0.378 | 0.007 | 0.371 | 0.007 |
| 6 | Vegetation Change | - | - | 0.350 | 0.006 | - | - | 0.218 | 0.007 | 0.227 | 0.007 |
| 7 | HH Amp. Difference | - | - | - | - | - | - | - | - | -0.115 | 0.004 |
| 8 | HV Amp. Difference | - | - | 0.193 | 0.004 | - | - | 0.146 | 0.004 | 0.179 | 0.005 |
| 9 | Geology | - | - | - | - | 0.693 | 0.005 | 0.659 | 0.005 | 0.664 | 0.005 |
| 10 | Landslide Probability | - | - | - | - | 0.230 | 0.006 | -0.334 | 0.006 | -0.342 | 0.006 |
| 11 | Slope | - | - | - | - | 0.951 | 0.005 | 0.889 | 0.005 | 0.903 | 0.005 |
| 12 | DEM | - | - | - | - | - | - | - | - | -0.160 | 0.014 |
| 13 | Aspect | - | - | - | - | - | - | - | - | 0.068 | 0.005 |
| 14 | Curvature | - | - | - | - | -0.250 | 0.004 | -0.248 | 0.004 | -0.235 | 0.004 |
| 15 | Precipitation | - | - | - | - | -0.311 | 0.005 | -0.258 | 0.005 | -0.092 | 0.015 |

The ROC curves showing testing accuracy of the preferred classification Model, as well as other case Models are shown via Figure 12. According to the ROC curve analysis derived from the testing phase, the Models 1 to 5 show good performance with the best accuracies achieved by Models 4 and 5. Considering landslide class as the positive class, the preferred Model 4 had a TPR of 85.1%, TNR of 88%, Precision of 88.4 %, Recall of 84.6 %, and $F_1$ score of 86.46 %, which indicates a balanced high accuracy of prediction in both classes. Investigating at the validation and testing accuracy indices in Table 4, no sign of either overfitting or underfitting was observed, as the validation and testing accuracies were close.

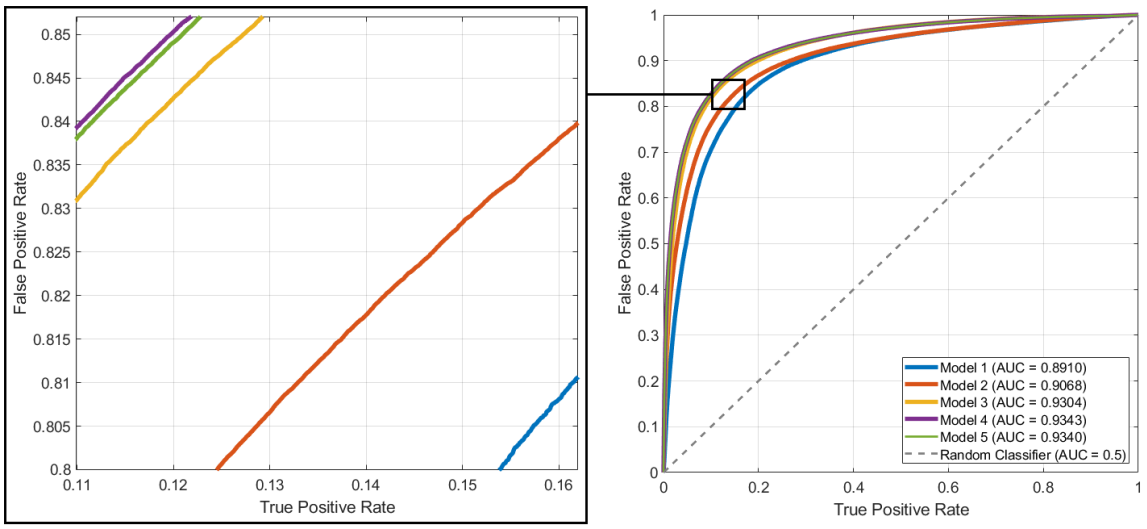



Figure 2-12. ROC plots showing the superior performance of the Model 4. The preferred Model 4 has an AUC of 0.9343 which is slightly higher than Model 5.

### 2.5.3 Feature Importance

To assess the feature importance in the logistic regression model, the coefficients are multiplied by the standard deviation of the data of the corresponding continuous variables. Since the model was developed via normalized data with standard deviation of 1 for all of the continuous variables, the feature importance index will be the same as the logistic

regression model coefficients presented via Table 4. Thus, the results show that slope has the highest importance among the selectable features (geospatial and change information), followed by geology and grayscale change. Among the RGB color channels, the green band is the most important, followed by red.

### 2.5.4 Final Mapping and Visualization

Although accuracy of the recommended Model 4 is acceptably high, the visualizations show that there is slight over-prediction of landslide across the study area. However, the extent of the predictions is still considerably better than the RGB-based Model 1, which shows the improvement by visual inspection. To avoid the overprediction of the landslides, the probability threshold of the logistic regression model output to define the classes was investigated, and 0.70 was found to be the optimal threshold based on a trade-off between the visual inspection of the landslide extent on the resulted maps and the accuracy of the final map, compared to the ground-truth inventory provided by NIED (2016). Figure 13 shows the change in the accuracy of the model on final map prediction for each class.

As can be seen in Figure 13, the optimal probability threshold considering the accuracy of predictions in both classes together is ~53.5 % with overall prediction accuracy of ~85.7 % on the final map. However, increasing the probability threshold to ~70 %, increases the overall prediction accuracy to ~91.5 %, while decreasing the landslide class prediction accuracy to ~78.3 %. Although in leads to less overprediction of landslide features based on the visual inspections as shown in Figure 14.

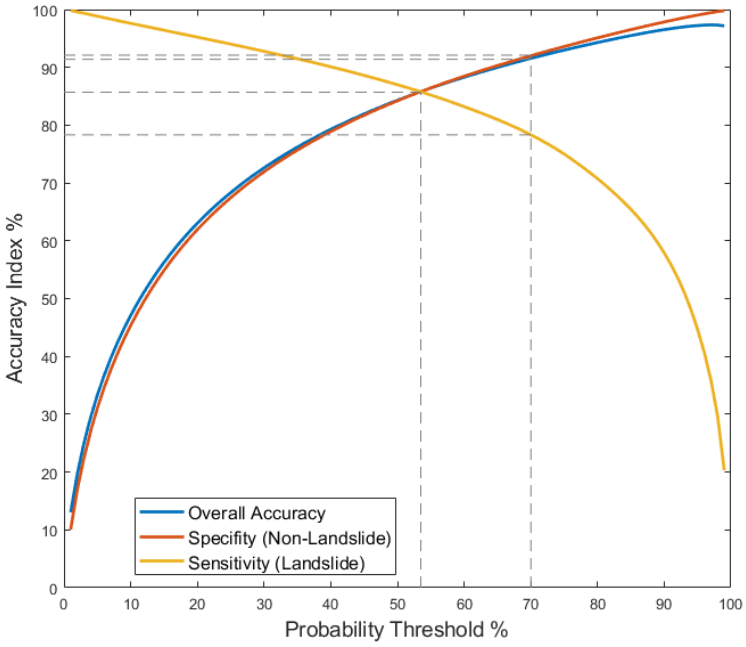



Figure 2-13. Optimal probability threshold selection plot. Increasing the probability threshold from ~53.5 % to ~70 %, increases the overall prediction accuracy from ~85.7 % to ~91.5 %.

Figure 14 shows the outputs of Models 1 (Figure 14-a) and 4 (Figure 14-b) to visualize the impact of adding change and geospatial information to the imagery for classification. Additionally, the final output of the Model 4 based on the 0.70 threshold is provided in the bottom of the Figure 14-c. The modeled maps include colors red (true positive: correctly mapped landslides), yellow (false positive: wrongly mapped landslides), and blue (false negative: missing landslide features, predicted as non-landslide). True negative pixels (correctly predicted non-landslide) are not visualized.

The impact of the post-processing on the final map accuracy was investigated by calculating the predicted map's accuracy and comparing it with the ground-truth label. It is concluded that the median filtering post-processing of the resulting landslide map increases the map accuracy from 85.7% to 89.4% considering Model 4 output, and from 91.5% to 94 % when the 70% probability threshold is applied to the model output.

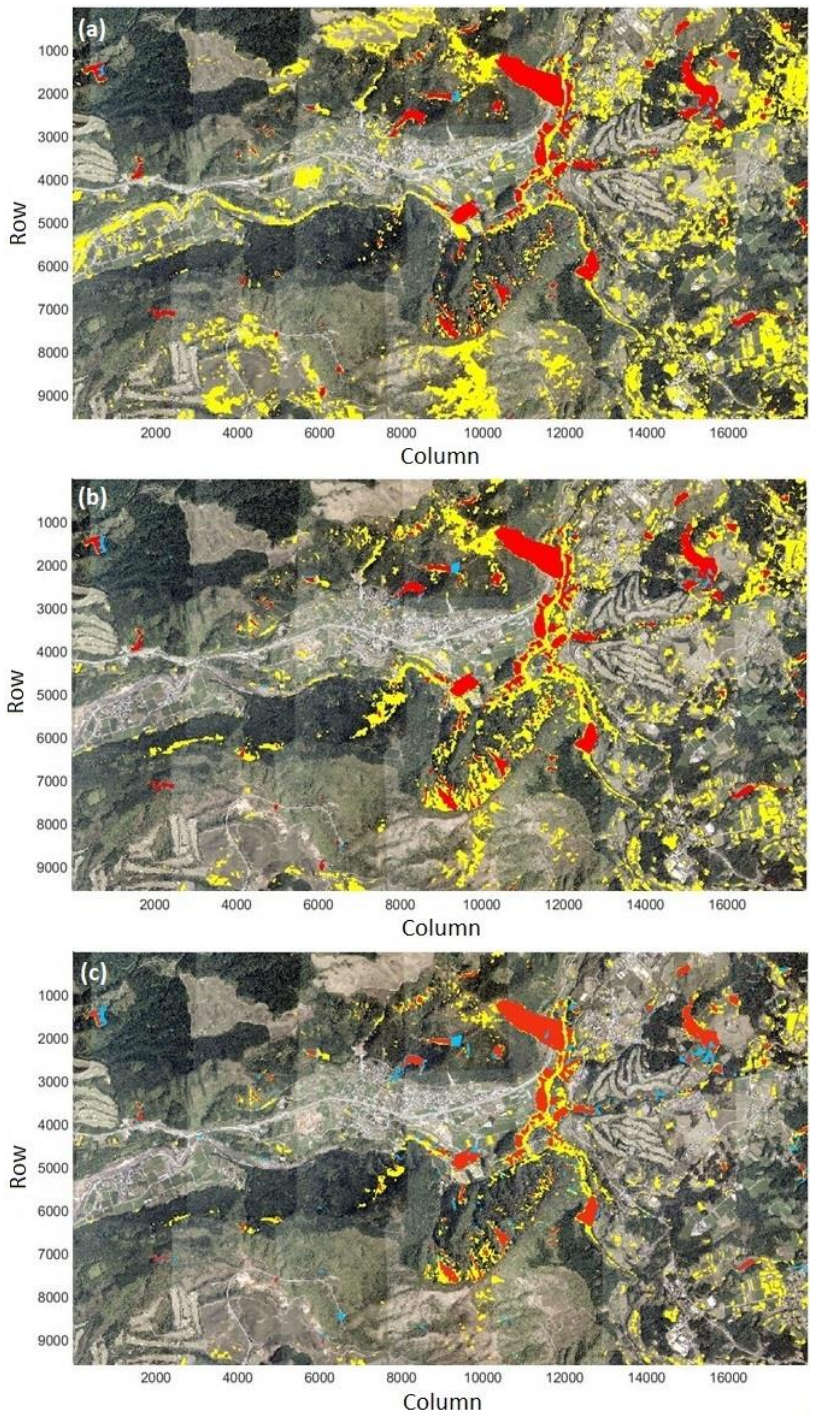


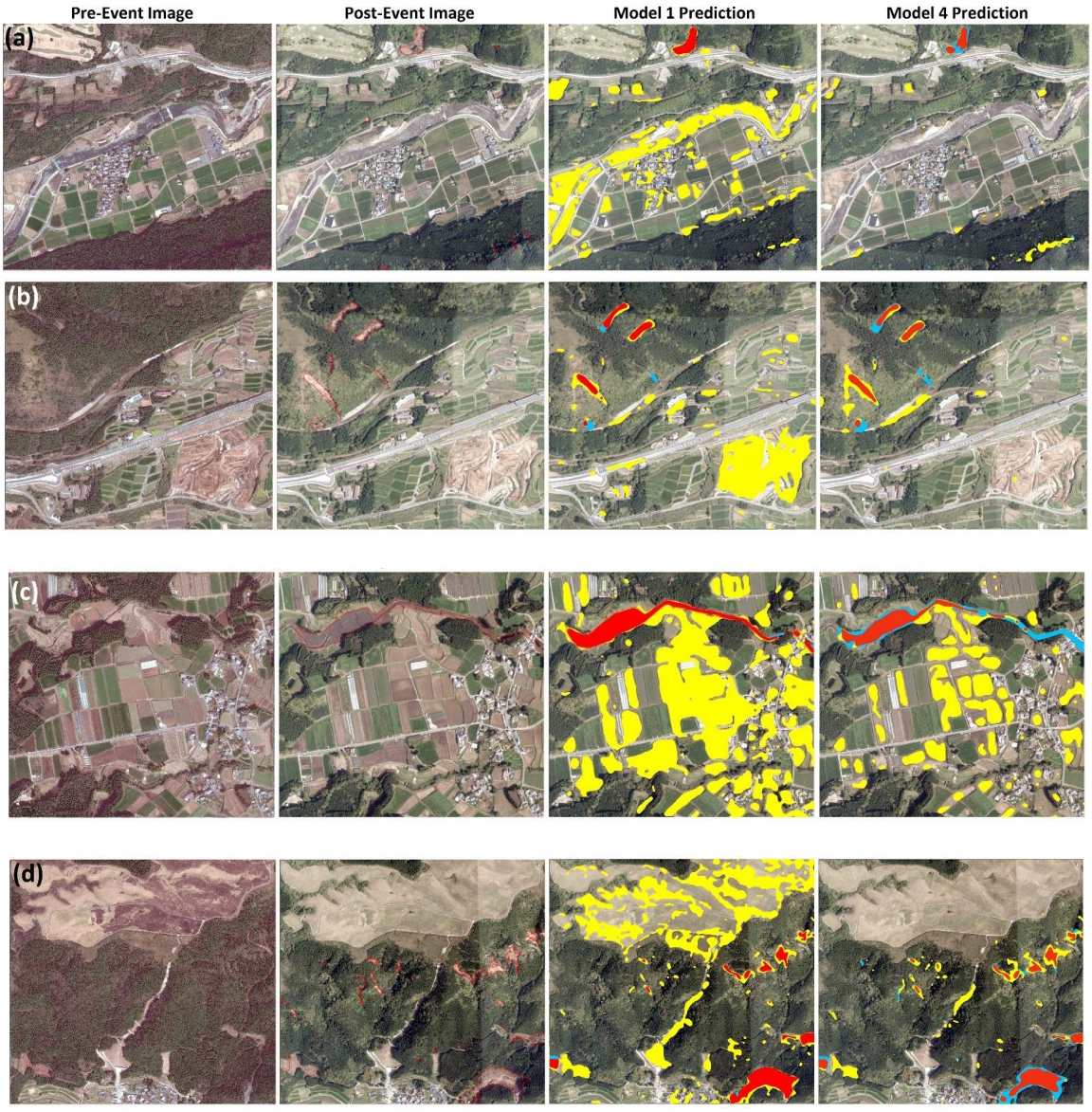

Figure 2-14. Comparative maps of the landslides by Models 1 (a), Model 4 (b) and the Model 4 with 70% probability threshold (c). The maps show post-event imagery with predicted labels. The modeled maps include colors red (true positive: correctly mapped landslides), yellow (false positive: wrongly mapped landslides), and blue (false negative: missing landslide features, predicted as non-landslide). Rows and columns as axes titles are indicators of pixel number (pixel size is 0.5 m).

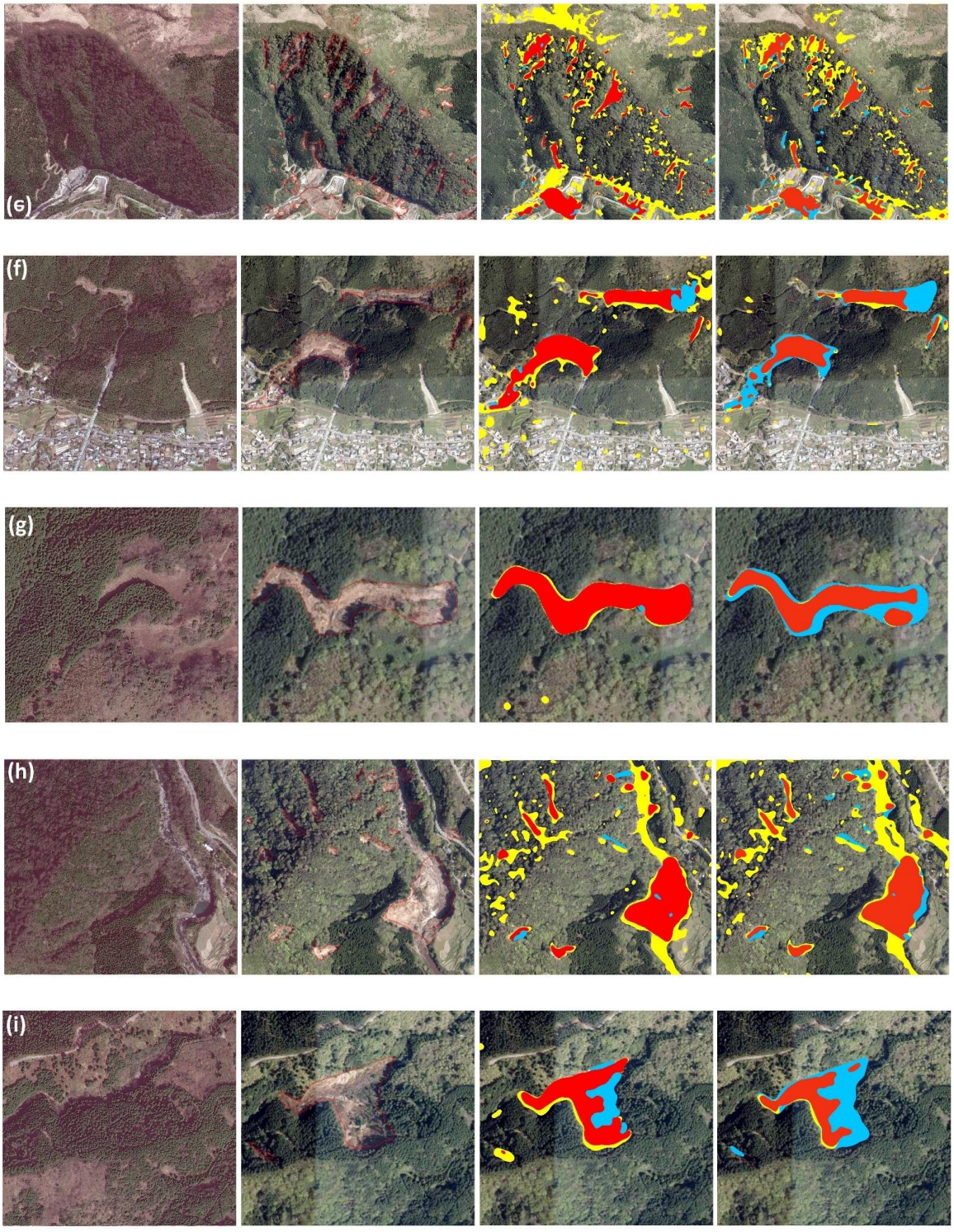
Figure 15 shows detailed views of the landslide mapping approach presented in this study. The examples show pre-event imagery, post-event imagery with ground-truth NIED landslide labels via red polygons, the RGB-derived Model 1 output, and the final Model 4 output maps. The final Model 5 (with threshold) slightly tends to underpredict landslides (f, i, j) and overpredict non-landslides (c, h, j) in some regions, which is not necessarily a weakness, given the extensive corrections (a, b, d, j) of the landslide over-prediction. In addition, although the extent of landslides in some cases is smaller than the ground-truth polygons (c, f, I, j), the recognition of landslide features is considered acceptably good.

It is also visible among the detailed views that in some cases the deviation from the NIED polygons (f) is not necessarily a model error, but a correction of weakly drawn polygons. In addition, the visual quality checks performed on the polygon data showed that they are not perfectly matching the imagery, and there are slight shifts in different directions (g, h, i), which might impose some small error in the modeling process. However, since the model is pixel-based, and the training pixels were collected randomly, the feature boundary pixels with higher uncertainty remain minor in the training data, and the classifier tends to learn to predict the majority of data precisely, neglecting minor errors. This results in the model even correcting the shifted boundaries (g, h, i), which is a benefit of pixel-based mapping.

In some cases, the original labels were not delineated carefully and contained some portions of the non-landslide area (c, f, g). Thus, some of the errors could be related to that issue. Many incorrect predictions by the Model 1 are shown to be resolved by the Model 4 (a, b, c, d, j). The preferred model shows removal of many incorrectly detected landslides

by the 3-bands Model 1 (a, b, c, d, j). However, the preferred Model 4 still some wrongly detected landslide feature predictions, many of which are due to harvested farmlands (c, j, k), bare (non-vegetated) soils (d, e) and rivers when they change color due to containing soil particles (k).




(e)
(f)
(g)
(h)
(i)

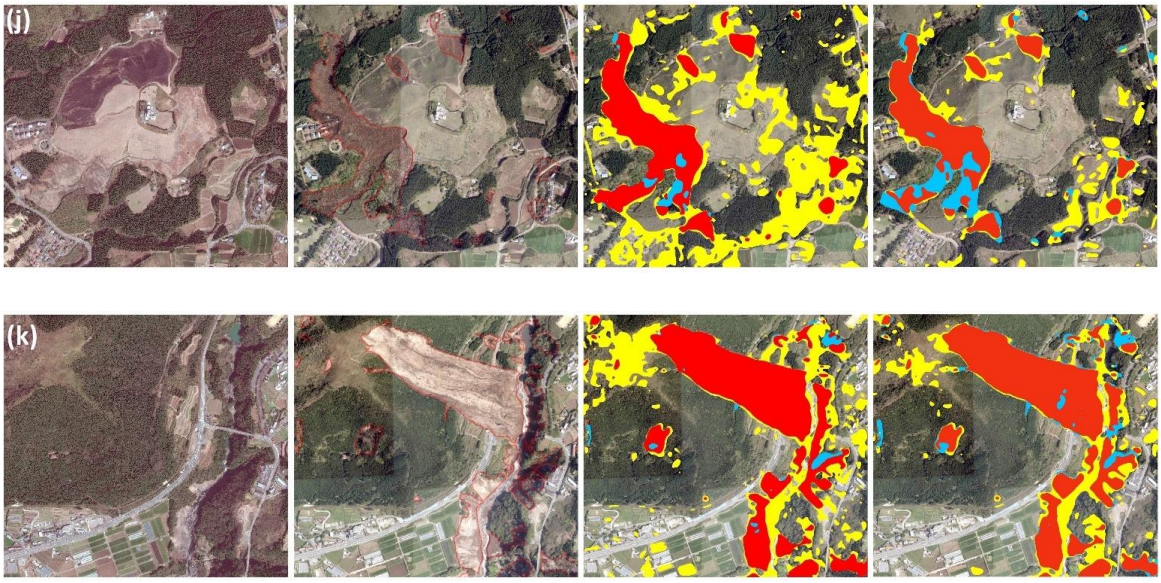



Figure 2-15. Different examples of the landslide mapping approach presented in this study. The examples show pre-event imagery, post-event imagery with ground-truth NIED landslide labels via red polygons, the RGB-derived Model 1 output, and the final Model 4 output maps. The modeled maps include colors red (true positive: correctly mapped landslides), yellow (false positive: wrongly mapped landslides), and blue (false negative: missing landslide features, predicted as non-landslide).

## 2.6 Conclusions and Future Works

This study introduces a data fusion machine learning approach for landslide inventory development. Optical and SAR imagery as well as readily available geospatial inputs for the region impacted by landslides due to the Kumamoto 2016 earthquakes were used within a logistic regression model. A visually drawn landslide inventory, created by NIED of Japan by investigating the post-event aerial imagery and satellite SAR data and pre-event optical imagery, was used for training, validation, and testing data. Optical imagery included pre-event satellite (DigitalGlobe) and post-event aerial (GSI) high-resolution imagery. SAR imagery was also available for pre- and post-event from ALOS-2 satellite's PALSAR-2 sensor. Several change indices were tested including grayscale index change,

vegetation index change, and temporal SAR amplitude change as additional predictive inputs for pixel-wise landslide classification.

Based on the ROC-based feature ranking results, grayscale change, vegetation change, and HV amplitude difference were included in the classification algorithm. To enhance the performance of the model, geospatial data of related variables were also evaluated including elevation, slope, aspect, curvature, landslide probability, precipitation, and geology. Based on the conducted feature ranking analysis, five geospatial features with the highest predictive ability were included in the classification algorithm, including, slope, curvature, landslide probability, precipitation, and geology. Model development (training and validation) data and testing data were randomly spatially sampled across distinct regions. Using 10% of available landslide pixels and 0.4% of the non-landslide pixels from the model development grid regions, a relatively small and balanced training dataset was created.

The logistic regression classification across five different scenarios was tested: 1) RGB; 2) RGB + Change; 3) RGB + Geospatial; 4) RGB + Change + Geospatial; and 5) all variables regardless of feature ranking results. Comparing the outcomes of the logistic regression classification models, the model trained and tested via 11 features (RGB plus selected change and geospatial variables) outperformed the other case models and was introduced as the preferred model. The feature importance analysis after training the classifier showed that the slope and geology variables are the most important among geospatial variables, green channel has the highest importance among RGB channels, and

the grayscale change index has the highest importance among the change information variables.

Given that this model was only developed by only 3.85% of the total pixels across the region, a landslide inventory can be developed using a relatively small training dataset with a short run-time due to the very high speed of logistic regression classification modeling.

It is suggested that the methodology proposed in this research be tested on other earthquake-triggered landslide events in Japan and other regions around the globe. Additional informative bands might be still beneficial to the landslide modeling, including landcover data, Near-Infrared and other bands of the multi-spectral sensors, etc. In addition, feature transformation and engineering, as well as dimensionality reduction techniques are worthy of investigation. The other direction of research could be to try model via deep learning algorithms such as convolutional neural networks (CNN) on the same datasets to analyze the comparative classification performance using post-event images with and without additional parameters used in this research.

## 2.7 Acknowledgements

The support of National Geospatial Intelligence Agency (NGIA) for this research through the NGIA Academic Research Program Grant #HM0476-20-0006 (NGA NURI Project: Benchmark data development to classify damage for natural disaster relief efforts) is greatly appreciated. The authors are also thankful for the support of U.S. Geological Survey (USGS) through research grant #G22AS00006 – Proposal 2022-0047 (Innovative data-driven frameworks for geospatial ground failure models). The authors would also like to

acknowledge the GSI, JAXA, Geological Survey of Japan and NIED organization of Japan for publishing their GIS datasets of landslide inventories, aerial imagery, and satellite SAR datasets related to the Kumamoto 2016 earthquake, plus the geology map of the study area. All other organizations which provided data resources are greatly appreciated, including DigitalGlobe for the satellite imagery used in this study, and NASA/METI and their collaborators for providing the digital elevation products.

# 3 Semi-Supervised Learning Method for the Augmentation of an Incomplete Image-based Inventory of Earthquake-Induced Soil Liquefaction Surface Effects

Adel Asadi [1], Laurie G. Baise [1], Christina Sanon [1], Magaly Koch [2], Snehamoy Chatterjee[3], and Babak Moaveni [1]

1 Geohazards Research Lab, Civil and Environmental Engineering Department, School of Engineering, Tufts University, Medford, MA 02155, USA; Laurie.Baise@tufts.edu

2 Center for Remote Sensing, Earth and Environment Department, Boston University, Boston, MA 02215, USA; mkoch@bu.edu

3 Geological and Mining Engineering and Sciences Department, Michigan Technological University, Houghton, MI 49931, USA; schatte1@mtu.edu

(The material contained in this chapter is being submitted for review and possible publication in Remote Sensing Journal – MDPI Publications)

## 3.1 Abstract

Soil liquefaction often occurs as a secondary hazard during earthquakes and can lead to significant structural and infrastructure damage. Liquefaction is most often documented through field reconnaissance and recorded as point locations. Complete liquefaction inventories across the impacted area are rare but valuable for developing empirical liquefaction prediction models. Remote sensing analysis can be used to rapidly produce the full spatial extent of liquefaction ejecta after an event to inform and supplement field investigations. Visually labeling liquefaction ejecta from remotely sensed imagery is time-consuming, and prone to human error and inconsistency. This study uses a partially labeled liquefaction inventory created from visual annotations by experts and proposes a pixel-based approach to detect unlabeled liquefaction using advanced machine learning and image processing techniques, and to generate an augmented inventory of liquefaction ejecta with high spatial completeness. The proposed methodology is applied to aerial

imagery taken from the 2011 Christchurch earthquake and considers the available partial liquefaction labels as high-certainty liquefaction features. This study consists of two specific comparative analyses: 1) to tackle the limited availability of labeled data and their spatial incompleteness, a semi-supervised self-training classification via linear discriminant analysis is presented, and the performance of the semi-supervised learning approach is compared with supervised learning classification; 2) post-event aerial image is used to extract color transformation bands, statistical indices, texture components from the RGB (red-green-blue) channels, and dimensionality reduction outputs, and performance of the semi-supervised model with different combinations of selected features from these four groups are compared. Building footprints are also used as the only non-imagery geospatial information to improve classification accuracy by masking out building roofs from the classification process. To prepare the multi-class labeled data, regions of interest (ROIs) were drawn to collect samples of seven landcover and land use classes. The labeled samples of liquefaction were also clustered into two groups (dark and light) using the Fuzzy c-means clustering algorithm to split the liquefaction pixels into two classes. Comparing the generated maps with fully and manually labeled liquefaction data, showed that the proposed semi-supervised method performs best, when selected high-ranked features of the two groups of statistical indices (gradient weight and sum of the band squares) and dimensionality reduction outputs (first and second principal components) are used. It also outperforms supervised learning, and can better augment the liquefaction labels across the image in terms of spatial completeness.

## 3.2 Introduction

Liquefaction is a type of ground failure in which increased pore pressure in granular soil reduces effective stress and can result in ejected silt to fine sand and water, settlement, lateral spreading resulting in cracking, and/or bearing capacity failure. Surface manifestations of liquefaction, including ejecta, settlements, and cracking due to lateral settlements, are often identified during field reconnaissance after earthquake events as evidence of liquefaction in the subsurface. These surface manifestations of liquefaction are associated with liquefiable sediments in the upper few meters (van Ballegooy et al. 2014). In developing semi-empirical and empirical models for liquefaction prediction such as those commonly used in practice and regional models used in rapid response and loss estimation, liquefaction inventories are used. These inventories generally provide locations as points or polygons of observed liquefaction occurrence. Most often, the liquefaction occurrence is equivalent to the existence of a surface manifestation of liquefaction.

Existing liquefaction inventories are generally presented as binary labels (either the location had evidence of liquefaction (1) or did not have evidence of liquefaction (0)) such as the Next Generation Liquefaction (NGL) project discussed by Brandenberg et al. (2020). Other efforts have proposed an alternate liquefaction inventory format, including several recent Geotechnical Extreme Events Reconnaissance (GEER) Association efforts which result in near spatially complete polygon-based maps of liquefaction occurrence (Stewart et al., 2019; Zimmaro et al., 2019; Ponti et al., 2020; Allstadt and Thompson, 2021). Rashidian et al. (2020) have also developed inventories that map the spatial extent of liquefaction by providing polygons around ejecta and surface cracking. A spatially complete and unbiased catalog of liquefaction occurrence is critical for validating and

refining existing regional models for liquefaction occurrence (Zhu et al., 2015 & 2017). Alternatively, light detection and ranging (LiDAR) datasets of surface settlements as a more quantitative measure of liquefaction have also been used (Sharifi-Mood et al., 2018; Kajihara et al., 2020). A benefit of these alternate liquefaction inventories is an estimate of the scale of liquefaction which is necessary to map impact and fragility. This paper examines an efficient way to develop a spatial inventory of liquefaction surface manifestations using aerial imagery.

Liquefaction (ejecta, cracking, etc.) can be identified at a regional scale using satellite optical imagery or high-resolution aerial imagery in a site-specific manner (Rathje and Franke, 2016). The liquefaction identification can be made by visual interpretation of post-event images (Ghosh et al., 2011; CGD, 2013; Rollins et al., 2014) or by inspection of both pre- and post-event images (Hamada et al., 1987). However, any visual process can be tedious even for an expert and take a significant amount of time, sometimes resulting in incomplete and spatially limited datasets. In addition, this approach can be highly inconsistent because it depends on the experience and judgment of human inspectors and other contributing factors such as data quality and resolution. Thus, automated liquefaction detection methods are needed to provide consistent, spatially complete liquefaction inventory datasets after major earthquakes.

Automated mapping methods can be developed based on pairs of pre-event and post-event data or post-event data only. Satellite multispectral near-Infrared (NIR) and short-wave Infrared (SWIR) bands provide essential information about landcover types, which can be helpful for identifying liquefaction ejecta (Ramakrishnan et al., 2006; Sengar et al.,

2012-b; Oommen et al., 2013; Morgenroth et al., 2016; Rathje and Franke, 2016; Baik et al., 2019; Rashidian et al., 2020). Change detection methods use pairs of pre-event and post-event imagery for liquefaction detection (Ramakrishnan et al., 2006; Saraf et al., 2010; Ishitsuka et al., 2012; Baik et al., 2019). These methods are seensitive to image acquisition orientations of satellites, different light and weather conditions, seasonal variations, resolution mismatch, cloud coverage of pre- or post-event images, etc. As an alternative, post-event optical imagery can be used to map liquefaction using supervised learning algorithms, including advanced machine learning (Bi et al., 2019; Njock et al., 2020) and deep learning-based methods (Hacıefendioğlu et al., 2021; Zhang et al., 2023). Successful deployment of supervised learning techniques depends on the availability, quantity, and quality of labeled (annotated) data.

In this research, semi-supervised learning as an alternative mapping strategy for developing liquefaction inventories is demonstrated. Semi-supervised self-training methods are helpful when datasets are partially labeled. Such algorithms alleviate the requirement of large-scale labeled data, by learning from the labeled data (like supervised learning algorithms) and from the unlabeled data to capture the patterns within the dataset. This improves the generalizability of the model to other feature types, which have limited or no presence in the labeled portion of the data.

Semi-supervised algorithms have been used for several applications, including semantic image segmentation (Feng et al., 2022), especially the segmentation of medical images (Chen et al., 2022). Semi-supervised learning has been proven to reduce data labeling labor work and to improve predictive accuracy with additional unlabeled data in

image classification frameworks (Wu et al., 2017; Chen et al., 2022). Oludare et al. (2021) have shown the superior performance of semi-supervised learning for post-disaster building damage assessment via pre- and post-event satellite imagery.

Within the semi-supervised algorithm, this research will leverage the benefits of various image transformations to the color channels. Texture analysis and statistical measurements of color channels have been used for feature extraction to detect raveling in asphalt pavements (Tong et al., 2018; Nhat-Duc and Van-Duc, 2023). Crack detection on asphalts based on texture properties is also investigated by Abdelkader (2021). The benefit of using color, morphological and textural indices is also reported for multi-class vision-based rock type classification (Chatterjee, 2012) and lithological mapping by optimal integration of such characteristics with multi-sensor remote sensing data (Kumar et al., 2021). Dumitru et al. (2015) have also used texture indices to flood damage mapping. The color and textural features of the image were used for liquefaction detection via the implementation of tile-based deep convolutional neural networks (Hacıefendioğlu et al., 2021; Zhang et al., 2023). This research will explore these methods in a pixel-based classification approach.

Rashidian et al. (2020) demonstrated that building roofs are highly variable in terms of colors and materials. As a result, they have a high spectral variation that makes it challenging to discriminate roofs from other classes in a pixel-wise classification framework due to the overlap with other classes. Using a building footprint mask to remove buildings from the classification process has increased accuracy significantly in the liquefaction mapping study by Rashidian et al. (2020). Thus, this study also uses building

footprints as prior known geospatial information to exclude buildings and improve classification results.

This study demonstrates a semi-supervised self-training classification for developing a spatially complete and consistent liquefaction inventory using a limited set of labeled liquefaction polygons and high-resolution aerial imagery. Because high-resolution aerial imagery only has three color bands (Red, Green, and Blue), data transformations across four categories (color transformations, texture analysis, statistical indices, and dimensionality reduction techniques) are leveraged to increase the information content. This method, which uses semi-supervised learning and imaged-derived information to detect liquefaction for the first time, is compared to a simpler supervised learning approach. Using the results of this study, a spatially complete and consistent liquefaction inventory can be created. The augmented liquefaction map of the studied area in this research is provided as a digital supplementary material of this paper.

## 3.3 Data and Methods

### 3.3.1 Datasets

#### *3.3.1.1 Region of Study*

The February 22, 2011 (MW6.2) Christchurch earthquake struck directly beneath Christchurch City. Although of smaller magnitude than the Darfield earthquake in 2010, the Christchurch earthquake caused more substantial damage in northern and eastern Christchurch (Cubrinovski et al., 2011), partly because of its proximity to the city but also because it was directly beneath the liquefaction-prone land. Examination of the aerial and satellite images and extensive ground surveys has confirmed the presence of widespread

liquefaction in rural areas and the Christchurch urban area (Green et al., 2014; Orense et al., 2014; Geyin et al., 2021).

#### *3.3.1.2 Imagery*

The aerial imagery used in this research study comprises three spectral (color) bands (RGB) with a very high spatial resolution (10 cm). On February 24, 2011 (2 days after the earthquake), they were captured by New Zealand Aerial Mapping (NZAM) at 1700 m above the earth's surface using the UltraCam XP digital aerial camera. The 1,785 ortho-rectified RGB GeoTIFF imagery published by LINZ (2011) represents a ground surface area of 33.3 ha in Christchurch. The size of each vertical image tile is 7200-by-4800 pixels (720-by-480 meters). Figure 1 shows the locations of the tiles used in this paper, and Figure 2 shows the tiles in a close view. The tile shown in Figure 2-a is used for model development and evaluation. Then, the same approach is applied to the tile in Figure 2-b.

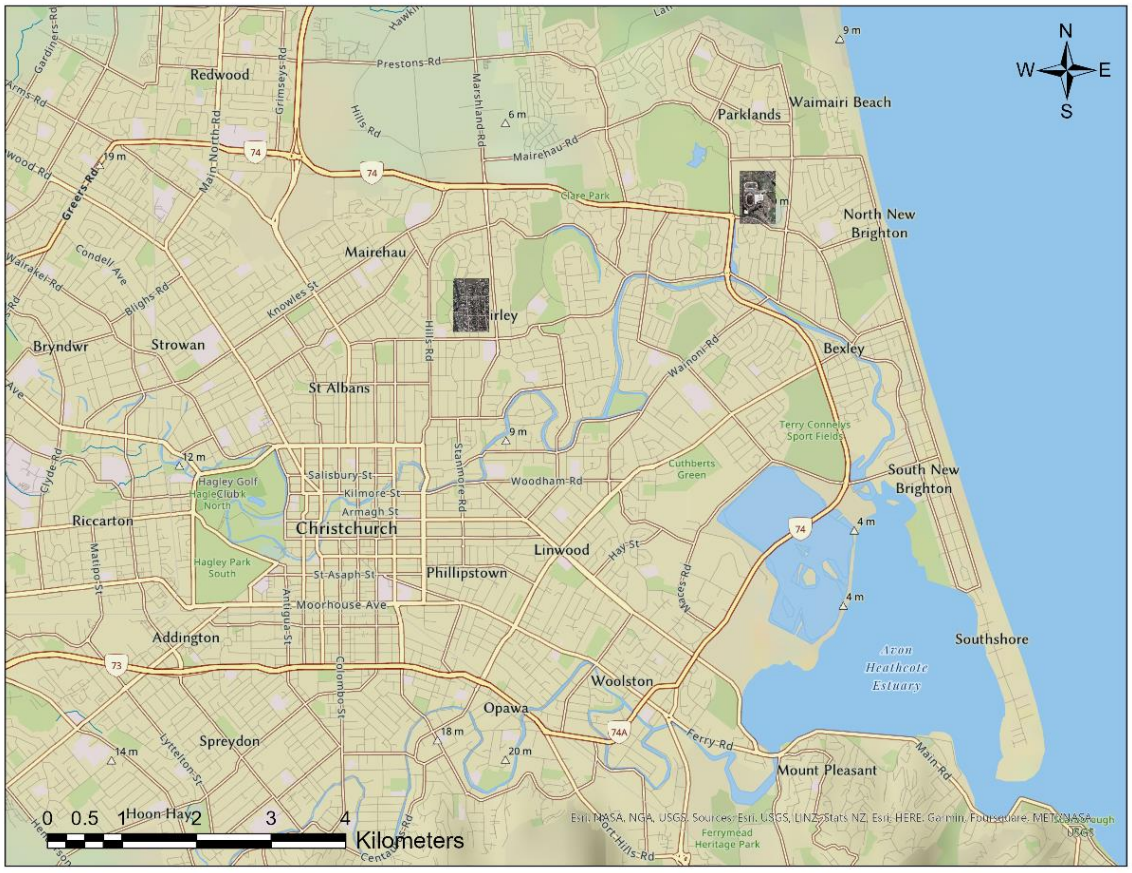



Figure 3-1. The location of the two aerial image tiles used in this study in the city of Christchurch, New Zealand. The resolution of the images is 10 centimeters, and each tile is 720-by-480 meters in size. The image tiles are all projected to the New Zealand Transverse Mercator (NZTM) system. The left and right images are shown in detail via Figures 2-a and 2-b, respectively.


Figure 3-2. **a)** The image tile used for model development and evaluation. **b)** The image tile used for further model application. Both tiles have a resolution of 10 centimeters, and each tile is 720-by-480 meters in size. Linear stretching is performed on the image for better visualization contrast.

#### *3.3.1.3 Liquefaction Labels*

Liquefaction inventories have historically been developed from field investigations (Stewart et al., 2019; Allstadt and Thompson, 2021) and visual labeling of satellite and aerial imagery (Townsend et al., 2016; Zimmaro et al., 2019). The lack of a standardized protocol for documenting earthquake-induced ground failure creates uncertainties in how surface manifestations are mapped. Liquefaction surface manifestations vary in shape, size, and color. Townsend et al. (2016) is a polygon-based liquefaction inventory that includes the 2011 Christchurch earthquake. The Townsend et al. (2016) inventory includes large-

scale polygon delineations to show the broad extent of liquefaction-induced impact but does not mimic the individual surface manifestations (e.g., ejecta).

Sanon et al. (2023) have developed an alternative polygon-based inventory that also includes the 2011 Christchurch earthquake. The Sanon et al. (2023) inventory is based on the delineation of individual liquefaction surface manifestation to use the inventory to train computer-vision-based algorithms. Sanon et al. (2023) discuss the challenges and limitations of creating a consistent and spatially complete inventory. One rule in the Sanon et al. (2023) inventory is only to include liquefaction labels that are identified with high certainty and sufficient visibility of boundaries. They have also only delineated liquefaction occurrences which appeared to be in their original form, not the piled liquefaction ejecta on the roadsides (Sanon et al., 2023). As a result, the Sanon et al. (2023) inventory is a partially labeled, spatially incomplete inventory of liquefaction surface manifestations with high quality and certainty.

An accuracy assessment of the model is performed on the tile in Figure 3-a, which shows the labels by Sanon et al. (2023) overlaying the image tile presented in Figure 2-a (model evaluation tile). To calculate the model accuracy in this study, the spatially complete labels are created by manually drawing polygons (Figure 3-b). As Sanon et al. (2023) discussed, visual labeling of liquefaction has inherent uncertainty due to the ambiguity of boundaries and the inconsistency between labelers. Thus, to minimize human inconsistency, a single expert labeled the entire model validation tile (Figure 3-b). This study aims to use the high-certainty liquefaction polygons labeled by Sanon et al. (2023) in Figure 3-a to map the unlabeled portion of the tile using semi-supervised learning

methods to augment the partially labeled liquefaction data into a spatially complete and consistent inventory.

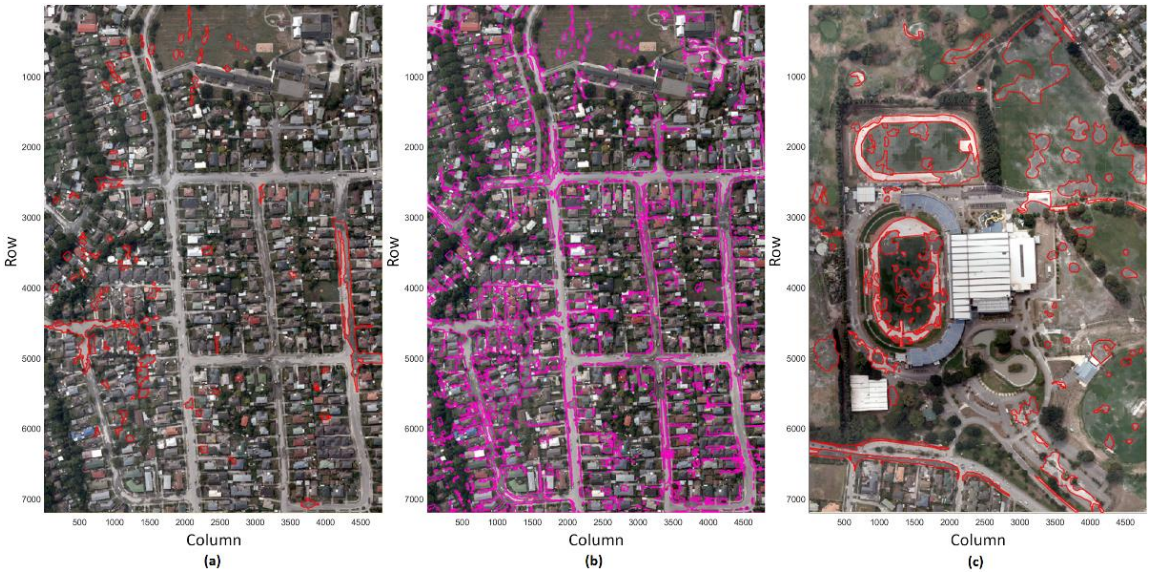



Figure 3-3. **a)** The image tile used for model development and evaluation with Sanon et al. (2023) liquefaction polygons (in red) overlayed. **b)** The image tile with a full label (in magenta) created for the purpose of model validation in this study. **c)** The image tile used for model application with Sanon et al. (2023) liquefaction polygons (in red) overlayed.

The Sanon et al. (2023) labels used in this study (Figure 3-a,c), as well as the validation labels created in this study (Figure 3-b), are summarized in Table 1. Model application is evaluated on a second tile, as shown in Figure 3-c with the labels from Sanon et al. (2023). The model application tile has been labeled by a different person than labeled the model evaluation tile (Sanon et al., 2023). The model application tile is used to apply the same method performed on the model development tile without validating the resulting map via ground-truth validation data. Visual inspection of the polygon labels in the two tiles (Figure 3-a,c) shows differences in the labeling strategies by different experts. This is an indicator of inconsistencies resulting from human labeling.

Table 3-1. Spatial coverage of the polygons used per image tile.

| Area Coverage | Model Evaluation Tile | | Model Application Tile |
|---|---|---|---|
| | Sanon et al. (2023) | Validation Label | Sanon et al. (2023) |
| Total Extent ($m^2$) | ~ 9,919 | ~ 44,513 | ~ 38,500 |
| Percentage of Tile | 2.87 % | 12.88 % | 11.14 % |

#### *3.3.1.4 Building Footprints*

This study uses building footprints as prior known geospatial information to exclude buildings and improve classification results, based on the recommendation made by Rashidian et al. (2020). The geospatial layer of building footprints used in this study was created by New Zealand Aerial Mapping (NZAM). The dataset has some errors (identified from visual inspection) resulting from change in land use since the building footprints were created, mismatch of building boundaries compared to the imagery used, or incompleteness of the building footprints (missing building polygons). For the tiles used in this research, and to enhance the validation process, a few building polygons were created to improve the accuracy and completeness of the building footprint. Figure 4 shows the building footprints used in this study to mask the buildings from the imagery.

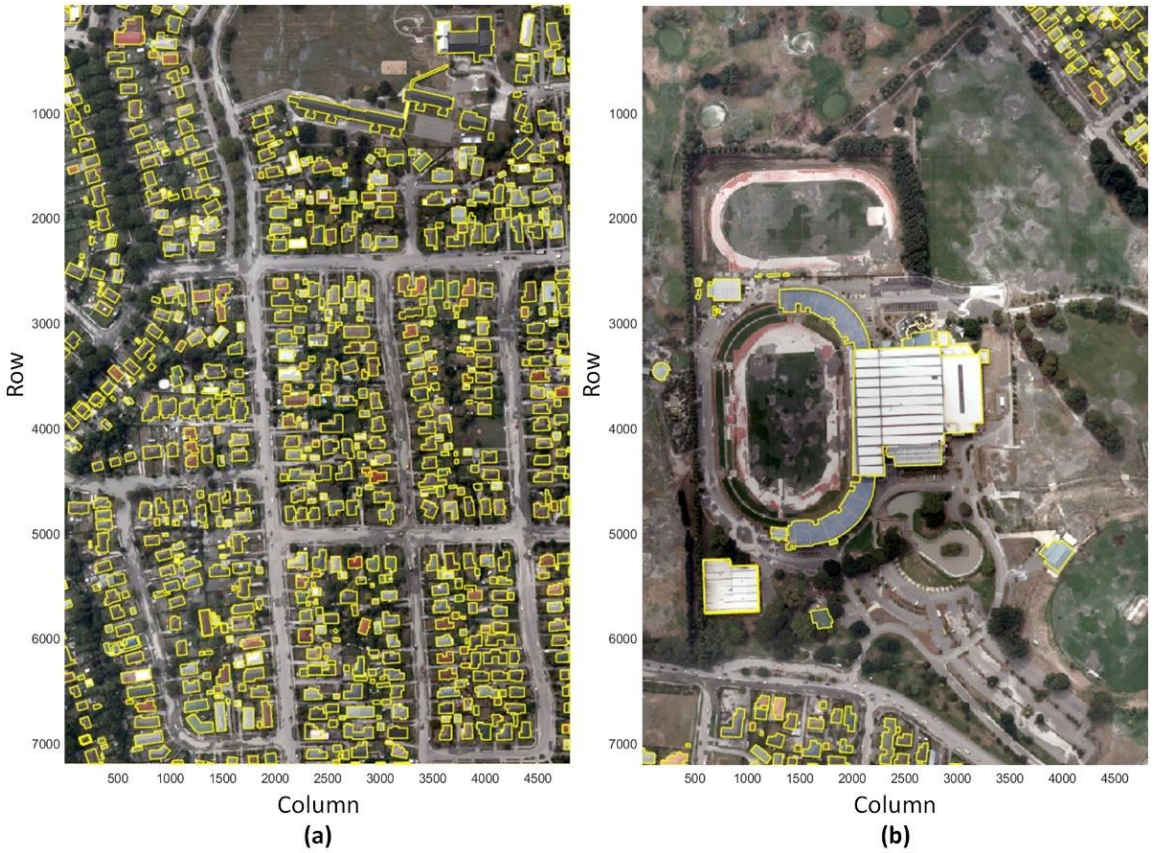



Figure 3-4. Building footprints (in yellow) created by NZAM, used to mask out the buildings in this study. The version showed in this figure was modified by adding a few polygons for missing buildings and removing some polygons for which no building was observed in the imagery.

#### *3.3.1.5 Land Use and Land Cover ROIs*

High-quality training data are collected by drawing region of interest (ROI) polygons over seven different land covers and land use, including trees, vegetation, soil, water, shadow, roads, and pavements/driveways. Drawing ROIs was done via ArcGIS Pro using the polygon drawing tool in a way that the boundary of the polygons is inside the boundary of the object of interest, so no mixed pixels or out-of-class pixels are sampled for that class. It is recommended that enough ROIs be drawn to cover the most possible range of pixel values for each class of objects in the tile.

### 3.3.2 Methodology

The proposed methodology presented in this study uses color, statistical, and texture transformations on high-resolution aerial imagery and dimensionality reduction techniques to create a multi-band dataset from a three-band color image. The flowchart of the methodology is provided in Figure 5. The high certainty spatially incomplete liquefaction labels combined with spatially incomplete labels for seven additional land cover and land use classes sampled by ROIs and a complete building footprints mask are used as input data. The building footprint is used to mask the building pixels, so they are not included in the training process. Rashidian et al. (2020) observed that liquefaction surface manifestations present as light-colored (dry sand) and dark-colored (wet sand). Thus, the approach also includes Fuzzy C-Means clustering to create two liquefaction label clusters. The Fuzzy approach is chosen to remove outliers and mixed pixels from the liquefaction classes based on an individual pixel having a low probability of belonging to either of the two liquefaction classes. The model is evaluated using accuracy statistics for the labels based on a second dataset of spatially complete liquefaction labels (Figure 3-b). The minimum Redundancy Maximum Relevance (MRMR) method is used to identify relevant color, statistical, and texture transformations from the imagery, plus the best outputs of dimensionality reduction techniques to separate the 9 classes (2 liquefaction and 7 landcovers). Finally, a semi-supervised linear discriminant analysis classifier is trained to classify the unlabeled portion of the partially labeled image. A few combinations of features are tested, and the preferred model with the best set of selected features is introduced. The steps of the methodology are discussed below.

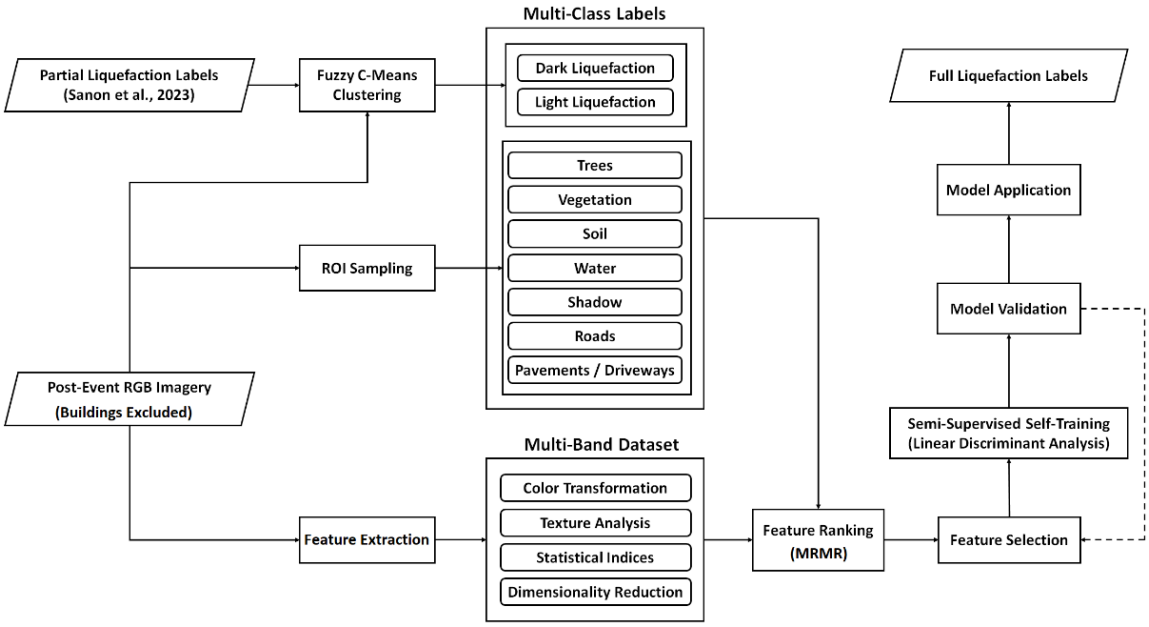



Figure 3-5. Flowchart of the proposed method, showing how multi-band multi-class labeled data are created and fed to the semi-supervised learning method to complete the partially labeled liquefaction data.

#### *3.3.2.1 Fuzzy C-Means Clustering of Liquefaction Data*

The liquefaction labels in the training dataset are visually developed based on expert knowledge. Rashidian et al. (2020) discovered that the surface effects of liquefaction are bimodal in terms of spectral signature and hence should be classed as either wet (darker) or dry (lighter) liquefaction. This is because the water content of these two types differs. Thus, we have considered creating two liquefaction labels using clustering methods based on the spectral characteristics of liquefaction. Sanon et al. (2023) also discussed that the liquefaction ejecta boundaries, which are hand drawn by the expert as polygons, are uncertain, resulting in imperfect labels, also known as out-of-class or mixed pixels. Particularly in coarse spatial resolution imagery, pixels may be a mix of two or more classes, introducing uncertainty and ambiguity into the labels. Soft classification approaches to aid in assessing uncertainties and ambiguity in transition zones between

different types of land cover (Sengar et al., 2012-a). Fuzzy logic is a soft classification approach which can be used to extract a single land cover class from out-of-class or mixed pixels. Thus, in this research, Fuzzy C-Means (FCM) clustering method is used to convert the liquefaction labels to a binary liquefaction label and to improve the labels for the out-of-class and mixed pixels.

Fuzzy C-Means (FCM) is a clustering technique where each data point belongs to a cluster with some degree specified by a membership grade, and the sum of the memberships for each pixel must be equal to 1 (Bezdek, 1981). FCM allows each data point to belong to multiple clusters with varying degrees of membership. By iteratively updating the cluster centers and the membership scores for each data point, FCM moves the cluster centers to the right location within a data set. This iteration is based on minimizing an objective function representing the distance from any given data point to a cluster center weighted by that data point's membership grade. The result of this step in the methodology is to have a more consistent set of training labels, including, two classes of liquefaction. After running the fuzzy clustering, a probability threshold of 0.55 was used to filter out the low certainty pixels from the labeled liquefaction data of each class (dark and light). The threshold is chosen to be low enough to only remove very low probability pixels in terms of belonging to either class.

#### *3.3.2.2 Feature Extraction from RGB Image*

RGB wavelengths are highly correlated and thus contain limited spectral information. In this research, applying a variety of image processing, analysis, and transformation techniques is considered to include additional information which can be beneficial in finding new liquefaction features. Table 2 provides a brief overview of the

feature extraction techniques used in this study, categorized into 4 groups of color transformations (3 features with 10 total bands), texture analysis (4 features with 7 total bands), statistical indices (7 features with 7 total bands) and dimensionality reduction techniques (3 features with 7 total bands).

Common applications of indices derived from multi-band remote sensing images have been to reduce topography and illumination differences within temporal remote sensing images, and to enhance a certain class of interest, such as minerals, water, and soil content. When generating indices from an image, remote sensing data dimensionality is indirectly reduced to one, for instance, signifying a specific feature or matter (Sengar et al., 2012-a). These additional bands are extracted information contained in the original RGB; they are showing uncorrelated information in a new set of bands.

The color transformation features extracted from the image include Hue-Saturation-Value (HSV), decorrelation stretch, and cyan-magenta-yellow-black (CMYK). Color transformations have been used previously to improve the mapping of objects of interest via remote sensing data, for instance, to map Maize plants (Gnädinger and Schmidhalter, 2017). The texture analysis techniques considered in this research include convolution and correlation filters, Gabor filters, and wavelet transformation. The statistical indices presented in Table 2 include entropy, gradient weight, sum of squares, mean absolute deviation, variance, standard deviation, and range filters. This study uses grayscale transformation, Minimum Noise Fraction (MNF) rotation transforms, and Principal Component Analysis (PCA) techniques as dimensionality techniques. Texture transformation, statistical analysis, and dimensionality reduction techniques have also been

used previously to detect objects of interest via remote sensing data, for instance, to map leafless trees in plantations (Khorrami et al., 2022).

Table 3-2. Extracted features used in the study.

| Category | No. | Feature | Bands | Description | Reference |
|---|---|---|---|---|---|
| Color Image | 1 | RGB | 3 | Red-Blue-Green components of the visible spectrum | NZAM (2011) |
| Color Transformations | 2 | HSV | 3 | Saturation (chroma) — the intensity or purity of a hue<br>Brightness (value) — the relative degree of black or white mixed with a given hue<br>Hue — another word for color | Smith (1978) |
| | 3 | Decorrelation Stretch | 3 | A linear, pixel-wise operation for visual enhancement<br>Output = T * (Input - mean) + mean<br>T = Transformation Matrix (Covariance-derived) | Stéphane (2009) |
| | 4 | CMYK | 4 | C, M, Y, and K represent the Cyan, Magenta, Yellow and Black components of the CMYK color space image. | Kang (1999) |
| Dimensionality Reduction | 5 | Grayscale | 1 | Single-band derivation of the RGB channels<br>Grayscale = (0.2898×R) + (0.5870×G) + (0.1140×B) | Burger and Burge (2010) |
| | 6 | PCA | 3 | The three variance-based principal components derived from RGB | Jolliffe (2002) |
| | 7 | MNF | 3 | Minimum Noise Fraction rotation transforms are used to determine the inherent dimensionality of image data, to segregate noise in the data, and to reduce the computational requirements for processing. | Boardman and Kruse (1994) |
| Texture Analysis | 8 | Gabor Filter | 4 | A filter bank representing a linear Gabor filter that is sensitive to textures with a specified wavelength and orientation. | Jain and Farrokhnia (1991) |
| | 9 | Wavelet Transform | 1 | Single-level discrete 2-D wavelet transform derived by taking the tensor products of the one-dimensional wavelet and scaling functions, leading to decomposition of 4 components. The approximation coefficients are used as a single band. | Meyer (1995) |
| | 10 | Convolution Filter | 1 | convolution is used to apply filters to an image. A filter (kernel function) is a small matrix of numbers, which is slid over an image, performing a mathematical operation at each position to create a new filtered output image. | Gonzalez et al. (2003) |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 11 | | Correlation Filter | 1 | GLCM-derived Correlation as a measure of how correlated a pixel is to its neighbor over the whole image. Range = [-1 1] | Haralick et al. (1973) |
| Statistical Indices | Local Spatial Statistics | 12 | Entropy | 1 | Local entropy of grayscale image (Intensity). Entropy is a statistical measure of Randomness | Gonzalez et al. (2003) |
| | | 13 | Gradient Weight | 1 | Calculates pixel weight for each pixel in the image based on the gradient magnitude at that pixel and returns the weight array. The weight of a pixel is inversely related to the gradient values at the pixel. | Gonzalez et al. (2003) |
| | | 14 | Std. Deviation Filter | 1 | Local standard deviation of the image pixels | - |
| | | 15 | Range Filter | 1 | Local range of image pixels (Min - Max) | - |
| | Pixel Statistics | 16 | Mean Abs. Deviation | 1 | Pixel-based mean absolute deviation of the color bands | - |
| | | 17 | Variance | 1 | Pixel-based variance of the color bands | - |
| | | 18 | Sum of Squares | 1 | Pixel-based sum of squares of the color bands | - |
| **Total number of indices (excluding RGB) = 17 Total number of bands (excluding RGB) = 31** | | | | | | |

#### *3.3.2.3 Feature Ranking and Selection*

The color, statistical, and texture transformations and dimensionality reduction outputs are stacked for the training image and sampled at the locations of updated liquefaction labels and the pixels with ROI labels, and the building pixels are removed. The multi-band stacked dataset is then normalized before being fed to the semi-supervised classification algorithm. Normalization significantly reduces the impact of data range on the final liquefaction prediction model and increases the processing and convergence speed (Rojas 1996). Using the mean (μ) and standard deviation (σ) of individual features, data of all features are normalized using Equation 1.

$$z_i = (x_i - \mu)/\sigma \quad (1)$$

The machine learning-based classification algorithm proposed in this study is designed to not only learn from the RGB channels of the imagery but also analyze the extracted color, texture, statistical and dimensionality reduction output bands. Due to a

large number of bands, a feature ranking/selection method is required to find the most relevant features. Feature selection is required to avoid redundancy and irrelevance by removing low-weight features. The higher the feature weight, the better the feature can separate the classes.

The methodology used herein implements the Minimum Redundancy Maximum Relevance (MRMR) method (Ding and Peng, 2005), which is used to identify important predictors for the multi-class classification problem. The MRMR algorithm finds an optimal set of mutually and maximally dissimilar features that can represent the response variable effectively. The algorithm minimizes the redundancy and maximizes the relevance of the feature set to the response variable. The algorithm quantifies the redundancy and relevance heuristically using the mutual information of variables, which means pairwise mutual information of the features and mutual information of each feature and the binary response. The mutual information between two variables measures how much uncertainty of one variable can be reduced by knowing the other variable. The MRMR algorithm uses this definition to compute the mutual information values used in its iterative process to rank the features by their classification importance (Darbellay and Vajda, 1999).

The MRMR algorithm in this study used all land use and landcover classes, plus the two liquefaction classes (9 classes in total), and ranks feature by the following steps: 1) Selecting the feature with the highest relevance to start creating the list of important features; 2) Finding features with non-zero relevance and zero redundancy complementing the feature chosen in step 1 and adding the most relevant feature to the list of selected features. If none of the features in this step has zero redundancy, the most relevant feature among none-zero redundancy features will be selected; 3) Repeating step 2 until no feature

with zero redundancy remains among the unselected ones; 4) Selecting the most relevant feature among none-zero redundancy features, and adding it to the selected list; 5) Repeating step 4 until all unselected features have zero relevance; and finally, 6) adding zero relevance features to the list of selected features in a random order as low-ranked features.

The MRMR feature ranking is done separately on each of the 4 categories of features extracted, and the selection of features to be used for the classification step is conditioned to being among the high weight features by investigating the change in the slope of the curve after plotting the feature weights, and choosing the features showing the considerable contribution to class separability. Additionally, to avoid redundancy in features used, a cap of two features per feature category is considered in this study. Then, the high-ranked features per category are added to the RGB channels to evaluate the performance of the classification model. Finally the preferred combination of features is introduced based on the accuracy results.

#### *3.3.2.4 Semi-Supervised Self-Training Classification via Linear Discriminant Analysis*

After feature selection, the semi-supervised classification model for liquefaction prediction is developed. The classifier was trained to distinguish liquefaction (two classes) and non-liquefaction (7 classes) pixels using the selected non-color features alongside the color imagery. Semi-supervised learning combines aspects of supervised learning, where all the training data is labeled, and unsupervised learning, where true labels are unknown.

Discriminant analysis is used as the semi-supervised classification method, which assumes that different classes generate data based on different Gaussian distributions

(Fisher, 1936). To train the classifier, the fitting function estimates the parameters of a Gaussian distribution for each class and finds the smallest calculated misclassification cost to predict classes of new (unlabeled) data. The linear discriminant analysis (LDA) used in this research computes the sample mean of each class. Then it computes the sample covariance by first subtracting the sample mean of each class from the observations of that class and taking the empirical covariance matrix of the result (Guo et al., 2007).

The semi-supervised algorithm proposed herein iteratively trains a LDAs classifier. This iterative process is called self-training. First, the function trains the LDA classifier on the labeled data alone and then uses that classifier to make label predictions for the unlabeled data. At each iteration, the algorithm calculates scores for the predictions. Then it treats the high-probability predictions as true labels (pseudo-labels) for the next training iteration of the classifier if the scores are above a certain threshold. This process repeats until the label predictions converge and all pixels are labeled (final labels).

Since the proposed method is pixel-based, a post-processing step is performed following the classification to avoid pixelated output maps, remove noise, and smooth the boundaries of predicted features. The median filtering approach is used in the post-processing step with a window size of 15-by15 pixels, which is considered as neither very small nor very large compared to the size of the liquefaction features. It should be noted that masking the buildings was a pre-processing step which masks building pixels from the beginning, while median filtering is a post-processing step in this study, which removes noise and isolated liquefaction predictions from the final binarized liquefaction/non-liquefaction map.

#### *3.3.2.5 Model Evaluation and Comparative Analysis*

After running the classification using the self-trained classifier on the whole image, except pixels within buildings footprints, the model is evaluated based on the binary liquefaction labels. The binary classification includes liquefaction as the positive class, and the confusion matrix specifies the predictions of each sample by one of the four categories, including true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The TP and TN represent that liquefaction (1) and non-liquefaction (0) are correctly classified, respectively, whereas FP represents that Non-Liquefaction is misclassified as Liquefaction, and FN represents the cases that Liquefaction is misclassified as Non-Liquefaction class.

The accuracy indices used to compare the different models include precision, recall, F1 score, and overall accuracy. The proportion of correctly classified observations in the positive (liquefaction) class is referred to as the Positive Predictive Values (PPV) or precision. A high precision indicates that the model has a high probability of correctly classifying positive samples. The recall or Negative Predictive Value (NPV) is the proportion of correctly classified observations in the negative (non-landslide) class. While the recall explains how sensitive the model is to identifying actual positive samples, it also quantifies the probability of detecting actual positive samples. Because precision and recall assess different aspects of the model, an index that combines the two is also used. The F1 score is the harmonic mean of the precision and recall, where an F1 score reaches its best value at 1 and worst at 0. The precision, recall, and F1 score formulas are provided via Equations 2 to 4.

$$Precision = TP/(TP + FP) \quad (2)$$

$$Recall = TN/(TN + FN) \quad (3)$$

$$F1\ Score = (2 \times Precission \times Recall)/(recission + Recall) \quad (4)$$

Different models are tested for liquefaction classification performance in this study. First, 4 different models based on high-ranked features extracted from the 4 categories of extracted features are evaluated and compared with the model only using RGB channels. Then, features from the categories with improved classification performance are combined, and the classification performance is again compared with the other models, including the one only using RGB. Finally, the preferred model is presented, and its performance is also compared with a supervised learning approach using the same linear discriminant analysis method to investigate the impact of the semi-supervised learning approach.

## 3.4 Results and Discussion

### 3.4.1 Training Data

The training data includes the two clusters of liquefaction pixels and the seven sets of ROI pixels. Figure 6 shows an example of the clustered liquefaction pixels, demonstrating how the Fuzzy approach helps detect and remove low-probability pixels. In Figure 6-a, the liquefaction ejecta has been traced within polygons as part of Sanon et al. (2023). The pixels within those polygons are included as training data and used in the Fuzzy-C Means clustering. In Figure 6-b, the results of the Fuzzy-C Means clustering are shown, the non-liquefaction pixels on the running path were omitted successfully. The liquefaction clusters are grouped into dark and light liquefaction ejecta. The performance of the algorithm in clustering the dark and light liquefaction pixels is visually acceptable. At the end of this step, the training data for liquefaction are more consistent and clustered based on pixel-based spectral properties.

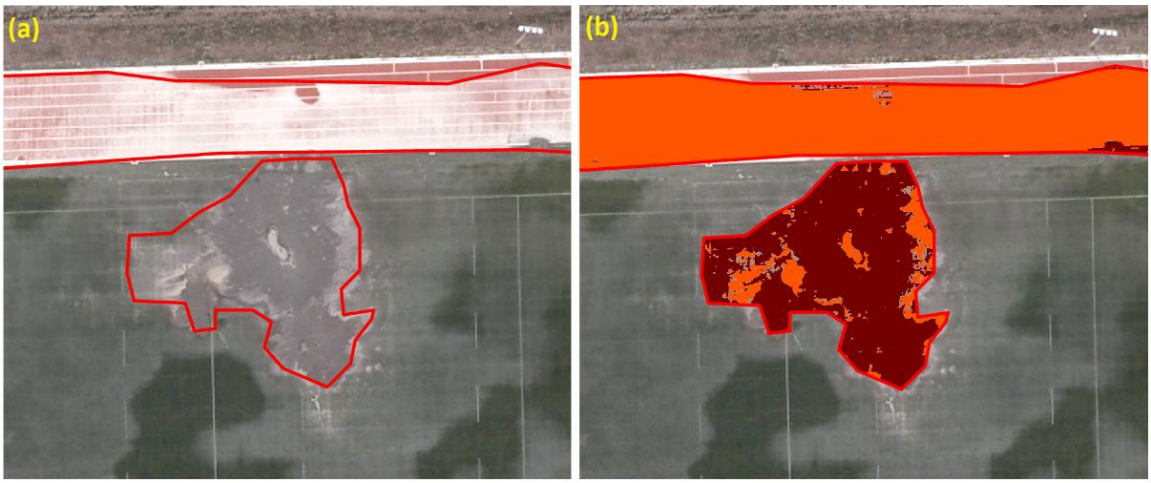



Figure 3-6. a) An example of the liquefaction ejecta polygons drawn by Sanon et al. (2023). b) Fuzzy C-Means clustering results for the example image tile with dark and light liquefaction. The dark red and orange colors are indicators of dark and light liquefaction, respectively. Uncertain pixels are removed from liquefaction training data based on their low probability of belonging to any of the dark or light liquefaction classes.

Representative ROIs (polygons) were drawn for 7 classes over the image tiles, as presented in Figure 7. In Figure 8, box plots of the RGB channels of collected data of these seven classes are compared with the clustered liquefaction pixels and pixels located within the building footprints (which are not used in the classification process). It is observed that the buildings have the widest range of values among the classes, overlapping with multiple classes. This reconfirms the benefit of masking out the buildings from training data. Shadow pixels, as expected, have the lowest mean among the classes and are the closest to the class of trees.

From a class separability point of view, considering the box plot quantiles and the RGB mean line, the closest classes are the classes of roads and pavements/driveways. Although the class of pavements/driveways has a wider range of values, and the median values of the individual RGB channels have some degree of deviation. At the same time, the RGB means to have a slight difference. It should be noted that even if these two classes

would have similarities, confusing the classifier in distinguishing them, it should not be harmful to the final performance of the model based on the purpose of this study, which is liquefaction detection, especially since the RGB mean of liquefaction classes (dark and light liquefaction) are sufficiently different from these two classes (roads and pavements/driveways)

There are some notable observations in Figure 9 about the two liquefaction classes. The light liquefaction class has the highest pixel values on average compared to the other classes. The median values of the individual RGB channels in the light liquefaction class are even higher than the 75th percentiles of all other classes, including the class of buildings in which bright roof pixels are common. The RGB mean of the dark liquefaction class is closest to the RGB mean of the soil class; however, looking at the individual box plots of their RGB channels, a difference is observed in terms of the distance between the medians of the channels within each class of dark liquefaction and soil. Another observation is that while for the trees and vegetation classes, the green channels have higher median values, the red channel dominates all the other classes in this sense. Since some liquefaction pixels are in vegetated areas, this might be a helpful distinguishment aspect.

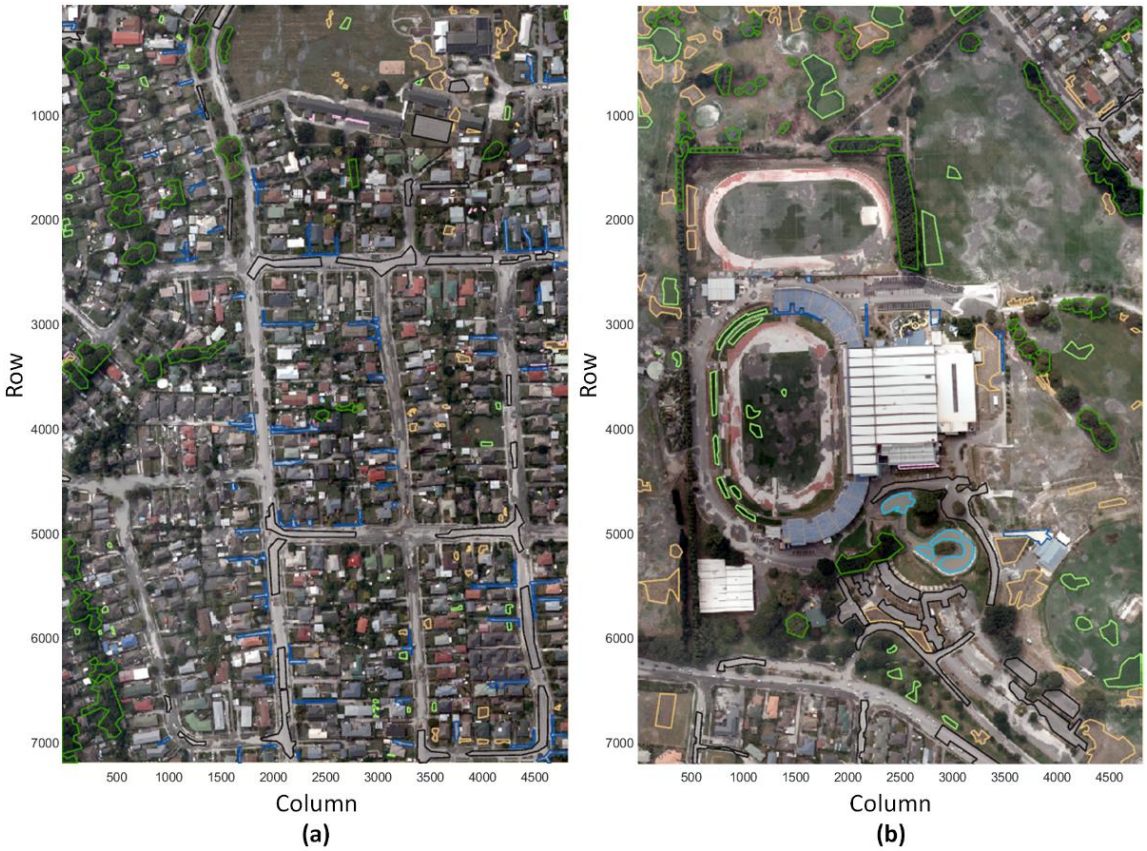



Figure 3-7. The drawn ROIs to collect training data for different land use and landcover classes on the two used tiles in this paper, including trees (dark green), vegetation (light green), soil (light brown), shadow (pink), water (light blue), roads (black), and pavements / driveways (dark blue).

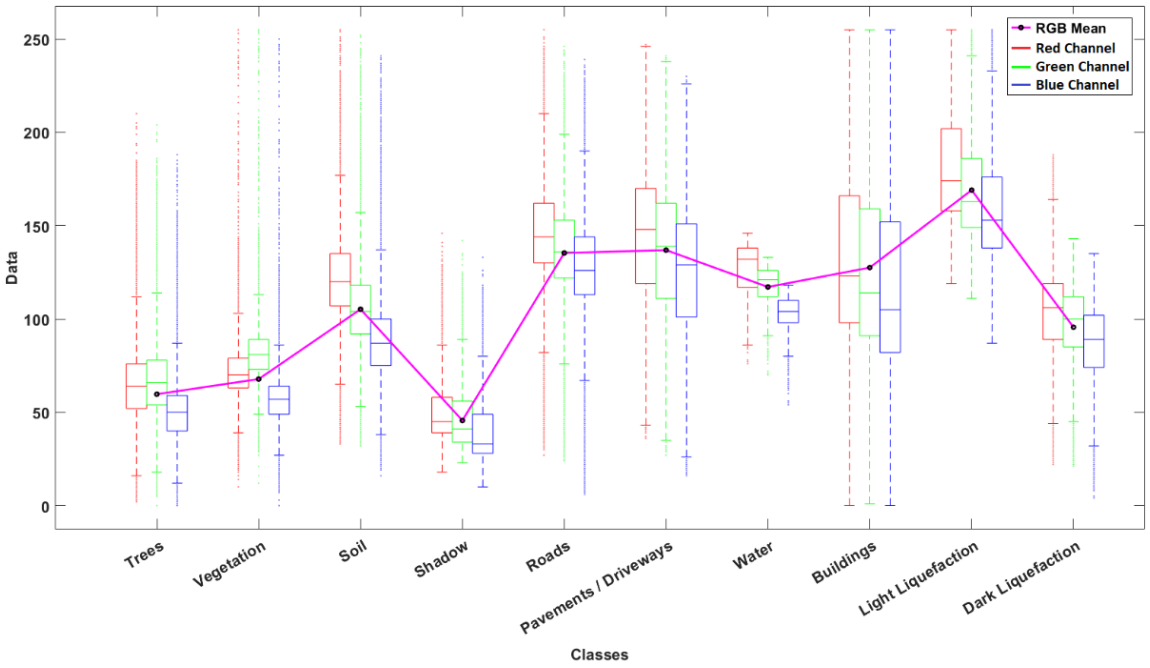



Figure 3-8. Comparative box plots of the individual RGB channels per class. The mean of RGB channels of each class is alco calculated and plotted on the figure using magenta color.

### 3.4.2 Feature Extraction and Ranking

Visualizations of the extracted features are provided via Figures 9 to 12 for a small region within the model development tile (Figure 3-a) categorized by color transformations (Figure 9), texture analysis (Figure 10), statistical indices (Figure 11), and dimensionality reduction techniques (Figure 12).

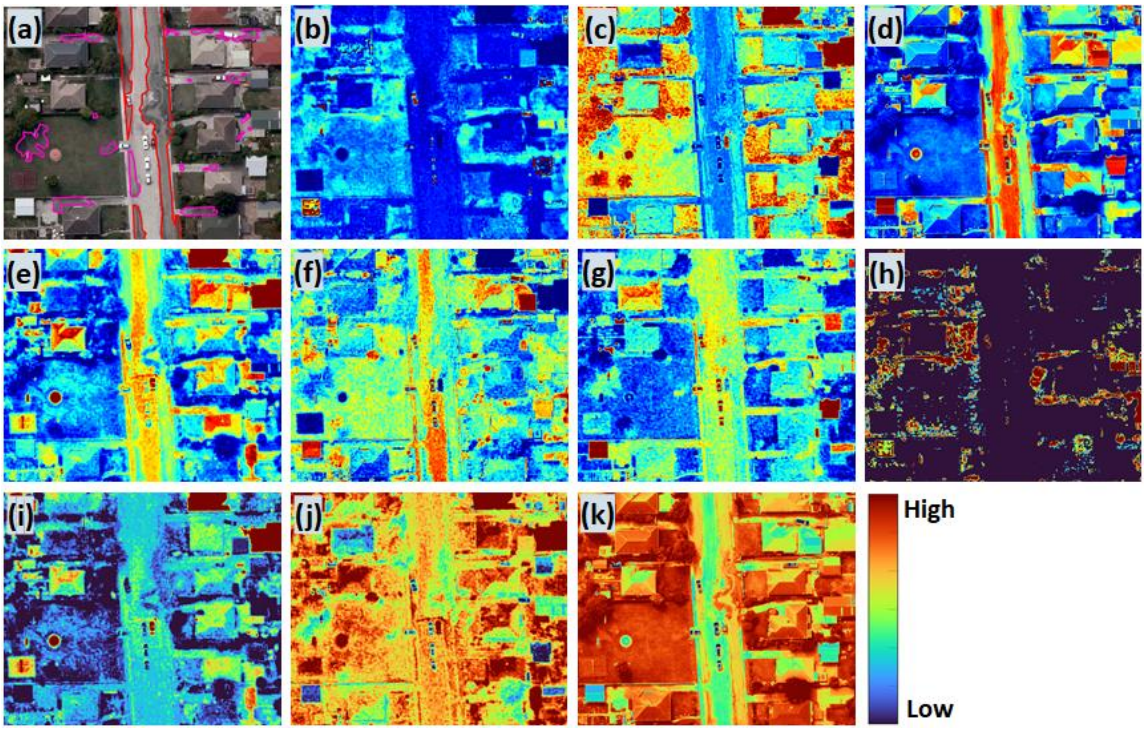



Figure 3-9. Features extraction outputs via color transformation techniques: a) The RGB image with Sanon et al. (2023) polygons (in red) and validation polygons used in this study (in magenta); b, c, d) Hue, Saturation, and Value bands of HSV transformation, respectively; e, f, g) Decorrelation stretch bands 1, 2 and 3, respectively; h) Cyan; i) Magenta; j) Yellow; and k) Black.

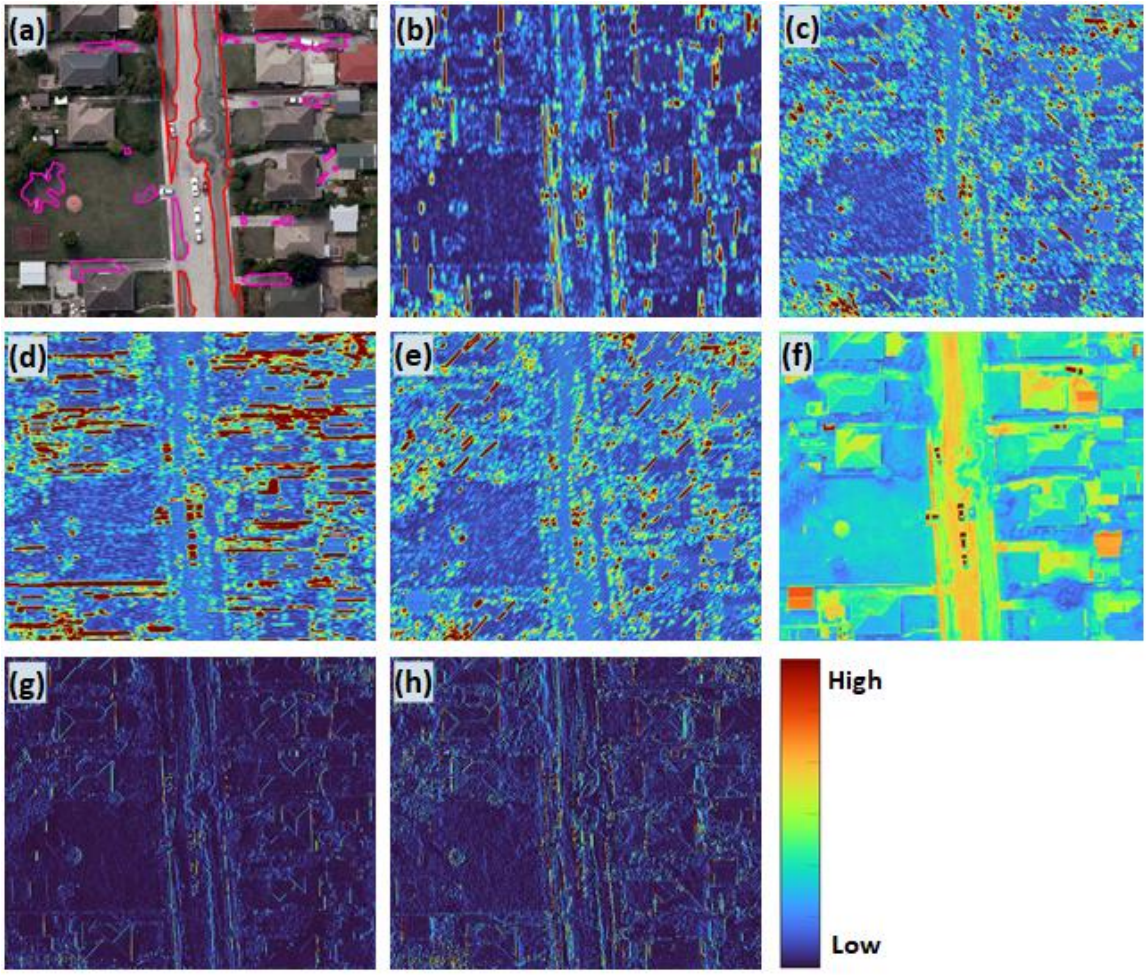



Figure 3-10. Features extraction outputs via texture analysis techniques: a) The RGB image with Sanon et al. (2023) polygons (in red) and validation polygons used in this study (in magenta); b, c, d, e) Gabor filters generated at 4 orientations (0, 45, 90, and 135 degrees, respectively) with a wavelength of 5 pixels/cycle; f) Approximation coefficients of the two-dimensional Haar wavelet transform with symmetric extension mode (half point): boundary value symmetric replication; g) Convolution filter; and h) Correlation filter.

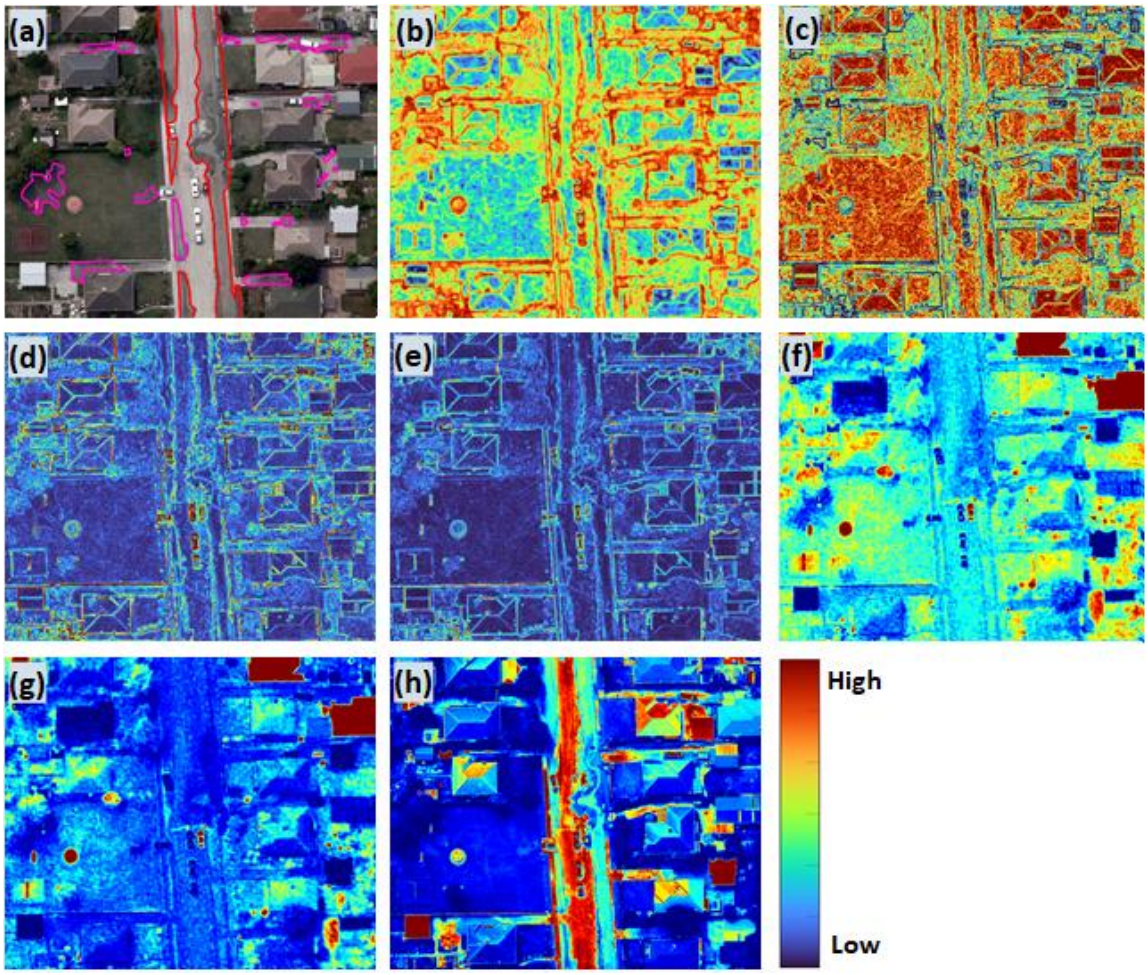



Figure 3-11. Features extraction outputs via statistical indices: a) The RGB image with Sanon et al. (2023) polygons (in red) and validation polygons used in this study (in magenta); b) Entropy filter; c) Gradient weight; d) Standard deviation filter; e) Range filter; f) Mean absolute deviation; g) Pixel variance; and h) Sum of squares.

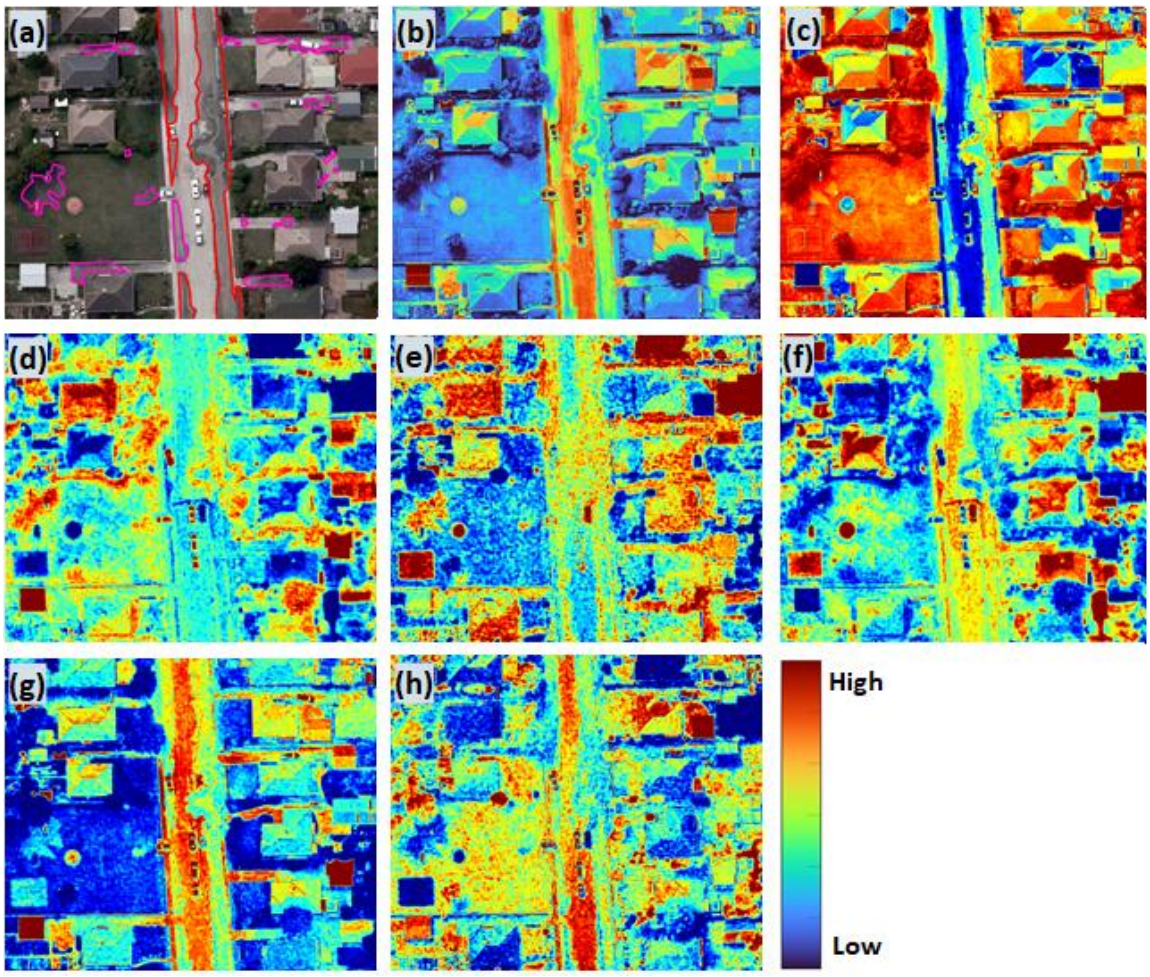



Figure 3-12. Features extraction outputs via dimensionality reduction techniques: a) The RGB image with Sanon et al. (2023) polygons (in red) and validation polygons used in this study (in magenta); b) Grayscale image; c, d, e) First, second, and third bands of PCA method's output, respectively; f, g, h) First, second, and third bands of MNF method's output, respectively.

Feature ranking was done by running the MRMR algorithm separately on each category of extracted features, and high-rank features of each category were considered for model development. The results of feature ranking are provided in Table 3. As discussed in Section 2.2.3, the minimum score of 0.5 was considered for selecting features, and the number of selected features per category was capped at 2 to avoid redundancy and classification model overfitting.

The results of feature ranking are presented in Table 3 and led to the selection of 7 high ranked features, including black and magenta channels of CMYK as two color transformation outputs, approximate coefficients of the wavelet transform as the only texture-derived band, Sum of the squares and gradient weight as two statistical indices, and the first two components of PCA analysis as the two bands from the dimensionality reduction outputs. A large score value indicates that the corresponding predictor is important. Also, a drop in the feature importance score represents the confidence of feature selection, meaning that the MRMR algorithm recognizes that adding additional features is not helpful to the classification. Based on this logic, the statistical indices seem more effective overall compared to the other three categories of features. In contrast, the texture components seem to be the least effective among the categories. It should be noted that the texture components might have been of higher importance if the buildings had not been masked as they have a texture that can be detected and learned by texture analysis techniques.

Comparing all the extracted features, regardless of the categories they belong to, the first component of the PCA has the highest score of 0.818, followed very closely by the sum of squares (statistical index) with the score of 0.813. The approximation coefficients of the wavelet transform (texture analysis) and the black channel (color transformation) also have very high feature scores, 0.794 and0.791, respectively.

Table 3-3. MRMR feature ranking results. The rank column in the table is indicator of rank within the category, not the overall rank.

| Feature Category | No. | Feature Description | Score | Rank | Selection |
|---|---|---|---|---|---|
| Color Transformation | 1 | Black (CMYK) | 0.791 | 1 | ✓ |
| | 2 | Magenta (CMYK) | 0.675 | 2 | ✓ |
| | 3 | Cyan (CMYK) | 0.652 | 3 | - |
| | 4 | Decorrelation Stretch (Band 1) | 0.409 | 4 | - |
| | 5 | Hue (HSV) | 0.404 | 5 | - |
| | 6 | Value (HSV) | 0.387 | 6 | - |
| | 7 | Decorrelation Stretch (Band 3) | 0.386 | 7 | - |
| | 8 | Saturation (HSV) | 0.304 | 8 | - |
| | 9 | Yellow (CMYK) | 0.271 | 9 | - |
| | 10 | Decorrelation Stretch (Band 2) | 0.198 | 10 | - |
| Texture Analysis | 11 | Approximation Coefficients (WT) | 0.794 | 1 | ✓ |
| | 12 | Gabor Filters (45 deg) | 0.232 | 2 | - |
| | 13 | Gabor Filters (135 deg) | 0.229 | 3 | - |
| | 14 | Gabor Filters (90 deg) | 0.215 | 4 | - |
| | 15 | Convolution Filter | 0.193 | 5 | - |
| | 16 | Gabor Filters (180 deg) | 0.185 | 6 | - |
| | 17 | Correlation Filter | 0.048 | 7 | - |
| Statistical Indices | 18 | Sum of Squares | 0.813 | 1 | ✓ |
| | 19 | Gradient Weight | 0.573 | 2 | ✓ |
| | 20 | Pixel Variance | 0.561 | 3 | - |
| | 21 | Range Filter | 0.553 | 4 | - |
| | 22 | Entropy Filter | 0.551 | 5 | - |
| | 23 | Mean Absolute Deviation | 0.412 | 6 | - |
| | 24 | Standard Deviation Filter | 0.158 | 7 | - |
| Dimensionality Reduction | 25 | PCA (Band 1) | 0.818 | 1 | ✓ |
| | 26 | PCA (Band 2) | 0.646 | 2 | ✓ |
| | 27 | MNF (Band 2) | 0.364 | 3 | - |
| | 28 | Grayscale Image | 0.260 | 4 | - |
| | 29 | MNF (Band 1) | 0.187 | 5 | - |
| | 30 | PCA (Band 3) | 0.103 | 6 | - |
| | 31 | MNF (Band 3) | 0.079 | 7 | - |

### 3.4.3 Semi-Supervised Classification

To understand the impact of each of the selected feature sets by category, the first step was to develop 4 classification models, each using a catogry of selected features, via the semi-supervised discriminant analysis classification algorithm presented in Section 2. A model only using RGB channels (Model 1) was also developed for comparative analysis. The reason was to understand which feature categories are better in distinguishing the two liquefaction classes (dark and light) from other classes of land cover and land use in practice. After running the multi-class classification, the labels were converted to binary

labels of liquefaction versus non-liquefaction to evaluate accuracy against the validation labels.

F1 score and overall accuracy indices were calculated to compare the performance of the evaluated models and presented in Figure 13. All models resulted in improved performance over the RGB-only model (Model 1) except for Model 3 (RGB + color transformation). One explanation for the relatively inferior performance of Model 3 can be the high correlation of black and magenta channels with the RGB color channels, leading to redundancy in data. The improvement in performance accuracy resulting from adding the texture component (Model 4) was only marginal. The highest model performance was for Model 2 (RGB + dimensionality reduction), with Model 5 (RGB + statistical indices) in second place.

Next, an integrated model (Model 6) was developed by using the combination of selected features, including Principal components and statistical indices, which showed superior performance among all developed models. Compared to the RGB-based model (Model 1), the preferred model (Model 6) has the highest overall accuracy of 88.73%. It means compared to the Model 1's accuracy of 84.53%, Model 6 is leading to ~1.45 million higher number of correctly predicted pixel labels across the image tile, resulting in significant corrections in the spatial extent of liquefaction over the image tile.

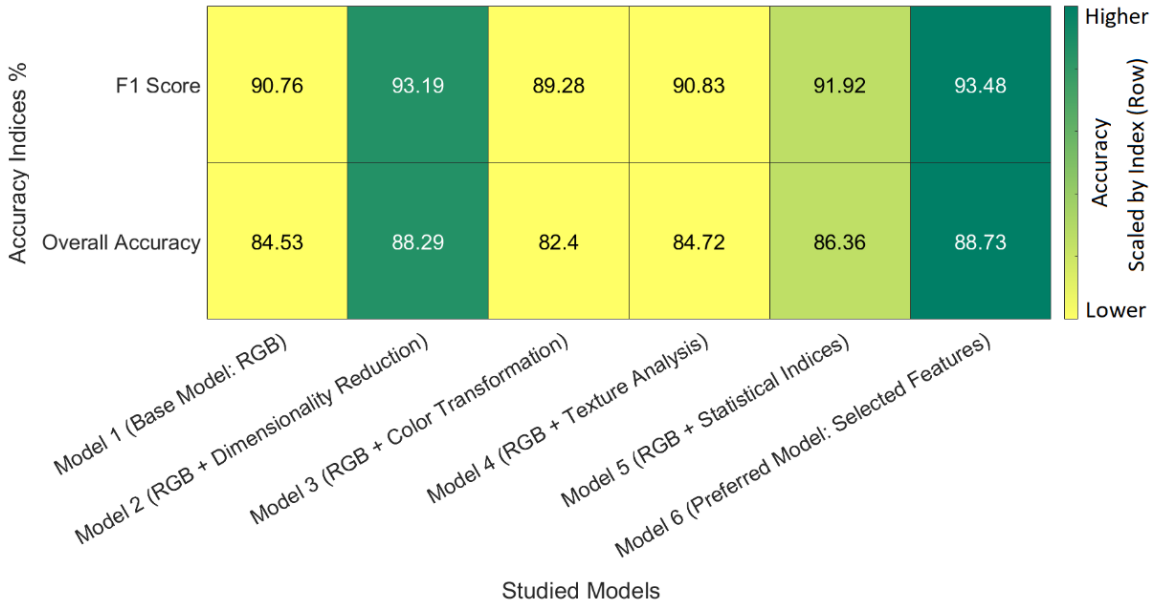



Figure 3-13. Binary classification accuracy results calculated by comparing the binarized classification labels with the validation liquefaction labels shown in Figure 3-a. The darker green color in the heatmap table indicates superior performance by the model.

Then, to analyze the benefit of the semi-supervised learning approach versus supervised learning, the same datasets and features used to develop Model 6 (preferred model) were used to train and validate a linear discriminant analysis classifier (same classifier used in the semi-supervised method). The results of the comparative analysis are provided in Figure 14. The results show that the proposed semi-supervised self-training classifier provides improved performance compared to supervised learning (higher precision, F1 score and overall accuracy, and very close recall score). This might be due to the benefits gained by learning not only from labeled data but also from unlabeled data. The overall accuracy was increased from 87.32% to 88.73%, leading to ~0.49 million higher number of correctly predicted pixel labels.

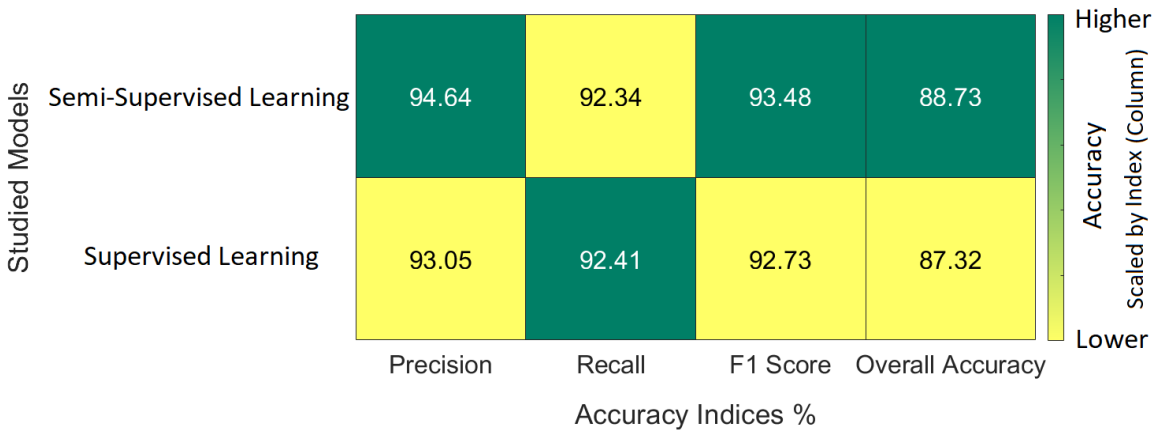



Figure 3-14. Binary classification accuracy results calculated by comparing the binarized classification labels with the validation liquefaction labels shown in Figure 3-a. The darker green color in the heatmap table indicates superior performance by the model.

### 3.4.4 Liquefaction Map Visualization and Model Application

Using Model 6, the classification was applied to the model development tile (Figure 15), showing the resulting map of liquefaction ejecta, and the predictions are also visually compared with Sanon et al. (2023) labels (Figure 15-a) and the validation labels (Figure 15-b). In addition, Figure 15-c provides the final map derived from the supervised learning approach. Figure 15-d also provides the map of liquefaction predicted via Model 1 as the base model using only RGB channels.

Figure 16 presents some examples of the Model 6 performance, compared with the base Model 1 with a closer view of different locations within the model development image tile. Figure 16-a is an example showing how the over-prediction of Model 1 is reduced by Model 6. In addition, Figure 16-a shows that the objects such as machinery are excluded from the model prediction. At the same time, they have been included within validation liquefaction polygons due to the difficulty of excluding them when drawing the polygon around such a large-scale liquefaction feature.

Figure 16-b is an example of boundary correction for the large-scale feature partially labeled by Sanon et al. (2023), perhaps due to the discontinuities or uncertainties in determining the extent of the feature. This example also shows some model over-predications on bare soils which was introduced in Section 2 as the most statistically similar class to the dark liquefaction classes. Figure 16-c also shows the power of the model in developing an almost spatially complete inventory of liquefaction occurrences using the proposed semi-supervised learning with partially labeled data.

Investigating individual features in Figure 16-d shows that Model 6 corrects the over-predictions made by Model 1 and predicts several missing liquefaction features which were not predicted by Model 1. Model 6 also improved the boundaries and the quality of the predictions within the liquefaction polygons even by correcting the labels where non-liquefaction areas are included within liquefaction polygons. Figure 16-d shows that small features of dark liquefaction within highly vegetated (darker green) areas are more likely to be missed by the model, rather than the ones in driveways with dark backgrounds.

Figure 16-e shows a relatively poorer performance of the model in the detection of light liquefaction compared to dark liquefaction features. The reason is that in the tile used for model development, there are significantly fewer light liquefaction pixels than dark liquefaction. This can cause the classification model not to learn enough from the limited data. The other notable point in this example is the tendency of the model to over-predict liquefaction features by expanding their boundary over the vegetated areas and detecting the very shallow traces of liquefaction, which were not included as ground-truth labels by validation liquefaction polygons. Figure 16-e also shows the confusion between dark asphalts and pavements with the dark liquefaction class, leading to some over-prediction.


Figure 3-15. a) Binary classification results of Model 6 (preferred model) with Sanon et al. (2023) liquefaction polygons (in red) overlaying the model prediction (in yellow); b) Model 6 output (in yellow) compared with validation labels (in magenta); c) Supervised classification output with validation labels overlayed; d) Model 1 (RGB-based model) classification output with validation labels overlayed.

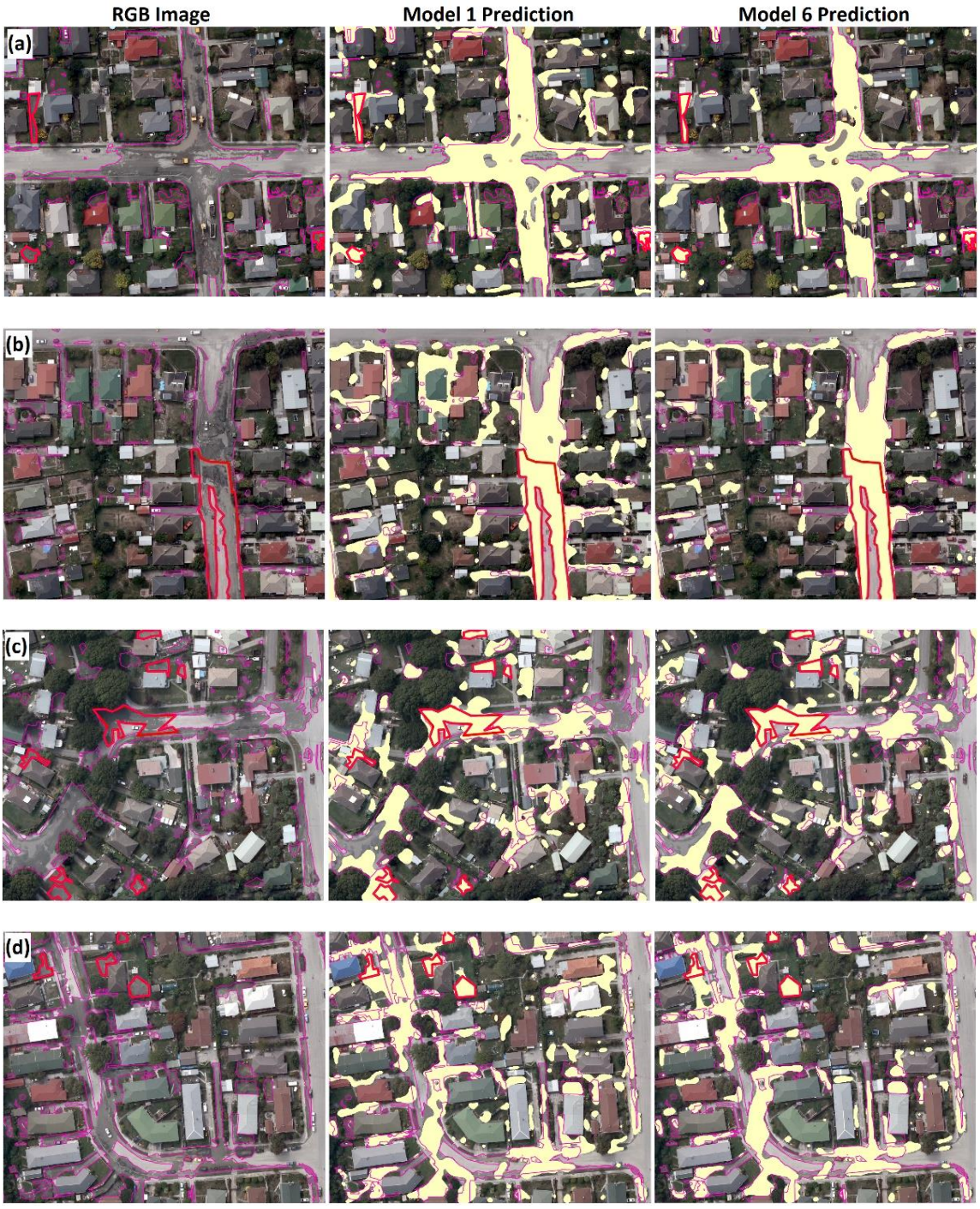

RGB Image
Model 1 Prediction
Model 6 Prediction
(a)
(b)
(c)
(d)


Figure 3-16. Binary classification results of Model 1 (RGB-based model in the middle column) Model 6 (preferred model in the right column) with Sanon et al. (2023) liquefaction polygons (in thick red) and validation labels (in magenta) overlaying the model prediction (in yellow).

The same approach was applied on another tile (Figure 2-b) to check the applicability of the approach on a new image. Finally, the classification output is compared with Sanon et al. (2023) labels via Figure 17, and Figure 18 shows a few examples of model predictions. In this image, contrary to the model development tile, a significant amount of light liquefaction pixels is available.

Visual inspection of the maps shown in Figure 18 shows that the algorithm works well in the detection of liquefaction features but with some over-predictions across the tile. Figure 18-a is an excellent example of the model performance in detecting liquefaction traces as liquefaction. In this case, since the partial labels by Sanon et al. (2023) include such pixels, it is an expected performance. The same observation applies to Figure 18-b, where a polygon is drawn by Sanon et al. (2023) over the liquefaction traces.

Figure 18-c shows an excellent performance of the model in detecting liquefaction occurrences across the region and completing the partially labeled inventory. However, Figure 18-d indicates that there are some inconsistent labels between asphalt and liquefaction, which can be due to the lack of availability of road pixels in this tile compared to the model development tile. The authors recommend that the area of training data

collection should be large enough for the classifier to capture as many patterns as possible within the data. It is also recommended that drawing ROIs should be done in multiple tiles to gather a variety of pixels per class to train the classifier better. Finally, it should also be noted that the applicability of this method depends on the capture of post-event imagery on time before the surface dries, and before human interference causes the removal of liquefaction-related surface effects. Additionally, other sources of surface moisture should not be present to interfere with liquefaction ejecta effects.

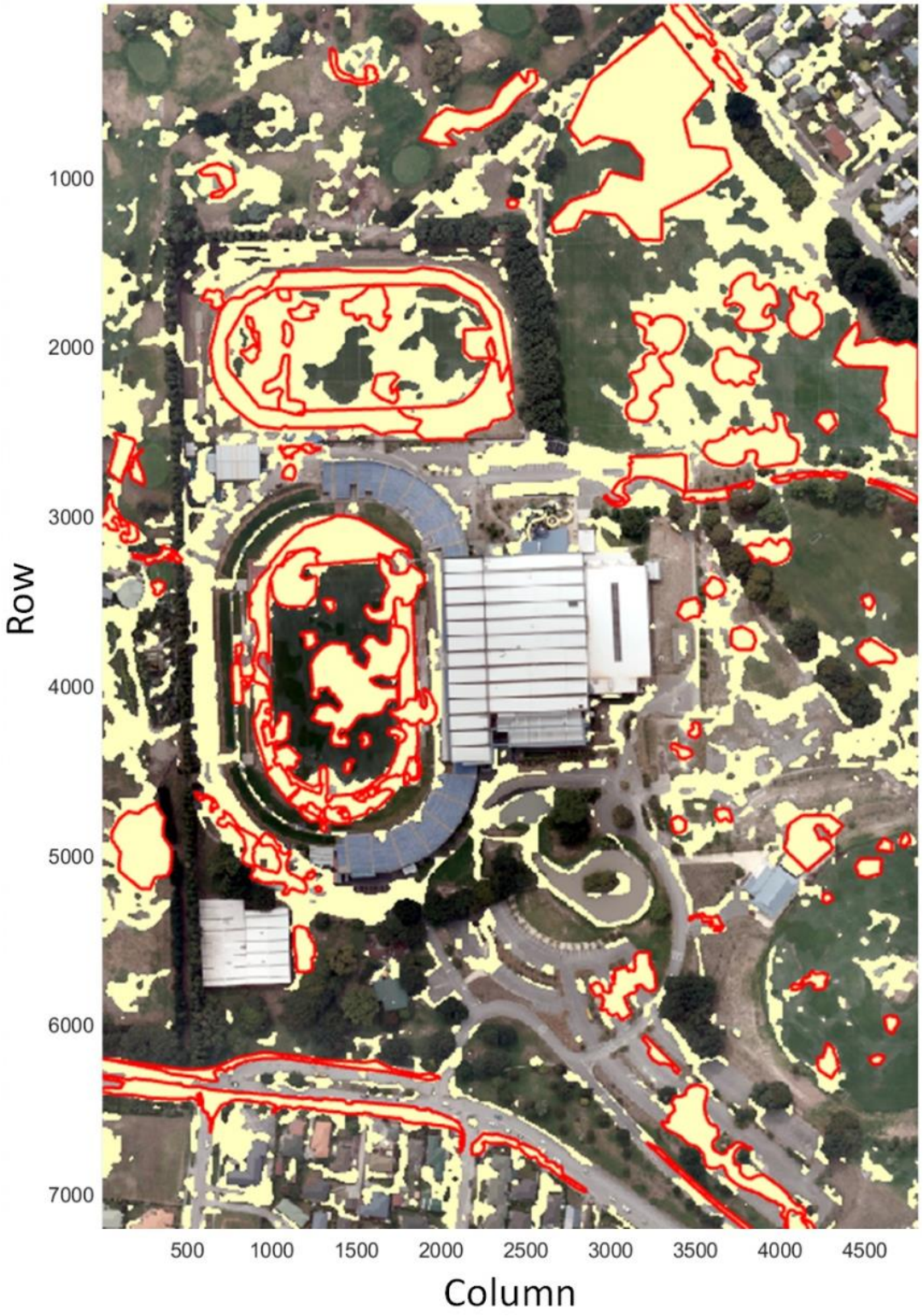



Figure 3-17. The model application image tile with binary classification results of Model 6 (preferred model in the right column) and Sanon et al. (2023) liquefaction polygons (in thick red).

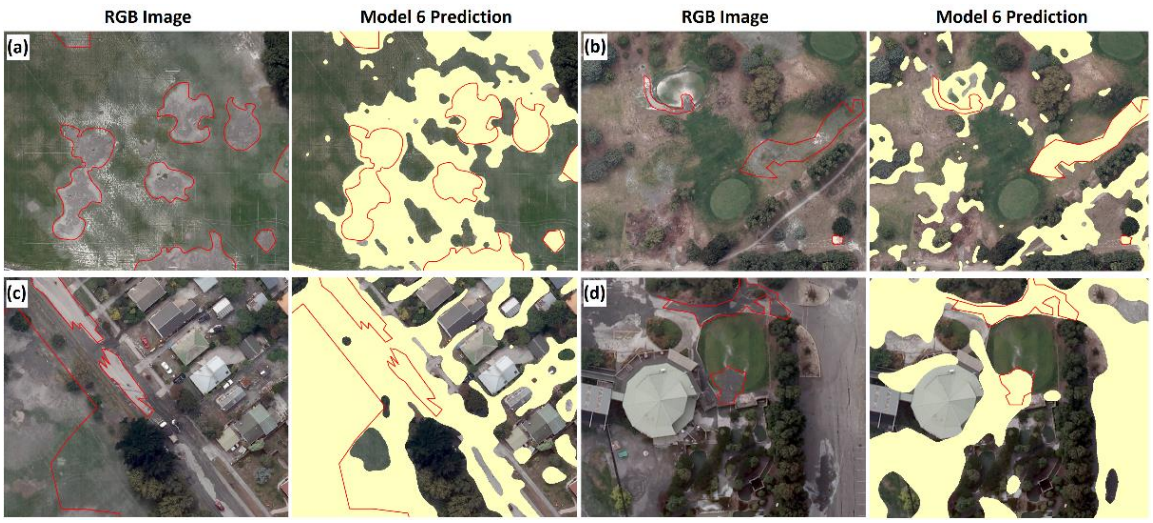



Figure 3-18. The model application example images with binary classification results of Model 6 (preferred model in the right column) and Sanon et al. (2023) liquefaction polygons (in thick red).

## 3.5 Conclusions

The idea of this paper was to generate spatially consistent and complete maps of liquefaction ejecta using a small number of visually labeled liquefaction surface effects using high-resolution aerial imagery with the aid of advanced computer vision and machine learning algorithms. A semi-supervised self-training classification method was presented and validated in this study and showed superior performance in the detection of liquefaction surface effects compared to the supervised learning approach. This paper also investigated using RGB image-derived products as inputs into the classification algorithm, such as transformations based on color, texture, statistics, and dimensionality reduction. The results showed enhanced performance of liquefaction labeling when selected statistical indices and dimensionality reduction techniques' outputs were used alongside RGB channels of the imagery. Texture and color transformation components did not impact the model performance considerably. The feature ranking was done via the MRMR algorithm, which reduces the redundancy and increases the relevance of the features to the classes of

interest. A Fuzzy-based clustering approach was proposed for clustering liquefaction pixels into two classes of dark and light liquefaction to increase the certainty of the pixels by removing outliers and mixed pixels from the labeled liquefaction training data. The advanced machine learning and image processing methods used in the study made it possible to spatially complete and augment the partial inventory of liquefaction labels in the image library provided by Sanon et al. (2023) with acceptable accuracy. The proposed methodology can be applied in other domains where labeled and high-quality training data is limited, and there is a need for semi-supervised algorithms to classify data due to the difficulty of labeling.

## 3.6 Supplementary Materials

The output augmented liquefaction maps resulted from this study are published as digital files in the GitHub repository of the first author at https://github.com/adel-asadi/Christchurch_Augmented_Liquefaction.

## 3.7 Acknowledgments

The support of National Geospatial Intelligence Agency (NGIA) for this research through the NGIA Academic Research Program Grant #HM0476-20-0006 (NGA NURI Project: Benchmark data development to classify damage for natural disaster relief efforts) is greatly appreciated.

# 4 Regional Landslide Mapping Model Developed by a Deep Transfer Learning Framework Using Post-Event Optical Imagery

Adel Asadi [a], Laurie Gaskins Baise [a], Snehamoy Chatterjee [b], Magaly Koch [c], and Babak Moaveni [a]

[a] Geohazards Research Lab, Department of Civil and Environmental Engineering, School of Engineering, Tufts University, Medford, MA 02155, USA.

[b] Geological and Mining Engineering and Sciences Department, Michigan Technological University, Houghton, MI 49931, USA.

[c] Canter for Remote Sensing, Earth and Environment Department, Boston University, Boston, MA 02215, USA.

(The material contained in this chapter is being submitted for review and possible publication in Georisk Journal – Taylor & Francis Publications)

## 4.1 Abstract

Landslides are a significant natural disaster typical of mountainous terrains that can be triggered by earthquakes and extreme rainfall events. Manual delineation of landslide features to create inventories by visual interpretations of optical remotely sensed imagery is a tedious and time-consuming process; therefore, post-disaster automatic detection of landslides is desirable. Deep convolutional neural networks are a state-of-the-art computer vision methods that can be used for landslide detection. A few recent studies use both pre-event and post-event RGB channels in a change detection framework for detection of landslides or they include different sources of data, including digital elevation model (DEM) products. Since the performance of deep learning depends on the quality and quantity of training data, it is challenging to use such methods in a timely manner after the natural hazard events. Previous studies have not well investigated and reported the deep learning landslide mapping models' transferability from a training event to a testing event.

In this study, the transferability of a trained deep convolutional network using the landslide inventory from the 2016 Kumamoto Earthquakes 6.6 Mw only based on post-event true-color optical imagery is analysed and a regional landslide mapping model is introduced. The landslide affected areas after the 2018 Hokkaido earthquake in Japan are used to test cross-event transferability. The landslide affected areas after the 2017 Asakura Rainfall event in Japan are used for a second cross-event transferability test with a different landslide triggering factor. The rationale for selecting these three case studies is the regional similarity of the events, all occurring in Japan, containing vegetated and forested mountainous areas, including urban settings, with comparable geology, topography and landcover. The presented deep transfer learning model uses a ResNet50-based DeepLabV3+ semantic segmentation method as an automated way to extract landslide features from post-event imagery instantly given the availability of the imagery without the need to fine-tune the model or to develop event-specific training data. High accuracy was achieved on the testing data from the 2016 Kumamoto Earthquakes (88.77 %), and on the case studies (unseen events) of the 2018 Hokkaido earthquake (85.75 %) and the 2017 Asakura Rainfall (86.84 %).

## 4.2 Introduction

Landslides are a type of significant natural disaster typical of mountainous terrains that can be triggered by earthquakes and extreme rainfall events (Keefer 1984; Guzzetti 2006). Due to climate change trends, the frequency of such disasters is increasing, and imposing threats to sustainable development in susceptible areas (Ghorbanzadeh et al., 2021). Landslides' spatial extent (area) range in size from less than a few square meters, for shallow soil slides,

to thousands of square kilometers, for large submarine failures (Cruden and Varnes 1996). Due to this scale, landslides are often easily visible in optical remotely sensed imagery, especially high-resolution imagery. This has led to numerous landslide inventories developed using visual inspection of imagery as well as automated machine learning and computer vision approaches (Dou et al. 2015; Zhong et al. 2020). However, landslides in different domains may happen due to different triggers, in different land cover types, with different shapes, and within soil surface characteristics, which makes it challenging to automate the landslide mapping process (Kurtz et al., 2014). Landslide mapping can be described as a pixel-wise classification process in which each pixel in the georeferenced remote sensing image is assigned to either a landslide or non-landslide class (Zhong et al., 2020). This process is also called semantic segmentation which is a branch in computer vision that assigns a semantic label to each pixel in the image. An alternative computer vision approach which could be used to detect and map landslides is object detection which is based on locating landslides as objects in the imagery.

Remote sensing data has been widely used to create landslide inventories, to map landslides, and to perform landslide susceptibility analysis (Zhong et al. 2020; Casagli et al. 2023). High-resolution imagery within multispectral (Martha et al. 2011; Hölbling et al. 2017; Mohan et al. 2021) and microwave (Solari et al. 2020; Mondini et al. 2021) domains of the electromagnetic spectrum have been used to create landslide inventories. Landslides can also be mapped using image processing techniques that extract surface changes from synthetic aperture radar (SAR) intensity images (Bai et al. 2017; Ji et al. 2018; Aimaiti et al., 2019; Liu and Yamazaki, 2020; Ohki et al., 2020). Advanced machine learning-based classification and computer vision techniques have been used in a wide range of

geotechnical applications, including landslide identification (mapping) via remotely sensed data (Moosavi et al. 2014; Fanos et al. 2018; Tehrani et al. 2021; Mohan et al. 2021), and landslide susceptibility modeling (Dou et al. 2020-a; Fang et al. 2020; Cui et al., 2021; Liu et al. 2021).

There are a number of studies that are using deep learning to map landslides, but most of those studies use multiple data sources beyond the post-event RGB channels for detection of landslides. They either include different sources of data, including multispectral data and digital elevation model (DEM) products (Huang and Xiang, 2018-a; Ghorbanzadeh et al., 2019; Ye et al., 2019; Dou et al., 2020-b; Qi et al., 2020; Su et al., 2021; Meena et al., 2022), or propose multi-temporal change detection analysis (Lei et al., 2019; Liu et al., 2020; Lv et al., 2020; Zhang et al., 2020-a;; Yu et al., 2021). LiDAR data as a DEM product has also been used for landslide mapping via deep learning (Fang et al., 2022), as well as SAR data (Boehm et al., 2022; Nava et al., 2022-b). Object detection-oriented deep learning models have also been introduced for landslide mapping (Lu et al., 2020; Fu et al., 2022; Yang et al., 2022). Li et al. (2022) have combined the deep learning-based object detection approach with segmentation in a two-phase model which firstly detects landslides in large scale, then segments the objects by boundary detection.

When optical imagery is used for landslide mapping, the approaches can be divided into two categories 1) those using both pre- and post-event imagery for change detection and 2) those using only post-event imagery. Change detection approaches require a pair of imagery ideally with matched spatial and spectral resolutions. As result, many change detection attempts at landslide mapping suffer from errors resulting from differences in

look angle or differences in atmospheric conditions or require significant processing. An approach that can be applied to the post-event image only and does not require significant post-processing is desirable. Nava et al. (2022-b) have used deep learning for landslide detection on the 2018 Hokkaido earthquake event via different combinations of synthetic aperture radar (SAR) data bands, optical images, elevation, and slope data. They concluded that optical imagery (Red-Green-Blue channels) had the best performance in landslide detection. A few other studies have used only RGB channels of the high-resolution (3 m) Planet Satellite imagery with four spectral bands (RGB plus Near-Infrared) in their deep learning-based landslide mapping of the 2018 Hokkaido earthquake (Zhang et al., 2020-b; Yang and Xu, 2022).

Many applications of deep learning to landslide mapping only present testing and validation of the model for the single event, typically by dividing a specific region into training and testing subregions. Therefore, the transferability of the model to new regions and events is not well investigated. Thus, transferable deep learning models have recently been trending in the literature (Huang et al., 2018-b; Tong et al., 2020). The development of a trained deep learning model for landslide detection that can quickly applied to new events is of great benefit and importance. Zhang et al. (2023) have developed a global inventory of landslides in different terrain conditions to tackle this issue; however, their trained deep learning model does not report model transferability to unseen events. Deep transfer learning has also been used recently for landslide mapping using post-event RGB in an object detection framework (Hacıefendioğlu et al., 2021), bi-temporal RGB (Xu et al., 2022), and multi-temporal multi-spectral (RGB + Near-Infrared) channels (Bhuyan et al., 2023). In another recent effort, Ghorbanzadeh et al. (2022) have developed an inventory

of reference benchmark landslide data from four earthquakes (2018 Hokkaido Iburi, 2018 Kodagu, 2015 Gorkha, and 2009 Taiwan) which contains 10 m resolution optical and multispectral layers from Sentinel-2, digital elevation and slope layers from ALOS PALSAR, and pixel-wise landslide labels. They have trained different types of deep learning algorithms for landslide detection on those events by partitioning the data of each earthquake into training and testing tiles but did not evaluate transferability of models to unseen events (Ghorbanzadeh et al., 2022).

A few studies have been conducted previously to assess earthquake-induced damage using satellite optical satellite imagery (Kim and Kim 2018; Xu et al. 2018), multi-temporal airborne LiDAR (Yamazaki et al. 2018; Konishi and Suga 2018; Liu et al. 2019) and SAR data (Nava et al. 2022-a) for landslide mapping after Kumamoto 2016 earthquake. UAV (Unmanned Aerial Vehicles) imagery-derived three-dimensional (3D) models for damage mapping (Yamazaki et al. 2017), and a combination of non-optical remote sensing data for surface displacement and landslide mapping (He et al. 2019; Uemoto et al. 2019) are also performed in the aftermath of the 2016 Kumamoto Earthquakes.

In this study, the landslide inventory developed for the 2016 Kumamoto Earthquakes 6.6 Mw and the co-registered post-event aerial true-color RGB imagery with 0.5 m resolution, taken by the Geospatial Authority of Japan (GSI), are used to train and test a deep transfer learning model. The proposed deep transfer learning model uses a ResNet50-based DeepLabV3+ semantic segmentation method as the base model. Then, the landslide affected areas from the 2018 Hokkaido earthquake in Japan and the 2017 Asakura Rainfall event in Japan are used to test cross-event transferability. The rationale for

selecting these three landslide case studies is the regional similarity of the events, plus availability of high-resolution optical imagery and ground-truth landslide labels with acceptable quality. The model is trained, evaluated, and applied on vegetated and forested mountainous areas, including urban settings in Japan, with comparable geological, topographic, landcover and land use settings. This research presents a regional deep transfer learning-based landslide mapping model as an automated way to extract earthquake-induced landslides from post-event imagery instantly after the availability of the imagery without the need to fine-tune the model or to perform supervised learning for the new event.

## 4.3 Datasets

The datasets used in this study include post-event true-color RGB aerial imagery and the corresponding landslide inventories related to two earthquake events and a rainfall event, all in Japan. The events include the 2016 Kumamoto Earthquakes which occurred in the Kyushu Province, the 2018 Hokkaido Iburi Earthquake which occurred in the Hokkaido Province, and the 2017 Asakura Rainfall event which occurred in the northern Kyushu province. Figure 1 shows the map of earthquake epicenters and the location of the rainfall event, and the areas considered with the landslide polygons overlaying on the post-event imagery.

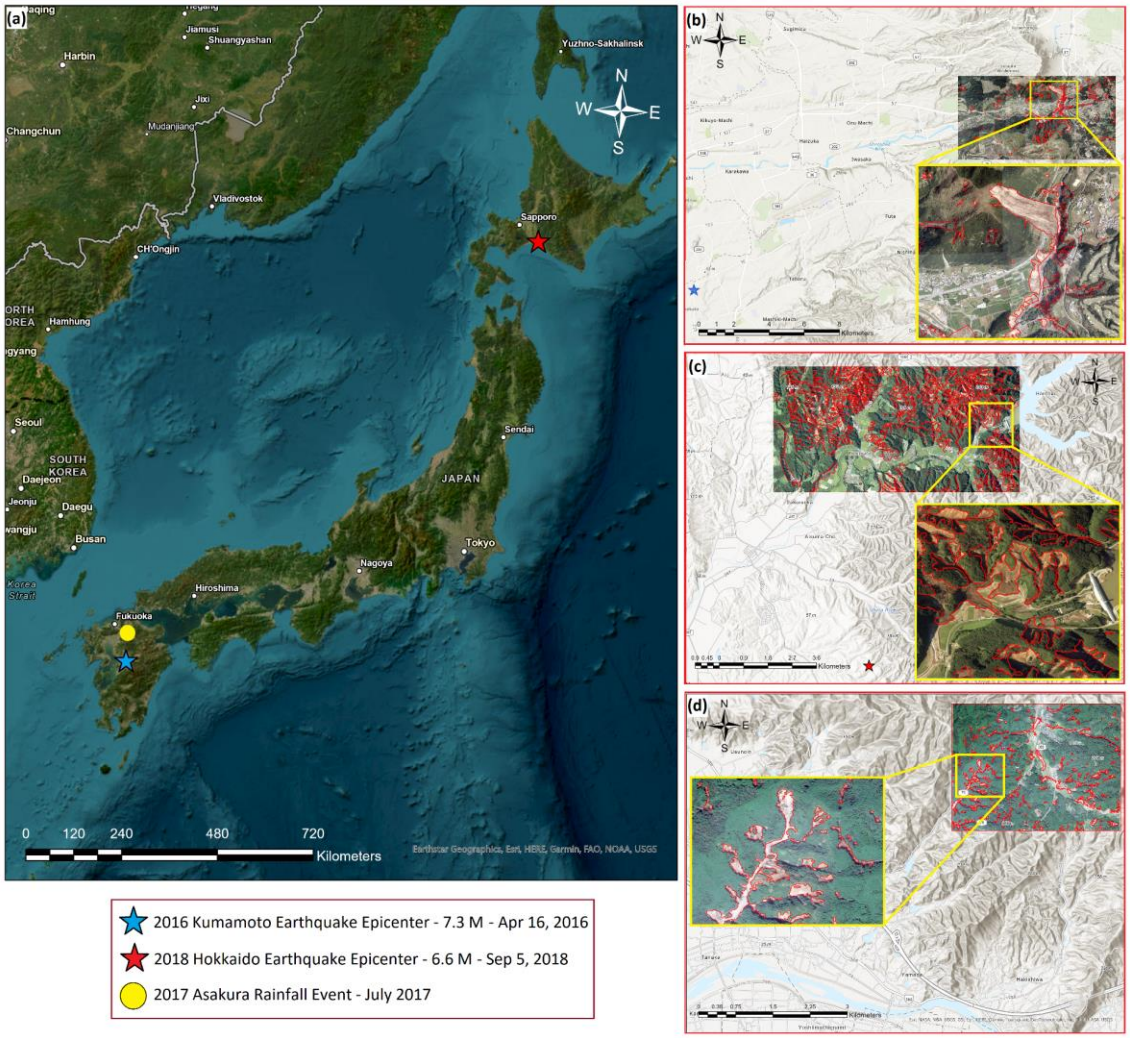



Figure 4-1. **a)** Satellite map of Japan showing the epicenters of the two historical earthquakes (2016 Kumamoto and 2018 Hokkaido) and the 2017 Asakura rainfall event. The imagery shown in this figure are the portion of available data used in this research: **b)** Post-event aerial image used in this study with the ground-truth NIED landslide polygons overlayed. The orthographic image resolution is 50 centimeters, and it is captured by GSI after the Kumamoto event on April 20th, 2016; **c)** Post-event aerial image used in this study with the ground-truth GSI landslide polygons overlayed. The orthographic image resolution is 3 m, and it is captured by GSI after the Hokkaido event on September 6th, 2018; and **d)** Post-event satellite optical image used in this study with the ground-truth landslide polygons overlayed. The orthographic image resolution is 0.59 m, and it is acquired from Google Earth after the Asakura rainfall event on July 13th, 2017.

### 4.3.1 2016 Kumamoto Earthquakes

The 2016 Kumamoto earthquakes in Kyushu Island, Japan, began with a magnitude 6.2 M earthquake on April 14, followed by another foreshock of magnitude 6.0 M on April 15, and a larger magnitude 7.0 M earthquake on April 16, 2016 beneath Kumamoto City, Kumamoto Prefecture on Kyushu, Japan (GEER Report 2016; GSI 2016-a). Figure 1-b shows the Kumamoto region analyzed in this study. Many natural slopes and road embankments failed ranging from small to medium-sized landslides, with 47% of the landslide feature having area of less than 103 $m^2$ (Chen et al. 2017-a). Landslides primarily occurred in volcanic ash, pumice and andosol, a highly porous dark-colored substance made up of volcanic ash combined with organic materials (Kiyota et al. 2017).

The National Research Institute for Earth Science and Disaster Resilience (NIED) of Japan have released a few versions of the landslide inventory for the 2016 Kumamoto events (NIED, 2016). NIED compared aerial photographs, satellite images, and airborne SAR images before and after the earthquake centering on Kumamoto to identify the locations of landslides and create a landslide inventory. This study uses the version published on April 21, 2016 (Figure 1-b), which was created by comparing the aircraft SAR image on the April 17th with the satellite image posted on Google Earth before the disaster, and by inspecting aerial imagery published by GSI (2016-b) on April 16, 19 and 20, 2016.

The NIED inventory is created based on defining a boundary of the landslide features as the area covering sediment movement, travel paths, and accumulation areas (Shinoda et al. 2019). Due to reasons such as insufficient field surveys and the human-

drawn nature of the boundaries (NIED 2016), the inventory should not necessarily be considered as absolutely true but is used as ground truth in this study. NIED set a size threshold of 100 m (moderate-size landslides) to identify and map landslide features (Chen et al. 2017-a), therefore, some moderate- or small-scale landslides were probably omitted (Xu et al. 2018).

Aerial images of the study area, taken on April 20, 2016, were collected from GSI repository (GSI, 2016-b). These post-event orthographic images taken near Nishihara 2 district have 0.5 m resolution. The cropped post-event image of the study area shown in Figure 1-b with landslide polygons overlayed covers an area of almost 42.71 $km^2$, consisting of approximately 16.61 million image pixels and 264 landslide polygons (Table 1).

### 4.3.2 2018 Hokkaido Earthquakes

The September 6th, 2018, Mw 6.6 earthquake struck the region of Iburi-Tobu, southern Hokkaido, Japan, covering a part of the Yubari Mountain range. The earthquake struck just one day after Typhoon Jebi had produced substantial rains in the region (Yamagishi and Yamazaki, 2018). The rainfall and the influence of Typhoon Jebi were found to be of secondary importance, compared to the earthquake (Wang et al., 2019). The topography of the region is characterized by gently rising and falling hills with elevations ranging from 100 to 800 m (Dou et al., 2020-b). The surficial geology is comprised of volcanic deposits of multiple layers formed by volcanic eruptions (Yamagishi and Yamazaki, 2018). These layers contain unconsolidated pumice, scoria, and coarse-grained volcanic ash particles (Dou et al., 2020-b). Additionally, the presence of weak and low-consolidated sedimentary

deposits makes the slopes prone to landslides (Ayalew et al., 2011).

The Geospatial Information Authority of Japan took and published the aerial ortho-photographs of the area affected by the earthquake (GSI, 2018-c). GSI has also published the interpreted landslide inventory, which are used as ground-truth labels in this study (GSI, 2018-b). According to GSI, the landslides were visually interpreted using high-resolution aerial photos taken between September 6th and 11th, 2018. It should be noted that deposits and sands on the roads and rivers may have been removed before image captures (GSI, 2018-b). Figure 1-c shows the portion of the aerial imagery used in this study, taken on September 6th, 2018, with 3 m resolution, 4.76 million pixels, 42.8 $km^2$ areal coverage, and with 344 landslide features overlayed (Table 1).

### 4.3.3 2017 Asakura Rainfall

Significant rainfall in the first week of July 2017 triggered floods, debris flows and large-scale landslide events on the mountainous regions around Asakura area of Kyushu province in Japan. Asakura experienced more than one and a half times its monthly average rainfall on July 5th, 2017 (Miura and Nagai, 2020). The 2017 Asakura rainfall event is not well investigated in the literature comparative to the other two events in this study, and there are very limited studies published on landslide detection methods for the event (Miura and Nagai, 2020; Zhang et al., 2023). The high-resolution (0.59 m) post-event satellite image used in this study was captured on July 13th, 2017, acquired from Google Earth. The imagery and corresponding landslide labels were published by Zhang et al. (2023) as a part of their large-scale dataset. The image tile from the 2017 Asakura event in northern Kyushu, which is used in this study, as shown in Figure 1-d, covers an area of 7.03 $km^2$,

consisting of approximately 20.19 million image pixels and 306 landslide polygons (Table 1).

### 4.3.4 Landslide Inventories

The basic statistics of the landslide features used in this study, limited by the area cropped for analysis, is provided via Table 1. As shown via the table, the number of landslide features in the three selected areas of studies are close. However, Hokkaido event has relatively larger landside features in terms of area, and it also includes significantly larger overall extent of landslides used in this study, based on the statistical analysis shown in Table 1. Figure 2 also indicates this, while showing the probability distribution of the areas of the landslide features behave similarly across the three events, with the Kumamoto event having the lowest number of landslide features and the higher frequency of small-size landslides. Although, as shown in Figure 2, like many other earthquake events, the frequency density of the landslide feature size decreases by increasing the landslide feature size.

Table 4-1. Summary statistics of the landslides used in this study.

| Landslide Event | Triggering Factor | Number of Landslides | Minimum Area ($m^2$) | Maximum Area ($km^2$) | Average Area ($m^2$) | Median Area ($m^2$) | Standard Deviation of Area ($m^2$) | Total Area ($km^2$) |
|---|---|---|---|---|---|---|---|---|
| 2016 Kumamoto | Earthquake | 264 | 23.4 | 0.23 | 5373 | 1049 | 18726 | 1.42 |
| 2018 Hokkaido | Earthquake | 344 | 168.3 | 1.22 | 32929 | 4777 | 100563 | 11.33 |
| 2017 Asakura | Rainfall | 306 | 46.6 | 0.06 | 1797 | 728.9 | 4065 | 0.55 |

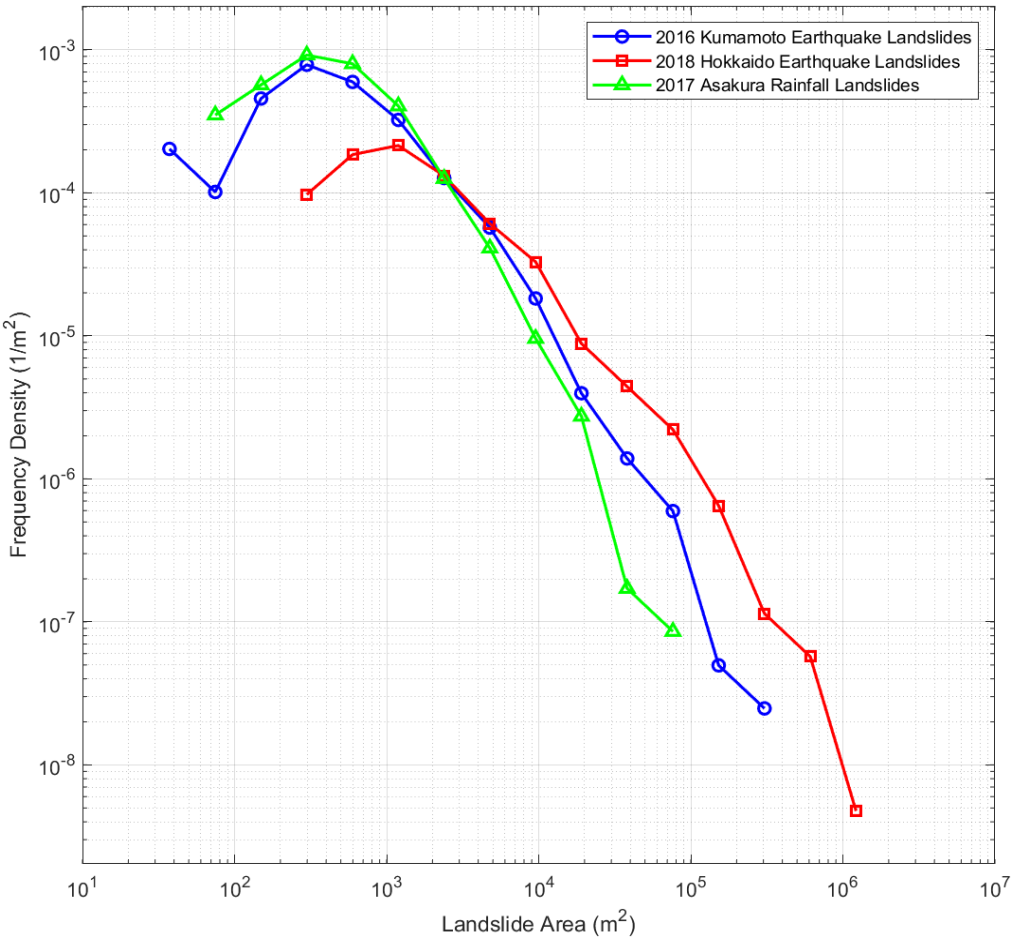



Figure 4-2. Comparative log-binned probability density distribution of the area of seismic-induced landslide features of the 2016 Kumamoto and the 2018 Hokkaido earthquake events, and the rainfall-induced landslides of the 2017 Asakura rainfall event, all in Japan.

## 4.4 Methodology

The goal of this study is to develop a deep transfer learning framework to map landslides induced by earthquakes and rainfall events, and to investigate the cross-event transferability of the model for future unseen events. The approach is illustrated in the flowchart presented in Figure 3.

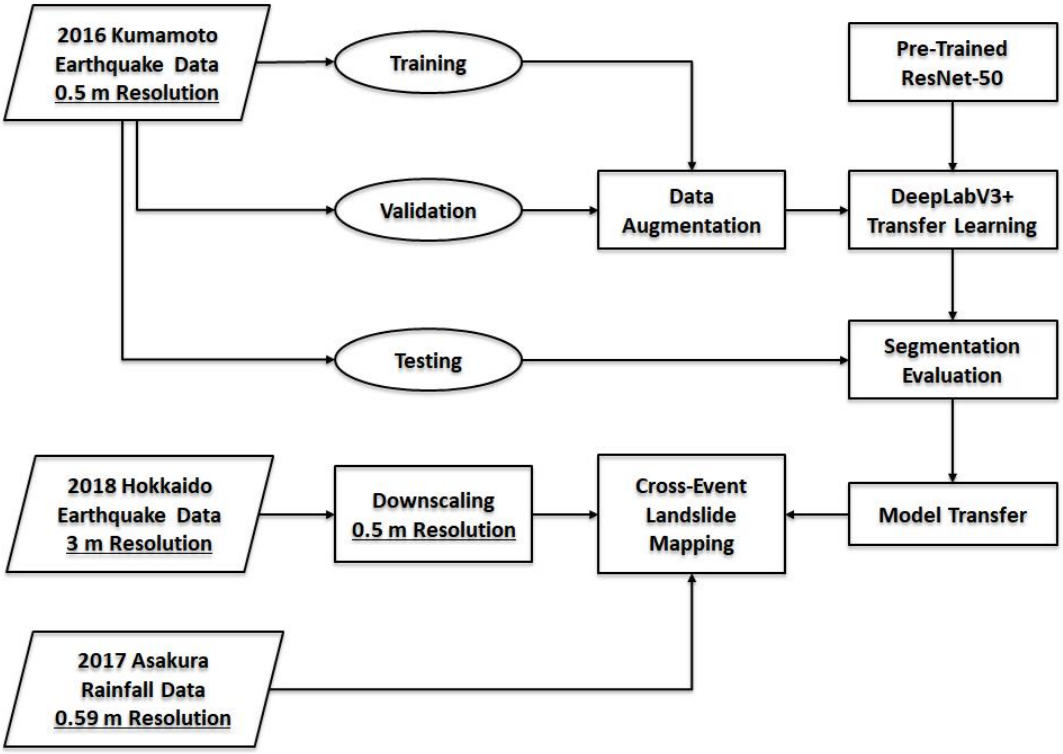



Figure 4-3. Flowchart of the proposed deep transfer learning framework for mapping landslides induced by earthquakes and rainfall events.

### 4.4.1 Deep Transfer Learning

The core algorithm in the approach is a convolutional neural network (CNN) as represented by transfer learning of the pre-trained ResNet-50 through the DeepLabV3+. A CNN is composed of an input layer, an output layer, and many hidden layers in between. These layers perform operations that alter the data with the intent of learning features specific to the data and the goal of the modeling. Three of the most common layers are convolution, activation, and pooling layers (Murphy, 2012). Convolution passes the input images through a set of convolutional filters, each of which recognizes certain characteristics and features from the images. Activation allows for faster and more effective training by turning the pixel values of the convolved layers into numbers that lead to the output labels

at the end. Pooling simplifies the output by performing nonlinear down-sampling. Up-pooling is done to inverse the down-sampling process by up-sampling the layers, for instance, in the case of encoder-decoder networks like U-Nets (Ronneberger et al., 2015).

Fine-tuning a pretrained CNN with transfer learning for a new task is typically much faster and easier than training a CNN from the scratch, and it requires considerably less amount of data and computational resources. Learning edges, forms, colors, and textures from earlier experiences on non-related datasets, and then re-training the model on landslide detection task, makes the process much more efficient and reduces the training time significantly. This is done by freezing the initial layers of the model that has already learned the simplest features. Training on top of the frozen layers allows learning new deeper features about the landslides from different regions without losing previously acquired detection capabilities.

In this work, transfer learning occurs by taking the encoder portion of a pre-trained deep convolutional neural networks (ResNet-50) and turning it into a semantic segmentation model (DeepLabV3+) for the new task of landslide mapping by adding a task-specific decoder part. The proposed computationally efficient DeepLabv3+ model has an encoding phase to extract the essential information, including multiscale contextual information, from the image using a CNN (pre-trained ResNet-50, in this study), and a decoding phase to reconstruct the output label of appropriate dimensions based on the information obtained from the encoder, and refines the segmentation results along boundaries (Chen et al., 2017-b & 2018). The encoder part of the pre-trained network (ResNet-50) is originally connected to a fully connected layer and a SoftMax layer designed for image classification. The role of those layers is to determine the probability

of the image belonging to each class of objects in the ImageNet database (Deng et al., 2009) of a million images with 1000 classes, based on the features learned from the encoder part of the network.

For the landslide mapping task, a decoder part is needed in addition to the pre-trained encoder in order to detect landslides. The decoder part up-samples the encoder output, resulting in pixel-wise probabilities of landslide occurrence in the input image. The SoftMax function is used here as the final activation layer determining the probability of belonging to the two classes of landslide and non-landslide, which can be turned into the binary output label by assigning a probability threshold. For classification problems, a SoftMax layer and then a classification layer usually follows the final fully connected layer. The SoftMax function is also used as the output unit activation function after the last fully connected layer for multi-class classification problems (Bishop, 2006).

The structure of the DeepLabV3+ model used in this study is shown via Figure 4, in which the pre-trained ResNet-50 model is located within the encoder portion of the model. Multiple down-sampling layers of a CNN lead to relatively small feature maps, resulting in lower prediction accuracy and loss of boundary information for semantic segmentation projects. In addition, aggregating context around a feature helps in segmenting it better, which is accomplished with the atrous convolutions designed as a standard part in DeepLabV3+ network. Atrous Convolution also called dilated convolution, is used in DeepLabV3+ to refine the effective field of view of the convolution process, meaning preserving the large-scale view while performing the convolution process resulting in lower resolution. It modifies the field of view using a parameter termed the atrous rate. Atrous convolution is used to explicitly control the resolution at which

feature responses are computed within deep CNNs, and the depth-wise separable convolution replaces all max pooling procedures in the DeepLabV3+ encoder. The proposed final DeepLabV3+ model in this study has 43.9 million total learnable parameters in its 206 layers.

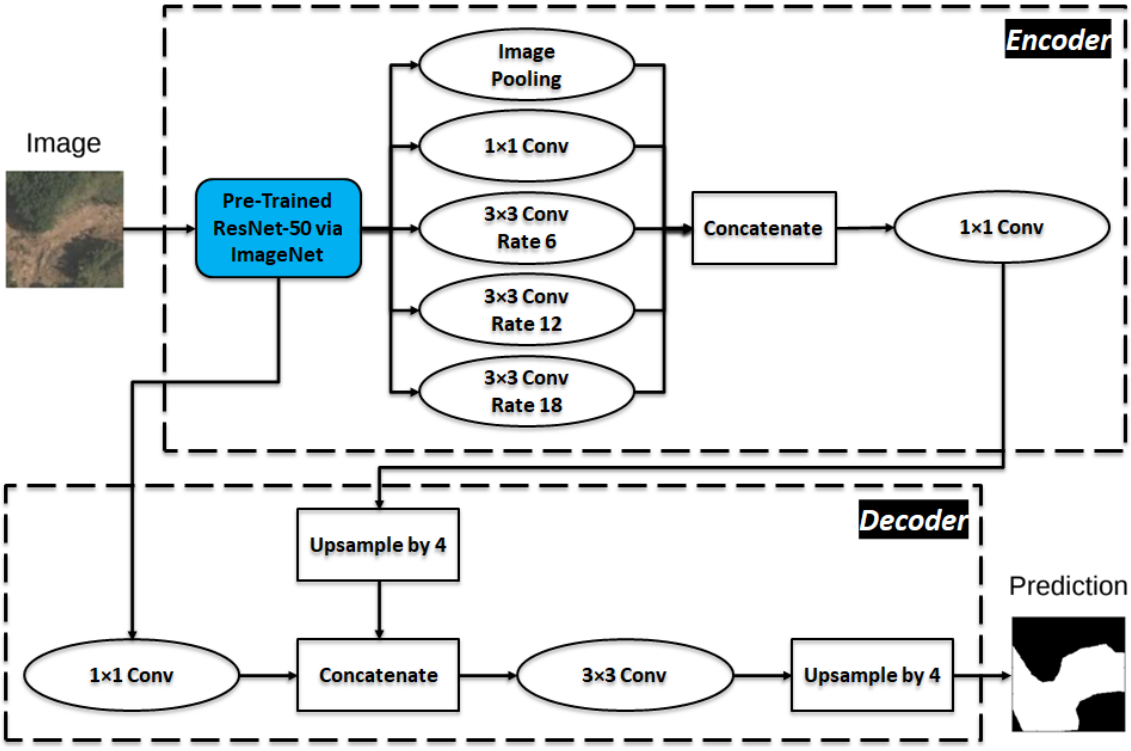



Figure 4-4. DeepLabV3+ structure using in this study to perform semantic segmentation of landslide occurrences. Pre-trained ResNet-50 via ImageNet database is used as a part of the encoder portion of the proposed deep transfer learning approach. The transfer learning occurs in the blue box where the weights of the pre-trained ResNet-50 are frozen, and used in the DeepLabV3+ model, whereas the remaining structure undergoes a training with the new landslide dataset of the 2016 Kumamoto earthquakes.

### 4.4.2 Data Partitioning, Augmentation, and Modeling

For the 2016 Kumamoto earthquake data, the tiles were divided into three groups of training (70%), validation (20%) and testing (10%). The tile size to be used to split the image was defined by a trial-and-error approach, testing three different tile sizes (112×112, 224×224, and 448×448). However, the tiles extracted by either of the three sizes, need to

be resized to 224×224 based on the standard input size of the pre-trained ResNet50 deep learning model which is used in the encoder part of the proposed model. The image tiles have three color channels (Red, Green and Blue: RGB), and were normalized to have continuous values between 0 and 1, instead of the original 0 to 255. The labels with binary 0-1 values were also split to have the same size as the image tiles. Data augmentation is essential to avoid overfitting and to ensure the robustness of the trained segmentation networks as shown by Soares et al. (2020) and Bhuyan et al. (2022) when fitting deep learning models for landslide mapping.

Several data augmentation schemes were performed on the training and validation tiles which had the landslide class present in their labels. A total of 11 techniques were used in this step, including image rotation (90, 180, and 270 degrees), flipping (horizontally and vertically), brightness change (darkening and brightening), contrast change (higher and lower contrast), gaussian blurring (Haddad and Akansu, 1991), and gaussian noise addition (Barbu, 2013) as illustrated in Figure 5. Each original tile is rotated with three angles, and flipped in two directions in order to generate a total of 6 different orientations. The six variously oriented tiles undergo six different pixel-wise augmentation techniques, including two brightness change, two contrast change, gaussian blurring, and gaussian noise addition techniques.

The data augmentation was applied to the training and validation data but not done on the testing dataset, to ensure having an unbiassed model evaluation with unseen data in the testing step. Since the training and validation accuracy are calculated alongside each other during the training process, the class weights should be approximately similar. Thus, performing data augmentation only on the training set, but not on the validation set, will

lead to possible over-fitting or under-fitting. However, in order to have unbiassed assessment during the training phase, the data augmentation is performed on independent training and validation sets after data partitioning, which means that augmented training tiles are not present in the validation set, and vice versa.

The binary classes for landslide mapping are highly imbalanced. For the case of the 2016 Kumamoto earthquake used for model development, only 6 % of the pixels belong to the landslide class (before performing data augmentation). After data augmentation, 14 % of the pixels in the training data belong to the landslide class. To tackle the class imbalance issue, different class weights are assigned in the loss function in the deep learning model during the training and validation process to calculate the accuracy in each training iteration.

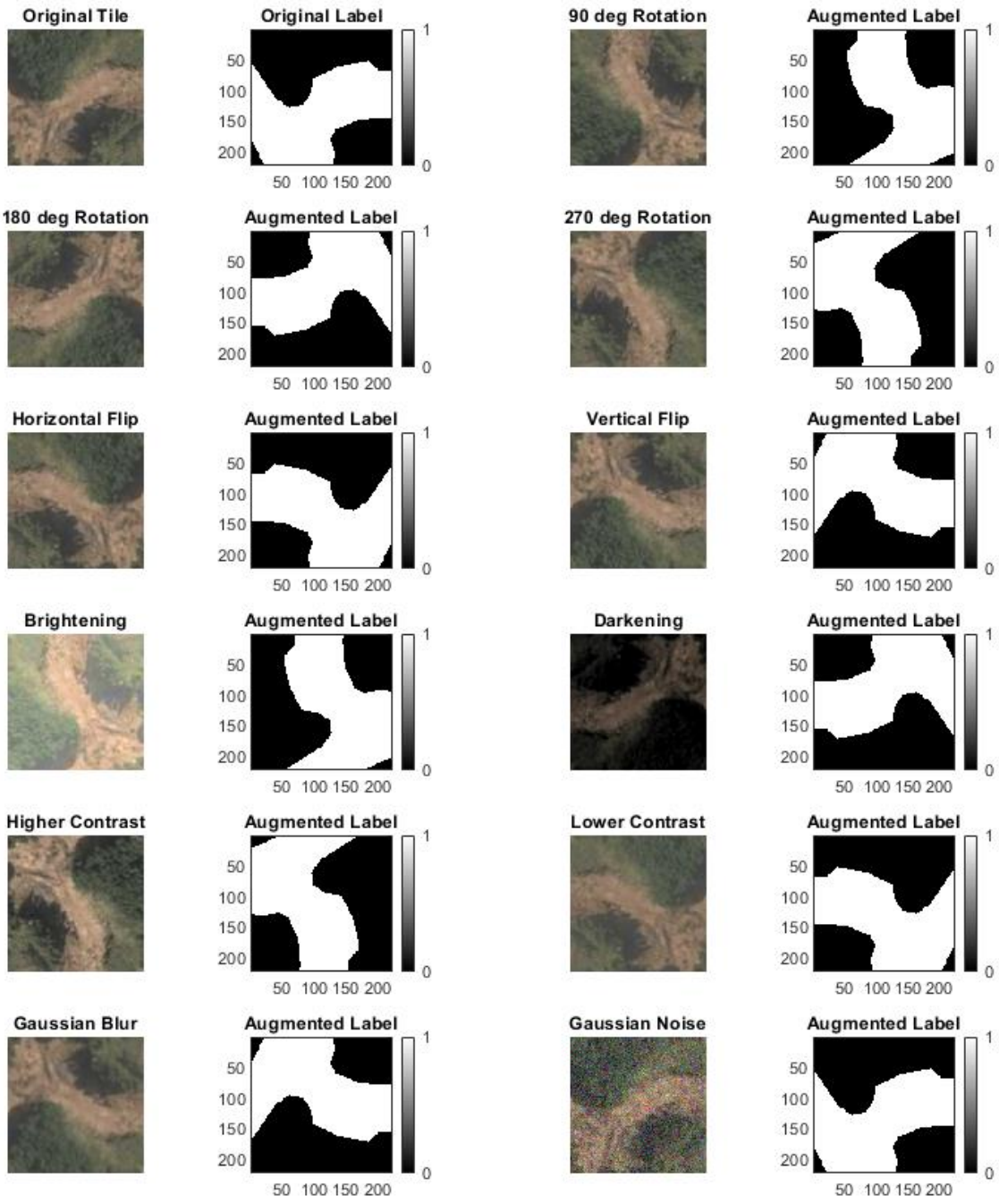



Figure 4-5. Data augmentation techniques used in this study, including flipping (horizontally and vertically), rotating (90, 180 and 270 degrees), brightness change (darkening and brightening), contrast change (higher and lower contrast), gaussian blurring, and gaussian noise addition.

Table 2 presents the resulting number of tiles before and after the data augmentation. Since 11 augmentation techniques are used in this study, even performing

the data augmentation on the tiles with landslide class present in their labels increases the number of tiles significantly for the training and validation datasets. Table 2 also presents the tile specifications per different scenarios evaluated in this study. Since the resolution of the 2017 Asakura event was close to the 2016 Kumamoto event, no down-sampling was performed, and the satellite optical image was used in its original resolution published by Zhang et al. (2023). It is clear in Table 2 that the definition of tile size and the resolution of imagery both have a significant impact on the number of tiles that can be extracted from the imagery in a specific domain of study. In the case of the 2018 Hokkaido event, increasing the resolution of the image led the total number of tiles to almost 43 times the lower resolution case.

Table 4-2. Number of tiles per partition for different scenarios and stages in the study.

| | Event | Image Resolution (meters) | Tile Size (Pixels) | Training Data Tiles | | Validation Data Tiles | | Testing Data Tiles | Total Number of Tiles | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Original | Augmented | Original | Augmented | | Original | Augmented |
| Model Development | 2016 Kumamoto Earthquake | 0.5 | 112×112 | 9,520 | 20,949 | 2,720 | 5,723 | 1,360 | 13,600 | 28,032 |
| | | | 224×224 | 2,352 | 7,049 | 672 | 2,179 | 336 | 3,360 | 9,564 |
| | | | 448×448 | 588 | 2,876 | 168 | 850 | 84 | 840 | 3,810 |
| Case Study | 2018 Hokkaido Earthquake | 3 | 224×224 | - | - | - | - | 78 | 78 | - |
| | | 0.5 | 224×224 | - | - | - | - | 3,321 | 3,321 | - |
| | 2017 Asakura Rainfall | 0.59 | 224×224 | - | - | - | - | 400 | 400 | - |

### 4.4.3 Model Evaluation

Based on the inverse frequency of the classes in the dataset calculated after performing data augmentation, the weights to be used for the binary cross-entropy loss function for the non-landslide and landslide classes were assigned by 1.05 and 7.15 values, respectively. This means that the deep learning model is being notified to pay more attention to the landslide (minority) class, which balances the loss calculation for the model to be fair in accuracy measurements on both classes. Otherwise, the model would be trained in favor of

the non-landslide (majority) class, neglecting the small error from misclassification of the landslide class.

Training parameters tuning of the proposed deep transfer learning method is done with a grid search approach, and best input parameters are reported, plus the runtime and computer system specifications. Batch-normalization was also performed to facilitate the training process (Ioffe and Szegedy, 2015). The training stop criteria is assigned by a validation patience of 4 iterations, which means that the training progress stops when 4 iterations lead to validation accuracy of less than the minimum loss value already achieved during the training process. This will ensure that the training stops when the model begins overfitting.

The trained model is then tested via unseen testing data of the 2016 Kumamoto event. There are several classification accuracy indices which were used in this study to make sure the model performs well enough on the testing data and case studies. The proportion of correctly classified observations per positive class (landslide class) is referred to as the Sensitivity or True Positive Rate (TPR). Specificity or True Negative Rate (TNR) is the rate of correct prediction in the negative class (non-landslide). The proportion of incorrectly classified observations per true class is referred to as the False Negative Rates (FNR). Precision (Positive Predictive Value), Recall, and F-1 score are used to evaluate the accuracy of the models as three of the most common classification accuracy indices. The F1 score is simply a way to combine the precision and recall, calculated by taking harmonic mean of the precision and recall.

Mean BF (Boundary F1) score computes the contour matching score between the predicted segmentation and ground-truth label. The score indicates how well the predicted

boundaries of classes align with the true boundaries. Intersection over Union (IoU) is the ratio of the intersection to the union of the predicate and the true, which in the confusion matrix is expressed as the ratio of the number of true positives to the sum of true positives, false negatives, and false positives. Formulations of the discussed accuracy indices are provided through Equations 1 through 6.

$$TPR = TP/(TP + FN) \tag{1}$$

$$FPR = FP/(TN + FP) \tag{2}$$

$$Precision = PPV = TP/(TP + FP) \tag{3}$$

$$Recall = TN/(TN + FN) \tag{4}$$

$$F_1\ Score = (2 \times Precision \times Recall)/(Precision + Recall) \tag{5}$$

$$IoU = TP\ /\ (TP + FP + FN) \tag{6}$$

Gradient-weighted class activation mapping (Grad-CAM) technique is also used to understand why the deep learning network makes its classification decisions. Grad-CAM, invented by Selvaraju et al. (2020), uses the gradient of the classification score with respect to the convolutional features determined by the network to understand which parts of the input images are most important for classification. Strongest activation layers in different depths are also visualized for investigation of different layers, and how they contribute to the landslide segmentation.

### 4.4.4 Cross-Event Transferability

Two independent case studies of 2018 Hokkaido earthquake and 2017 Asakura rainfall events were performed, and the prediction accuracies were reported to investigate cross-event transferability of the trained model. On the case studies, no augmentation is done, similar to the testing data for the 2016 Kumamoto event. However, since the resolution of the 2018 Hokkaido event was lower than the imagery used while training the model via the

2016 Kumamoto event's post-event imagery, a down-sampled version of the data matching the 0.5 m resolution of training data was also evaluated for testing of the 2018 Hokkaido event. The down-sampling was done using bilinear interpolation technique. As a part of the study, a sensitivity analysis to the implementation of data augmentation is also performed, both to analyze the model performance on the testing data of the Kumamoto event itself, and to analyze the impact of data augmentation on the cross-event transferability of the model.

## 4.5 Results and Discussion

### 4.5.1 Model Evaluation

As the first step in model validation, three different tile sizes of 112-by-112, 224-by-224 and 448-by-448 on the 2016 Kumamoto event were used to investigate the impact of tile size on model accuracy, as summarized in Table 3. The 224-by-224 tile size had the highest accuracy (F1 score) comparing the other two tile sizes (Table 3). The smaller tile size of 112-by-112 also led to acceptable accuracy. The larger tile size of 448-by-448 resulted in lower accuracy. One interpretation of this result could be that the deep learning model has difficulty learning landslide features when they are comparatively smaller relative to the tile size. Alternatively, with a larger tile size, the number of extracted tiles becomes much smaller (75% reduction in this case from 224-by-224 to 448-by-448), and therefore reduces the training data significantly. The accuracy indices for the different tile sizes are shown in Figure 6 using a spider chart. F-1 score was the main criteria of model comparison and tile size selection.

After finding the best tile size, a sensitivity analysis was performed by running two models, with and without data augmentation step. As shown in Table 3 and Figure 6, the

model without data augmentation performed relatively better in a few global accuracy indices, but the True Positive Rate (landslide detection) was lower than the model trained with data augmentation.

Table 4-3. Comparative accuracy results.

| Stage | Event | Model / Case Study | | Validation Accuracy % | Testing Accuracy % | | | Precision % | Recall % | F-1 Score % | Weighted IoU % | Mean BF Score % |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | TNR | TPR | Global | | | | | |
| Model Development | 2016 Kumamoto Earthquake | ResNet50-based DeepLabV3+ (Deep Transfer Learning) | 112×112 Tiles | 81.32 | 87.23 | 81.08 | 87.34 | 86.39 | 82.18 | 84.23 | 85.05 | 59.91 |
| | | | 224×224 Tiles (w/ Aug.) | 82.42 | 88.80 | 88.07 | 88.77 | 88.72 | 88.16 | 88.44 | 86.18 | 77.57 |
| | | | 224×224 Tiles (wo/ Aug.) | 94.31 | 95.46 | 86.25 | 95.15 | 95.00 | 87.41 | 91.04 | 93.11 | 87.72 |
| | | | 448×448 Tiles | 86.31 | 92.93 | 21.07 | 91.50 | 74.98 | 54.07 | 62.83 | 89.73 | 66.70 |
| Case Study | 2018 Hokkaido Earthquake | Original Resolution (3 m) | | - | 99.49 | 4.88 | 77.57 | 90.54 | 51.12 | 65.35 | 60.51 | 43.51 |
| | | Down-sampled Resolution (0.5 m) (w/ Training Aug.) | | - | 89.30 | 73.17 | 85.75 | 87.24 | 76.90 | 81.74 | 76.41 | 62.58 |
| | | Down-sampled Resolution (0.5 m) (wo/ Training Aug.) | | - | 93.83 | 45.67 | 83.21 | 88.09 | 63.33 | 73.69 | 71.67 | 48.36 |
| | 2017 Asakura Rainfall | Original Resolution (0.59 m) (w/ Training Aug.) | | - | 87.23 | 81.08 | 86.84 | 86.39 | 82.18 | 84.23 | 82.45 | 54.21 |

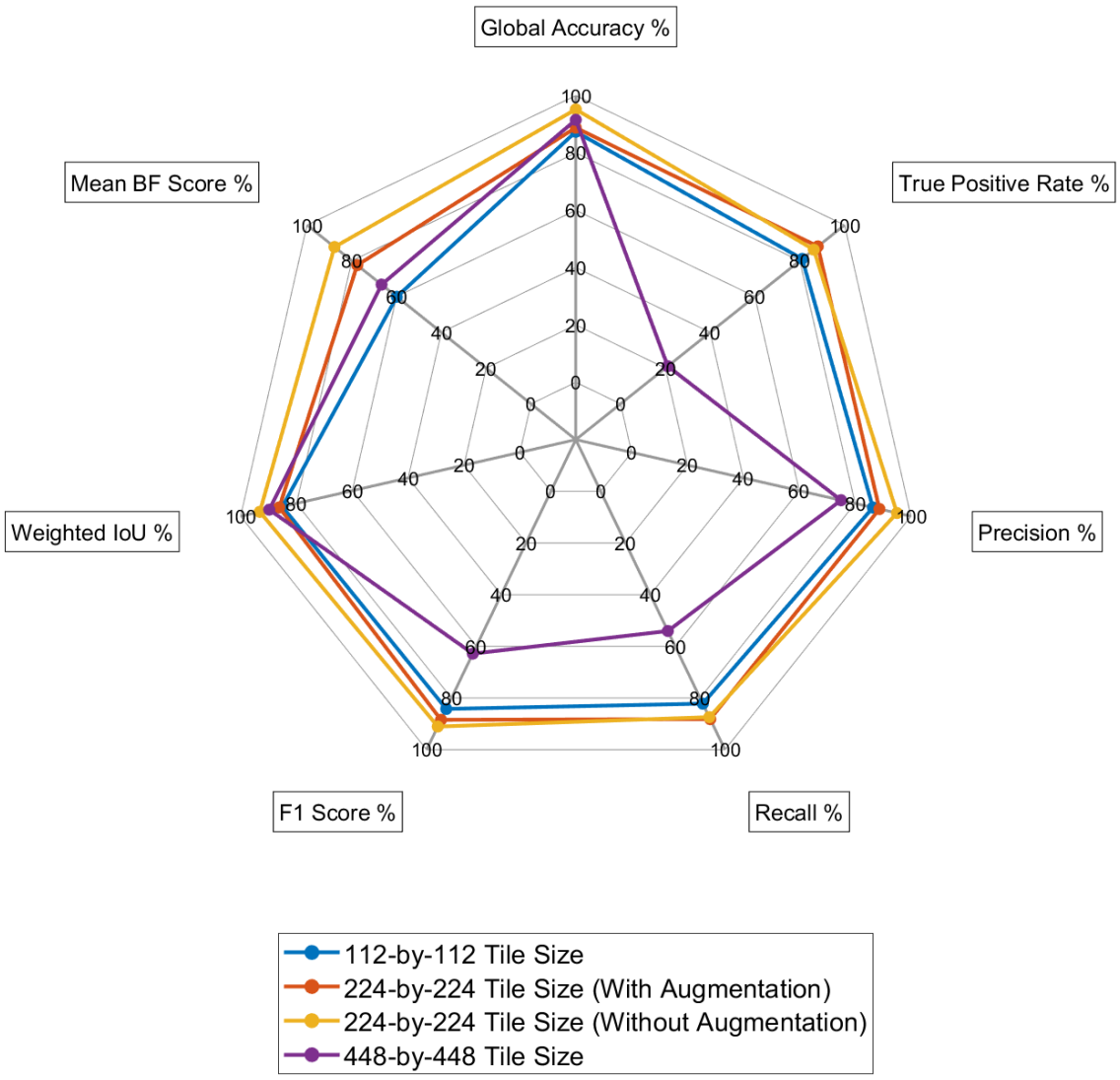



Figure 4-6. Spider chart showing different comparative analysis results by comparing seven major accuracy indices for model comparisons in different scenarios.

Using the 224-by-224 tiles and data augmentation, the best model was stopped at iteration 750 in epoch 3 when the validation accuracy criteria was met after 247 minutes of run time via MATLAB software on a PC with 64 GB RAM and 3.1 GHz Core i9 CPU. The model was trained on single CPU, using stochastic gradient descent optimizer algorithm with momentum of 0.9, learning rate of 1e-3, and L2 regularization parameter (weight decay) of 0.005 which helps in avoiding over-fitting. To avoid RAM limitation

errors, a mini-batch size of 16 was used to split the training data into smaller subsets. Each epoch during the training indicates one pass through the whole training data, and each iteration is an indicator of one pass through each training subset (mini-batch). To start each epoch, the training data were shuffled for better model learning and regularization.

Figure 7 shows the training progress, while reducing the loss and increasing the training and validation accuracy. The fluctuations of accuracy in epoch one is considered normal since the model it trying to adjust its parameters by learning from each mini-batch. The accuracy starts becoming smooth in the second epoch. After training and validation of the model, testing data from the 2016 Kumamoto event were used to evaluate the performance of the model on the unseen tiles from the same event. Continuous reduction of loss in Figure 7 and acceptable testing accuracy shown in Table 3 are indicators that the model is not overfitting to the training and validation data and can regularize well on testing data.

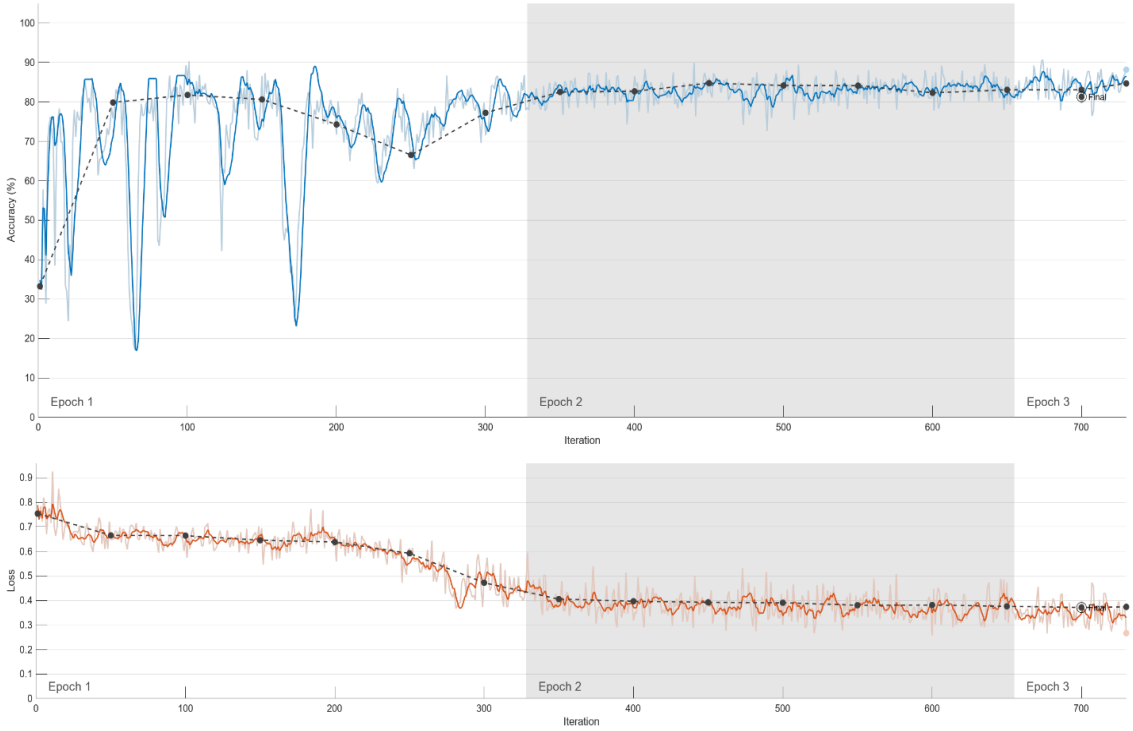



Figure 4-7. Training and validation accuracy graphs during the training progress. The solid blue line on the top plot shows the smoothed training accuracy through iterations while the model is being trained, and the lighter blue line plots the exact training accuracy over time (passing iterations). The red lines in the bottom plot are showing the loss value reduction while the model is being trained. The dotted black line plots the validation accuracy at each 50 iterations step.

Visualizations of a few tiles from the testing data of the 2016 Kumamoto event are provided in Figure 8, showing good delineation of landslide features compared to the ground-truth labels. The examples show post-event imagery, post-event imagery with ground-truth landslide labels, model prediction and the prediction errors by class. The modeled maps include colors white (true positive: correctly mapped landslides), green (false positive: wrongly mapped landslides), magenta (false negative: missing landslide features, predicted as non-landslide), and black (true negative: correct non-landslide prediction). As seen in Figures 8-a and 8-c, some over-predictions of the landslide features can be related to the inaccurate ground-truth delineation of features which are corrected by the model.

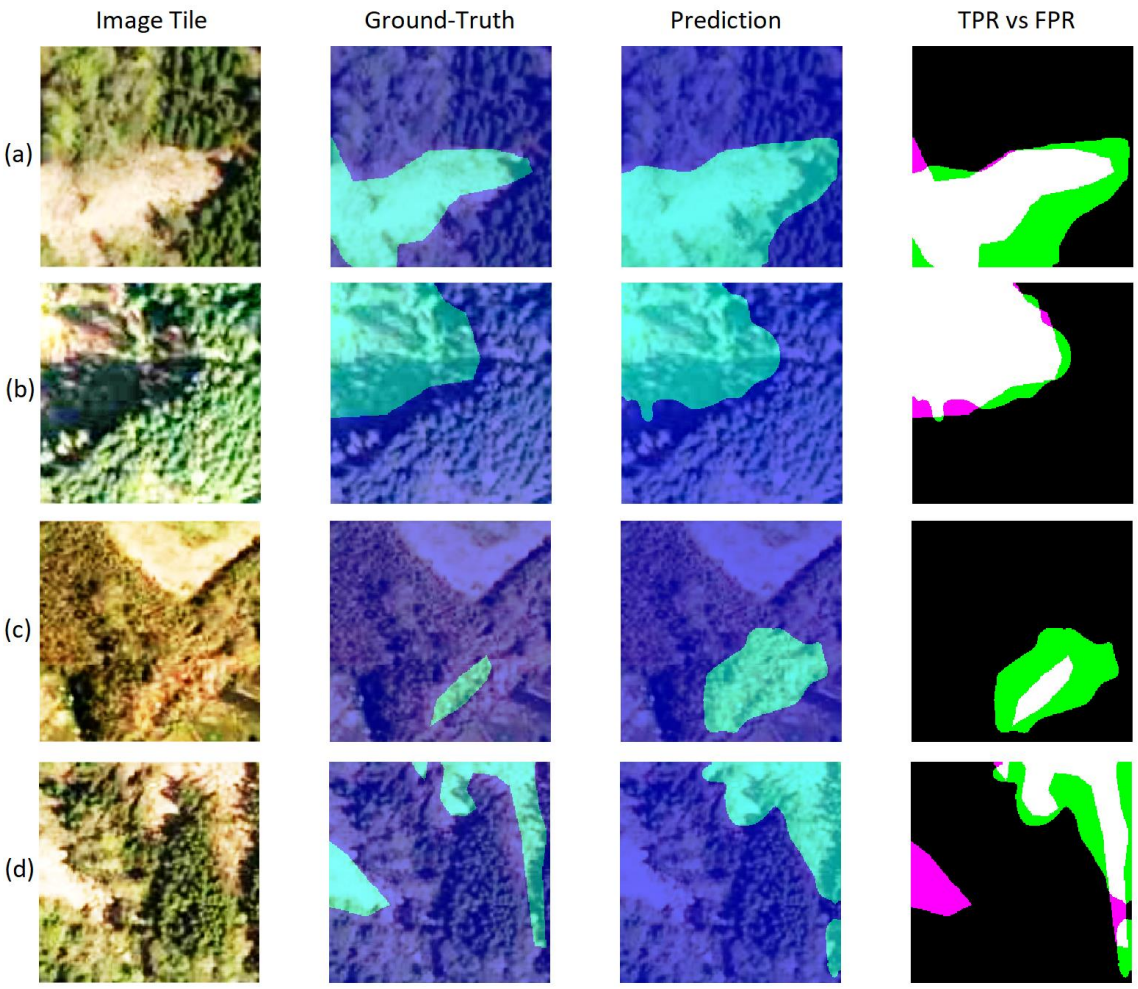



Figure 4-8. Visualizations of the testing tiles and prediction outputs for the 2016 Kumamoto earthquake. White, green, and magenta are indicators of true positives, false positives and false negatives, respectively.

To better understand the information extracted by the deep learning models in their different layers, the gradient of the classification scores of convolutional features were further visualized via the Grad-CAM technique in Figure 9, and strongest activations in the same layers via Figure 10. The layers are selected and visualized from very shallow to very deep layers of the DeepLabV3+ structure. These figures nicely show how the deeper layers of the model extracted features with more detailed spatial information about potential landslide areas, especially in the boundary regions.

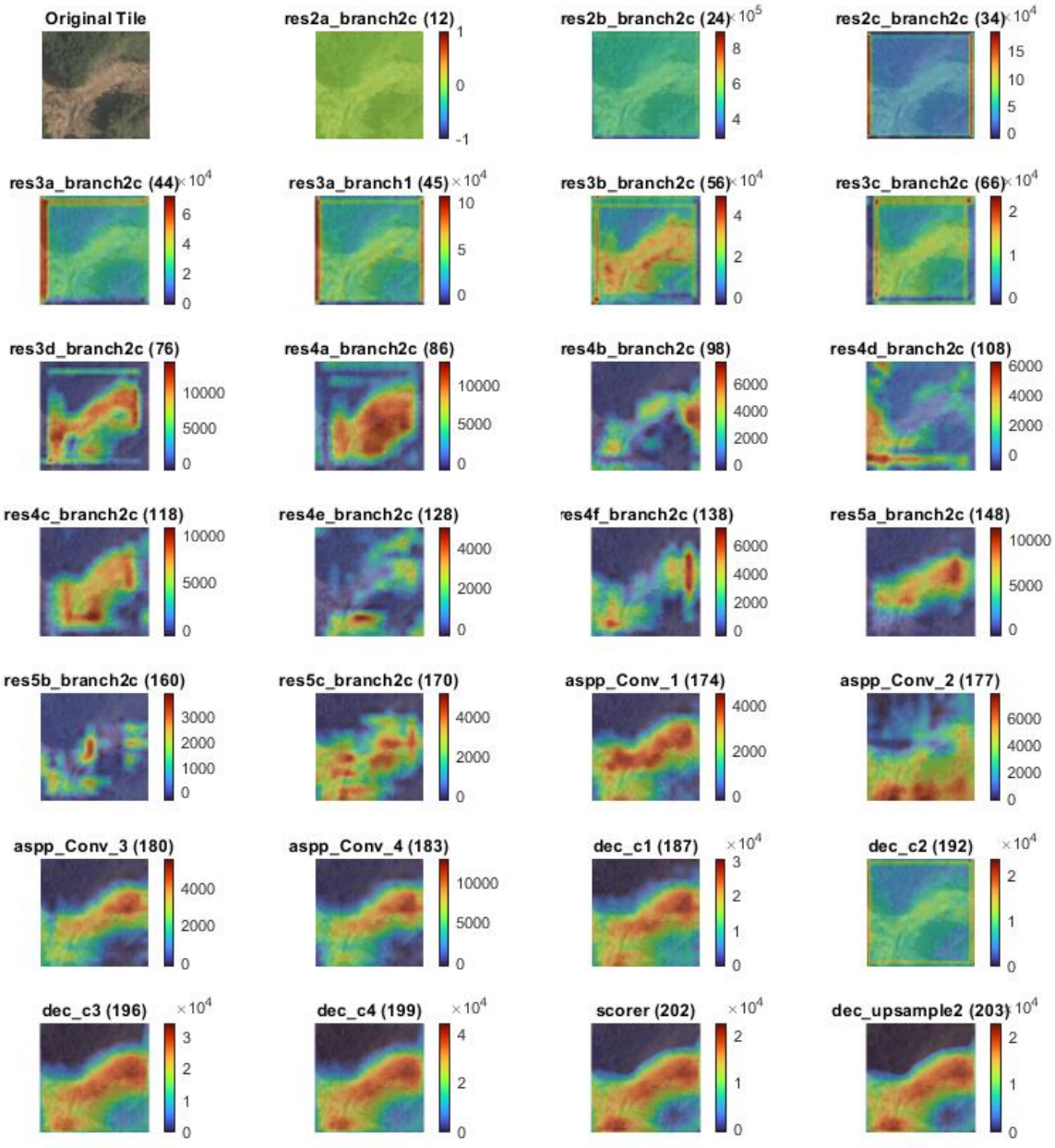



Figure 4-9. Grad-CAM weight maps of different layers in DeepLabV3+ structure using an example tile. The layer number (depth) is written in parentheses on the plot titles.

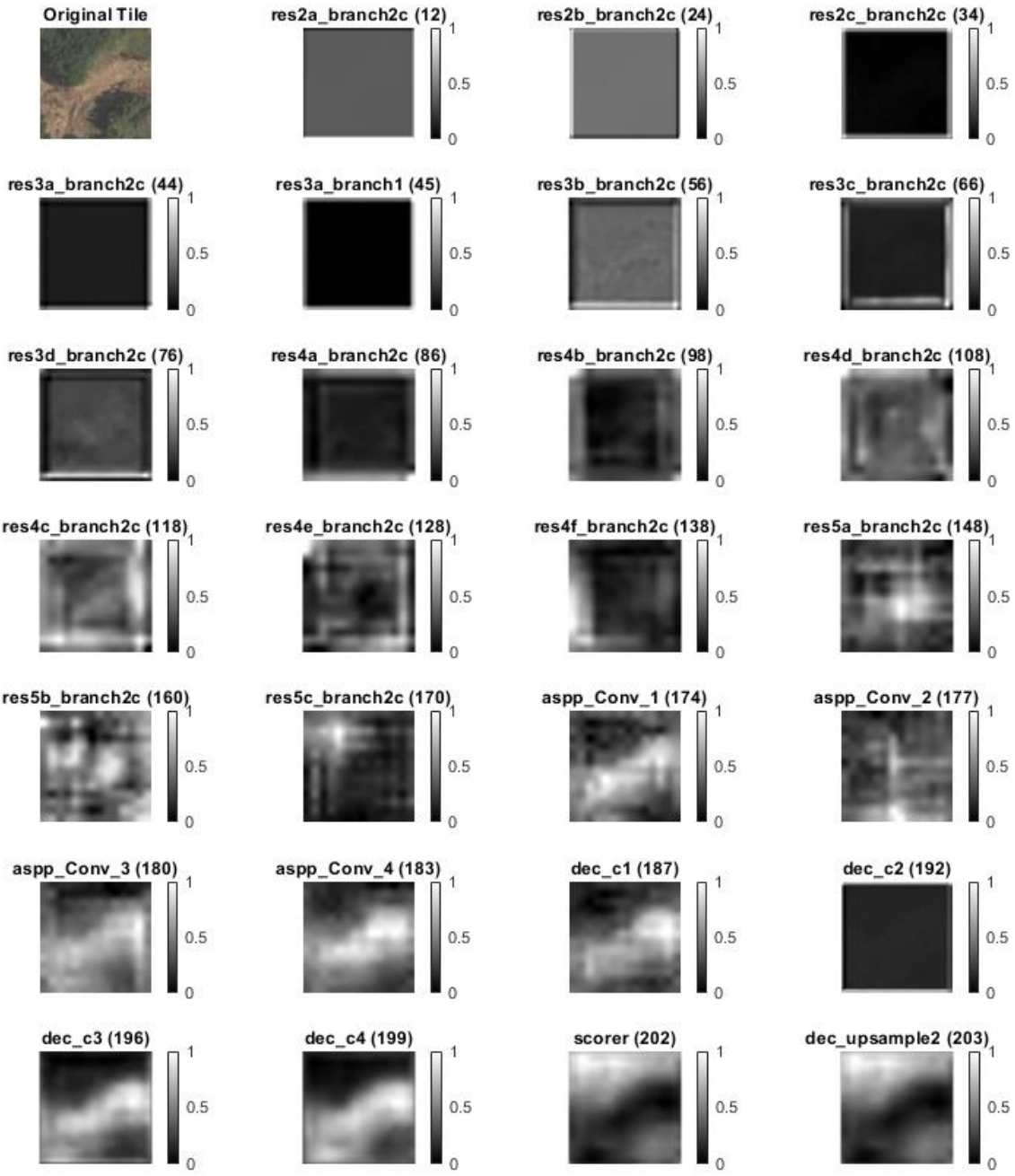



Figure 4-10. Strongest activation layer maps of different layers in DeepLabV3+ structure using an example tile. The layer number (depth) is written in parentheses on the plot titles.

### 4.5.2 Cross-Event Transferability

The next step model validation was to assess the performance of the model in completely new events to evaluate the transferability of the trained and tested model to new events. The cross-event transferability was tested with data from the two case studies (2018 Hokkaido and 2017 Asakura events). In case of the 2018 Hokkaido earthquake, the original 3m resolution tiles were fed to the model resulting in low accuracy. However, a simple biliner interpolation was run to increase the resolution of the tiles to 0.5 m (equal to the 2016 Kumamoto data), and the model performed well on down-sampled data. Yu et al. (2021) also tried using imagery with different spatial resolutions while testing their model, and they also reported lower accuracy of their proposed model when running on higher resolution data. The results of the case studies showed that the model performs acceptably well when the resolution of the target event is equal or close to the model development data.

Table 3 and the spider chart in Figure 11 show the accuracy comparison explained above. Looking at the spider chart in Figure 11, comparing the case studies with the model development event, it seems that six of the seven accuracy indices have maintained relatively high values, while the drop in the Mean BF score is relatively higher. This indicates that the boundaries of the landslide features are less aligned with the ground-truth polygons in the 2018 Hokkaido and the 2017 Asakura events, compared to the 2016 Kumamoto event which was using for training purpose.

To analyze the impact of running the data augmentation in the model development phase on the cross-event transferability, the model trained on the Kumamoto event without data augmentation was also used to map landslides on the Hokkaido event. However, as

shown via Table 3 and Figure 11, the accuracy decreased considerably, especially for TPR (landslide detection) while the difference was not considerable for testing data of the Kumamoto event (Figure 6). This indicated the importance of data augmentation to create models which can generalize well when encountering unseen events.

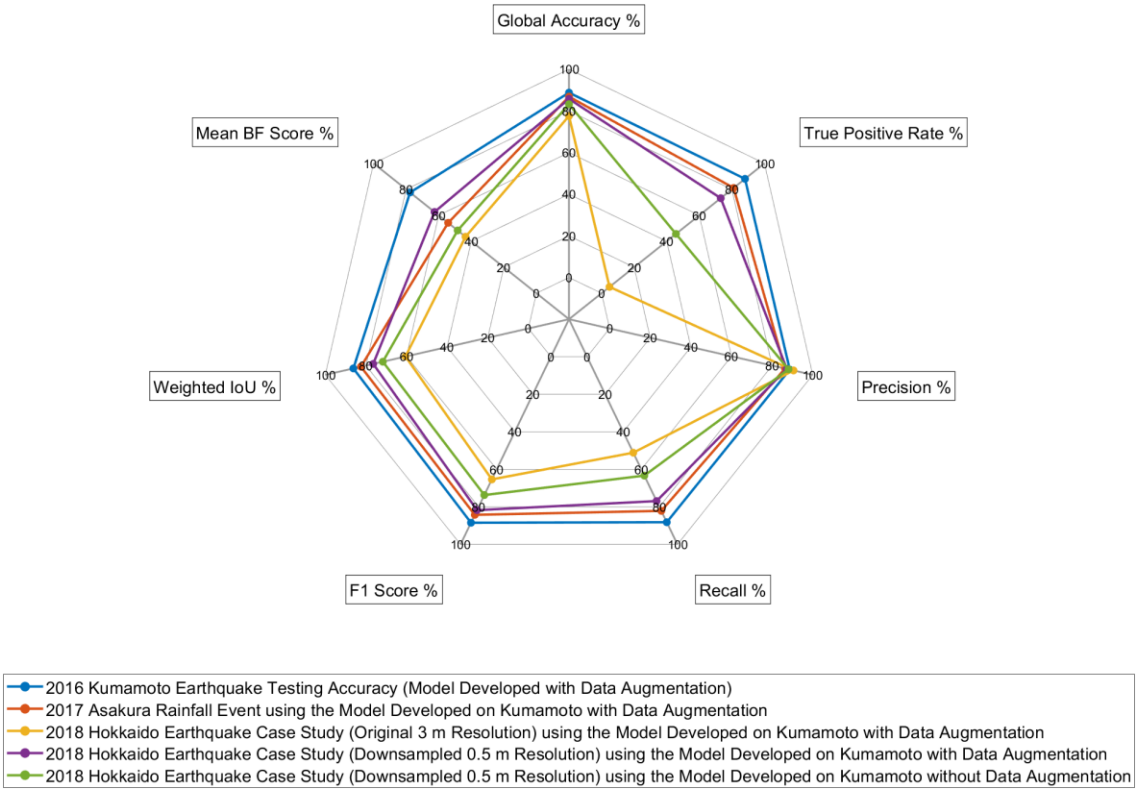



Figure 4-11. Spider charts showing different comparative analysis results by comparing seven major accuracy indices for model comparisons in different scenarios.

Visualizations of the test tiles with their corresponding predictions from the two case studies of the 2018 Hokkaido and the 2017 Asakura events are provided via Figures 12 and 13, respectively. Figure 12-b shows that a relatively large landslide feature is missed in the prediction; however, looking at the image tile, it seems that the under-prediction is due to a wrongly delineated landslide feature. Figure 12-c shows the power of the model in predicting landslide features in shadow conditions. Figure 12-d proves the power of the

model in finding landslide features with complex shapes and patterns accurately. Looking into the results shown in Figure 13-d, the model shows superior performance in detecting the boundary of the landslide features in comparison with the ground-truth labels provided by Zhang et al. (2023), while minimal over-predictions are observed in Figures 13-a, 13-b, and 13-c.

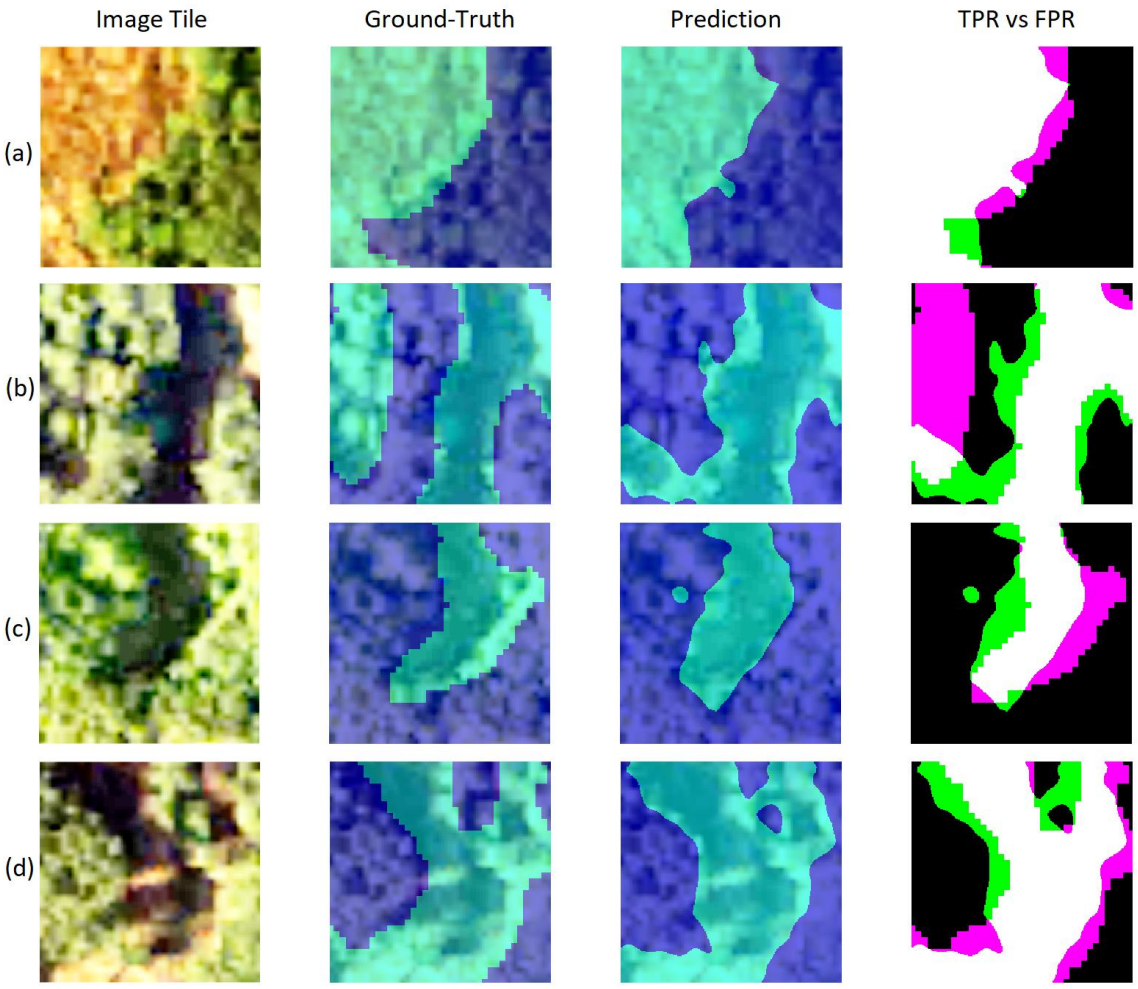



Figure 4-12. Visualizations of the test tiles and prediction output for the 2018 Hokkaido earthquake. White, green, and magenta are indicators of true positives, false positives, and false negatives, respectively.

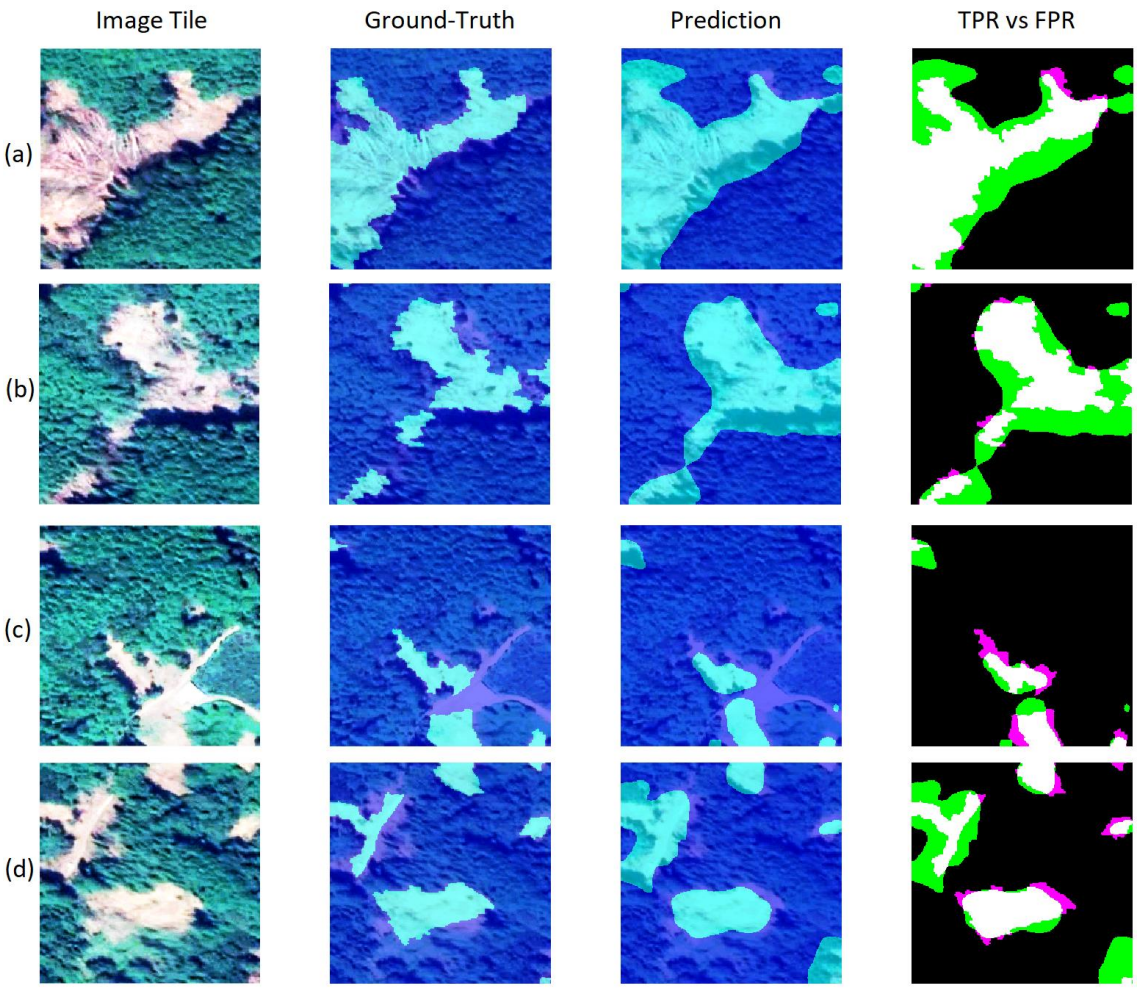



Figure 4-13. Visualizations of the test tiles and prediction output for the 2017 Asakura rainfall event. White, green, and magenta are indicators of true positives, false positives, and false negatives, respectively.

Visual inspection of the final landslide probability maps of the 2018 Hokkaido event in Figures 14 (full area coverage) and 15 (an example zoomed area), and the 2017 Asakura event in Figure 16, shows that the recognition (detection) of landslide features is acceptably good. It is also visible among the detailed views that in some cases the deviation from the ground-truth polygons is not necessarily a model error, but a correction of weakly-drawn or missing landslide polygons. In some cases, the original labels were not delineated carefully and contained some portions of the non-landslide area. Thus, some of the

calculated errors could be related to such issues. Ohki et al. (2020) excluded the boundary area (areas within 10 m from the boundary of the landslide polygons) from their accuracy validation step in order to avoid the influence of polygon error, since surveys may have minor errors regarding the shapes of the polygons. Cui et al. (2021) also explain how the type of imagery used, and the interpretation procedure can impact the landslide inventories, showing significantly different number of polygons derived by different studies as the ground-truth landslide inventories.

The non-landslide regions are generally assigned very low landslide probability values in predicted maps, which indicates the model output is not much pixelated, which is a common issue in pixel-based classification methods. Ghorbanzadeh et al. (2022) reported that although DeepLabV3+ can detect most landslide regions well, it tends to yield a high false-alarm rate, and many non-landslide pixels are misclassified as the landslide category. However, in this study our observation does not comply with their conclusion, which might be due to the significant difference in the type and resolution of the datasets used.

As a final point, it should be emphasized that although the selected regions affected by landslides differ in various aspects such as topography, triggers, and landslide shape and size, most of them have high vegetation coverage, which is the main reason for the often-limited model transferability to new areas without vegetation land cover. Then, the transferability of models trained on this dataset alone remains a challenge for unseen regions that are less vegetated.

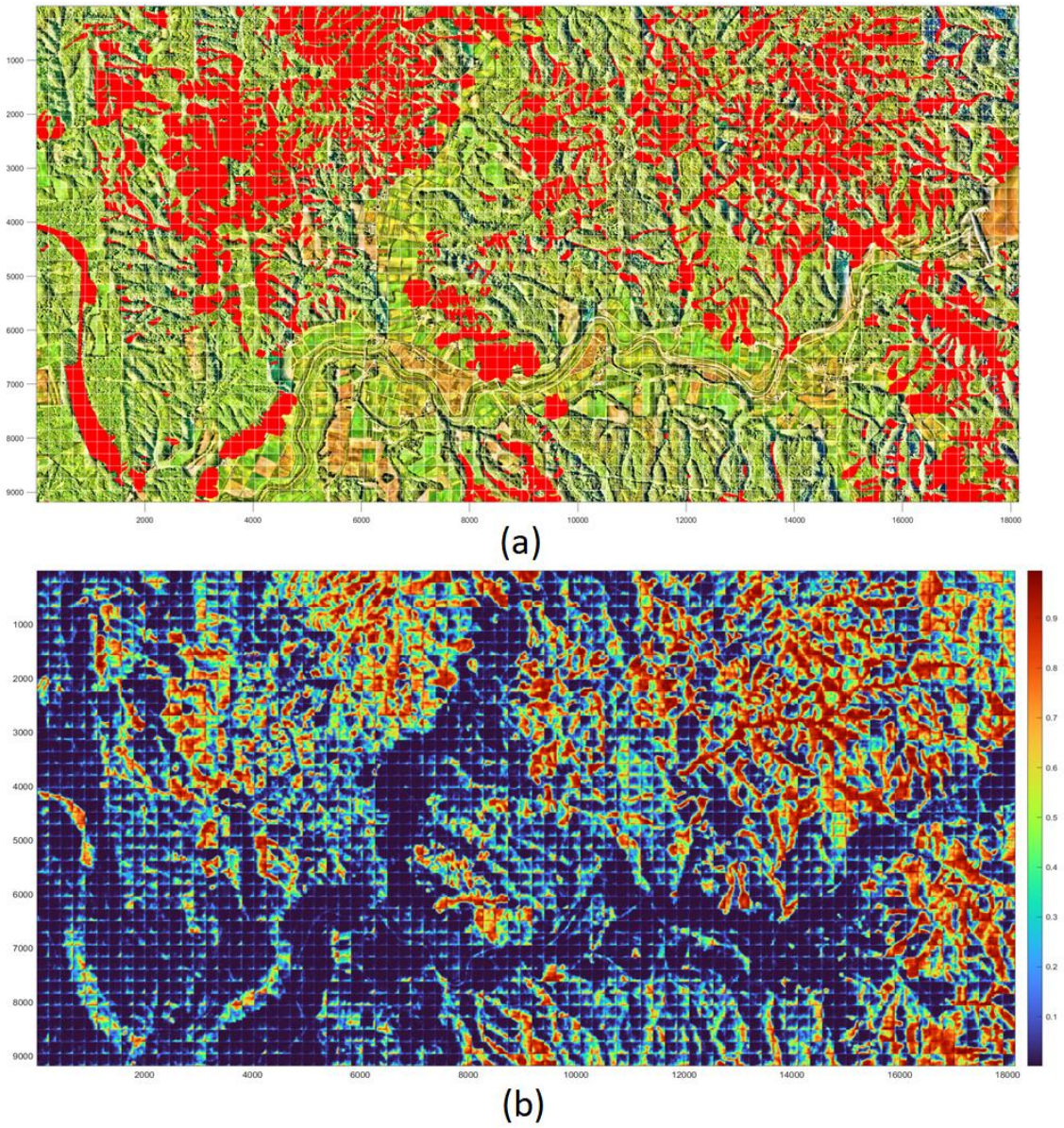



Figure 4-14. Probability maps of the proposed deep learning model output on the 2018 Hokkaido earthquake, compared to the ground-truth GSI labels, with full coverage of the studied area. Higher probability (closer value to 1) means higher possibility of the pixels belonging to landslide class.

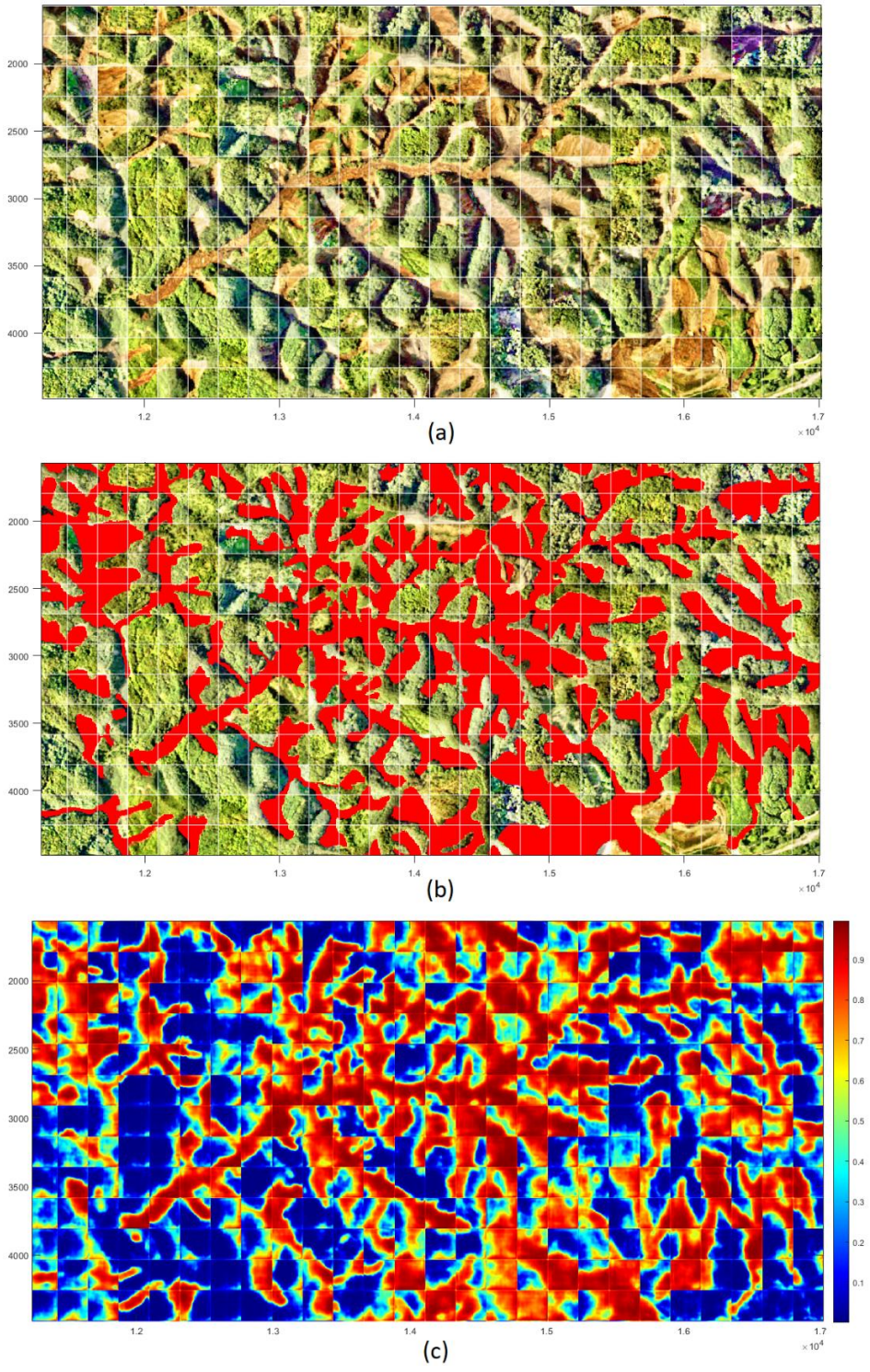



Figure 4-15. Probability maps of the proposed deep learning model output on the 2018 Hokkaido earthquake, compared to the ground-truth GSI labels, with an example zoomed area coverage. Higher probability (closer value to 1) means higher possibility of the pixels belonging to landslide class.

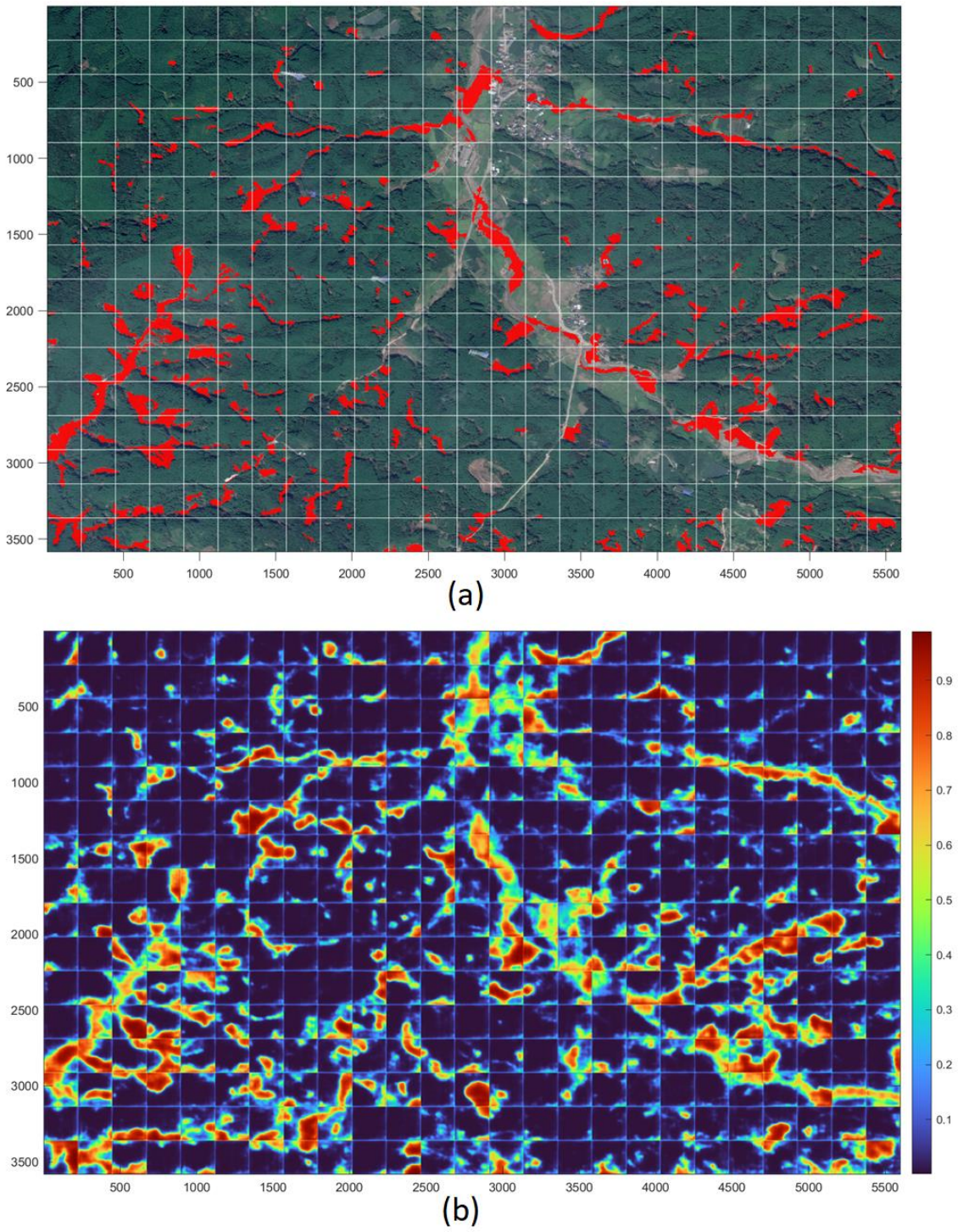



Figure 4-16. Probability maps of the proposed deep learning model output on the 2017 Asakura rainfall, compared to the ground-truth labels published by Zhang et al. (2023). Higher probability (closer value to 1) means higher possibility of the pixels belonging to landslide class.

## 4.6 Conclusions and Future Works

In this work, we trained a deep learning CNN for mapping landslides from aerial or satellite optical RGB imagery using transfer learning from the DeepLabV3+ with the pre-trained ResNET50 base model. The model was trained, validated, and tested using the landslide inventory and aerial imagery (0.5m resolution) from the 2016 Kumamoto earthquake induced landslides. The trained model was then applied on two separate events (landslide impacted region from the 2017 Hokkaido earthquake and landslide impacted region from the 2019 Asakura rainfall events). For the 2017 Hokkaido event, the optical RGB imagery was collected by xx and had an original resolution of 3m. The pretrained model had acceptable accuracy only when the Hokkaido imagery was down-sampled to 0.5m. The 2019 Asakura rainfall event had the same resolution as the Kumamoto event and resulted in similar accuracy. The proposed transfer learning model is applicable for post-event landslide mapping with good cross-event transferability. It should be emphasized that although the selected regions affected by landslides differ in various aspects such as topography, triggers, and landslide shape and size, most of them have high vegetation coverage, which is the main reason for the often-limited model transferability to new areas without vegetation land cover. A few recent studies have tried to extend the transferability of deep learning-based landslide detection methods by training the models via data from multiple events (Yu et al., 2021; Ghorbanzadeh et al., 2022; Zhang et al., 2023). However, this scheme is just in its beginning phase, and it still has a significant capacity to improve and expand in order to achieve globally applicable models.

## 4.7 Acknowledgements

The support of National Geospatial Intelligence Agency (NGIA) for this research through the NGA Academic Research Program Grant #HM0476-20-0006 (NGA NURI Project: Benchmark data development to classify damage for natural disaster relief efforts) is greatly appreciated. The authors are also thankful for the support of U.S. Geological Survey (USGS) through research grant #G22AS00006 – Proposal 2022-0047 (Innovative data-driven frameworks for geospatial ground failure models). The authors would also like to acknowledge the GSI, JAXA and NIED organizations of Japan for publishing their GIS datasets of landslide inventories, and aerial imagery related to the 2016 Kumamoto and 2018 Hokkaido earthquakes. The authors would also like to appreciate the study done by Zhang et al. (2023) which provided the satellite imagery acquired by Google Earth, and the ground-truth landslide labels for the 2017 Asakura rainfall event.

## 4.8 Code Availability

Name of the code/library: Deep Transfer Larning for Landslide Mapping

Contact: adel.asadi@tufts.edu

Hardware requirements: The model was run on single 3.1 GHz Core i9 CPU on a PC with 64 GB RAM

Program language: MATLAB

Software required: MATLAB and ArcGIS

Program size: Less than 1 GB

The source codes are available for download at the GitHub repository of the first author of the manuscript:

https://github.com/adel-asadi/Transfer_Learning_Landslide_Mapping.

# 5 Regionally Informed Global Geospatial Modeling of Earthquake-Induced Soil Liquefaction Using a System of Voting Machine Learning Classifiers

Adel Asadi [1,2], Laurie Gaskins Baise [1], Snehamoy Chatterjee [2], Weiwei Zhan [1,3], Alexander Chansky [1], Babak Moaveni [1]

[1] Geohazards Research Lab, Department of Civil and Environmental Engineering, School of Engineering, Tufts University, Medford, MA 02155, USA.

[2] Geological and Mining Engineering and Sciences Department, Michigan Technological University, Houghton, MI 49931, USA.

[3] Texas Advanced Computing Center, Civil, Architectural and Environmental Engineering Department, The University of Texas at Austin, TX 78758, USA.

(The material contained in this chapter is being submitted for review and possible publication in Engineering Geology Journal – Elsevier Publications)

## 5.1 Abstract

Data-driven geospatial liquefaction models are valuable tools for real-time post-event impact and regional seismic hazard assessments. Geospatial liquefaction models are based on liquefaction occurrence inventories, widely available geospatial variables, and earthquake-specific parameters. This research uses an updated inventory with geospatial data from non-liquefaction and liquefaction occurrence locations in 53 earthquakes around the world, including categorical and continuous variables representing proxies for soil saturation, soil density, and earthquake loading. This study evaluates the performance of advanced machine learning (ML) algorithms in learning complex nonlinear patterns in the large dataset to predict liquefaction susceptibility, as an alternative approach to previously published logistic regression classifiers. The proposed methodology starts with an exploratory data analysis and feature selection scheme on the sampled data across both the liquefaction and non-liquefaction classes. The reason is to remove redundant and less

relevant features, run data transformations, and perform statistical analysis to define some classification thresholds. The class and event imbalance issue is treated innovatively by under-sampling large events and distributing the datasets over several balanced subsets. Instead of training a single classifier, a system of voting machine learning classifiers is designed to achieve higher accuracy and certainty. Different classification techniques are used to balance the performance as a single classification technique cannot be considered the best for all earthquake events. The voting classifiers use different portions of the data, representing global, coastal/non-coastal, and regional information, to optimize the predictive power of the proposed model by taking the majority votes of six different classifiers. The classifiers include three global classifiers, two coastal/non-coastal classifiers, and a regional classifier. For each data subset, the ML-based binary classification model is trained and validated via a K-fold cross-validation approach. The voting system is then tested based on the leave-one-out approach, by totally excluding individual earthquake events one at a time for a fair accuracy, reliability, and applicability assessment. The final class assignment is generated by a majority vote of the system of 6 voting classifiers trained by the individual data subsets. The results of this study are compared with a model developed using logistic regression to investigate the benefits and limitations of the proposed voting approach.

## 5.2 Introduction

Liquefaction is a well-known phenomenon of ground failure due to earthquake-induced cyclic loading. Liquefaction hazard maps are necessary for earthquake risk assessment and loss evaluation, and they are important for both pre-event planning and post-event response and mitigation. Liquefaction susceptibility of soil deposits has been

evaluated at different scales of global (Zhu et al., 2017), continental (Meisina et al., 2022), regional (Berov et al., 2017) and local (Brankman and Baise, 2008), and using a variety of inputs, including geotechnical (Boulanger and Idriss, 2014), geological (Youd and Perkins 1978), and geospatial (Zhu et al., 2015) input parameters. Machine learning (ML) has also been successfully exploited in modeling liquefaction phenomena using such in-situ experimental data (Xie at al., 2020) and has recently been used in liquefaction hazard mapping (Geyin et al., 2022; Todorovic and Silva, 2022).

Geospatial methods can predict liquefaction spatial extent quickly for earthquakes with existing global datasets and have therefore been integrated into the USGS Ground Failure Products generated for recent earthquakes (Allstadt et al., 2022). Geospatial liquefaction methods build on prior studies of liquefaction susceptibility, soil amplification, and hydrology, which are based on the geological and depositional conditions (Youd and Perkins, 1978; Knudsen and Bott, 2011), topographic characteristics (Wald and Allen, 2007), and topographic controls on hydrology (Beven and Kirkby, 1979). These methods have created the baseline for subsequent studies of liquefaction hazard modeling via geospatial variables.

Logistic regression models have been developed for global and regional liquefaction prediction maps, using a variety of geospatial explanatory variables (Zhu et al., 2015; Zhu et al., 2017; Rashidian and Baise, 2020; Bozzoni et al., 2020; Baise et al., 2021). Recently, advanced ML algorithms have also proved their power in liquefaction hazard modeling (Geyin et al., 2022; Todorovic and Silva, 2022). Advanced ML algorithms would allow more geospatial predictor variables to be used easily, compared to

traditional logistic regression models, with greater potential for those explanatory variables to be fully exploited. The conventional models require the developer to define the importance of features, while ML can learn from the data more automatically and efficiently based on the goal of increasing the accuracy of predictions (Geyin et al., 2022). Algorithmic ML also enables better exploitation of complex nonlinear relationships that cannot be fully exploited by traditional regression approaches (Liu and Tesfamariam, 2012; Geyin et al., 2022). Zhang and Wang (2021) and Preethaa et al. (2022) have implemented a weighted average voting system and aggregated classifiers using advanced ML algorithms to classify liquefaction and non-liquefaction data based on data from SPT (Standard Penetration Test) and CPT (Cone Penetration Test) tests. However, such an approach has not been previously proposed for geospatial liquefaction hazard modeling, to the best of the authors' knowledge.

This study develops an efficient data-driven geospatial liquefaction model using advanced ML algorithms to develop real-time post-event liquefaction maps. A system of voting machine learning classifiers is proposed to improve the performance of geospatial liquefaction models by incorporating regional information into the decision-making process using a global dataset of earthquake-induced liquefaction occurrences. Combining ML algorithms and implementing a voting classifier system reduces the risk of misclassification. Furthermore, the proposed voting classifier system takes advantage of regional data, when available, in parallel with optimal global models. For the first time, the model proposed in this study uses a voting classification system to predict liquefaction hazards by exploiting several geospatial proxies and advanced ML algorithms to improve

the performance of liquefaction hazard modeling, and reduce the uncertainty associated with individual methods.

## 5.3 Geospatial Liquefaction Model Background

### 5.3.1 Logistic Regression Models

Zhu et al. (2015) introduced a logistic regression-based geospatial model to predict the probability of liquefaction after an earthquake on a regional scale (4 earthquakes from Japan and New Zealand). The used a set of geographic explanatory variables as proxies for ground motion intensity, soil density, and degree of saturation. The Zhu et al. (2015) dataset was limited to spatially complete inventories so that the probabilities from the model could represent the spatial extent of liquefaction. Yilmaz et al. (2020) performed a large-scale liquefaction risk assessment of Portugal based on the Zhu et al. (2015) approach. Matsuoka et al. (2015) followed a similar approach to a Japanese liquefaction dataset. Still, their results are based on the Japanese geomorphological classification map, which does not have global coverage (Todorovic and Silva, 2022).

Zhu et al. (2017) updated their previous logistic regression model (Zhu et al., 2015) by expanding both the liquefaction database (26 earthquakes from New Zealand, China, Taiwan, Japan, and the United States) and the number of geospatial explanatory variables, and by proposing two distinct models for coastal and non-coastal regions. Their non-coastal Global Geospatial Liquefaction Model (GGLM) is currently used by USGS (US Geological Survey) in their “PAGER” system as the preferred liquefaction model, which results in a liquefaction probability interpreted as liquefaction spatial extent (LSE). The current USGS implementation of Zhu et al. (2017) is available as a product on the overview

page for each earthquake on the USGS Earthquake Hazard Program website (Allstadt et al., 2022). Coastal events were considered where liquefaction had occurred within ~20 km of the coast or where the epicentral distance is less than 50 km to the coast for earthquakes with little or no liquefaction (Zhu et al., 2017).

Their developed database was finally biased toward coastal earthquakes (Zhu et al., 2017). Eighteen geospatial features were analyzed for GGLM as proxies of soil density, soil saturation, and dynamic loading of the earthquake. The non-coastal model uses ShakeMap's PGV, topography-based Vs30, distance to a waterbody, water table depth, and annual precipitation as explanatory variables, whereas the coastal model used PGV, Vs30, distance to coast, distance to the river, and annual precipitation. In a later study, the effectiveness of GGLM (Zhu et al., 2017) was improved by Rashidian and Baise (2020), with a few threshold definitions on the explanatory variables of PGA, PGV, Vs30, and annual precipitation.

Baird et al. (2018) have investigated the sensitivity of the geospatial liquefaction models developed by Zhu et al. (2015 & 2017) to the quality, accuracy, and resolution of the explanatory variables, especially for regions where high-quality data is accessible, such as New Zealand in the case of Canterbury Earthquakes (2010-2016). Their findings indicated that except for the case of Vs30, the standard or high-quality data would have a negligible positive impact on the model performance. However, they strongly recommended developing and using a high-quality and high-resolution Vs30 layers as the input to the geospatial models (Baird et al., 2018). Lin et al. (2021) also evaluated the performance of Zhu et al. (2017) models on the Canterbury earthquake sequence (4 events)

that occurred in 2010 and 2011 in New Zealand. They concluded that the geospatial models were unable to account for increased liquefaction severity caused by pre-shocks. They also investigated the impact of replacing the global input variables with regional New Zealand datasets, which did not significantly improve the prediction accuracy of the models but led to improved spatial accuracy (Lin et al., 2021). Geyin et al. (2020) also concluded that the geospatial liquefaction model modified by Rashidian and Baise (2020) defines the condition susceptible to liquefaction if the ground is flat, saturated, and near water. They suggested that mapped geological data can be used to reduce geologic errors due to soils that aren't susceptible to liquefaction (e.g., clays, peats, or gravels) (Geyin et al., 2020). Geyin et al. (2020) tested the Zhu et al. (2017) model against 18 CPT-based models using field data and verified the geospatial model's potential while reporting less accurate performance compared to CPT-based models.

Bozzoni et al. (2020) also computed liquefaction potential maps via logistic regression after pre-filtering the datasets to exclude the regions that were incompatible with liquefaction geologically (e.g., rocky formations). They also filtered regions that were impacted by a ground-shaking severity too low to initiate liquefaction (using a threshold of 0.1 g for horizontal peak ground acceleration). They have selected their optimal geospatial variables via the Luco and Cornell methodology based on the criteria of efficiency, practicality, and proficiency (Cornell and Luco, 2001). Bozzoni et al. (2020) used ln(Vs30), CTI and ln(PGAm), based on their feature selection and transformation approaches, which is in agreement with the proposed variables by Zhu et al. (2015).

Baise et al. (2021) enhanced the performance of the GGLM model (Zhu et al., 2017; Rashidian and Baise, 2020) by adding more recent earthquakes (51 earthquakes in total, including 6 non-liquefaction events), and a few explanatory variables including elevation above water body, soil and sediment thickness (Pelletier et al. 2016), topographic roughness index, and topographic position index. They re-modeled the logistic regression using the improved database, and introduced two best-fit models (GGLM-21-a and GGLM-21-b) among their test models. The best-fit model uses PGV, TRI, distance to closest coast, distance to nearest river, and elevation above the closest water body as the explanatory variables. The second best-fit model uses slope-based Vs30 instead of TRI, and other variables remain the same. They also performed a regional evaluation and reported that the performance was best for Japan, North America, and Oceania and relatively weaker for Europe and South America, probably due to a smaller number of earthquake events included in the database (Baise et al., 2021). Based on the results of the study by Baise et al. (2021), their recommended updated global geospatial liquefaction's logistic regression model (GGLM-21) was described by Equation 1.

$$P(x) = \frac{1}{1+e^{-X}} \quad (1)$$

Where, P(x) is the probability of liquefaction which lies between zero and 1; and X includes explanatory variables that describe density, saturation and loading conditions which is given by Equations 2 and 3 leading to the two introduced models of GGLM-21-a and GGLM-21-b, respectively. The coefficients are provided in the report by Baise et al. (2021).

$$X_1 = \alpha_0 + \alpha_1.ln(PGV) + \alpha_2.(TRI)^{0.5} + \alpha_3.ln(d_c + 1) + \alpha_4.ln(d_r + 1) + \alpha_5.(Z_{wb})^{0.5} \quad (2)$$

$$X_2 = \beta_0 + \beta_1.ln(PGV) + \beta_2.(Vs_{30})^{0.5} + \beta_3.ln(d_c + 1) + \beta_4.ln(d_r + 1) + \beta_5.(Z_{wb})^{0.5} \quad (3)$$

### 5.3.2 Analytical Hierarchy Process (AHP) Methods

Bozzoni et al. (2021) developed a geospatial methodology for zoning the earthquake-induced soil liquefaction risk at a European continental scale using Analytical Hierarchy Process (AHP), a multi-criteria knowledge-based decision analysis (Saaty, 1980). They established a Geographic Information System (GIS) environment to couple data-driven and knowledge-driven approaches for their liquefaction susceptibility mapping. Since liquefaction is a phenomenon of soil instability that occurs at a very local spatial scale, mega-zonation of the liquefaction risk at a continental or global scale is a great challenge, and relatively low accuracy is tolerated (Bozzoni et al., 2020 & 2021). AHP is considered strong, robust tool for knowledge-based modeling in data-scarce environments, or in discontinuous data series in time and/or space dimensions (Castellanos et al., 2007; Gorsevski and Jankowski, 2010; Pisano et al., 2017; Zumpana et al., 2022).

AHP was also used by Zumpana et al. (2022) for regional-scale geospatial liquefaction mapping of Foggia Province, Italy, validated using local borehole geological and geotechnical data. Zumpana et al. (2022) used PGA×Vs30 as a geospatial predictor instead of using each independently, which resulted in the highest weight among their eight predictor variables, followed by Lithology and Water Table Depth. Karpouza et al. (2021) introduced an AHP-based GIS approach for simultaneous hazard zonation mapping of earthquake-induced secondary effects. Their implemented methodology is supposed to categorize and discriminate regions prone to landslides, liquefaction, or both, using several

geospatial variables selected as influential factors. However, the expert subjectivity of AHP in the feature ranking and weighing process is a drawback of such methods.

### 5.3.3 Advanced Machine Learning Algorithms

Even with the recent rich literature in global and regional geospatial liquefaction models, liquefaction map accuracy is still limited. Regional biases persist, necessitating more quality research studies with innovative solutions. The previously introduced models, e.g., Zhu et al. (2017), Bozzoni et al. (2020), and Baise et al. (2021) use parametric models (Logistic Regression), which may need re-calibration of the parameters when applied to a region where they are not well trained on (Todorovic and Silva, 2022). Recent advances in ML with algorithms that do not require a strong assumption about their functional form, in addition to the availability of large liquefaction datasets, provide an opportunity to achieve higher accuracy of geospatial liquefaction models.

Geyin et al. (2022) used field test data to develop geospatial models driven by algorithmic learning of AI (Artificial Intelligence) methods but pinned to a physical framework of liquefaction mechanics. They used remotely predicted subsurface cone penetration test (CPT) measurements in their AI-driven model.

A few studies have used the Random Forest algorithm (Breiman, 2001) to create geospatial liquefaction models (Durante and Rathje, 2021; Todorovic and Silva, 2022). It consists of multiple independent decision trees to solve classification and regression problems. The overall average of the prediction by those trees will be the final response of the random forest model. Todorovic and Silva (2022) have recently introduced a non-parametric Random Forests model which can learn any functional form from the training

dataset, thus not requiring the model developer to formulate the function. Todorovic and Silva (2022) have reported higher predictive accuracy of their Random Forests model compared to the logistic regression model (GGLM-17) published by Zhu et al. (2017). The optimal variables Todorovic and Silva (2022) used in their machine-learning model include strain proxy (a ratio between PGV and Vs30), slope, topographic roughness index, water-table depth, average precipitation, and distance to the closest water body. Similar to Bozzoni et al. (2020), Todorovic and Silva (2022) have also used the method proposed by Cornell and Luco (2001) for their optimal feature selection step. In addition, they used the proficiency parameter (Padgett et al., 2008) to rank their candidate geospatial variables. Proficiency is the ratio between efficiency (Luco and Cornell, 2007) and practicality (Mackie and Stojadinovic, 2002). The lower the value of this metric, the better the correlation of the variable with liquefaction occurrence (Todorovic and Silva, 2022).

## 5.4 Datasets

Training ML models require a large amount of data points. Therefore, a global inventory was built from liquefaction manifestations that occurred in 47 earthquakes on different continents, in addition to 6 earthquakes with only non-liquefaction samples (53 earthquakes in total), as summarized in Baise et al. (2021). The earthquakes were selected based on the quality and quantity of available information in terms of seismological data, geospatial information, and well-documented liquefaction occurrences (precise location, etc.). The goal was to sample a variety of earthquakes in terms of magnitude and geologic settings to ensure a well-distributed database. The earthquake events are listed via Table A-1 in the Appendix section. The earthquakes are labeled as coastal or non-coastal with the same approach used by Zhu et al. (2015 & 2017) and Baise et al. (2021). We also add

six regional categories to group events geographically. Figure 1 presents a map showing the geographical distribution of earthquake data around the globe on an aerial map grouped by region. Figure 2 represents the historical timeline of the earthquakes. The oldest event in the inventory is the 1949 Puget Sound earthquake; the most recent earthquake occurred in 2016.

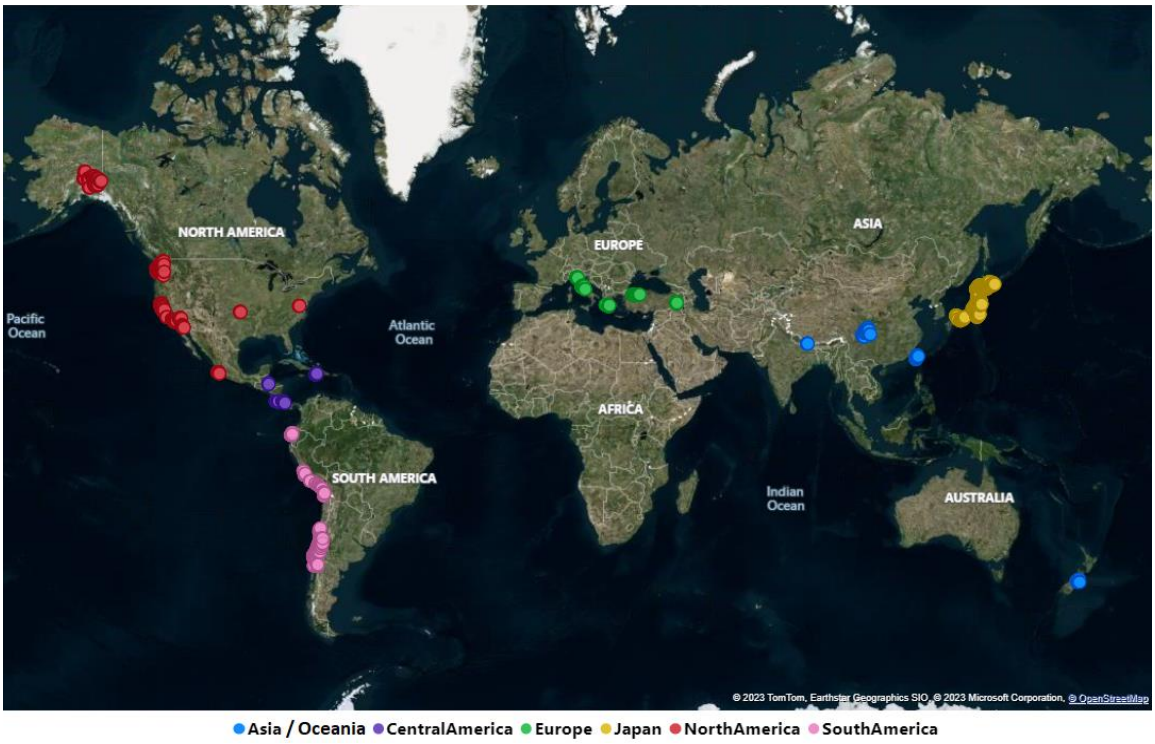



Figure 5-1. World aerial map showing the geographical distribution of earthquake data used in this study.

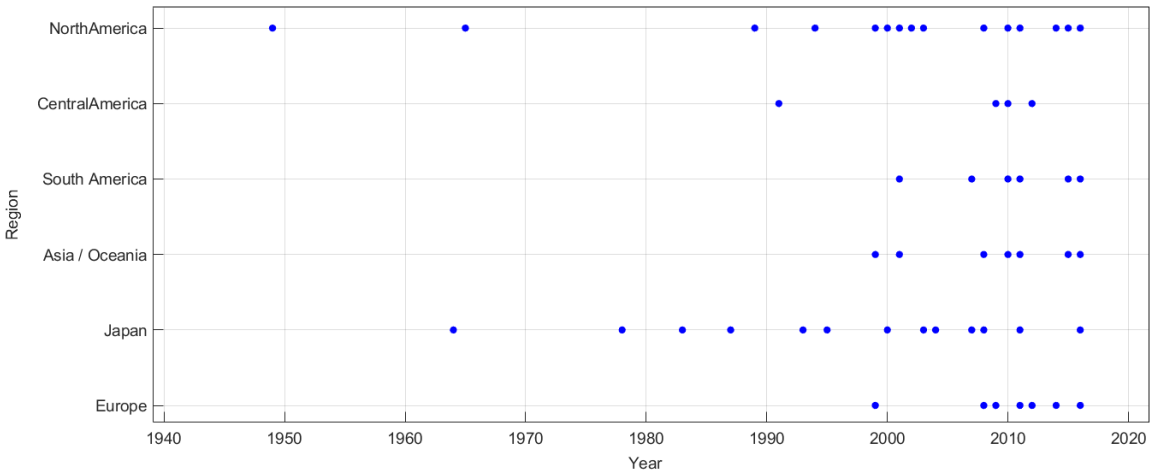



Figure 5-2. Historical earthquakes used in the inventory by region.

Figure 3 shows the comparative histogram of the frequency of earthquake samples per class of liquefaction and non-liquefaction. A few earthquakes, including the Canterbury earthquake sequence 2010-2011 (Christchurch and Darfield) in New Zealand and a few earthquakes in Japan (Kobe, Nihonkai, Tohoku, and Nigata earthquakes) dominate the liquefaction samples in the inventory. Denali, Wenchuan, Maule, and Puget Sound earthquakes contain the majority of the non-liquefaction data samples. The average number of samples per class (2085 for non-liquefaction and 2577 for liquefaction) across the events is also plotted as a horizontal line in Figure 3. The total number of samples in the inventory is 247,114, with 136,600 liquefaction samples (55.28 % of the inventory).

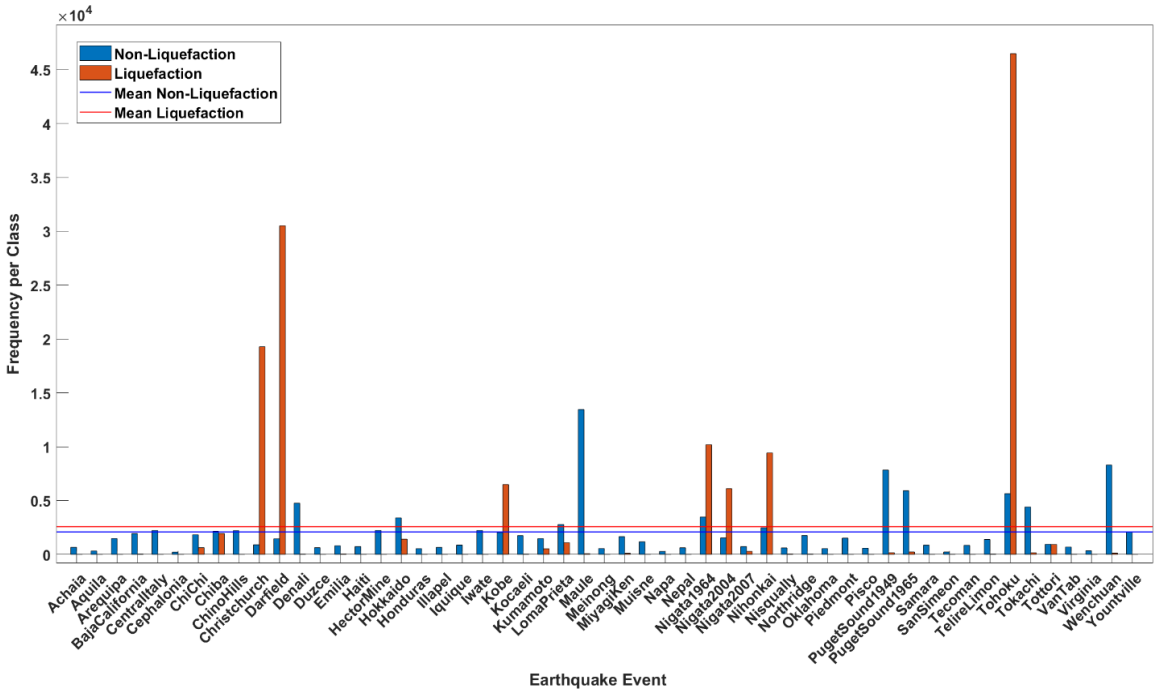



Figure 5-3. Histogram of the inventory samples by earthquake and class. The horizontal lines are indicative of the average number of samples per class across all earthquake events.

Increasing the number of samples, the number and regional diversity of earthquake events, and sampling a broader range of explanatory variables improves the generalization of the proposed model in terms of future predictions (Hastie et al., 2001; Zhu et al., 2017). Zhu et al. (2017) demonstrated how the addition of non-liquefaction earthquake events helps to distinguish liquefaction versus non-liquefaction classes better.

All the samples included in the inventory are discrete points, and no spatial polygons are used in this study. The spatial reference (latitude and longitude) of the liquefaction points were either derived from the reconnaissance reports, mapped using their geographic coordinates, or digitized through images and maps within the published literature, similar to Zhu et al. (2017). It should be noted that it is possible that some instances of liquefaction phenomena might have not been recorded, and there might be

uncertainty in the recorded liquefaction data. For instance, The Bhuj earthquake's samples were recently removed from the inventory used by Baise et al. (2023). The data from the Bhuj earthquake were derived from a remote sensing study (Singh et al., 2002) that lacks validation.

The dataset used in this study was collected by Zhan et al. (2023). Following the procedure used by previous researchers (Zhu et al.,2017; Bozzoni et al.,2020; Baise et al., 2021; Todorovic and Silva, 2022), two circular buffers (with a radius of 1 km and 15 km) were defined around each liquefaction point (manifestation observations) for the spatially incomplete data. The non-liquefaction sampling was performed in the area outside the 1 km but within the 15 km buffers on a grid with 1.5 km spacing (Zhan et al., 2023). Figure 4 shows the sampling strategy adopted in this research. A similar sampling approach has been implemented previously for landslide susceptibility mapping (Van Den Eeckhaut et al., 2012).

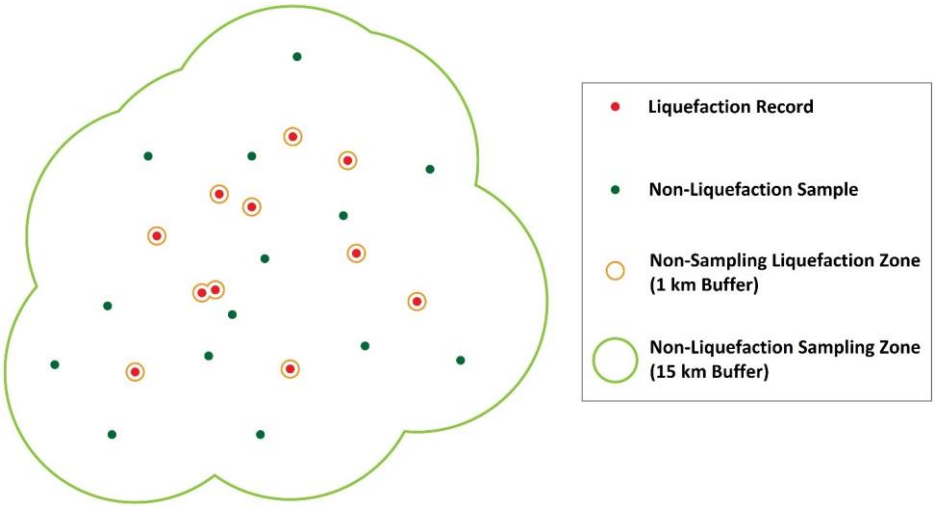



Figure 5-4. Non-liquefaction record sampling strategy used to develop the inventory. The red points are liquefaction occurrence records in the inventory, and the blue points are non-liquefaction sampled points within the buffer zone.

A total of 28 geospatial explanatory variables were used in this study, including two categorical and 26 continuous variables, falling within three categories of soil density (11 variables), saturation (14 variables), and earthquake loading (3 variables). Table 1 lists the variables with their brief description. The complete explanations with references and sources of the datasets used are provided in Table A-2 in the Appendix section. Each model variable is available as a global layer in the World Geographic Coordinate System 1984 (WGS1984) with a resolution of 30 arc-sec (~900 meters on the Equator).

Table 5-1. List of explanatory geospatial variables used in the study (***** is the indicator of categorical variables).

| No. | Name | Unit | Description | Category | | |
|---|---|---|---|---|---|---|
| | | | | Density | Saturation | Load |
| 1 | $Vs_{30}$ | m/s | Time-averaged shear-wave velocity over the first 30 m | | | |
| 2 | Elevation | m | Elevation | | | |
| 3 | Slope | deg | Topographic slope | | | |
| 4 | Thickness | m | Soil and sedimentary deposit thickness | | | |
| 5 | Landcover | - | Uplands vs. Lowlands (binary) | * | | |
| 6 | TPI | m | Topographic Position Index | | | |
| 7 | TRI | m | Topographic Roughness Index | | | |
| 8 | Roughness | m | Surface roughness | | | |
| 9 | CTI | - | Compound Topographic Index | | | |
| 10 | Landform | - | Landform and physiographic (land facet) patterns (15 classes). | * | | |
| 11 | Aspect | deg | Slope aspect | | | |
| 12 | PGA | g | Peak Ground Acceleration | | | |
| 13 | PGV | cm/s | Peak Ground Velocity | | | |
| 14 | Magnitude | Mw | Earthquake magnitude | | | |
| 15 | DC | km | Distance to coast (source 1) | | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 16 | DC2 | km | Distance to coast (source 2) | | | |
| 17 | DR | km | Distance to river (source 1) | | | |
| 18 | DR2 | km | Distance to river (source 2) | | | |
| 19 | DL | km | Distance to Lake | | | |
| 20 | DWB | km | Distance to the nearest waterbody (alternative source) | | | |
| 21 | DWB2 | km | Distance to nearest waterbody (source 1: minimum of DL, DR and DC) | | | |
| 22 | WBE | m | Elevation of the nearest waterbody | | | |
| 23 | HWB | m | Height above nearest surface waterbody (alternative source) | | | |
| 24 | HWB2 | m | Height above nearest waterbody (source 1: Elevation – WBE) | | | |
| 25 | HAND | m | Height above nearest drainage | | | |
| 26 | WTD | m | Water Table Depth | | | |
| 27 | Precipitation | mm | Mean annual precipitation | | | |
| 28 | AI | - | Aridity Index | | | |

## 5.5 Methodology

In this research, a novel system of voting ML classifiers is proposed as an alternative to logistic regression based geospatial liquefaction models. The system of voting machine classifiers allows for the inclusion of multiple global models, coastal and noncoastal models, and geographic regional models. The dataset is balanced across events by considering a cap for the number of samples per class. The sampled imbalanced data (in terms of binary classes) is distributed into subsets of balanced training data: 3 global models, 2 coastal models, 2 non-coastal models, and 6 geographic regional models. The

system of classifiers is implemented through both majority voting and soft voting of the classifiers. The model performance is assessed with a leave-one-out testing approach, which means that data of individual earthquake events are excluded from the training data in a loop over all events. The flowchart of the proposed system is presented in Figure 5.

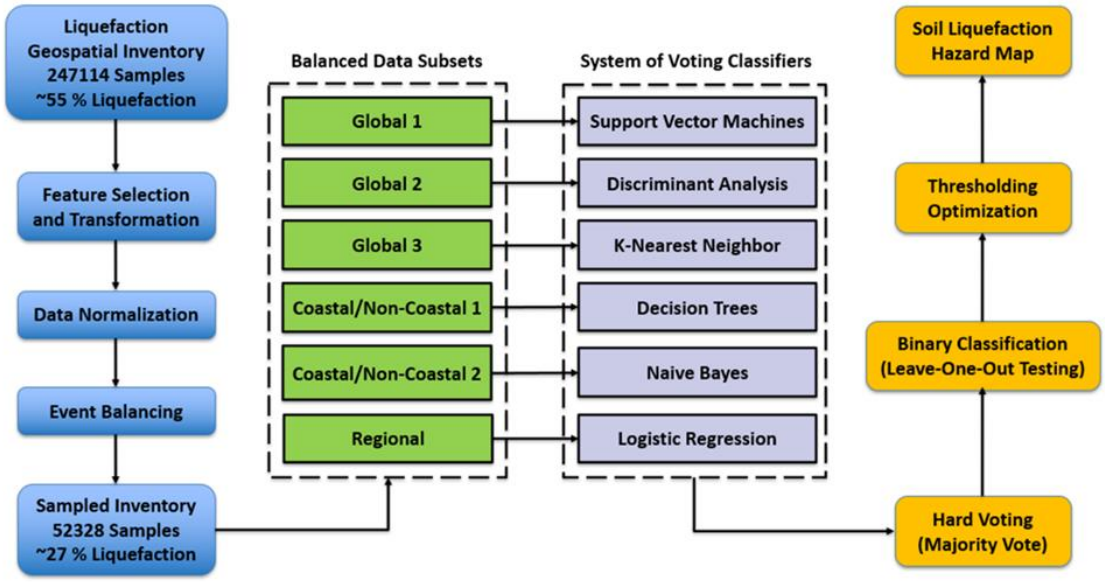



Figure 5-5. Flowchart of the proposed voting classification approach. Since 6 regions are considered in the inventory, 6 different regional models will be utilized based on the individual earthquake being tested.

### 5.5.1 Exploratory Data Analysis

Different types of exploratory data analysis (EDA) techniques were used to evaluate the features in terms of redundancy and relevance to the binary classification problem in this research. Various univariate (histograms, box plots, bar plots, etc.) and bi-variate (correlation analysis, etc.) analysis methods were considered to evaluate the parameters' class separability and their relationship. Feature transformations were proposed to improve the performance of several features in distinguishing the classes of liquefaction and non-liquefaction. After the EDA process, the data of all features ($X_i$), except the categorical

variables are normalized using the Z-score method to have a mean of zero and standard deviation of one, as presented via Equation 4, where μ is the mean, and σ is the standard deviation of each variable. Since most ML algorithms perform better when variables have a Gaussian distribution (Geyin et al., 2022), all predictors were transformed and normalized.

$$Z = \frac{X-\mu}{\sigma} \tag{4}$$

### 5.5.2 Data Imbalance

Data imbalance can be problematic for ML classification models as they are designed for balanced datasets; otherwise, their predictions will be biased toward the majority classes (Weiss and Provost, 2001). The spatially incomplete nature of the liquefaction inventory database and the associated class imbalance strongly influence the model's probabilities (Oommen et al., 2011; Thompson et al., 2016; Zhu et al., 2015 & 2017). One solution to encounter this issue is to under-sample the majority class by random sampling (Fernandez et al., 2019). The other solution could be to over-sample the minority class using synthetic data generation techniques. Bozzoni et al. (2020 & 2021) resolved the issue of class imbalance by trying commonly used synthetic data generation (over-sampling) algorithms, including SMOTE (Chawla et al., 2002) and ADASYN (He et al., 2008). However, such methods can be uncertain in terms of losing valuable information in under-sampling and adding noise and unrealistic samples to the data in the over-sampling process. Zhu et al. (2015) chose to use a highly imbalanced dataset (almost 13 non-liquefaction samples per each liquefaction record) because they aimed to develop a probability estimator that predicted the areal extent of liquefaction and wanted the probability to

correlate to the spatial extent. However, in their later version, Zhu et al. (2017) balanced their database after performing the non-liquefaction sampling process. Similarly, Baise et al. (2021) used a balanced sampling strategy, where the more significant liquefaction events are under-sampled so that they don't dominate the model. They capped the number of liquefaction and non-liquefaction points from each earthquake to 2000 and 1000 points, respectively.

In this study used an alternative approach. The dataset is balanced across the events by considering a cap for the number of samples per class (2085 for non-liquefaction and 2577 for liquefaction) based on the information presented in Figure 3. Then, the resulting dataset which contains ~27 % liquefaction samples are distributed over six various balanced subsets, including three sets of global data, two sets of coastal/non-coastal data (depending on the earthquake being tested), and a set of six geographical regional datasets, based on the regional categorization of the events, presented via Table A-1 and Figure 1. Because the method is demonstrated with the leave-one-out approach to model evaluation, the data of the tested earthquake is not included in any of the data subsets.

### 5.5.3 Feature Selection via Machine Learning

After removing a few features (explanatory geospatial variables) through the EDA process, a feature selection scheme is implemented to remove any features with zero or near-zero weight in the classification process. The strategy in this research is to keep explanatory variables as long as they are informative. The goal of feature selection is to reduce the risk of model over-fitting, to improve accuracy, and reduce training time. Neighborhood component analysis (NCA) is used in this study as the feature selection technique. NCA is

a ML algorithm for metric learning that learns in a supervised manner to improve the classification accuracy of a stochastic nearest neighbors' rule in the transformed space (Goldberger et al. 2005). The algorithm best estimates feature importance for distance-based supervised models that use pair-wise distances between observations to predict the response (Guyon and Elisseeff, 2003). NCA feature selection is performed with regularization to learn feature weights to minimize an objective function that measures the average leave-one-out classification loss over the training data (Yang et al., 2012). The regularization term (λ) in NCA's objective loss function derives the weights of irrelevant predictors to zero. As the number of observations n increases, the chance of overfitting decreases, and the required amount of regularization also decreases. λ is considered 1/n, where, n is the number of samples, as an approximately good selection for the regularization parameter. The stochastic gradient descent algorithm was used as the solver to estimate feature weights.

### 5.5.4 Classification Modeling

The logistic regression models developed in the studies by Zhu et al. (2015), Zhu et al. (2017), and Baise et al. (2021), were based on an intentional strategy to develop a simple model that uses parameters that reflect the underlying physics of the problem, and can be used easily and quickly for future events. In this work, the goal is to increase the predictive accuracy through the use of a complex ML model, and to exploit as many features as possible, if beneficial.

#### *5.5.4.1 Classification Thresholds*

Thresholds have commonly been used in geospatial liquefaction models to prevent overfitting and to include thresholds related to physical limits to liquefaction occurrence. The GGLM model (Zhu et al., 2017) used a PGV minimum threshold of 3 cm/s. Rashidian and Baise (2020) modified the GGLM model by recommending a PGA minimum threshold of 0.1 g as an additional threshold to better characterize the step-function behavior of the PGA plot. A threshold of 0.1 g for horizontal peak ground acceleration (PGA) was also used by Bozzoni et al. (2020) to reduce the spatial extent of liquefaction from prior geospatial liquefaction models, based on experimental data showing that liquefaction triggering is highly unlikely for PGA values of less than 0.10 g (Green and Bommer 2019). Based on the inventory used herein, liquefaction has occurred in the 2012 M6 Emilia Earthquake in Italy when PGA < 0.1 g indicating that this threshold may be too high.

The US Geological Survey (USGS) has also implemented a magnitude scaling factor (MSF) to the PGV term in the GGLM equation to reduce the over-prediction for low-magnitude earthquakes (M < 6). Rashidian and Baise (2020) have left the decision on using MSF to the GGLM model users, since they didn't observe any considerable positive or negative impact in their comparative analysis for high-magnitude earthquakes where over-prediction is more likely to happen, although they confirmed that the spatial extent of liquefaction will be moderated for low-magnitude events when using this factor.

Zhu et al. (2017) also implemented a maximum threshold of 620 m/s for the Vs30, which means that any point with a Vs30 value of greater than the threshold is automatically considered as non-liquefaction. In the current inventory, this Vs30 threshold leads to under-

prediction in for the 1993 M7.7 Hokkaido earthquake in Japan. Rashidian and Baise (2020) also recommended the use of an upper bound for GGLM (Zhu et al., 2017), when annual precipitation of a region exceeds 1700 mm, which was the upper quartile of annual precipitation in their development database of earthquakes, since overprediction of liquefaction was observed in regions with high precipitation. The annual precipitation threshold reduced the over-prediction of liquefaction in the 14 affected earthquake events in their study.

The use of thresholds has also been used in the study by Jena et al. (2022) in their deep learning model of liquefaction hazard for Australia. They have applied thresholds for different input variables including magnitude of Mw6. Green and Bommer (2019) concluded that earthquakes as small as Mw 4.5 can trigger liquefaction in extremely susceptible soil deposits, based on their field observations and a simple parametric study. However, they have indicated that for soil profiles which are appropriate for building structures, the minimum Mw 5 can be considered for liquefaction triggering (Green and Bommer, 2019). In this study, the threshold value of Mw 5 is considered as the minimum liquefaction triggering earthquake magnitude, although the lowest magnitude in the inventory, leading to liquefaction occurrence was 5.8 Mw.

In the current version of the GGLM models (GGLM-21-a and GGLM-21-b), Baise et al. (2021) do not retain the thresholds established by Zhu et al. (2017) and Rashidian and Baise (2020) around PGA, PGV, precipitation, and Vs30, since only 1.5% of the liquified data points in their inventory violate those thresholds. In the current study, we use the magnitude judgement-based thresholds and then rely on data-driven thresholds. Data

driven thresholds are only used for the features that are used in the final model, and do not violate the physics-based criteria of liquefaction triggering. It should be noted that the thresholds are designed to optimize the model performance as a post-processing step after the classification modeling, and no sample bypasses the classification step.

#### *5.5.4.2 System of Voting Classifiers*

This study proposes a system of voting machine learning classifiers to create liquefaction hazard maps. There are several benefits in implementing such approach, including 1) Reduction in sensitivity to under-sampling the dataset for the purpose of balancing, and taking advantage of all available valuable data, instead of under-sampling the majority class; 2) Reduction in model errors/uncertainties and inaccurate predictions by taking average of several model predictions, leading to stabilization of predictions and better generalizability and consistency of hazard modeling; 3) Benefiting from different type of classifiers, since not a single classifier is the best to predict all events; 4) Regional and coastal/non-coastal optimization of the classification decision by including these special classifiers in addition to the global classifiers; and 5) The ability to generate probability maps in a continuous form (soft voting) rather than 0-1 binary (hard voting) prediction.

Six different algorithms are used in this study, as shown in the model flowchart via Figure 5. These algorithms are widely used ML algorithms for classification purposes. They include support vector machines (SVM), discriminant analysis, K-nearest neighbor (KNN), decision trees, Naïve Bayes, and logistic regression, which are used as three global, two coastal/non-coastal, and a regional classifier, respectively. The k-fold cross-validation was used within the training step for all 6 classifiers to avoid overfitting (Kohavi, 1995).

The method divides the training dataset into k (k=5 in this study) smaller folds with equal number of samples points for the optimization of parameters. The proposed classification algorithms' hyperparameters and architecture were optimized separately with a grid search approach during the training process to achieve the best performance. A brief description of the classification algorithms used in this study is provided in the following sub-sections.

#### *5.5.4.3 Support Vector Machines*

Support vector machines are widely used and reported among the best-performing binary classifiers, although they are also used to solve multi-class problems. The objective of the SVM algorithm is to find a hyperplane that, to the best degree possible, separates the data points of one class from those of another class via the hyperplane with the largest margin between the two classes. Margin means the maximal width of the slab parallel to the hyperplane that has no interior data points. The algorithm can find such a hyperplane only for linearly separable problems. For most practical problems, the algorithm maximizes the soft margin allowing a small number of misclassifications. Support vectors refer to a subset of the training observations that identify the location of the separating hyperplane (Vapnik, 1995; Christianini and Shawe-Taylor, 2000; Hastie et al., 2009). In this study, the efficient linear SVM is used as a classifier for one of the global models in the voting classification system, which is less prone to over-fitting.

#### *5.5.4.4 Discriminant Analysis*

Discriminant analysis is used as a supervised classification method, which assumes that different classes generate data based on different Gaussian distributions (Fisher, 1936). To train the classifier, the fitting function estimates the parameters of a Gaussian distribution for each class and finds the smallest calculated misclassification cost to predict classes of

new (unlabeled) data. The linear discriminant analysis (LDA), which is less prone to overfitting, is used in this research computes the sample mean of each class. Then it computes the sample covariance by first subtracting the sample mean of each class from the observations of that class and taking the empirical covariance matrix of the result (Guo et al., 2007).

#### *5.5.4.5 K-Nearest Neighbors*

The K-nearest neighbors' algorithm is a non-parametric, supervised learning classifier which uses proximity to make classifications or predictions about the grouping of an individual data point. Categorizing query points based on their distance to points in a training data set can be a simple yet effective way of classifying new points. K-nearest neighbor (KNN) classifiers can use various metrics (Euclidean distance, in this study) to determine the distance between a set of data and query points. KNN classifier simply ranks the points by their distance to the query points in order to label them as the closest class, assuming that similar points can be found close to each other. A class label is assigned based on a majority vote, meaning the most frequently represented label around a given data point is used (Cover and Hart, 1967). The value of K is optimized to avoid over-fitting and under-fitting on the global data.

#### *5.5.4.6 Decision Trees*

Decision trees, or supervised non-parametric classification trees, predict responses to data inputs by following the decisions in the tree from the root (beginning) node down to a leaf (end) node, which contains the response (Breiman et al., 1984; Rokach and Maimon, 2008). In tree structures, which are famous for their simplicity and easy interpretability, leaves represent class labels, and branches represent conjunctions of features that lead to those

class labels. Classification trees give nominal responses, such as liquefaction or non-liquefaction. A tree is built by splitting the source set, constituting the root node of the tree, into subsets that constitute the successor children. The splitting is based on a set of splitting rules based on classification features (Shalev-Shwartz and Ben-David, 2014). This process is repeated on each derived subset in a recursive manner called recursive partitioning. The recursion is completed when the subset at a node has all the same values of the target variable or when splitting no longer adds value to the predictions. The cost of using the tree (i.e., predicting data) is logarithmic in the number of data points used to train the tree. Thus, the classifier is assigned to the coastal/non-coastal part of the model, which can deal with a low number of data when the tested event is non-coastal.

#### *5.5.4.7 Naïve Bayes*

Naïve Bayes is a classification algorithm that applies density estimation to the data. The naïve Bayes classifier is designed based on the naïve assumption that predictors are independent of one another within each class (Hastie et al., 2009). While the class-conditional independence between predictors is not true in general, research shows that this optimistic assumption works well in practice. It classifies data in two steps: 1) Using the training data, the method estimates the parameters of a probability distribution, assuming predictors are conditionally independent given the class; then, 2) For any unseen test data, the method computes the posterior probability of that sample belonging to each class. The method then classifies the test data according to the largest posterior probability (Manning et al., 2008). This study uses normal (Gaussian) predictor conditional distribution in the training function. This assumption of class-conditional independence of the predictors allows the naive Bayes classifier to estimate the parameters required for accurate

classification while using less training data than many other classifiers. That is why Naïve Bayes classifier is assigned to the coastal/non-coastal part of the model, which can deal with low data when the tested event is non-coastal. This makes it particularly effective for data sets containing many predictors.

#### *5.5.4.8 Logistic Regression*

Logistic Regression is a supervised ML algorithm that is commonly used for binary class classification problems. Logistic regression is currently one of the most well-known and widely used ML algorithms (James et al., 2013). Logistic regression operates using a sigmoidal function for values between 0 and 1. A logistic regression algorithm takes as its input a feature vector and outputs a probability, where the feature vector represents an object belonging to the positive class. Since the logistic regression model is very simple, compared to the other classifiers used in this study, it is assigned to the regional datasets, as these datasets are expected to have more common characteristics, compared to global and coastal/non-coastal models.

### 5.5.5 Model Validation

#### *5.5.5.1 Leave-One-Out Testing*

The validation of the system is done by excluding data from an individual earthquake from the training process, and independently evaluating the inventory for that event in a leave-one-out testing approach. This process is done to fairly assess the performance of the model for new unseen earthquakes. This approach is different than the validation scheme followed by previous studies (Zhu et al., 2017; Todorovic and Silva, 2022), where they use all or a large portion of samples from all earthquakes in the training and model development process. This approach leads to a better understanding of the

performance of the system in an unbiased way and provides insights into system of classifiers.

Figure 6 provides a visual explanation of the model assessment and implementation and provides an example of how the soft and hard voting systems work with data from an unseen earthquake event. In this study, hard voting considering the majority votes of the classifiers. If at least 3 classifiers label a sample with liquefaction (class 1), the point will be classified as liquefaction. As an alaternative, the soft voting can generate probability-based liquefaction maps that take into account epistemic uncertainty. The soft voting decision is made by taking the average of the class labels predicted by the classifiers in the voting system. In this research, since multiple classifiers are used to automatically balance each other's outputs, the majority hard voting (3+ votes in favor of either class) is considered as the suitable boundary between the two classes and is similar to the optimal threshold used by Bozzoni et al. (2020) in their logistic regression model (0.57).

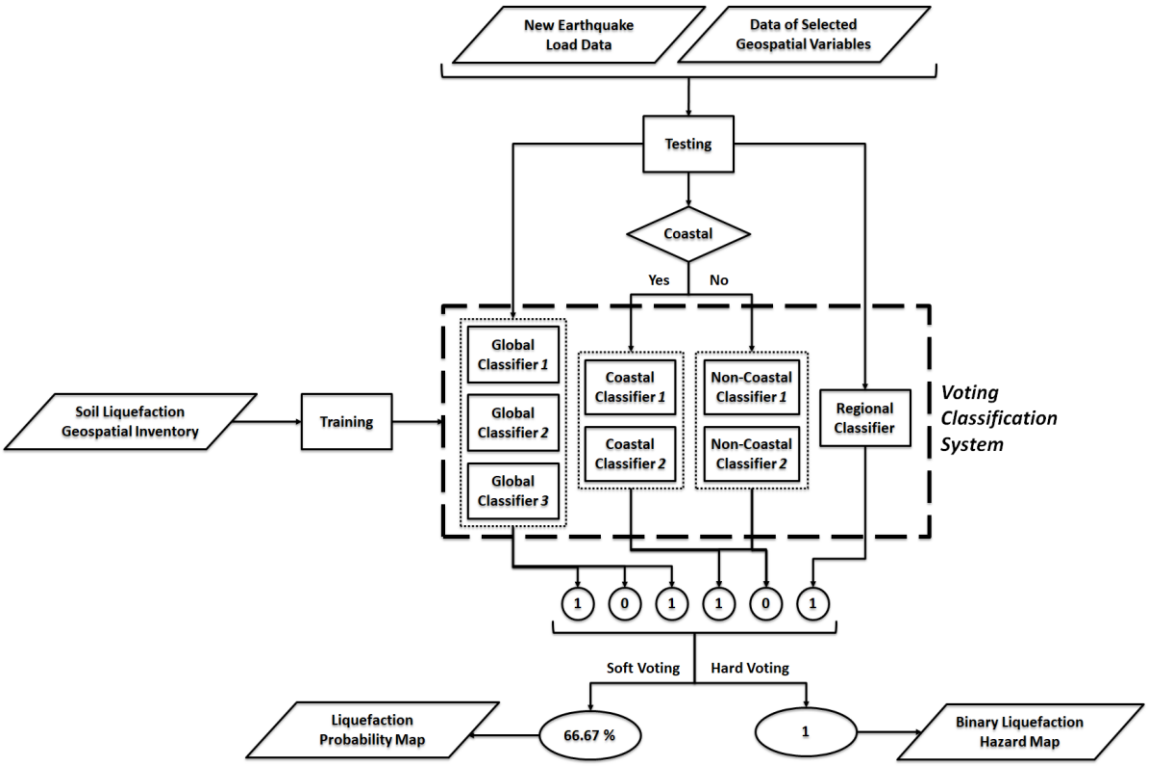



Figure 5-6. The leave-one-out testing approach used in this study, plus an example of how the final classification decision is made via hard voting or soft voting procedures. Global classifiers 1, 2 and 3 are SVM, discriminant analysis, and KNN, respectively. Coastal/non-coastal classifiers 1 and 2 are decision trees and Naïve Bayes, respectively, and the regional classifier uses logistic regression.

#### *5.5.5.2 Accuracy Indices*

In order to evaluate accuracy, we use the hard voting with majority voting so that each pixel is labelled as 1 (liquefaction or positive class), or 0 (non-liquefaction or negative class). The performance of the model was assessed via ROC (Receiver Operating Characteristics) analysis and expressed via the Area Under Curve (AUC) parameter. In a ROC curve, the Sensitivity or True Positive Rate (TPR) is plotted against the Specificity or False Positive Rate (FPR). The TPR and FPR are calculated via Equations 5 and 6.

$$\text{Sensitivity} = TPR = \frac{TP}{TP+FN} \quad (5)$$

$$Specificity = TNR = \frac{TN}{TN+FP} \quad (6)$$

Where, TP (True Positive) is liquefaction predicted as liquefaction, FP (False Positive) is non-liquefaction predicted as liquefaction, TN (True Negative) is non-liquefaction predicted as non-liquefaction, FN (False Negative) is liquefaction predicted as non-liquefaction. The AUC ranges between 0.5 (random classifier) and 1 (100% accurate classifier). These accuracy indices (Sensitivity, Specificity, and AUC) are provided for the models alongside the Prevalence, which shows the percentage of the specific earthquake dataset belonging to the positive class (liquefaction). The AUC values are compared with previous works and among the models presented in this study.

## 5.6 Results and Discussion

### 5.6.1 Exploratory Data Analysis

In the exploratory data analysis (EDA) step, two main goals were pursued: 1) EDA-based and knowledge-based feature removal; and 2) Implementation of appropriate feature transformations. As a result, several features were removed or transformed before undergoing the ML-based feature selection. Table 2 provides the results of the 11 features removed and the reasons behind the decisions. As an example, roughness was removed because of very high correlation to TRI, leading to redundancy in features (Figure 7-a). Similarly, aspect variable was removed because of the lack of separability between the two classes of liquefaction and non-liquefaction (Figure 7-b).

PGA is frequently used in liquefaction modeling methods (Seed and Idriss, 1971; Zhu et al., 2015; Bozzoni et al., 2020) because it is proportional to the maximum shear stress induced in the sediment (Terzaghi et al., 1996). However, PGV has shown a stronger correlation with liquefaction occurrence than PGA in several studies (Orense, 2005; Zhu

et al., 2017; Baise et al., 2021). This implies that soil liquefaction could be more sensitive to the low-frequency components of the ground motion, and liquefaction is more dependent on input kinetic energy, which is well-characterized by the PGV (Kotoda et al., 1986; Midorikawa and Wakamatsu, 1988; Liang et al., 1995; Orense, 2005; Bardet and Liu, 2009). Additionally, Bardet and Liu (2009) observed nonlinear behavior between the PGV and the empirical probability of liquefaction. As PGV increases, the probability of liquefaction rapidly increases and then reaches a plateau when PGV is greater than a threshold of 10 cm/s. Therefore, PGA was removed from the features based on the previous knowledge in published literature, and PGA was PGV was retained as the loading feature.

Table 5-2. EDA-based and knowledge-based feature removal results.

| Number | Removed Feature | Reasoning |
|---|---|---|
| 1 | DR2 | Repetitive information from alternative source |
| 2 | DC2 | Repetitive information from alternative source |
| 3 | Precipitation | ○ Correlation with AI (redundancy)<br>○ Previous knowledge from literature on over-prediction issues (Rashidian and Baise, 2020) |
| 4 | PGA | ○ Previous knowledge from the literature on the superiority of PGV in correlating with liquefaction occurrences (Zhu et al., 2017)<br>○ Correlation with PGV |
| 5 | DWB2 | Repetitive information from alternative source |
| 6 | HWB2 | Repetitive information from alternative source |
| 7 | Slope | Correlation with TRI (redundancy) |
| 8 | Roughness | Correlation with TRI (redundancy) |
| 9 | Aspect | No observable separability of classes in univariate and bi-variate analysis with other features (irrelevance) |
| 10 | Landform | High number of classes leading to several categorical dummy variables in the classification modeling |
| 11 | Landcover | Binary categorical data, not desirable based on the designed voting classification method |

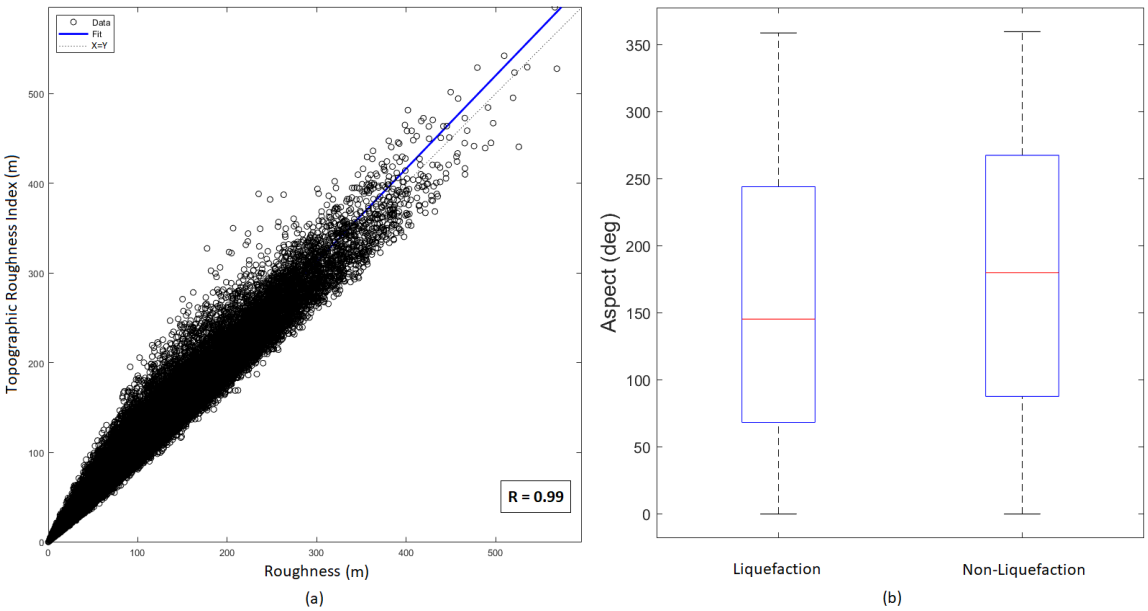



Figure 5-7. a) High linear correlation with regression coefficient of R=0.99 between the two variables pf roughness and topographic roughness index (TPI); and b) Comparative box plots of the aspect variable per class, showing almost similar behavior of data in the two classes.

Feature transformations were applied when the transformation resulted in increased ability to distinguish the two classes. Table 3 presents the feature transformations implemented in this study. The shaking parameters (PGV and PGA) are well-represented by a lognormal distribution; thus, they were transformed by taking their natural logarithm. The same transformation was applied to the Vs30 values. Several previous studies have used these transformations (Zhu et al., 2017; Bozzoni et al., 2020; Baise et al., 2021; Todorovic and Silva, 2022). The distributions of some variables were skewed (distance to the water body and water table depth); thus, they were transformed by taking the square root of the original values. This transformation was also used by Todorovic and Silva (2022). TPI and TRI were transformed in the study by Baise et al. (2021) by taking the square root of the absolute value. Figure 8 illustrates the benefit of feature transformations

on the TPI variable. Baise et al. (2021) have also used square-root transformation for distance to the closest waterbody, elevation above the closest waterbody, and elevation.

Table 5-3. Feature transformation techniques used in this study. “Ln”, “SQRT”, “ABS”, and “x”, stand for natural logarithm, square root, absolute value, and the original value, respectively.

| Number | Feature | Transformation Technique |
|---|---|---|
| 1 | PGV | Ln (x) |
| 2 | DL | SQRT (x) |
| 3 | DWB | SQRT (x) |
| 4 | HWB | SQRT (ABS (x)) |
| 5 | WTD | SQRT (x) |
| 6 | HAND | SQRT (x) |
| 7 | $V_{S30}$ | Ln (x) |
| 8 | TPI | SQRT (ABS (x)) |
| 9 | TRI | SQRT (ABS (x)) |
| 10 | Elevation | SQRT (x) |

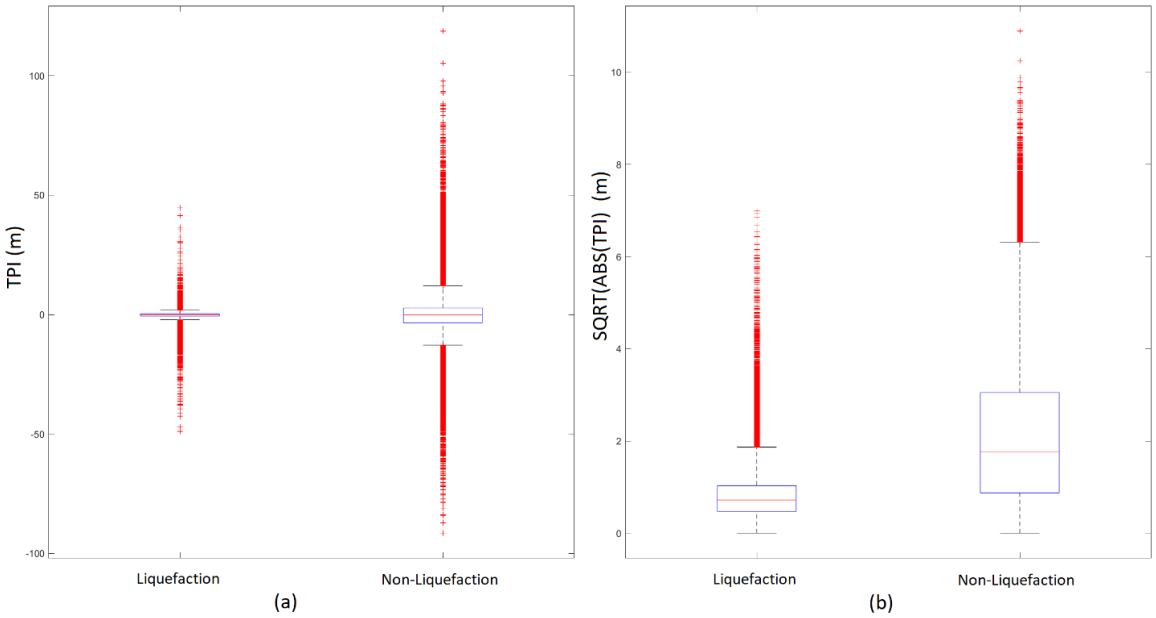



Figure 5-8. a) Comparative box plots of the original TPI data of the two classes, versus b) Transformed TPI data (square root of the absolute value), showing the shift in their comparative means, and better distinguishment between the two classes.

### 5.6.2 Feature Selection via Machine Learning

Neighborhood component analysis (NCA) feature selection was used to compare the features' weight, as combined, and to remove the features with zero or near-zero weight. However, based on the results provided via Table 4, all the features showed contribution to the classification. Thus, it was decided to keep all the remaining 17 variables from the previous steps in the classification modeling. To improve the certainty of NCA performance, and because of the stochastic manner of NCA, the datasets were partitioned into five subsets, and the average weight of features on the NCA analysis of the subsets is reported in Table 4, ranked from high to low. It should be noted that the weights are based on the combined use of the variables, as NCA does not evaluate features independently, which is one of the reasons for using this algorithm at this feature selection step. As an example, on the importance of integrated feature selection, Zhu et al. (2017) reported in

interaction assessment (bivariate analysis) that distance to the river does not appear to be a good predictor when it is evaluated alone. Still, it becomes valuable when combined with distance to the coast.

Table 5-4. NCA feature selection results.

| Rank | Feature | Weight |
|---|---|---|
| 1 | AI | 5.702 |
| 2 | PGV | 5.409 |
| 3 | Magnitude | 5.300 |
| 4 | DL | 4.558 |
| 5 | DC | 4.350 |
| 6 | DWB | 3.413 |
| 7 | Elevation | 3.388 |
| 8 | $Vs_{30}$ | 3.273 |
| 9 | DR | 3.254 |
| 10 | HWB | 2.966 |
| 11 | HAND | 2.961 |
| 12 | Thickness | 2.452 |
| 13 | TRI | 2.386 |
| 14 | WTD | 1.761 |
| 15 | CTI | 1.249 |
| 16 | TPI | 0.936 |
| 17 | WBE | 0.713 |

## 5.6.3 Binary Classification Modeling

### *5.6.3.1 Data-Driven Classification Thresholds*

Based on the basic descriptive statistics of the selected features, presented in Table 5, a few data-driven thresholds were chosen for post-classification optimization of the results. The selected thresholds are shown in red in the Table 5. These data-driven thresholds, plus the Mw 5 threshold for liquefaction triggering (10 variables in total), results in pre-defined non-liquefaction labels for 9418 samples out of 247114 total samples (3.8 % of the samples). For instance, no liquefaction observation is available in the inventory in any location with an aridity index of higher than 27659.90. Thus, this value is used as a post-processing threshold higher than which the tested samples by the proposed model cannot be labeled as liquefaction.

Table 5-5. Descriptive data statistics of the continuous geospatial variables by class.

| Variable | Class | Min | Max | Mean | Median | Mode | STD | Skew |
|---|---|---|---|---|---|---|---|---|
| **$Vs_{30}$** | Liq. | 137.22 | 900.00 | 232.29 | 219.32 | 260.20 | 78.85 | 2.93 |
| | Non-Liq. | 98.00 | 900.00 | 460.36 | 434.02 | 641.30 | 193.00 | 0.48 |
| **Elevation** | Liq. | 0.00 | 3649.99 | 11.96 | 3.82 | 0.00 | 37.40 | 26.11 |
| | Non-Liq. | 0.00 | 4130.48 | 346.85 | 184.57 | 0.00 | 413.91 | 1.99 |
| **Thickness** | Liq. | 0.00 | 50.00 | 44.71 | 50.00 | 50.00 | 11.49 | -2.37 |
| | Non-Liq. | 0.00 | 50.00 | 13.52 | 3.25 | 1.00 | 18.06 | 1.21 |
| **TPI** | Liq. | **-49.03** | **44.57** | -0.19 | -0.08 | 0.00 | 1.73 | -3.95 |
| | Non-Liq. | -91.50 | 118.75 | 0.07 | -0.12 | 0.00 | 11.73 | 0.40 |
| **TRI** | Liq. | 0.00 | **322.23** | 6.57 | 4.07 | 0.00 | 9.82 | 7.46 |
| | Non-Liq. | 0.00 | 596.15 | 64.17 | 40.44 | 0.00 | 69.03 | 1.60 |
| **CTI** | Liq. | 505.46 | 2731.87 | 1018.14 | 961.09 | 928.39 | 219.13 | 1.73 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Non-Liq. | 386.38 | 2471.00 | 829.72 | 797.58 | 857.00 | 211.61 | 1.16 |
| **PGV** | Liq. | **1.01** | 158.15 | 45.46 | 33.79 | 28.54 | 28.31 | 0.96 |
| | Non-Liq. | 0.04 | 166.81 | 21.64 | 17.26 | 20.73 | 17.97 | 2.02 |
| **Magnitude** | Liq. | 5.80 | 9.10 | 7.66 | 7.70 | 7.70 | 0.96 | 0.26 |
| | Non-Liq. | 4.00 | 9.10 | 7.40 | 7.60 | 7.90 | 1.02 | -0.53 |
| **DC** | Liq. | 0.00 | 356.15 | 27.00 | 12.19 | 0.00 | 29.35 | 1.18 |
| | Non-Liq. | 0.00 | 370.99 | 58.69 | 28.12 | 78.00 | 76.16 | 1.95 |
| **DR** | Liq. | 0.00 | 1287.21 | 25.32 | 15.18 | 0.00 | 26.30 | 7.04 |
| | Non-Liq. | 0.00 | 1301.71 | 76.89 | 32.12 | 0.00 | 134.18 | 5.54 |
| **DL** | Liq. | 0.00 | 133.07 | 10.57 | 8.70 | 0.00 | 8.45 | 1.19 |
| | Non-Liq. | 0.00 | 142.27 | 18.11 | 13.53 | 0.00 | 16.53 | 2.51 |
| **DWB** | Liq. | 0.00 | **40.39** | 4.60 | 2.56 | 0.00 | 4.62 | 1.27 |
| | Non-Liq. | 0.00 | 109.33 | 10.03 | 7.41 | 0.00 | 11.37 | 4.35 |
| **WBE** | Liq. | 0.00 | **1651.00** | 8.38 | 0.80 | 0.00 | 45.01 | 17.49 |
| | Non-Liq. | 0.00 | 2703.00 | 196.00 | 25.00 | 0.00 | 317.63 | 2.52 |
| **HWB** | Liq. | **-1476.11** | **2342.99** | 3.73 | 1.93 | 0.00 | 16.98 | 25.62 |
| | Non-Liq. | -2800.10 | 3212.48 | 105.10 | 51.68 | 0.00 | 340.80 | -0.87 |
| **HAND** | Liq. | 0.00 | **1810.75** | 5.85 | 4.00 | 4.00 | 10.75 | 61.37 |
| | Non-Liq. | 0.00 | 4138.25 | 108.08 | 43.75 | 2.00 | 161.71 | 3.10 |
| **WTD** | Liq. | 0.00 | **216.13** | 1.69 | 0.63 | 0.00 | 4.47 | 9.23 |
| | Non-Liq. | 0.00 | 488.48 | 31.35 | 20.19 | 0.00 | 36.89 | 2.35 |
| **AI** | Liq. | 0.00 | **27659.90** | 7978.74 | 6159.70 | 11242.63 | 5595.91 | 0.47 |
| | Non-Liq. | 0.00 | 39830.00 | 10264.86 | 10444.00 | 0.00 | 6376.04 | 0.42 |

### 5.6.4 Voting Classification System

The voting classification system was developed by training individual classifiers and aggregating them through hard and soft voting. The individual earthquake events were tested via the leave-one-out approach. The accuracy results, including overall accuracy, sensitivity, specificity, and area under the ROC curve (AUC), are provided in Table 6 for all earthquakes in the inventory, ranked by the validation's AUC value, except for the non-liquefaction events which come at the bottom of the table, ranked by their overall accuracy. The prevalence column in Table 6 shows the percentage of liquefaction points in the inventory for each event. The accuracy of predictions for the 2003 San Simeon and the 2010 Haiti earthquakes were the best among the earthquakes with liquefaction observations, followed by the 2010 Tottori. Figure 9 shows the predicted liquefaction maps for these events.

Table 5-6. Accuracy statistics of the leave-one-out approach. All values are in percentage (%).

| Rank | Earthquake | Prevalence | Overall Accuracy | Specificity | Sensitivity | AUC |
|---|---|---|---|---|---|---|
| 1 | **San Simeon** | 7.79 | 96.31 | 99.11 | 63.16 | 98.77 |
| 2 | **Haiti** | 2.00 | 83.71 | 83.38 | 100.00 | 97.36 |
| 3 | **Tottori** | 49.43 | 91.96 | 84.63 | 99.45 | 96.65 |
| 4 | **Christchurch** | 95.59 | 93.22 | 83.15 | 93.69 | 95.99 |
| 5 | **Illapel** | 0.77 | 98.93 | 99.23 | 60.00 | 95.64 |
| 6 | **Samara** | 1.48 | 86.64 | 86.56 | 92.31 | 95.49 |
| 7 | **Loma Prieta** | 28.16 | 82.32 | 75.79 | 98.99 | 94.82 |
| 8 | **Cephalonia** | 5.31 | 96.46 | 99.53 | 41.67 | 94.68 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 9 | **Nihonkai** | 79.22 | 92.40 | 80.90 | 95.41 | 93.86 |
| 10 | **Kobe** | 76.11 | 91.36 | 67.70 | 98.78 | 92.40 |
| 11 | **Kocaeli** | 1.96 | 95.24 | 96.17 | 48.57 | 92.17 |
| 12 | **Honduras** | 2.40 | 54.90 | 53.79 | 100.00 | 92.14 |
| 13 | **Nisqually** | 8.45 | 64.40 | 61.78 | 92.86 | 91.67 |
| 14 | **Hokkaido** | 29.42 | 89.41 | 93.13 | 80.48 | 90.96 |
| 15 | **Nepal (Gorkha)** | 1.91 | 93.80 | 94.81 | 41.67 | 90.50 |
| 16 | **Muisne** | 1.10 | 85.40 | 85.32 | 92.31 | 89.65 |
| 17 | **Nigata 1964** | 74.61 | 88.62 | 68.80 | 95.36 | 89.60 |
| 18 | **Miyagi ken** | 5.93 | 65.56 | 63.52 | 98.08 | 89.42 |
| 19 | **Darfield** | 95.45 | 91.13 | 67.03 | 92.28 | 88.92 |
| 20 | **Tokachi** | 3.23 | 69.35 | 68.62 | 91.16 | 88.88 |
| 21 | **Napa** | 0.72 | 56.16 | 55.84 | 100.00 | 88.69 |
| 22 | **Tohoku** | 89.19 | 90.72 | 54.91 | 95.06 | 86.86 |
| 23 | **Oklahoma** | 0.74 | 92.96 | 93.25 | 50.00 | 86.47 |
| 24 | **Kumamoto** | 26.09 | 81.18 | 84.32 | 72.31 | 85.89 |
| 25 | **Nigata 2004** | 80.00 | 77.66 | 85.47 | 75.70 | 85.35 |
| 26 | **Chiba** | 47.37 | 72.47 | 54.91 | 91.99 | 84.84 |
| 27 | **Duzce** | 0.48 | 92.82 | 92.95 | 66.67 | 84.80 |
| 28 | **Nigata 2007** | 27.81 | 66.47 | 53.55 | 100.00 | 83.19 |
| 29 | **Tecoman** | 1.05 | 69.44 | 69.35 | 77.78 | 81.84 |
| 30 | **Puget Sound 1949** | 1.92 | 92.31 | 93.03 | 55.56 | 81.62 |
| 31 | **Telire Limon** | 3.00 | 59.53 | 58.35 | 97.67 | 81.48 |
| 32 | **Pisco** | 2.50 | 97.84 | 99.83 | 20.00 | 80.06 |

| | | | | | | |
|---|---|---|---|---|---|---|
| 33 | **Denali** | 0.83 | 94.34 | 95.05 | 10.00 | 75.29 |
| 34 | **Achaia** | 1.06 | 85.89 | 86.35 | 42.86 | 73.38 |
| 35 | **Meinong** | 1.99 | 39.17 | 37.94 | 100.00 | 73.36 |
| 36 | **Emilia** | 6.74 | 51.86 | 48.88 | 93.10 | 72.99 |
| 37 | **Baja California** | 2.52 | 51.66 | 50.62 | 92.00 | 67.56 |
| 38 | **Puget Sound 1965** | 3.73 | 84.14 | 85.97 | 37.12 | 65.18 |
| 39 | **Northridge** | 2.33 | 97.67 | 99.72 | 11.90 | 62.42 |
| 40 | **Arequipa** | 0.74 | 98.52 | 99.25 | 0.00 | 62.18 |
| 41 | **Maule** | 0.53 | 94.21 | 94.62 | 16.67 | 58.58 |
| 42 | **Chi-Chi** | 2.56 | 70.27 | 94.23 | 0.48 | 50.00 |
| 43 | **Aquila** | 1.23 | 98.77 | 100.00 | 0.00 | - |
| 44 | **Iquique** | 1.25 | 98.75 | 100.00 | 0.00 | - |
| 45 | **Wenchuan** | 1.38 | 98.48 | 99.86 | 0.00 | - |
| 46 | **VanTab** | 2.84 | 97.16 | 100.00 | 0.00 | - |
| 47 | **Virginia** | 0.57 | 99.43 | 100.00 | 0.00 | - |
| 1 | **Chino Hills** | 0.00 | 100.00 | - | - | - |
| 1 | **Hector Mine** | 0.00 | 100.00 | - | - | - |
| 1 | **Piedmont** | 0.00 | 100.00 | - | - | - |
| 1 | **Yountville** | 0.00 | 100.00 | - | - | - |
| 5 | **Central Italy** | 0.00 | 99.78 | - | - | - |
| 6 | **Iwate** | 0.00 | 89.43 | - | - | - |

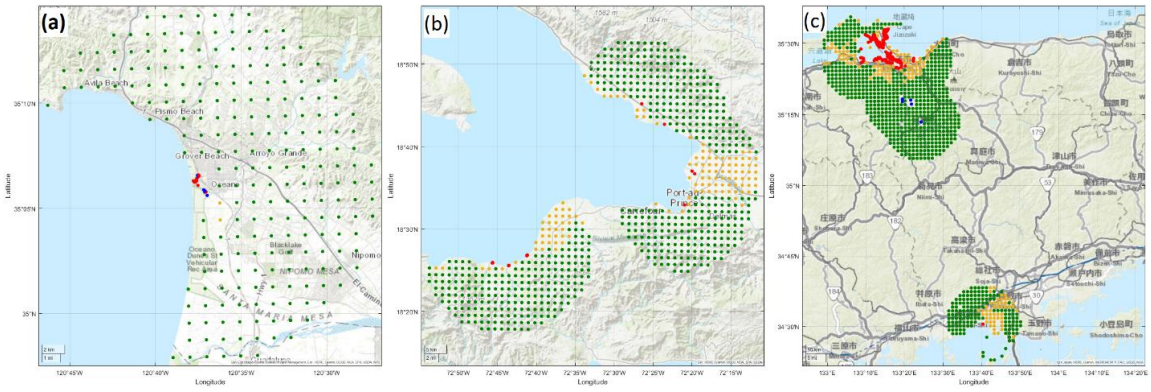



Figure 5-9. a) Predicted hazard map of the 2003 San Simeon; b) 2010 Haiti; and c) 2010 Tottori earthquakes. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

The 2011 Christchurch earthquake has the fourth highest accuracy among the liquefaction-observed earthquakes in the inventory, while the 2010 Darfield event's accuracy is relatively lower (ranked 19th among the liquefaction-observed events). Figure 10 provides the liquefaction maps for these two events (Canterbury 2010-11 events in New Zealand). The difference in the predicted maps' accuracy was also observed in the results of Todorovic and Silva (2022) for their random forest (decision tree-based) model but with a more significant difference. An interesting observation is that the decision tree classifier (coastal classifier 1) in this study also under-performed (AUC = 57.40) compared to the other five classifiers in the voting classification system. The voting system has a higher accuracy (add accuracy) than all individual 6 classifiers, which demonstrates the integrated power of the proposed method in providing more accurate predictions. Table 7 shows the accuracy (AUC = 88.92) of individual classifiers, in addition to the final voting system for all tested earthquakes.

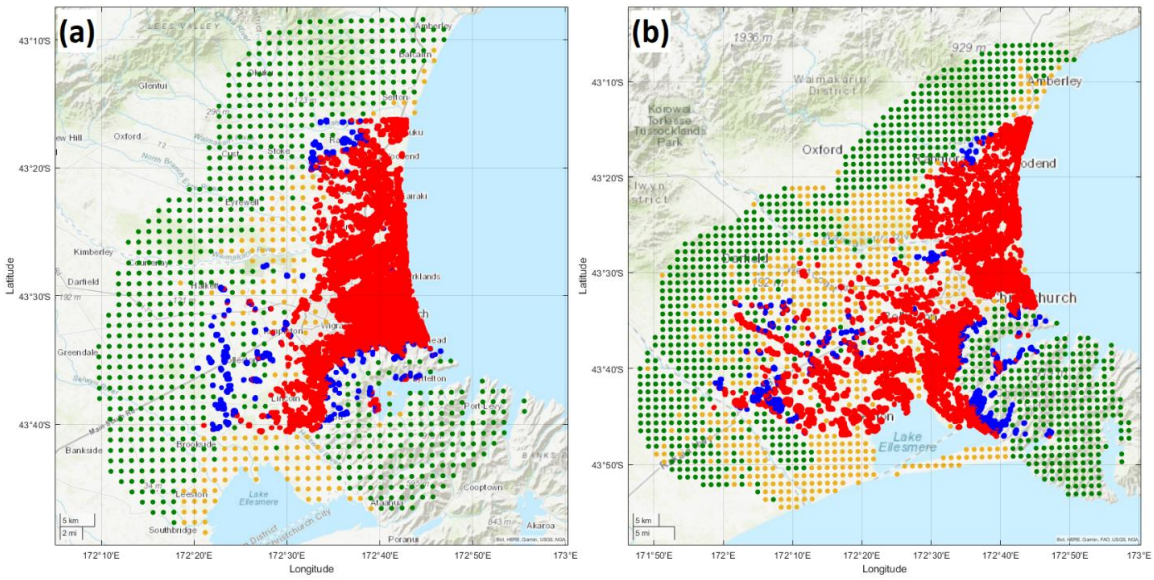



Figure 5-10. a) Predicted hazard map of the 2011 Christchurch; and b) 2010 Darfield earthquakes. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

According to Table 7, the AUC value for the proposed voting system is higher than all individual classifiers in 88.4 % of the tested earthquakes for which AUC is calculated. For the majority of other events, it's very close to the highest AUC among the classifiers. This is evidence of the benefit of the proposed method, especially compared to the three global classifiers, which represent statistically optimal global models. Additionally, the spatial extent of liquefaction is moderated, even in the cases where the voting system shows a relatively lower AUC than the highest performing individual classifier. A very important observation is that not a single global classifier can be considered as best-performing for all the events. According to the results provided in Table 7, among the 3 global models, global classifiers 1, 2, and 3, perform the best on 30 %, 35 % and 30 % of the events, respectively, while they perform equally poor on the remaining 5 % of the events. The

voting classification system performs better by taking into account the decisions made by coastal/non-coastal and regional classifiers in addition to the global models.

The 1999 Chi-Chi earthquake was among the six earthquakes which performed worst in terms of liquefaction detection by global models as well as by the voting system. The soft voting system is used with a probability threshold of 15 %, to investigate if at least one classifier among the voting classifiers predicted the liquefaction. Figure 11 shows the hard voting liquefaction map for the 1999 Chi-Chi event alongside the soft voting prediction with lower threshold for comparison, which shows a much better performance. The soft voting approach could be used to map regions with low probability of liquefaction. In the case of the Chi-Chi earthquake, the non-coastal classifier 2 performs the best. For the 1999 Chi-Chi event, Zhu et al. (2017) have also reported that many observed liquefaction points had medium probability, and their model overpredicted for coastal areas, where very few liquefaction occurrences were observed.

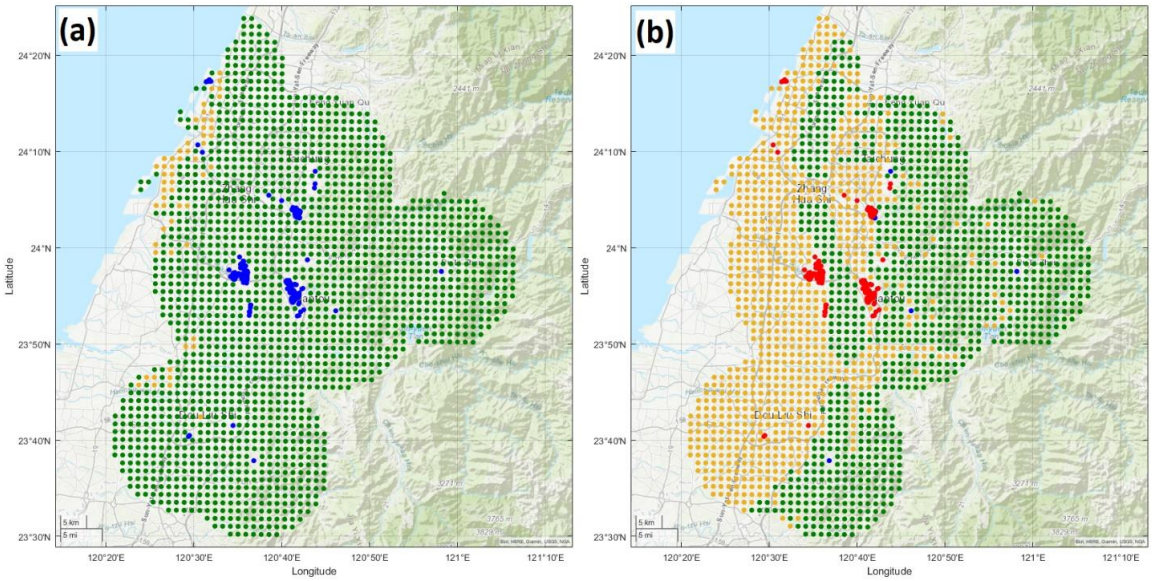



Figure 5-11. a) Predicted hazard map of the 1999 Chi-Chi earthquake with the voting classification system using hard voting; and b) using low-probability threshold via soft voting. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

Table 5-7. Comparative accuracy of single classifiers versus the proposed system of voting classifiers. All values are AUC in percentage (%). The events are listed in alphabetical order.

| Earthquake | Global 1 | Global 2 | Global 3 | Coastal / Non-Coastal 1 | Coastal / Non-Coastal 2 | Regional | Proposed Method |
|---|---|---|---|---|---|---|---|
| **Achaia** | 66.37 | 75.51 | 68.44 | 64.76 | 51.93 | 60.23 | 73.38 |
| **Arequipa** | 50.00 | 50.00 | 50.00 | 53.94 | 50.00 | 61.82 | 62.18 |
| **Baja California** | 50.00 | 68.93 | 50.00 | 50.00 | 59.19 | 69.34 | 67.56 |
| **Cephalonia** | 62.27 | 70.83 | 54.17 | 89.80 | 62.50 | 86.33 | 94.68 |
| **Chi-Chi** | 50.00 | 50.00 | 50.00 | 59.18 | 75.77 | 50.00 | 73.95 |
| **Chiba** | 72.86 | 65.23 | 77.82 | 70.06 | 58.82 | 69.15 | 84.84 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Christchurch** | 88.57 | 83.26 | 82.76 | 62.77 | 69.90 | 89.84 | 95.99 |
| **Darfield** | 78.80 | 73.72 | 75.66 | 57.40 | 71.54 | 72.97 | 88.92 |
| **Denali** | 50.00 | 50.00 | 52.67 | 72.35 | 67.11 | 51.46 | 75.29 |
| **Duzce** | 50.00 | 70.03 | 50.00 | 50.00 | 75.80 | 80.05 | 84.80 |
| **Emilia** | 55.08 | 58.36 | 72.75 | 49.24 | 55.94 | 61.60 | 72.99 |
| **Haiti** | 89.99 | 90.26 | 89.51 | 89.72 | 93.12 | 97.82 | 97.36 |
| **Hokkaido** | 89.20 | 85.84 | 72.86 | 67.68 | 85.95 | 77.21 | 90.96 |
| **Honduras** | 78.88 | 78.98 | 86.74 | 76.52 | 76.70 | 78.50 | 92.14 |
| **Illapel** | 50.00 | 88.92 | 50.00 | 79.30 | 50.00 | 82.53 | 95.64 |
| **Kobe** | 81.80 | 82.68 | 88.06 | 86.86 | 80.62 | 88.59 | 92.40 |
| **Kocaeli** | 68.40 | 84.20 | 54.20 | 77.71 | 50.00 | 85.82 | 92.17 |
| **Kumamoto** | 83.73 | 81.77 | 66.06 | 74.63 | 76.02 | 79.89 | 85.89 |
| **Loma Prieta** | 86.59 | 87.67 | 93.13 | 50.00 | 85.53 | 87.89 | 94.82 |
| **Maule** | 58.12 | 57.18 | 54.13 | 55.09 | 54.54 | 50.00 | 58.58 |
| **Meinong** | 69.06 | 69.15 | 67.03 | 66.38 | 66.67 | 62.25 | 73.36 |
| **Miyagi ken** | 79.25 | 80.10 | 84.11 | 82.65 | 76.38 | 82.89 | 89.42 |
| **Muisne** | 70.27 | 82.79 | 67.74 | 89.11 | 87.71 | 53.92 | 89.65 |
| **Napa** | 78.10 | 79.74 | 85.58 | 79.93 | 74.45 | 79.93 | 88.69 |
| **Nepal (Gorkha)** | 84.60 | 59.74 | 50.00 | 78.88 | 69.29 | 50.00 | 90.50 |
| **Nigata 1964** | 81.85 | 82.29 | 81.82 | 82.50 | 80.53 | 83.46 | 89.60 |
| **Nigata 2004** | 82.28 | 80.64 | 56.77 | 62.49 | 80.43 | 80.87 | 85.35 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Nigata 2007** | 76.09 | 75.20 | 78.31 | 76.19 | 76.52 | 77.39 | 83.19 |
| **Nihonkai** | 87.70 | 89.32 | 70.31 | 85.05 | 87.19 | 88.25 | 93.86 |
| **Nisqually** | 75.93 | 77.02 | 87.09 | 88.04 | 74.16 | 63.34 | 91.67 |
| **Northridge** | 50.00 | 62.31 | 50.00 | 54.79 | 50.00 | 56.80 | 62.42 |
| **Oklahoma** | 71.55 | 59.79 | 50.00 | 58.58 | 93.84 | 50.00 | 86.47 |
| **Pisco** | 50.00 | 79.32 | 50.00 | 74.53 | 50.00 | 59.15 | 80.06 |
| **Puget Sound 1949** | 77.85 | 74.23 | 66.27 | 62.51 | 74.82 | 73.93 | 81.62 |
| **Puget Sound 1965** | 61.61 | 60.31 | 59.90 | 59.82 | 58.88 | 64.29 | 65.18 |
| **Samara** | 88.74 | 82.32 | 92.85 | 88.97 | 89.14 | 92.10 | 95.49 |
| **San Simeon** | 50.00 | 95.11 | 55.26 | 71.24 | 73.46 | 97.33 | 98.77 |
| **Tecoman** | 74.57 | 80.60 | 74.22 | 63.69 | 72.79 | 76.57 | 81.84 |
| **Telire Limon** | 77.69 | 76.95 | 64.97 | 77.31 | 77.94 | 60.00 | 81.48 |
| **Tohoku** | 72.96 | 67.93 | 70.18 | 69.59 | 70.21 | 76.21 | 86.86 |
| **Tokachi** | 78.15 | 77.21 | 85.65 | 83.21 | 78.12 | 83.98 | 88.88 |
| **Tottori** | 90.67 | 90.87 | 94.87 | 91.99 | 91.83 | 94.12 | 96.65 |
| **Wenchuan** | 50.00 | 50.00 | 50.00 | 52.60 | 63.79 | 50.00 | 65.41 |

As presented in Table 6, 4 out of the 6 non-liquefaction events resulted in 100% accurate predictions, followed by a negligible error of 0.22 % for the 2016 Central Italy event. However, the 2008 Iwate earthquake is predicted to have extensive liquefaction (10.57 % of tested samples). Rashidian and Baise (2020) have also reported that the performance for the 6.9 Mw Iwate earthquake in Japan was inconsistent with no

liquefaction observation (although lateral spreads were observed). Figure 12 shows predicted maps of the two non-liquefaction earthquakes with predicted liquefaction.

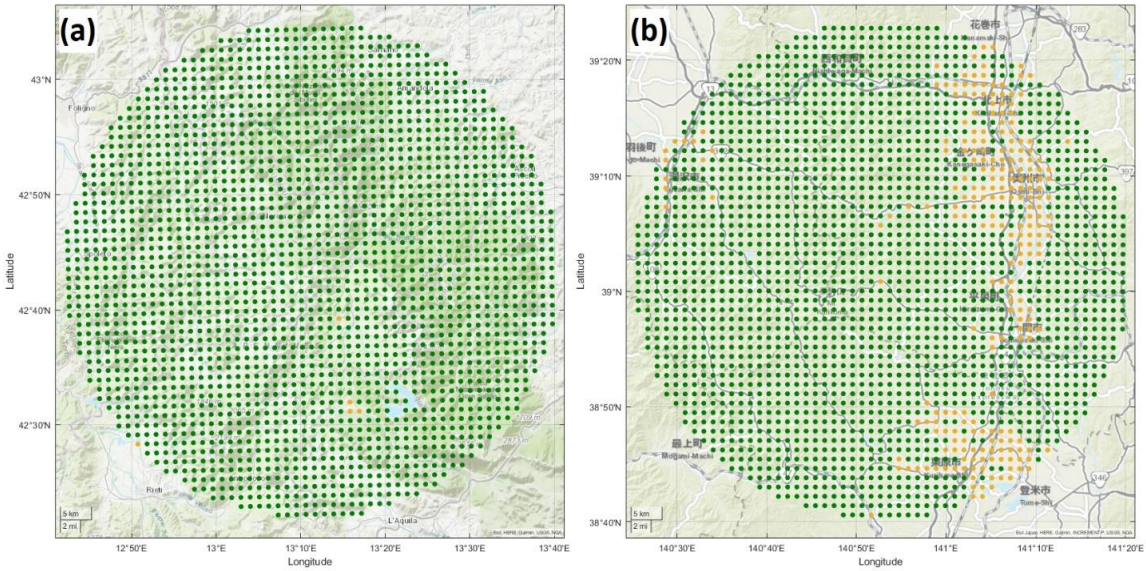



Figure 5-12. a) Predicted hazard map of the 2016 Central Italy; and b) 2008 Iwate earthquakes. Green is indicator of correctly predicted non-liquefaction. Yellow is indicator of wrongly predicted liquefaction.

The 2015 Nepal (Gorkha) Earthquake has been discussed in several studies of geospatial liquefaction modeling (Rashidian and Baise, 2020; Moss et al., 2015; Todorivic and Silva 2022). Rashidian and Baise (2020) noticed over-prediction of liquefaction due to the existence of a large flat area with low Vs30 values. Todorovic and Silva (2022) also observed lower accuracy of their model for the M 7.5 Gorkha earthquake and attributed the misfit to seasonal variations in the groundwater table or underestimation/overestimation of soil stiffness (like Moss et al. (2017) and Pokhrel et al. (2022)). They suggested the implementation of a seasonal element to inform the model about variation in depth through time (Todorovic and Silva, 2022). Moss et al. (2015) also studied the discrepancy between some model predictions and field observations of liquefaction by the GEER reconnaissance

teams (Hashash et al., 2015; Moss et al., 2015), which resulted in extensive ground failures of different types (landslides, rockfalls, liquefaction, etc.). Liquefaction spatial extent in the Kathmandu Valley (an ancient lakebed now dominated by river systems) was limited, while model predictions (UNDP, 1994; JICA, 2002; Piya, 2004), and historical data from the region reported higher liquefaction impacts. Their study found that pumping has depressed the groundwater table significantly in the last few decades. Additionally, high near-surface shear wave velocity was the other reason for over-prediction of liquefaction, since the Vs30 values were higher than average liquefaction sites around the world (Moss et al., 2015).

Figure 13 shows the map of liquefaction for the Nepal event by the voting system method proposed in this study, with overall accuracy of 94.81 % and 41.67 % liquefaction detection. An interesting observation after performing a feature importance analysis for this event, by removing one feature at a time, showed that there is a single critical feature which helps detecting liquefaction hazard areas on this event, which is the elevation of the nearest waterbody (WBE). Removal of this feature from the feature set, resulted in zero sensitivity (liquefaction record detection), compared to the 41.67 % when the feature is used (Table 6). It should be noted that WBE has the lowest weight among the features used, as shown via Table 4, and this result highlights the importance of using all relevant features in the liquefaction models.

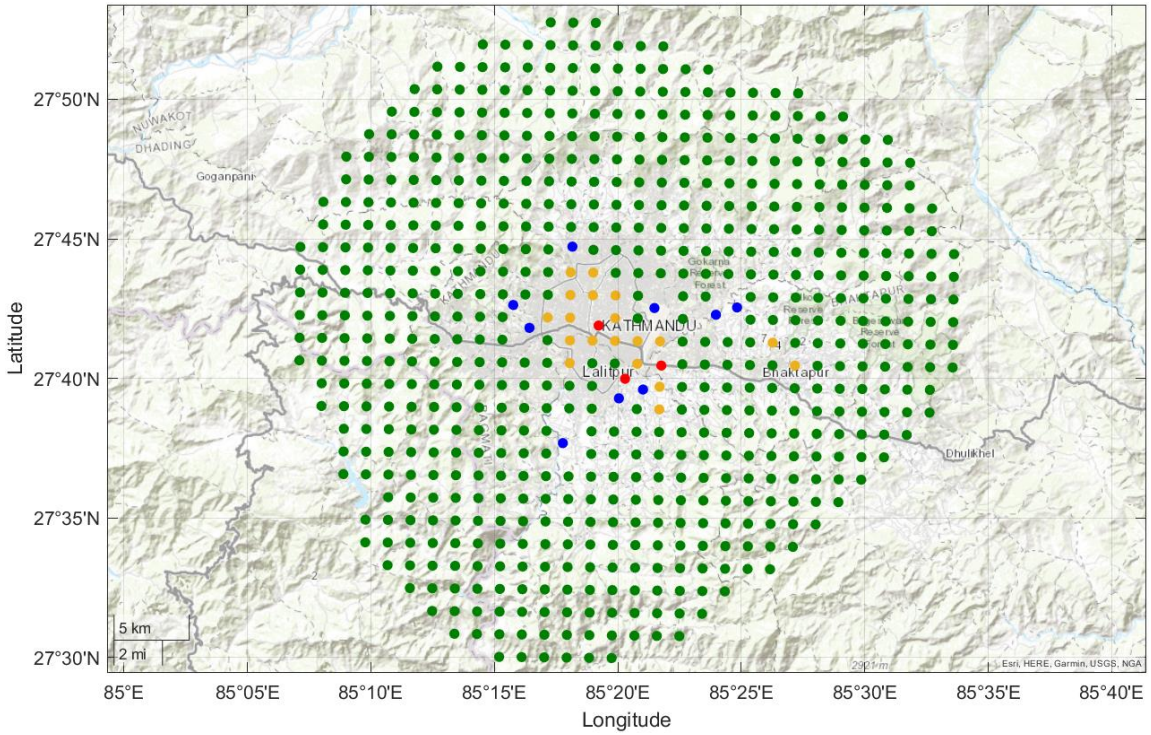



Figure 5-13. Predicted hazard map of the 2015 Nepal (Gorkha) earthquake with the voting classification system. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

The 9.1 Mw Tohoku earthquake in Japan is the largest-magnitude earthquake tested in this study, leading to moderately good accuracy (AUC = 86.86 %) with high liquefaction detection (95.06 % sensitivity), according to Table 6. Some over-prediction of liquefaction is observed for this event, but mostly in the areas where liquefaction has occurred in the vicinity. Zhu et al. (2017) also reported that for earthquakes with very large magnitudes, such as the 2011 Tohoku earthquake, their model predicts larger areas of high liquefaction probabilities than the area where liquefaction was observed. Figure 14 shows the visualized prediction map of the Tohoku event, in addition to another five major events in Japan, including the 1995 Kobe and the 1983 Nihonkai earthquakes in Japan, as a few examples significantly impacted by liquefaction.

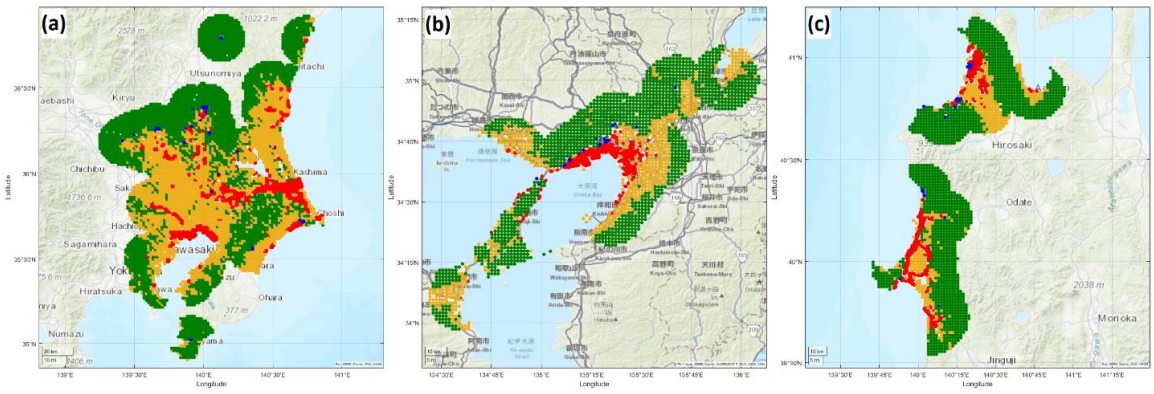



Figure 5-14. a) Predicted hazard map of the 2011 Tohoku; b) 1995 Kobe; and c) 1983 Nihonkai earthquakes. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

Figure 15 plots the resulting maps of the 6 classifiers (3 global, 2 coastal, and 1 regional), and the final map of the 1989 Loma Prieta earthquake in Northern California, derived by the voting system. It is observed that the final model provides the best AUC value compared to all 6 individual classifiers. In addition, it is among the best in terms of detecting liquefaction points. The voting system's prediction (Figure 15-g) has also balanced the over-prediction (Figures 15-a, 15-b, 15-e, and 15-f) and under-prediction (Figures 15-c and 15-d) of the individual models.

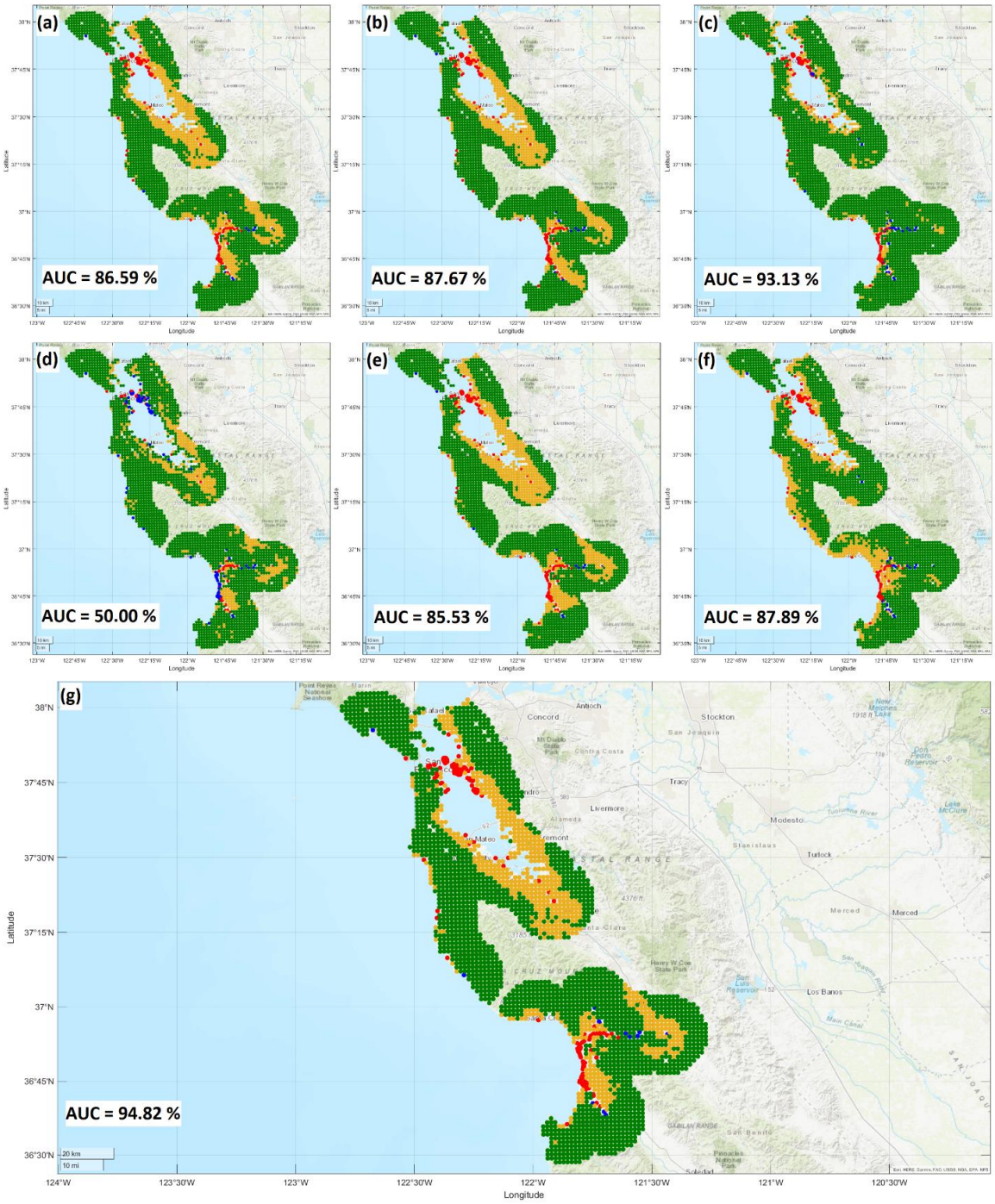



Figure 5-15. Predicted hazard map of the 1989 Loma Prieta earthquake by: a) Global classifier 1; b) Global classifier 2; c) Global classifier 3; d) Coastal classifier 1; e) Coastal classifier 2; f) Regional classifier; and g) System of voting classifiers. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

The three earthquakes which have occurred in Nigata, Japan in 1964 (7.6 M), 2004 (6.6 M), and 2007 (6.6 M), were analyzed as a part of the model validation. Based on the field reconnaissance reports, the 7.6 M Nigata 1964 earthquake event has resulted in more extensive and widespread liquefaction. The difference between this event and the other two is only in the earthquake shaking parameters' values (magnitude and PGV) since the values for the other geospatial variables are the same. Figure 16 provides the model predictions for these three events, which shows that for the 1964 event, the detection of hazardous areas in the coastal area is very high, although the model could just partially recognize the farther part from the coast as liquefaction-prone (Figure 16-a). Among the three events, the 1964 event showed the highest accuracy (AUC = 89.60 %).

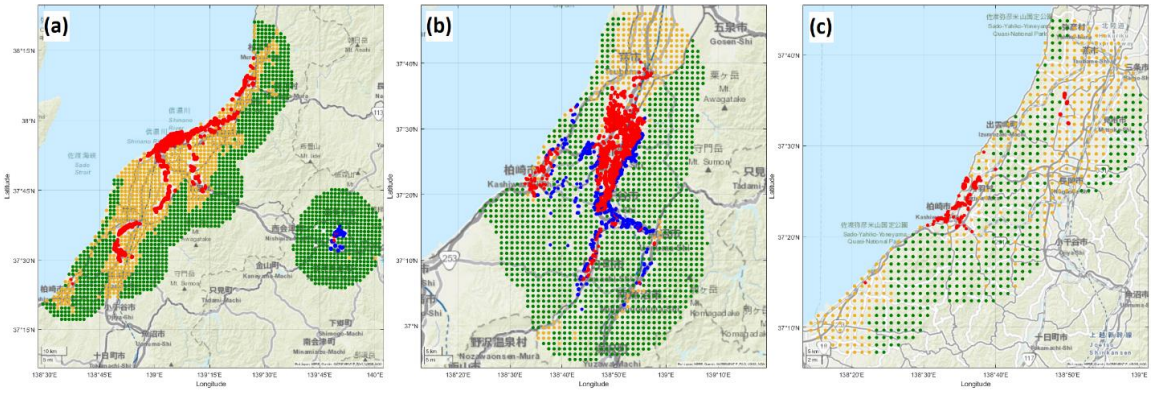



Figure 5-16. Predicted hazard map of the three earthquakes which have occurred in Nigata, Japan, in: a) 1964; b) 2004; and c) 2007. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

There are a few cases where the performance of an individual classifier might be best (better than the voting system's prediction). However, that does not happen consistently for a specific classifier through all such events, and for any specific event, either of the classifiers might be the best. For instance, as shown via Figure 17, the

prediction by the global classifier 2 is considerably better than all other 5 classifiers, leading to comparatively lower accuracy of the voting system, which appears to under-predict liquefaction in the Granada Hills area on the north of the San Fernando Valley where many ground failures were observed, matching the low probabilities predicted by Zhu et al. (2017) model, because of the relatively deep water-table depth (>10 m). The observed ground failures in the area might be a result of dynamic ground compaction of loose unsaturated surface material, not liquefaction (Stewart et al., 1994). Although, Zhu et al. (2017) have reported over-prediction in the southern part, while that is not the case here with the voting system, whereas the global classifier 2 over-predicts the same area reported by them (Figure 17).

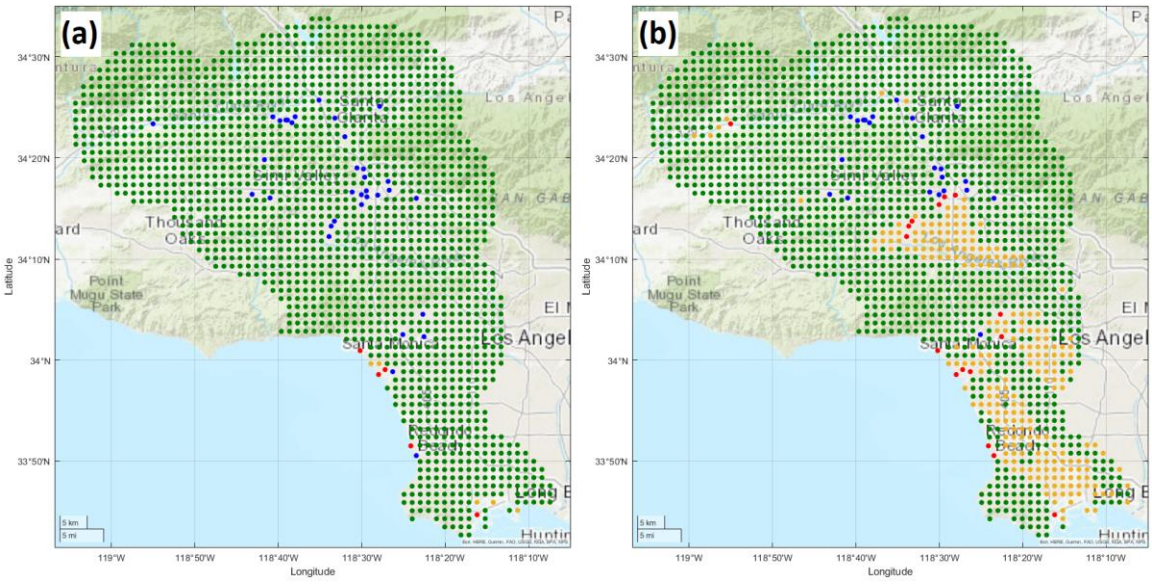



Figure 5-17. Predicted hazard map of the 1994 Northridge earthquake by: a) System of voting classifiers; and b) Global classifier 2. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

The calculated average AUC value of the non-coastal models provided in Table 7 is 64.91 %, whereas the average AUC for the coastal models is 73.10 %. The accuracy of the coastal model should be considered more important since most tectonically active regions are near coastlines. Since the proposed model is built via a relatively low-spatial-resolution database, the model may fail to capture more localized, small-scale features of liquefaction. The problem is also present in other geospatial models developed before (Rashidian and Baise, 2020; Geyin et al., 2022). The spatial resolution of the global model output proposed in this study, like other related studies, is relatively low considering the extremely local nature of the earthquake-induced soil liquefaction manifestations.

There is also some uncertainty regarding the regional performance of such models in under-represented regions worldwide since the database does not have a balanced global coverage. It should also be noted that the investigation accuracy, the covered area of investigation, and the datasets from which the values for explanatory variables were extracted should not be considered perfect. Inherently, the accuracy of liquefaction predictions is sensitive to the accuracy of model inputs, although this study did not consider uncertainties associated with measurements and input parameters' modeling.

For further visual investigation of the model performance, Figure 18 provides the visualizations of liquefaction hazard maps for an additional 6 earthquake inventories tested using the proposed method.

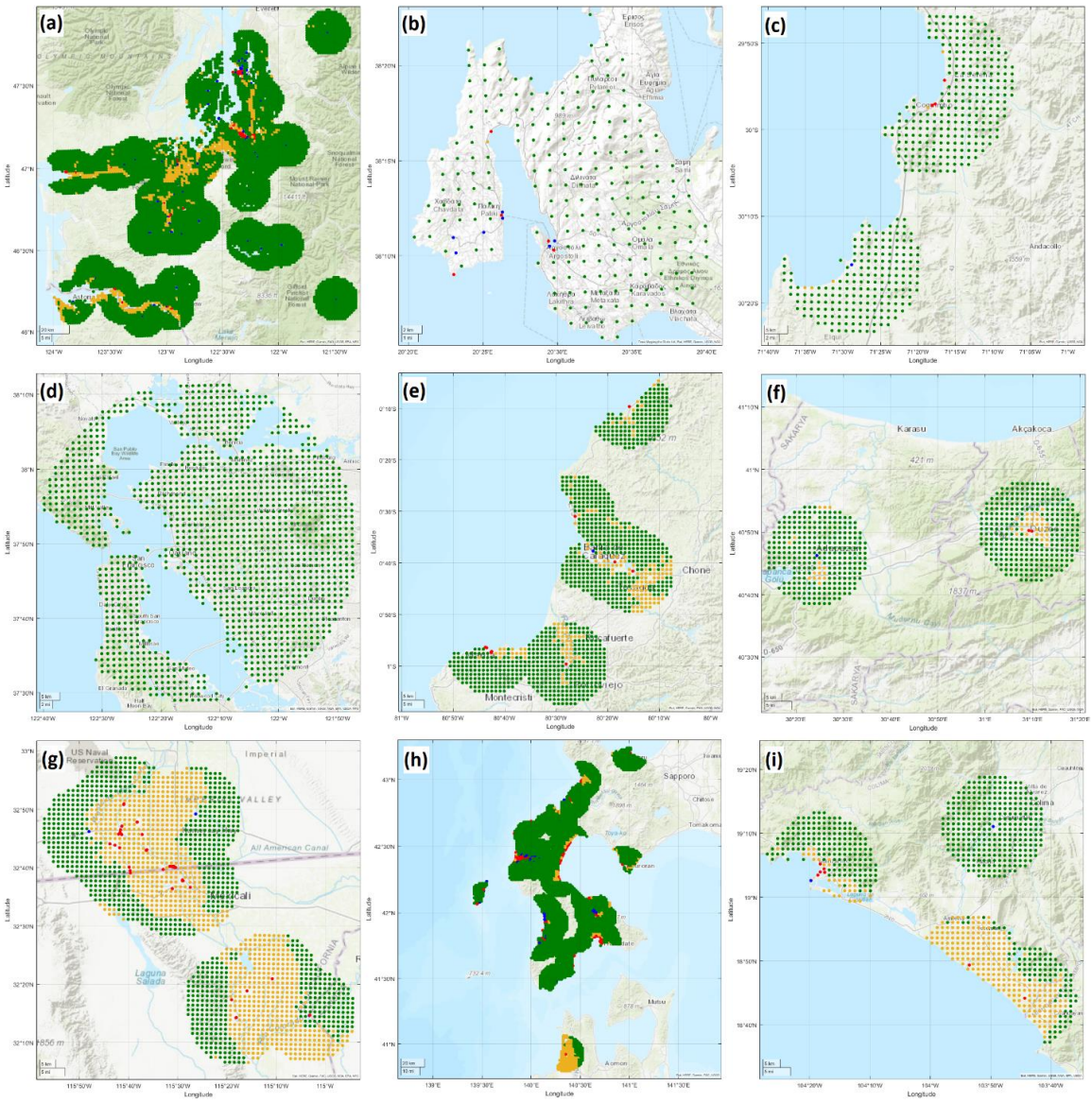



**Figure 18.** Predicted hazard maps of: a) 1949 Puget Sound; b) 2014 Cephalonia; c) 2015 Illapel; d) 2015 Piedmont; e) 2016 Muisne; f) 1999 Duzce; g) 2010 Baja California; h) 1993 Hokkaido; and i) 2003 Tecoman earthquakes. Red is correctly predicted liquefaction. Green is correctly predicted non-liquefaction. Yellow is wrongly predicted liquefaction. Blue is wrongly predicted non-liquefaction.

#### *5.6.4.1 Comparative Analysis*

Finally, it should be noted that the leave-one-out testing approach, results in lower accuracies but less bias than (Zhu et al., 2017; Baise et al., 2021) when the data of all earthquakes are partitioned into training and testing via hold-out partitioning (Todorovic and Silva, 2022) due to overfitting. To provide a fair comparative analysis, a hold-out partitioning approach is also used, and the accuracy (AUC) is compared with the USGS preferred model (Zhu et al., 2017) via Table 8. As shown in the table, for all the earthquake events reported in Zhu et al. (2017) study, the accuracy of the proposed method in this study, with hold-out partitioning testing approach, is significantly higher. However, with the leave-one-out testing, the AUC of one event (Chi-Chi earthquake) is lower than the AUC reported by Zhu et al. (2017), and for the other events, the values are moderated and look more realistic. The results of the proposed method even using the leave-one-out approach significantly outperform the Zhu et al. (2017) model, with 80 % of the earthquakes in Table 8 having AUC of higher than 0.8, compared to only 20 % reported by Zhu et al. (2017). This study highly recommends that geospatial liquefaction models should be developed based on the leave-one-out accuracy assessments, not data partitioning. The observations of this study, especially for events such as the 2015 Nepal and 1999 Chi-Chi earthquakes, indicate that validation strategies in previous research studies can be subject to uncertainties and bias, and the applicability of such models for new events should be re-evaluated and further investigated.

Table 5-8. Comparative accuracy statistics. All values are AUC. The AUC values for the Zhu et al. (2017) study are presented as reported in their original paper, although the datasets are not completely similar.

| Earthquake | Zhu et al. (2017) Model 1<br><br>(Coastal) | Zhu et al. (2017) Model 2<br><br>(Non-Coastal) | Proposed Method<br><br>(Hold-Out Partitioning) | Proposed Method<br><br>(Leave-One-Out Testing) |
|---|---|---|---|---|
| **Loma Prieta** | 0.886 | 0.848 | 0.985 | 0.948 |
| **Kobe** | 0.862 | 0.821 | 0.971 | 0.942 |
| **Tottori** | 0.806 | 0.775 | 0.982 | 0.967 |
| **Christchurch** | 0.801 | 0.770 | 0.977 | 0.960 |
| **Miyagi ken** | 0.791 | 0.801 | 0.944 | 0.894 |
| **San Simeon** | 0.757 | 0.775 | 0.996 | 0.988 |
| **Darfield** | 0.725 | 0.692 | 0.942 | 0.889 |
| **Nisqually** | 0.703 | 0.751 | 0.947 | 0.917 |
| **Nigata 2004** | 0.678 | 0.700 | 0.964 | 0.854 |
| **Tohoku** | 0.677 | 0.656 | 0.904 | 0.867 |
| **Hokkaido** | 0.673 | 0.720 | 0.973 | 0.910 |
| **Nigata 1964** | 0.667 | 0.628 | 0.932 | 0.896 |
| **Nihonkai** | 0.643 | 0.649 | 0.964 | 0.939 |
| **Chiba** | 0.608 | 0.657 | 0.901 | 0.848 |
| **Chi-Chi** | 0.603 | 0.660 | 0.950 | 0.500 |
| **Puget Sound 1949** | 0.601 | 0.610 | 0.950 | 0.816 |
| **Tokachi** | 0.561 | 0.571 | 0.943 | 0.889 |
| **Puget Sound 1965** | 0.558 | 0.548 | 0.886 | 0.652 |
| **Northridge** | 0.535 | 0.547 | 0.932 | 0.624 |
| **Wenchuan** | 0.527 | 0.551 | 0.929 | 0.654 |

## 5.7 Conclusions and Future Works

This study aimed at developing a ML geospatial liquefaction model. Such geospatial liquefaction models are very useful for global implementation, especially when local geotechnical data and surficial geologic maps are not available. Effort was spent to balance the dataset for liquefaction model development in a manner to exploit the data as much as possible. The class and event imbalance issue are treated in an innovative manner by under-sampling large events and distributing the datasets over several balanced subsets. Instead of training a single classifier, a system of voting machine learning classifiers is designed to achieve higher accuracy and certainty, using different classification techniques to balance the performance, as not a single classification technique can be considered the best for all earthquake events.

The results showed superior performance compared to previously published logistic regression-based liquefaction probability models. The use of several relevant features in this study, through advanced machine learning algorithms was found to be helpful in detecting liquefaction-prone regions. The leave-one-out testing approach based on the total exclusion of individual earthquake events from the training data, revealed the uncertainties associated with accuracy assessments reported in the literature for liquefaction susceptibility models. This study highly recommends that geospatial liquefaction models should be developed based on the leave-one-out accuracy assessments, not data partitioning.

An issue with any advanced machine learning paradigm is the lack of a defined analytical expression that can be provided for further use. Simple interpretations of model structure are also generally impossible. However, trained AI-based models should be delivered as codes or software, in a manner that is accessible to a broad community of users and researchers for further implementation of the proposed scheme.

Several organizations would benefit from the outcomes of this study, such as civil protection organizations, insurance companies, extreme events first responders, post-event reconnaissance teams, urban planners, and infrastructure constructors. The accuracy of the global geospatial models at the local scale is not very high, but the availability of large-scale susceptibility maps of liquefaction hazard can assist policymakers and infrastructure administrators in prioritizing territories where should be investigated further to assess the seismic risk of structures and infrastructures associated with the occurrence of earthquake-induced soil liquefaction.

## 5.8 Acknowledgements

This work was funded by the U.S. Geological Survey (USGS), U.S. Department of the Interior, under USGS Award Number G22AS00006 FY2022, titled “Innovative Data-Driven Frameworks for Geospatial Ground Failure Models”. This support is gratefully acknowledged by the authors of this paper.

## 5.9 Data and Codes Availability

The MATLAB codes and the inventory database of earthquake-induced soil liquefaction samples alongside the data of geospatial parameters, leading to the results published in this study are provided at the GitHub repository of the first author at https://github.com/adel-

asadi/Global_Geospatial_Liquefaction_Model. Since all these geospatial layers are freely available for download from their referenced sources in Table A-2, every researcher can use the data and the proposed approach to reproduce the presented liquefaction hazard maps, or to generate liquefaction hazard maps for other/future events.

## 5.10 Appendix

Table 5-9. Information on the earthquake events considered in this study.

| No. | Earthquake | M/D/Y | Mw | Coastal | Liq. | Region | Reference |
|---|---|---|---|---|---|---|---|
| 1 | Achaia | 06/08/2008 | 6.5 | 1 | 1 | Europe | Pavlides (2013) |
| 2 | Aquila | 04/06/2009 | 6.3 | 0 | 1 | Europe | Monaco (2011) |
| 3 | Arequipa | 06/23/2001 | 8.4 | 1 | 1 | South America | Gomez et al. (2005) |
| 4 | Baja California | 04/04/2010 | 7.2 | 0 | 1 | North America | Stewart and Brandenberg (2010) |
| 5 | Cephalonia | 01/26/2014 | 6.1 | 1 | 1 | Europe | Papathanassiou (2016) |
| 6 | Chi-Chi | 09/21/1999 | 7.7 | 1 | 1 | Asia & Oceania | Yuan et al. (2004) |
| 7 | Chiba | 12/17/1987 | 6.7 | 1 | 1 | Japan | Wakamatsu (2011) |
| 8 | Christchurch | 02/22/2011 | 6.1 | 1 | 1 | Asia & Oceania | Cubrinovski et al. (2011) |
| 9 | Darfield | 09/04/2010 | 7.0 | 1 | 1 | Asia & Oceania | Cubrinovski et al. (2010) |
| 10 | Denali | 11/03/2002 | 7.9 | 0 | 1 | North America | Kayen et al. (2004) |
| 11 | Duzce | 11/12/1999 | 7.2 | 0 | 1 | Europe | Sucuoğlu (2000) |
| 12 | Emilia | 05/20/2012 | 6.0 | 0 | 1 | Europe | Papathanassiou (2012) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 13 | Haiti | 01/12/2010 | 7.0 | 1 | 1 | Central America | Olson et al. (2011) |
| 14 | Hokkaido | 07/12/1993 | 7.7 | 1 | 1 | Japan | Miura et al. (1995) |
| 15 | Honduras | 05/28/2009 | 7.3 | 1 | 1 | Central America | Luna (2010) |
| 16 | Illapel | 09/16/2015 | 8.3 | 1 | 1 | South America | Candia et al. (2015) |
| 17 | Iquique | 04/01/2011 | 8.2 | 1 | 1 | South America | Rollins et al. (2014) |
| 18 | Kobe | 01/17/1995 | 6.9 | 1 | 1 | Japan | Hamada et al. (1996) |
| 19 | Kocaeli | 08/17/1999 | 7.6 | 1 | 1 | Europe | Rathje et al. (2004) |
| 20 | Kumamoto | 04/14/2016 | 6.2 | 1 | 1 | Japan | Mukunoki et al. (2016) |
| 21 | Loma Prieta | 10/17/1989 | 6.9 | 1 | 1 | North America | Holzer (1998) |
| 22 | Maule | 02/27/2010 | 8.8 | 1 | 1 | South America | Verdugo (2011) |
| 23 | Meinong | 02/06/2016 | 6.3 | 1 | 1 | Asia & Oceania | Sun et al. (2016) |
| 24 | Miyagi ken | 06/12/1978 | 7.7 | 1 | 1 | Japan | Tohno and Yasuda (1981) |
| 25 | Muisne | 04/16/2016 | 7.8 | 1 | 1 | South America | Nikolau et al. (2016) |
| 26 | Napa | 08/24/2014 | 6.0 | 1 | 1 | North America | Bray et al. (2014) |
| 27 | Nepal (Gorkha) | 04/25/2015 | 7.8 | 0 | 1 | Asia & Oceania | Hashash (2015) |
| 28 | Nigata 1964 | 06/16/1964 | 7.6 | 1 | 1 | Japan | Ishihara and Koga (1981) |
| 29 | Nigata 2004 | 10/23/2004 | 6.6 | 1 | 1 | Japan | Rathji et al. (2006) |
| 30 | Nigata 2007 | 07/16/2007 | 6.6 | 1 | 1 | Japan | Kayen et al. (2007) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 31 | Nihonkai | 05/26/1983 | 7.7 | 1 | 1 | Japan | Tohno and Shmoto (1985) |
| 32 | Nisqually | 02/28/2001 | 6.8 | 1 | 1 | North America | Bray et al. (2001) |
| 33 | Northridge | 01/17/1994 | 6.7 | 0 | 1 | North America | Stewart et al. (1994) |
| 34 | Oklahoma | 09/03/2016 | 5.8 | 0 | 1 | North America | Clayton et al. (2016) |
| 35 | Pisco | 08/15/2007 | 7.9 | 1 | 1 | South America | Taucer et al. (2009) |
| 36 | Puget Sound 1965 | 04/29/1965 | 6.7 | 1 | 1 | North America | Chleborad and Schuster (1990) |
| 37 | Puget Sound 1949 | 04/13/1949 | 7.1 | 1 | 1 | North America | Chleboard and Schuster<br><br>(1990) |
| 38 | Samara | 09/05/2012 | 7.6 | 1 | 1 | Central America | Rollins et al. (2013) |
| 39 | San Simeon | 12/22/2003 | 6.5 | 1 | 1 | North America | Holzer et al. (2005) |
| 40 | Tecoman | 01/22/2003 | 7.5 | 0 | 1 | North America | Wartman et al. (2005) |
| 41 | Telire Limon | 04/22/1991 | 7.6 | 0 | 1 | Central America | Yasuda et al. (1993) |
| 42 | Tohoku | 03/11/2011 | 9.1 | 1 | 1 | Japan | Bhattacharya et al. (2011)<br><br>& Goto et al. (2012) |
| 43 | Tokachi | 09/26/2003 | 8.3 | 1 | 1 | Japan | Wakamatsu (2011) |
| 44 | Tottori | 10/6/2000 | 6.7 | 1 | 1 | Japan | Kiyono et al. (2007) |
| 45 | VanTab | 10/23/2011 | 7.1 | 0 | 1 | Europe | Erdik (2012) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 46 | Virginia | 08/23/2011 | 5.8 | 0 | 1 | North America | Carter and Maurer (2011) |
| 47 | Wenchuan | 05/12/2008 | 6.3 | 0 | 1 | Asia & Oceania | Huang and Jiang (2010) |
| 48 | Central Italy | 08/24/2016 | 6.2 | 0 | 0 | Europe | Zimmaro and Stewart (2016) |
| 49 | Chino Hills | 07/29/2008 | 5.4 | 1 | 0 | North America | Hauksson et al. (2008) |
| 50 | Hector Mine | 10/16/1999 | 7.1 | 0 | 0 | North America | USGS (2000) |
| 51 | Iwate | 06/14/2008 | 6.9 | 0 | 0 | Japan | Kayen et al. (2008) |
| 52 | Piedmont | 08/17/2015 | 4.0 | 1 | 0 | North America | USGS (2015) |
| 53 | Yountville | 09/03/2000 | 4.9 | 1 | 0 | North America | Miranda et al. (2014) |

Table 5-10. Complementary information about the geospatial variables used in this study.

| Geospatial Variable | Definition / Information / Source / Reference |
|---|---|
| $Vs_{30}$ | The weighted-mean shear-wave velocity in the top 30 m ($Vs_{30}$), which was adopted as a proxy of soil stiffness since soft sandy soils are more susceptible to liquefaction (they are looser). The US Geological Survey provided the global topographic slope-based $Vs_{30}$ map.<br><br>US Geological Survey $Vs_{30}$ map: https://earthquake.usgs.gov/data/vs30/.<br><br>The global Vs30 layer has been estimated according to the method described in Wald and Allen (2007), which is based on elevation-derived slope. The $Vs_{30}$ layer used in the development of GGLM15 (Zhu et al., 2015) was based on GTOPO30. In GGLM17 (Zhu et al., 2017), $Vs_{30}$ was updated by deriving it from the GMTED, a global DEM with enhanced quality. The GTOPO30 was initially developed in 1996 by the USGS, and the GMTED10 was recently developed by the USGS and NGA (National Geospatial |

| | |
|---|---|
| | Intelligence Agency) collaboratively to replace GTOPO30 as the elevation dataset for global application (Danielson and Gesch 2010). |
| Roughness | Roughness is the largest intercell difference between a central pixel and its eight surrounding cells. |
| CTI | As a proxy for soil saturation, the compound topographic index (CTI) is a measure of the relative propensity for the soil to become saturated to the surface. The compound topographic index (CTI) layer was downloaded from the USGS Hydrologic Derivatives for Modeling and Analysis (HDMA), which has a resolution of 3 arc-seconds in most regions. |
| PGA | The ShakeMap provides PGA and PGV estimates for each event. ESRI Raster Files were downloaded from the USGS ShakeMap version 4, with resolutions of 30 arc-seconds. |
| PGV | The ShakeMap provides PGA and PGV estimates for each event. ESRI Raster Files were downloaded from the USGS ShakeMap version 4, with resolutions of 30 arc-seconds. |
| AI | The aridity index (AI) is derived from the Consultative Group on International Agriculture Research-Consortium for Spatial Information (CGIAR-CSI) Global Aridity dataset (Global Aridity and PET dataset: http://www.cgiar-csi.org/data/global-aridity-andpet-database). |
| Elevation | GMTED10 is used as the DEM. The Digital Elevation Model (DEM) for the globe has been accessed through Global Multi-resolution Terrain Elevation Data (Danielson and Gesch, 2010). The Digital Elevation Model (DEM) for the globe (Elevation) was downloaded through Global Multi-resolution Terrain Elevation Data (2010), at a resolution of 7.5 arc-seconds. |
| DR | The distance to river (and lake) layer has been computed globally by using the hydrological data from SHuttle Elevation Derivatives at multiple Scale (Hydro-SHEDS) which is based on topography (Lehner et al. 2008). The development of this database includes a sequence of extensive hydrologic conditioning procedures to incorporate the locations of known rivers and lakes. The river network provided in this database is in vector format. Because distance computation using vector data at a global scale is inefficient, we derived a raster river network using the same threshold (100 |

| | |
|---|---|
| | upstream cells) and perform a raster distance calculation. Distance is in number of pixels and is a Cartesian distance.<br><br>For the 2010 Haiti earthquake only, a local dataset was used to calculate the distance to river: https://haiti.caribbeangeoportal.com/datasets/caribgeoportal::haiti-rivers/about. |
| DC | The distance to coast layer has been computed from a global dataset by National Aeronautics and Space Administration (NASA)'s Ocean Color Group. This data set was generated with the Generic Mapping Tools (GMT) software using its intermediate-resolution coastline. It was first computed at a spatial resolution of 0.04 degrees and then interpolated to 0.01 degrees. In the original dataset, negative distances represent locations over land (including land-locked bodies of water), while positive distances represent the ocean. For our application, we further process the data so that locations on land have positive distances and on ocean have zero distance. |
| WTD | Simulated water table depth is from a global dataset by Fan et al (2013). The authors simulate groundwater flow at 30 arc-seconds spacing using a model constrained by climate, terrain and sea level. Hydraulic conductivity of soil is estimated from a soil map and assuming a steady decay over depth. They digitized over 1.5 million published records of water table depths. The results have been adjusted using actual water-table depths as a means of calibration across climate zones on all inhabited continents. This model predicts natural patterns of water table depth and so does not account for any pumping or irrigation carried out by humans. The data is at a spatial resolution of 30 arc-seconds. |
| DWB | The distance to the nearest water body layer is defined as the minimum of distance to coast and distance to river and lake layers. |
| Precipitation | The global precipitation layer has been created by bilinearly interpolating more than 40000 stations across the world averaging over 1959-2000 from WorldClim database (https://www.worldclim.org/). Hijmans et al. (2005) interpolated average monthly precipitation from weather stations across the world (47,554 locations) on a 30 arc-second resolution grid and averaged over the 1959- 2000 time periods Getting the average value for decades of precipitation will give a better pattern of precipitation in a specific region. |

| | |
|---|---|
| Landform | The global ALOS Landforms dataset provides landform classes created by combining the Continuous Heat-Insolation Load Index (ALOS CHILI) and the multi-scale Topographic Position Index (ALOS mTPI) datasets. It is based on the 30m "AVE" band of JAXA's ALOS DEM (available in EE as JAXA/ALOS/AW3D30_V1_1). The Conservation Science Partners (CSP) Ecologically Relevant Geomorphology (ERGo) Datasets, Landforms and Physiography contain detailed, multi-scale data on landforms and physiographic (aka land facet) patterns (Theobald et al., 2015). Data downloaded from: https://developers.google.com/earth-engine/datasets/catalog/CSP_ERGo_1_0_Global_ALOS_landforms#description. |
| HAND | Global 30m Height Above the Nearest Drainage (HAND) dataset with 30 m cell resolution and 1000 number of river head threshold cells is used as a variable (Donchyts et al., 2016). HAND is a digital elevation model normalized using the nearest drainage is used for hydrological and more general-purpose applications, such as hazard mapping, landform classification, and remote sensing. Downloaded from: https://gena.users.earthengine.app/view/global-hand. |
| DR2 | Free-Flowing Rivers (FFRs) data is used to calculate distance to rivers (Grill et al., 2019). Only rivers with high levels of connectivity (CSI ≥ 95%) throughout their entire length are considered FFRs by Grill et al. (2019). Four dimensions (connectivity dimensions) are considered to determine the CSI of river reaches: (1) longitudinal (connectivity between up- and downstream), (2) lateral (connectivity to floodplain and riparian areas), (3) vertical (connectivity to groundwater and atmosphere) and (4) temporal (connectivity based on seasonality of flows). |
| Soil Thickness | For the model development, Pelletier et al. (2016) explicitly mapped global landform types and then estimated the thicknesses of sedimentary deposit using models optimized for each landform type and the best available data for topography, climate, and geology as input. The sedimentary deposit thickness is within the range of 0-50 m, and areas with predicted sedimentary thickness greater than 50 m are assigned a value of 50 m. The dataset is archived at the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC) accessible at: https://daac.ornl.gov/SOILS/guides/Global_Soil_Regolith_Sediment.html. |

| | |
|---|---|
| Slope | Data was derived from the DEM using the gdaldem command in the Geospatial Data Abstraction Library in 7.5 arc-second resolutions. |
| Aspect | Derived from DEM layer using GIS tools. |
| Landcover | The landform type (uplands/lowlands) and sedimentary thickness data are from a high-resolution gridded global data set developed by a multidisciplinary team at the University of Arizona (Pelletier et al., 2016). The data set is developed for regional and global land surface modeling and has a spatial resolution of 30 arcsec (~1 km). For the model development, Pelletier et al. (2016) explicitly mapped global landform types and then estimated the thicknesses of sedimentary deposit using models optimized for each landform type and the best available data for topography, climate, and geology as input. The sedimentary deposit thickness is within the range of 0-50 m, and areas with predicted sedimentary thickness greater than 50 m are assigned a value of 50 m. The dataset is archived at the Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC) accessible at: https://daac.ornl.gov/SOILS/guides/Global_Soil_Regolith_Sediment.html. |
| TPI | TPI is the difference between a central pixel and the mean of its eight surrounding cells. Data was derived from the DEM using the gdaldem command in the Geospatial Data Abstraction Library in 7.5 arc-second resolutions. |
| TRI | TRI is a proxy used to differentiate soil and rock layers and was calculated as the square root of the elevation differences between the central cell and the neighboring eight cells (Wilson et al., 2007). Data was derived from the DEM using the gdaldem command in the Geospatial Data Abstraction Library in 7.5 arc-second resolutions. |
| Magnitude | Historical USGS reports and records in the literature (Table A-1) |
| WBE | Elevation of the nearest waterbody. |
| DWB2 | Distance to the closest waterbody. |
| DL | The distance to lake layer has been computed from a global dataset by National Aeronautics and Space Administration (NASA)'s Ocean Color Group. This data set was generated with the Generic Mapping Tools (GMT) software using its intermediate-resolution coastline. It was first computed at |

| | |
|---|---|
| | a spatial resolution of 0.04 degrees and then interpolated to 0.01 degrees. In the original dataset, negative distances represent locations over land (including land-locked bodies of water), while positive distances represent the ocean. For our application, we further process the data so that locations on land have positive distances and on ocean have zero distance. |

# 6 Slope Stability and Monitoring in Open Pit Mining Industry

Because of factors such as geology, weather, seismic activity, and mining techniques, open-pit mining operations are at danger of slope stability. Stability issues necessitate a quick response to keep mining operations running, prevent further damage, and keep employees safe. In this section, geomechanically crucial considerations in pit design are introduced, and slope monitoring techniques used to predict slope failure are presented.

## 6.1 Pit Slope Stability

Unexpected rock movement, whether beneath or on the surface, can cause casualties and potentially end in a mine disaster. Slope failure can have disastrous implications even in comparatively shallow strip mines and quarries. Open-pit mines are deeper than ever before, and because assessing rock mass strength at such huge scales is difficult, major slope failure can occur with little warning. Slope instability is one of the major concerns in open pit mining. A significant collapse of the pit wall can result in injuries or fatalities, damage to mining equipment, interruptions to production, and potential loss of reserves.

### 6.1.1 Geomechanics in Pit Slope Design

The selection of optimal slope of an open pit mine is crucially influenced by the slope geometry, quality of rock mass and presence of geological features and their properties. (Abdellah et al., 2022). It is required to determine the factor of safety associated with the rock slope and its possible failure pattern. The slope angle influences the stripping ratio (the ratio of tonnage or volume of overburden to be removed to tonnage or volume of ore

to be extracted) and, as a result, mining profitability (Bye and Bell, 2001). Reduced stripping ratio (i.e., less waste rock to be removed) necessitates keeping the total slope angle as steep as possible (i.e., increased ore recovery). To do this, successful pit planning is required, which necessitates precise information regarding geology and site characteristics (e.g., slope geometry, rock qualities, groundwater conditions, and associated discontinuities). The final design of the best slope angle is influenced not only by ore grade distribution and operational cost, but also by the overall attributes of the rock mass. As a result, it is advised that the possibility of failure be factored into the final open pit design (Abdellah et al., 2022).

There are numerous ways for assessing the stability of rock slopes, each with advantages and disadvantages. Limit equilibrium techniques (LEMs); analytical and finite-element methods; finite differences and discrete element methods are examples of these methods. LEMs (for example, the method of slices) are typical methods of analysis in which the factor of safety is calculated by dividing the soil/rock mass above the hypothetical failure surface into multiple vertical slices. The factor of safety (FOS) is defined as the ratio of soil/rock shear strength at the failure to the mobilized shear stress on the surface of failure (Abdellah et al., 2022).

### 6.1.2 Predicting Time to Failure

Inverse Velocity method is often used to predict the time of failure. Generally, a trend line is fitted to the inverse velocity (1/V) derived from monitoring systems which record the slope movements, and when the velocity approached infinity (1/V approaches to zero), the failure is highly expected. Fukuzono (1985) found that the time to accelerating creep failure

under gravity loading was inversely proportional to the deformation rate (velocity). By plotting the inverse-velocity versus time curve, one can estimate the time of failure by extrapolating the curve to the x-axis (inverse-velocity = 0). Extrapolating inverse-velocity to the x-axis is much easier than extrapolating a hyperbolic curve to a vertical asymptote when using the deformation rate.

### 6.1.3 Hydrology and Dewatering

The existence of groundwater reduces the shear strength of the materials which causes a decrease in the stability of pit slopes. Groundwater modeling is used to predict groundwater level on the pit slope which helps in determining the slope stability. Dewatering efforts in mines to reduce the slope failure risk include horizontal drilling on the face of the high walls to target areas of saturation, especially when seepages are observed. Vertical water well drilling is also done to dewater the pit, which has relatively higher costs, but can pump more water out of the pit, if the well is drilled in a well-studied spot with optimal depth. Instabilities of an open pit slope can be caused by the increasing of slope load and the changing of pore pressure due to groundwater level, and slopes can fail after seepage increases the pore pressures.

The effect of groundwater on slope stability is assumed to be harmful, via diminishing rock or soil strength, changing mineral components in the rock through chemical reactions or dissolving, changing rock or soil density, and producing erosion (Lee, 2002). The greater the rise in groundwater level, the greater the rise in pore water pressure. This entails lowering the slope mass's shear resistance, particularly in soil material (Ariana, 2014). The groundwater level affects slope stability by both decreasing

the strength of the slope and increasing the load on the slope (Johansson and Edeskär, 2014). Data and warnings from weather stations are also very important in mining operations, including monthly/yearly rainfall historical records, daily/weekly weather forecasts.

### 6.1.4 Geology and Faults

Geological structures determine the slope failure mechanism associated with rock mass and its pattern, e.g., planar, wedge, toppling and/or circular failure (Abdellah et al., 2022). The geology of the area can play a crucial role in the stability of the slope. The type and properties of the rock formation can affect the stability of the slope, and some formations are more prone to instability than others. For example, sedimentary rocks such as shale and sandstone tend to be less stable than igneous and metamorphic rocks due to their weak structure and low resistance to weathering.

In addition to the geology, the fault network can also play a significant role in slope stability. Faults are zones of weakness in the rock mass where movement can occur, leading to rock deformation and instability. Faults can act as channels for water to flow through, which can cause erosion and weaken the slope. Faults can also increase the likelihood of rock falls and landslides, as they can create zones of weakness where rocks are more likely to detach and slide down the slope. To mitigate the risk of slope instability, mining engineers use a range of techniques, including slope reinforcement and monitoring. Reinforcement techniques can include the use of rock bolts, mesh, and shotcrete to provide additional support to the slope. Slope monitoring techniques can include the use of geotechnical instruments such as inclinometers, piezometers, and crack gauges to detect

any movement or deformation in the slope. By monitoring the slope, engineers can identify any potential instability issues and take proactive measures to address them before they become a safety hazard.

### 6.1.5 Micro-Seismic Blast Vibrations

Micro-seismic data that has been properly analyzed can be used to supplement surface monitoring systems in identifying possible instability and the subsequent collapse. Micro-seismic monitoring techniques have been identified as an effective tool for mapping the evolution of rock fracture networks for slope stability assessment (Hardy and Kimble, 1994; Lynch and Malovichko, 2006). Many seismic events related with slope changes are weak, and they rarely trigger a micro-seismic monitoring system that is set up to record an event in a trigger mode. The results of a research study demonstrated that weak seismic events are just as significant as strong seismic events for assessing slope stability, and they should not be overlooked in open pit mining micro-seismic monitoring (Luo et al., 2018).

## 6.2 Pit Slope Monitoring

Slope stability monitoring is regarded as one of the most important aspects of safety in open-pit mines, owing to the substantial danger of slope displacement caused by mining operations. Slope monitoring is an essential component of slope management in open pit mines. It provides information for recognizing potentially unstable ground, analyzing slope design performance, which includes identifying any slope instability and/or failure causes that arise. The probability of slope displacement and its effects can be significantly decreased if the failure mechanisms are understood, and the slopes are adequately

monitored. This enables optimal mining conditions that are both safe for people and operating equipment. Slope monitoring efforts include: 1) Detection of movement; 2) Displacement measurement; and 3) displacement trend determination (uniform, decelerating or accelerating).

### 6.2.1 Visual Inspections

The most simple and basic method of slope monitoring is visual observation through field inspections and camera systems implemented in open pits. In such inspections, crack monitoring plays an important role. Small cracks on top of the pit or individual benches are early warning signs of instability. Because this displacement cannot be seen from the pit floor, it is critical to monitor the crests of highwalls above active work locations on a regular basis. At all times, safe access to the areas directly above the active mining should be maintained. During periods of heavy precipitation or spring run-off, as well as after large blasts, frequent inspections may be required.

The other important aspect of field observations includes hanging rocks/boulders on the walls and bench edges (crests). Information about the more advanced monitoring equipment and tools are provided through the following subsections. Fresh rubble near the pit's toe or on the pit floor is a clear indication that instability has occurred. It must be determined which piece of the slope failed and whether further material may fail. An overhang is one of the most dangerous circumstances that can arise.

### 6.2.2 Surveying with Prisms

Prisms, with millimeter precision, are the most widely used slope monitoring system in open pit mines, due to their cheap costs. The Geodetic (prism) monitoring system is used to detect long-term slope movement trends and to predict where slope failure is likely to occur. This technology gives highly accurate and long-term 3D displacement data, but only for the area immediately surrounding the placed prism because there are no data in-between prisms. Even though prisms can be securely fixed with protective casing fitted around the prism, many prisms are damaged or lost during blasts owing to rockfall, slope failure, etc. Prism installation can be dangerous in some regions, as well as time consuming and costly. When a prism becomes damaged or dusty, it inhibits the system's monitoring capability, necessitating replacement or cleaning. Prism monitoring is similarly limited in that it measures the movements of single, widely dispersed points on a slope.

### 6.2.3 Borehole Inclinometers

Inclinometers are geotechnical tools that measure horizontal displacements along a borehole's numerous spots. As a result, they are sometimes known as borehole inclinometers. These are useful for long-term, precise monitoring of a borehole's position along its entire length. It is also possible to track the rate of movement by taking a series of readings over time. The borehole must be extended below the depth of movement so that readings taken at the end of the hole are referenced to a stable base. The depth of the failure surface is determined by the depth at which shear movement is observed by the slope inclinometer.

### 6.2.4 Time Domain Reflectometry (TDR)

Time Domain Reflectometry is a technique that involves sending electronic pulses down the length of a coaxial cable that has been grouted in a drillhole. When a deformation or a break in the cable is met, a signal is reflected that provides information on the deformation of the subsurface rock mass. While inclinometers are more commonly used to detect subsurface displacements, TDR cables are becoming increasingly popular and have various advantages over regular inclinometers, including lower installation costs, deeper possibility of monitoring, etc. (Kane, 1998).

### 6.2.5 Slope Stability Radars

Slope stability radars measure sub-millimeter distances between antenna and continuous points on slopes over a set scanned area. Their range can be up to 3.5 km, and they are known for their rapid tactical installation and commissioning. There are two kinds of radar systems used for monitoring. Although radar devices can provide accurate range measurements, the area footprint of each measurement is huge when compared to a laser system. The SSR (Real Aperture Radar) is a mobile surveillance system that uses a radar dish. The scan time is determined by the size of the slope area chosen as well as the distance from the slope. The hardware components require frequent maintenance due to the large number of moving parts as it scans. A radar antenna moves side to side along a rail in the SAR (Synthetic Aperture Radar) system. The system can be configured to be mobile, semi-permanent, or permanent. The technology requires less maintenance and consumes little power. It also has a high level of accuracy.

### 6.2.6 Laser (LiDAR)

LiDAR (Light Detection and Ranging) is another surveying technology that can be used for slope monitoring in mines. It uses speed of light to measure distances from instrument and can be done via by aerial or ground surveys. LiDAR uses laser pulses to create a detailed 3D map of a terrain or structure, including slopes in mines. In slope monitoring, LiDAR can be used to create high-resolution images of slopes and to measure their movement over time. LiDAR sensors are typically mounted on aircraft or ground-based vehicles and scan the surface of the slope with laser beams. The reflected laser beams are then used to create a detailed 3D map of the slope's surface. The resulting LiDAR data can be used to create detailed digital elevation models (DEMs) of slopes, which can help identify potential failure zones and other geotechnical hazards. LiDAR data can also be used to create accurate and detailed slope displacement maps, which can help track changes in slope position over time.

Overall, LiDAR is a valuable tool for slope monitoring in mines, as it can provide highly detailed and accurate data over large areas. LiDAR data can be combined with other monitoring technologies, such as total stations and GNSS, to provide a comprehensive understanding of slope behavior and to help prevent slope failures and other geotechnical hazards in mines. Because it does not use prisms to measure points or faces, laser monitoring may fill in the gaps where prism monitoring cannot. This system allows the user to set monitoring areas and frequencies, as well as group specific points to be monitored. As a result, it is also utilized for long-term trend monitoring and detecting high-risk locations. While the laser is not as accurate as the radar system, the setup is

relatively comparable. Its primary advantages are that it is a simple automated operation that is simple to configure and that it provides medium to long range operation.

### 6.2.7 InSAR

Interferometric synthetic aperture radar (InSAR) is a satellite remote sensing technology used for mapping ground displacement across a number of sectors. Satellite InSAR is well-known in the open pit mining industry. The excellent precision and vast coverage of InSAR make it an indispensable tool for surface mining operations and planning. Open pit mining poses substantial geotechnical difficulties. These include building and maintaining pit walls at ideal angles and berm heights. Operators can optimize this balance for each slope and respond promptly when displacement rates alter thanks to precision displacement monitoring. Satellite InSAR can detect displacement on all active and inactive slopes, including pits, stockpiles, dumps, tailings facilities, and haul routes, in the same wide-area footprint as existing ground-based instruments. InSAR datasets provide 1-2 mm resolution coverage of ground movement for most of the earth surface with various satellites in operation. The limitation of InSAR is the limited measurement frequency (2-12 days intervals). Thus, it should be a supplementary monitoring method, not a replacement to local monitoring in the mine.

### 6.2.8 Total Stations

Total stations are a type of surveying instrument commonly used in slope monitoring in mines. They are used to measure horizontal and vertical angles as well as distances between points. In slope monitoring, total stations are used to measure the position and movement

of slopes, which can help detect potential slope failures or other geotechnical hazards. Total stations are typically set up on stable ground and are used to measure the positions of prisms installed on the slopes being monitored. The total station sends a beam of light to the prism, which reflects it back to the instrument. By measuring the angle and distance to the prism, the total station can calculate the position of the prism and hence the slope being monitored. The data collected by the total station can be used to create detailed slope displacement maps and to track changes in slope position over time. This information can be used to identify potential slope failures and to develop strategies for mitigating or avoiding them.

### 6.2.9 Global Positioning System (GPS)

GPS monitoring as a surface monitoring method, are based on satellites orbiting the Earth can provide real-time positioning at any location, 24 hours a day, in any weather. It has the potential to be employed in many engineering projects involving high cut slopes, huge open pit mining, subsidence, and landslides to measure ground displacements over a vast region (Chrzanowski et al., 1989). Several researchers have already investigated GPS applications in ground displacements. GPS was developed by scientists from many fields who indicated that the technology could measure displacements virtually as precisely as total station surveying. Relative positions (3D coordinate differences) between receivers can be measured with an accuracy of a few millimeters to about 20 mm over distances up to several kilometers, and about 1 ppm (parts per million) over distances up to several hundred kilometers, when two or more receivers work simultaneously in a so-called differential mode (Mark et al., 2009).

### 6.2.10 GNSS

GNSS (Global Navigation Satellite System) is another type of surveying technology commonly used in slope monitoring in mines. GNSS uses a network of satellites to determine the location of a receiver on the Earth's surface with high precision. In slope monitoring, GNSS receivers are used to measure the position and movement of prisms installed on slopes being monitored. The GNSS receiver receives signals from multiple satellites and uses triangulation to calculate the receiver's position. By measuring the positions of prisms installed on the slopes being monitored, GNSS can calculate the slope's position and movement. GNSS is particularly useful in slope monitoring because it can provide high-precision data over a large area. This makes it possible to monitor a wide range of slopes simultaneously, which can be important in large mines or areas with many slopes. GNSS data can be combined with other monitoring technologies, such as total stations, to provide a more comprehensive understanding of slope behavior. Overall, GNSS is a valuable tool for slope monitoring in mines, as it provides precise and accurate data on slope movement and can help identify potential slope failures and other geotechnical hazards.

### 6.2.11 Wireline Crack Extensometers

Wireline extensometers consist of a series of anchor points, a central rod, and a series of measuring wires or cables. The anchor points are installed at the ends of the borehole and provide a fixed reference point. The central rod is suspended between the anchor points and is equipped with measuring wires or cables that are attached to the rock mass at specific locations. As the rock mass deforms or moves, the distance between the anchor points

changes, which is recorded by the measuring wires or cables. In open pit mining, wireline extensometers are often installed in critical locations on slopes and walls, to monitor movement and deformation over time. The data collected by wireline extensometers can be used to create deformation maps and to identify potential failure zones. This information can be used to develop strategies for slope stabilization and to help prevent slope failures and other geotechnical hazards. Overall, wireline extensometers are an important tool for monitoring rock mass movement and deformation in open pit mining and can help identify potential geotechnical hazards before they occur.

### 6.2.12 Borehole Extensometers

The magnet extensometer measures settlement and heave in excavations, foundations, dams, and embankments. It can also be used behind retaining structures like sheet piles and slurry walls, as well as above underground apertures like tunnels and shafts. Extensometer data show the depths at which settlement occurred as well as the total amount of settlement. Wireline extensometers are instruments used in open pit mining to monitor the deformation or movement of rock masses, including slopes, walls, and floors. They are typically installed in boreholes drilled into the rock mass and are used to measure the distance between two reference points in the borehole.

### 6.2.13 Seismic Geophones

Micro-seismic monitoring systems detect seismic signals caused by rock breakage or movement within an open pit. Because micro-seismic sensors can remotely measure rock fracture events, these techniques can map the 4D pattern and dynamic development of rock

fractures before any deformation on the slope surface is visible. A micro-seismic system for open pit monitoring is typically comprised of a geophone network with a number of geophones implanted in the ground at 100-200 m intervals. The seismic signals must be recorded by four or more geophone sites to acquire the 4D pattern of seismicity. The system is typically used in a trigger mode, with short-time averages (STA) and long-time averages (LTA) algorithms. The system is triggered, and the event is recorded when at least four stations have a STA/LTA ratio greater than the set threshold (Mendecki, 1997; Lou et al., 2018).

### 6.2.14 Piezometers

Data from hydrogeological conditions are very useful in building a better slope stability analysis model, because it can be used to indicate the mechanism of slope failure and correction of slope failure boundary. Piezometers are instruments used in open pit mining to measure water pressure in the ground. They are used to monitor groundwater levels, which can be important in determining the stability of slopes and the potential for landslides and other geotechnical hazards. Piezometers are typically installed in boreholes drilled into the ground and are designed to measure the pressure of water in the surrounding soil or rock. The water pressure measured by the piezometer can be used to calculate the depth of the groundwater table, which is an important factor in slope stability analysis. In open pit mining, piezometers are often installed around the perimeter of the pit and in key locations on slopes, to monitor changes in groundwater levels and to detect potential hazards. The data collected by piezometers can be used to create groundwater flow models and to develop strategies for managing water in the mine, such as dewatering.

Overall, piezometers are an important tool for monitoring groundwater levels in open pit mining and can help identify potential geotechnical hazards before they occur. Unexpected variations in piezometer readings or changes in steady flow from dewatering wells may indicate subsurface movement that has cut through a perched water table or intersected a water bearing structure. Changes in water pressure caused by drain channel blockage can potentially cause slope failures. Water can also enter fissures and hasten weathering processes. Freeze-thaw cycles cause water-filled joints to expand and loosen highwall material. During cold weather, increased scaling may be required.

# 7 Summary and Future Works

## 7.1 Pixel-based Landslide Mapping Model

This study introduces a data fusion machine learning approach for landslide inventory development. Optical and SAR imagery as well as readily available geospatial inputs for the region impacted by landslides due to the Kumamoto 2016 earthquakes were used within a logistic regression model. A visually drawn landslide inventory, created by NIED of Japan by investigating the post-event aerial imagery and satellite SAR data and pre-event optical imagery, was used for training, validation, and testing data. Optical imagery included pre-event satellite (DigitalGlobe) and post-event aerial (GSI) high-resolution imagery. SAR imagery was also available for pre- and post-event from ALOS-2 satellite's PALSAR-2 sensor. Several change indices were tested including grayscale index change, vegetation index change, and temporal SAR amplitude change as additional predictive inputs for pixel-wise landslide classification. Based on the ROC-based feature ranking results, grayscale change, vegetation change, and HV amplitude difference were included in the classification algorithm. To enhance the performance of the model, geospatial data of related variables were also evaluated including elevation, slope, aspect, curvature, landslide probability, precipitation, and geology. Based on the conducted feature ranking analysis, five geospatial features with the highest predictive ability were included in the classification algorithm, including, slope, curvature, landslide probability, precipitation, and geology. Model development (training and validation) data and testing data were randomly spatially sampled across distinct regions. Using 10% of available landslide pixels and 0.4% of the non-landslide pixels from the model development grid regions, a relatively small and balanced training dataset was created. The logistic regression classification

across five different scenarios was tested: 1) RGB; 2) RGB + Change; 3) RGB + Geospatial; 4) RGB + Change + Geospatial; and 5) all variables regardless of feature ranking results. Comparing the outcomes of the logistic regression classification models, the model trained and tested via 11 features (RGB plus selected change and geospatial variables) outperformed the other case models and was introduced as the preferred model. The feature importance analysis after training the classifier showed that the slope and geology variables are the most important among geospatial variables, green channel has the highest importance among RGB channels, and the grayscale change index has the highest importance among the change information variables. Given that this model was only developed by only 3.85% of the total pixels across the region, a landslide inventory can be developed using a relatively small training dataset with a short run-time due to the very high speed of logistic regression classification modeling. It is suggested that the methodology proposed in this research be tested on other earthquake-triggered landslide events in Japan and other regions around the globe. Additional informative bands might be still beneficial to the landslide modeling, including landcover data, Near-Infrared and other bands of the multi-spectral sensors, etc. In addition, feature transformation and engineering, as well as dimensionality reduction techniques are worthy of investigation. The other direction of research could be to try model via deep learning algorithms such as convolutional neural networks (CNN) on the same datasets to analyze the comparative classification performance using post-event images with and without additional parameters used in this research.

## 7.2 Semi-Supervised Liquefaction Mapping Model

The idea of this paper was to generate spatially consistent and complete maps of liquefaction ejecta using a small number of visually labeled liquefaction surface effects using high resolution aerial imagery with the aid of advanced computer vision and machine learning algorithms. A semi-supervised self-training classification method was presented and validated in this study and showed superior performance in detection of liquefaction surface effects, compared to the supervised learning approach. This paper also investigated the use of RGB image-derived products as inputs into the classification algorithm such as transformations based on color, texture, statistics, and dimensionality reduction. The results showed enhanced performance of liquefaction labelling when selected statistical indices and dimensionality reduction techniques' outputs are used alongside RGB channels of the imagery. Texture and color transformation components did not impact the model performance considerably. The feature ranking was done via MRMR algorithm which reduces the redundancy and increases the relevance of the features to the classes of interest. A Fuzzy-based clustering approach was proposed for clustering liquefaction pixels into two classes of dark and light liquefaction to increase the certainty of the pixels by removing outliers and mixed pixels from the labeled liquefaction training data. The advanced machine learning and image processing methods used in the study, made it possible to spatially complete and augment the partial inventory of liquefaction labels in the image library provided by Sanon et al. (2023) with an acceptable accuracy. The proposed methodology can be applied in other domains where labeled and high-quality training data is limited, and there is a need for semi-supervised algorithms to classify data due to difficulty of labeling.

## 7.3 Regional Tile-based Landslide Mapping Model

In this work, we trained a deep learning CNN for mapping landslides from aerial or satellite optical RGB imagery using transfer learning from the DeepLabV3+ with the pre-trained ResNET50 base model. The model was trained, validated, and tested using the landslide inventory and aerial imagery (0.5m resolution) from the 2016 Kumamoto earthquake induced landslides. The trained model was then applied on two separate events (landslide impacted region from the 2017 Hokkaido earthquake and landslide impacted region from the 2019 Asakura rainfall events). For the 2017 Hokkaido event, the optical RGB imagery was collected by xx and had an original resolution of 3m. The pretrained model had acceptable accuracy only when the Hokkaido imagery was down-sampled to 0.5m. The 2019 Asakura rainfall event had the same resolution as the Kumamoto event and resulted in similar accuracy. The proposed transfer learning model is applicable for post-event landslide mapping with good cross-event transferability. It should be emphasized that although the selected regions affected by landslides differ in various aspects such as topography, triggers, and landslide shape and size, most of them have high vegetation coverage, which is the main reason for the often-limited model transferability to new areas without vegetation land cover. A few recent studies have tried to extend the transferability of deep learning-based landslide detection methods by training the models via data from multiple events (Yu et al., 2021; Ghorbanzadeh et al., 2022; Zhang et al., 2023). However, this scheme is just in its beginning phase, and it still has a significant capacity to improve and expand in order to achieve globally applicable models.

## 7.4 Global Geospatial Liquefaction Hazard Model

This study aimed at developing a global liquefaction hazard model, while soil liquefaction is a complex local phenomenon. Such global liquefaction hazard maps are very useful, especially when local geotechnical data and surficial geologic maps are not available. A great effort was spent to balance the dataset for liquefaction prediction model development in a manner to exploit the data as much as possible. The class and event imbalance issue are treated in an innovative manner by under-sampling large events and distributing the datasets over several balanced subsets. Instead of training a single classifier, a system of voting machine learning classifiers is designed to achieve higher accuracy and certainty, using different classification techniques to balance the performance, as not a single classification technique can be considered the best for all earthquake events. The results showed superior performance compared to previously published logistic regression-based liquefaction probability models. The use of several relevant features in this study, which was made possible by the aid of advanced machine learning algorithms was found to be very helpful in detecting liquefaction-prone regions. The leave-one-out testing approach based on the total exclusion of individual earthquake events from the training data, revealed the uncertainties associated with accuracy assessments reported in the literature for liquefaction susceptibility models. This study highly recommends that geospatial liquefaction hazard models should be developed based on the leave-one-out accuracy assessments, not data partitioning. Future directions of research could be designing even more advanced and complex systems and finding ways to assign different weights to the individual classifier in the proposed voting system to optimize the final decision. Addition of other geospatial variables or the same but higher quality products might also be effective,

especially for the events where the model did not perform well. Some of the geospatial predictors used are temporally variable, and future research can yield additional insights by considering fluctuations over time, and their effects on liquefaction potential. Including more information retrieved from liquefaction historical data, especially for under-represented areas, might help to improve models on a regional level. In addition, combining geospatial and geotechnical methods could be considered to address the limitations of the geospatial liquefaction models. An issue with any advanced machine learning paradigm is the lack of a defined analytical expression that can be provided for further use. Simple interpretations of model structure is also generally impossible. While traditionalists may resist using these models, they will finally open their ways into the world of geotechnics and geohazards, given their demonstrated capabilities in learning from huge datasets. However, trained AI-based models should be delivered as codes or software, in a manner that is accessible to a broad community of users and researchers for further implementation of the proposed scheme. Several organizations would benefit from the outcomes of this study, such as civil protection organizations, insurance companies, extreme events' first responders, post-event reconnaissance teams, urban planners, and infrastructure constructors. Although the accuracy of the global geospatial models at the local scale is not very high, the availability of large-scale susceptibility maps of liquefaction hazard can assist policymakers and infrastructure administrators in prioritizing territories where should be investigated further to assess the seismic risk of structures and infrastructures associated with the occurrence of earthquake-induced soil liquefaction.

## 7.5 Slope Stability in Open Pit Mining

In summary, different commonly used slope monitoring techniques in the open pit mining industry was presented, in addition to the geomechanical, hydrological and geological considerations of pit slope design briefly. The instrumentations vary from surface to subsurface and airborne. Visual investigations are also a crucial part of the monitoring which is complementary to the remote sensing data.

# 8 Reference List

## 8.1 Chapter 2


Allstadt KE, Thompson EM, Jibson RW, et al. The US Geological Survey ground failure product: Near-real-time estimates of earthquake-triggered landslides and liquefaction. Earthquake Spectra. 2022;38(1):5-36. doi:10.1177/87552930211032685.

Bai, Y.; Adriano, B.; Mas, E.; Koshimura, S. Machine learning based building damage mapping from the ALOS-2/PALSAR-2 SAR imagery: Case study of 2016 Kumamoto earthquake. J. Disaster Res. 2017, 12, 646–655.

Burger, Wilhelm; Mark J. Burge (2010). Principles of Digital Image Processing Core Algorithms. Springer Science & Business Media. pp. 110–111. ISBN 978-1-84800-195-4.

Cabinet Office of Japan. Summary of Damage Situation in the Kumamoto Earthquake Sequence. 2016; (In Japanese). Available online: http://www.bousai.go.jp/updates/h280414jishin/index.html.

Carrara, A., Cardinali, M., Guzzetti, F., 1992. Uncertainty in assessing landslide hazard and risk. ITC Journal 2, 172–183.

Casagli, N., Intrieri, E., Tofani, V. et al. Landslide detection, monitoring and prediction with remote-sensing techniques. Nat Rev Earth Environ 4, 51–64 (2023). https://doi.org/10.1038/s43017-022-00373-x.

Chen C-W, Chen H, Wei L-W, Lin G-W, Iida T, Yamada R. 2017b. Evaluating the susceptibility of landslide landforms in Japan using slope stability analysis: a case study of the 2016 Kumamoto earthquake. Landslides 14: 1793–1801. https://doi.org/10.1007/s10346-017-0872-1.

Chen F, Yu B, Xu C, Li B (2017a) Landslide detection using probability regression, a case study of Wenchuan, Northwest of Chengdu. Appl Geogr 89:32–40.

Collett, D. Modeling Binary Data. New York: Chapman & Hall, 2002.

Cruden DM, Varnes DJ (1996) Landslides: investigation and mitigation. Chapter 3: Landslide Types and Processes. Transportation Research Board Special Report, 247.

Dai FC, Xu C, Yao X, Xu L, Tu XB, Gong QM. 2011. Spatial distribution of landslides triggered by the 2008 Ms 8.0 Wenchuan earthquake, China. Journal of Asian Earth Sciences 40: 883–895. https://doi.org/10.1016/j.jseaes.2010.04.010.

DigitalGlobe GeoEye-1 sensor documentation: https://gbdxdocs.digitalglobe.com/docs/geoeye-1.

Dou, J., K. T. Chang, S. Chen, A. Yunus, J. K. Liu, H. Xia, and Z. Zhu. 2015. “Automatic Case-Based Reasoning Approach for Landslide Detection: Integration of Object-Oriented Image Analysis and a Genetic Algorithm.” Remote Sensing 7 (4): 4318–4342. doi:10.3390/rs70404318.

Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu Z, Chen C-W, Han Z, Pham BT (2020) Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 17(3):641–658.

Ding, Q. Zhang, X. Zhou and B. Dai, "Automatic recognition of landslide based on CNN and texture change detection," 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China, 2016, pp. 444-448, doi: 10.1109/YAC.2016.7804935.

Dobson, A. J. An Introduction to Generalized Linear Models. New York: Chapman & Hall, 1990.

Fan, X., Scaringi, G., Korup, O., West, A.J., van Westen, C.J., Tanyas, H., Hovius, N., Hales, T.C., Jibson, R.W., Allstadt, K.E., Zhang, L., Evans, S.G., Xu, C., Li, G., Pei, X., Xu, Q., Huang, R., 2019. Earthquake-induced chains of geologic hazards: patterns, mechanisms, and impacts. Rev. Geophys. 57, 421–503. https://doi.org/10.1029/2018RG000626.

Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Comput. Geosci. 2020, 139, 104470.

Fanos AM, Pradhan B, Mansor S, Yusoff ZM, Abdullah AFb (2018), A hybrid model using machine learning methods and GIS for potential rockfall source identification from airborne laser scanning data. Landslides 15(9):1833–1850.

Fawcett, T. “ROC Graphs: Notes and Practical Considerations for Researchers”, Machine Learning 31, no. 1 (2004): 1–38.

Fick, S.E. and R.J. Hijmans (2017). WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37 (12): 4302-4315.

Friedman, Jerome H. “Greedy function approximation: A gradient boosting machine.” Annals of statistics (2001): 1189-1232.

Frolova JV, Gvozdeva IP, Kuznetsov NP (2015) Effects of Hydrothermal Alterations on Physical and Mechanical Properties of Rocks in the Geysers Valley (Kamchatka Peninsula) in Connection with Landslide Development. In: Proceedings World Geothermal Congress 2015, pp 1–6.

GDEM-003, 2018, NASA/METI/AIST/Japan Space systems, and U.S./Japan ASTER Science Team. ASTER Global Digital Elevation Model V003. 2018, distributed by NASA EOSDIS Land Processes DAAC, https://doi.org/10.5067/ASTER/ASTGTM.003.

GEER Report, 2016; Geotechnical Aspects of the 2016 MW 6.2, MW 6.0, and MW 7.0 Kumamoto Earthquakes GEER Association; geerassociation.org; Kumamoto Earthquakes, Version 1.0, July 2016. Kayen R, Dashti S, Kokusho T, Hazarika H, Franke K, Oettle N, Wham B, Calderon JR, Briggs D, Guillies S, Cheng K, Tanoue Y, Takematsu K, Matsumoto D, Morinaga T, Furuichi H, Kitano Y, Tajiri M, Chaudhary B, Nishimura K, Chu C (2016) Geotechnical aspects of the 2016 Mw 6.2, Mw 6.0, and Mw 7.0 Kumamoto earthquakes. Geotechnical Extreme Events Reconnaissance Association, Version 1.0, July 2016. https://pubs.er.usgs.gov/publication/70185571.

Godt, J.W., Sener, B., Verdin, K.L., Wald, D.J., Earle, P.S., Harp, E.L. and Jibson, R.W., 2008, Rapid Assessment of Earthquake-induced Landsliding: Proceedings of the First World Landslide Forum, United Nations University, Tokyo, Japan, p. 392-395.

Geological Survey of Japan (2015); Seamless Digital Geological Map of Japan by the Geological Survey of Japan (2015): https://gbank.gsj.jp/seamless/.

Geospatial Information Authority of Japan (GSI). (In Japanese). Available online: https://www.gsi.go.jp/BOUSAI/H27-kumamoto-earthquake-index.html.

Geospatial Information Authority of Japan (GSI) – in Japanese: https://maps.gsi.go.jp/development/ichiran.html#fukkyukizu

Godt, J.W., Sener, B., Verdin, K.L., Wald, D.J., Earle, P.S., Harp, E.L. and Jibson, R.W., 2008, Rapid Assessment of Earthquake-induced Landsliding: Proceedings of the First World Landslide Forum, United Nations University, Tokyo, Japan, p. 392-395.

GSI; Global Map by the Geospatial Information Authority of Japan (GSI): https://www.gsi.go.jp/kankyochiri/gm_japan_e.html.

GSI (2016). “平成２８年熊本地震に関する情報”. Geospatial Information Authority of Japan, http://www.gsi.go.jp/BOUSAI/H27-kumamoto-earthquake-index.html.

Guzzetti F, Carrara A, Cardinali M, Reichenbach P (1999) Landslide hazard evaluation: a review of current techniques and their application in a multi-scale study. Central Italy Geomorphology 31:181–216. https://doi.org/10.1016/S0169-555X(99)00078-1.

Guzzetti, Fausto, 2006, Ph.D. dissertation on Landslide Hazard and Risk Assessment, at the Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität University of Bonn, Bonn, Germany, defended on July 2006.

Guzzetti, Fausto; Alessandro Cesare Mondini, Mauro Cardinali, Federica Fiorucci, Michele Santangelo, Kang-Tsung Chang, Landslide inventory maps: New tools for an old

problem, Earth-Science Reviews, Volume 112, Issues 1–2, 2012, Pages 42-66, ISSN 0012-8252, https://doi.org/10.1016/j.earscirev.2012.02.001.

Hajeb, M.; Karimzadeh, S.; Matsuoka, M. SAR and LIDAR Datasets for Building Damage Evaluation Based on Support Vector Machine and Random Forest Algorithms—A Case Study of Kumamoto Earthquake, Japan. Appl. Sci. 2020, 10, 8932. https://doi.org/10.3390/app10248932.

Harp, E.L., Jibson, R.L., 1995. Inventory of landslides triggered by the 1994 Northridge, California earthquake. U.S. Geological Survey Open File Report, pp. 95–213.

Hasi, B., Ishii, Y., Maruyama, K., Terada, H., Suzuki, S. and Nakamura, A. (2011). "Distribution and scale of landslides induced by recent reverse-fault earthquakes in Japan", Journal of the Japan Landslide Society, 48(1), 23-38. (in Japanese)

He, Ping; Yangmao Wen, Caijun Xu, Yunguo Chen, Complete three-dimensional near-field surface displacements from imaging geodesy techniques applied to the 2016 Kumamoto earthquake, Remote Sensing of Environment, Volume 232, 2019, 111321, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2019.111321.

He, Qian; Ming Wang, Kai Liu, Rapidly assessing earthquake-induced landslide susceptibility on a global scale using random forest, Geomorphology, Volume 391, 2021, 107889, ISSN 0169-555X, https://doi.org/10.1016/j.geomorph.2021.107889.

Hervás, J.; Barredo, J.; Rosin, P.L.; Pasuto, A.; Mantovani, F.; Silvano, S. Monitoring landslides from optical remotely sensed imagery: The case history of Tessina landslide, Italy. Geomorphology 2003, 54, 63–75.

Hölbling, D. et al. Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. Geosciences 7, 37 (2017).

Hong, Y., Adler, R. & Huffman, G. Use of satellite remote sensing data in the mapping of global landslide susceptibility. Nat Hazards 43, 245–256 (2007). https://doi.org/10.1007/s11069-006-9104-z.

Huang, Q.;Wang, C.; Meng,Y.; Chen, J.; Yue, A. Landslide monitoring using change detection in multitemporal optical imagery. IEEE Geosci. Remote Sens. Lett. 2019, 1–5.

Jaboyedoff, M., T. Oppikofer, A. Abellan, M.-H. Derron, A. Loye, R. Metzger, and A. Pedrazzini. 2012. "Use of LiDAR in Landslide Investigations: A Review." Natural Hazards 61 (1): 5–28. doi:10.1007/ s11069-010-9634-2.

James, Gareth • Daniela Witten • Trevor Hastie Robert Tibshirani; An Introduction to Statistical Learning with Applications in R, Springer New York Heidelberg Dordrecht London, 2013, ISBN 978-1-4614-7138-7 (eBook), DOI 10.1007/978-1-4614-7138-7.

JAXA, 2016, Japanese Aerospace Exploration Agency: https://satpf.jp/spf_atl/article/view/273?lang=en

Ji, Y.; Sumantyo, J.T.S.; Chua, M.Y.; Waqar, M.M. Earthquake/tsunami damage level mapping of urban areas using full polarimetric SAR data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2018, 11, 2296–2309.

Jolliffe, I. T. Principal Component Analysis. 2nd ed., Springer, 2002.

Juang CS, Stanley TA, Kirschbaum DB (2019) Using citizen science to expand the global map of landslides: introducing the cooperative open online landslide repository (COOLR). PLoS One 14:e0218657.

Kalantar, Bahareh; Biswajeet Pradhan, Seyed Amir Naghibi, Alireza Motevalli & Shattri Mansor (2018) Assessment of the effects of training data selection on the landslide susceptibility mapping: a comparison between support vector machine (SVM), logistic regression (LR) and artificial neural networks (ANN), Geomatics, Natural Hazards and Risk, 9:1, 49-69, DOI: 10.1080/19475705.2017.1407368.

Kazmi, Wajahat; Francisco Jose Garcia-Ruiz, Jon Nielsen, Jesper Rasmussen, Hans Jørgen Andersen; Detecting creeping thistle in sugar beet fields using vegetation indices, Computers and Electronics in Agriculture, Volume 112, 2015, Pages 10-19, ISSN 0168-1699, https://doi.org/10.1016/j.compag.2015.01.008.

Kim, Jae Sung and KyoHyouk Kim "Analysis of 2016 Minamiaso landslides using remote sensing and geographic information system," Journal of Applied Remote Sensing 12(3), 036001 (4 July 2018). https://doi.org/10.1117/1.JRS.12.036001.

Kiyota, T., Ikeda, T., Konagai, K. and Shiga, M. (2017). Geotechnical Damage Caused by the 2016 Kumamoto Earthquake, Japan. International Journal of Geoengineering Case histories, http://casehistories.geoengineer.org, Vol.4, Issue 2, p.78-95. doi: 10.4417/IJGCH-04-02-01.

Konishi, Tomohisa; and Yuzo Suga "Landslide detection using polarimetric ALOS-2/PALSAR-2 data: a case study of 2016 Kumamoto earthquake in Japan", Proc. SPIE 10788, Active and Passive Microwave Remote Sensing for Environmental Monitoring II, 107880P (9 October 2018); https://doi.org/10.1117/12.2324030.

Koyanagi, Kenta; Takashi Gomi; Roy C. Sidle; 2020; Characteristics of landslides in forests and grasslands triggered by the 2016 Kumamoto; EARTH SURFACE PROCESSES AND LANDFORMS; Earth Surf. Process. Landforms 45, 893–904 (2020); DOI: 10.1002/esp.4781. © 2019 John Wiley & Sons, Ltd.

Kyoto University, 2016, Occurrence of slope fluctuations in Minamiaso Village due to the 2016 Kumamoto Earthquake 2nd report (finding points as of 19:00, 2016/4/18); Disaster Prevention Research Institute at Kyoto University:

(http://www.slope.dpri.kyotou.ac.jp/disaster_reports/2016KumamotoEq/2016KumamotoEq2.html).

Liu, Wen; Fumio Yamazaki and Yoshihisa Maruyama; 2019; Detection of Earthquake-Induced Landslides during the 2016 Kumamoto Earthquake Using Multitemporal Airborne Lidar Data; Remote Sens. 2019, 11, 2292; doi:10.3390/rs11192292.

Liu Z, Gilbert G, Cepeda JM, Lysdahl AOK, Piciullo L, Hefre H, Lacasse S (2021) Modelling of shallow landslides with machine learning algorithms. Geosci Front 12(1):385–393.

Lopes, A., R. Touzi, and E. Nezry. "Adaptive Speckle Filters and Scene Heterogeneity." IEEE Transactions on Geoscience and Remote Sensing 28, No. 6 (1990): 992-1000.

Ma, Z., Mei, G. & Piccialli, F. Machine learning for landslides prevention: a survey. Neural Comput & Applic 33, 10881–10907 (2021). https://doi.org/10.1007/s00521-020-05529-8

Malamud, B.D., Turcotte, D.L., Guzzetti, F., Reichenbach, P., 2004. Landslide inventories and their statistical properties. Earth Surface Processes and Landforms 29 (6), 687–711.

Mao, W.; Wang, Y.; Wang, Y. Real-time Detection of Between-row Weeds Using Machine Vision. In Proceedings of the 2003 ASAE Annual International Meeting, American Society of Agricultural and Biological Engineers, Las Vegas, NV, USA, 27–30 July 2003; p. 031004.

Martha, T. R., Kerle, N., Van Westen, C. J., Jetten, V. & Kumar, K. V. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS J. Photogramm. Remote Sens. 67, 105–119 (2011).

McCullagh, P., and J. A. Nelder. Generalized Linear Models. New York: Chapman & Hall, 1990.

Meyer, G.E.; Mehta, T.; Kocher, M.F.; Mortensen, D.A.; Samal, A. Textural Imaging and Discriminant Analysis for Distinguishing weeds for Spot Spraying. Trans. Asae 1998, 41, 1189–1197.

Mohan, A, Singh, AK, Kumar, B, Dwivedi, R. Review on remote sensing methods for landslide detection using machine and deep learning. Trans Emerging Tel Tech. 2021; 32:e3998. https://doi.org/10.1002/ett.3998.

Mondini, A. C. et al. Landslide failures detection and mapping using synthetic aperture radar: past, present and future. Earth Sci. Rev. 216, 103574 (2021).

Moosavi, V., Talebi, A., Shirmohammadi, B., 2014. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. Geomorphology 204, 646–656.

Motohka, T.; O. Isoguchi, M. Sakashita, and M. Shimada, "ALOS-2 PALSAR-2 Cal/Val Updates", JAXA/EORC Joint PI Meeting of Global Environment Observation Mission FY2017, Jan. 24, 2018.

Nadim, F., Kjekstad, O., Peduzzi, P. et al. Global landslide and avalanche hotspots. Landslides 3, 159–173 (2006). https://doi.org/10.1007/s10346-006-0036-1.

NIED, 2016, Distribution map of mass movement by the 2016 Kumamoto earthquake, edited by National Research Institute for Earth Science and Disaster of Japan (in Japanese), National Research Institute for Earth Science and Disaster Resilience (NIED). Retrieved from: http://www.bosai.go.jp/mizu/dosha.html.

National Research Institute for Earth Science and Disaster Prevention, NIED (2016-b). "Rupture process of the Mj7.3, April 16, mainshock of the 2016 Kumamoto earthquake obtained from strong-motion data", http://www.kyoshin.bosai.go.jp/kyoshin/topics/Kumamoto_20160416/inversion/index_en.html.

Nava, L.; Bhuyan, K.; Meena, S.R.; Monserrat, O.; Catani, F. Rapid Mapping of Landslides on SAR Data by Attention U-Net. Remote Sens. 2022, 14, 1449. https://doi.org/10.3390/rs14061449.

Nowicki, M.A., Wald, D.J., Hamburger, M.W., Hearne, M., and Thompson, E.M., 2014, Development of a globally applicable model for near real-time prediction of seismically induced landslides: Engineering Geology, v. 173, p. 54–65.

Nowicki Jessee, M.A., Hamburger, H.W., Allstadt, K.E., Wald, D.J., Robeson, S.M., Tanyas, H., Hearne, M., Thompson, E.M., 2018, A Global Empirical Model for Near Real-time Assessment of Seismically Induced Landslides, Journal of Geophysical Research: Earth Surface, Volume 123, Issue 8, p. 1835-1859, AGU. https://doi.org/10.1029/2017JF004494.

NZSEE (2016). "Learning from Earthquakes Mission: Kumamoto Earthquake 2016, Japan." In-Country Report No 2 of 14th May 2016, New Zealand Society for Earthquake Engineering Inc.

Parise, Mario; Randall W. Jibson, A seismic landslide susceptibility rating of geologic units based on analysis of characteristics of landslides triggered by the 17 January,1994 Northridge, California earthquake, Engineering Geology, Volume 58, Issues 3–4, 2000, Pages 251-270, ISSN 0013-7952, https://doi.org/10.1016/S0013-7952(00)00038-7.

Rashidian, V.; Baise, L.G.; Koch, M. Using High Resolution Optical Imagery to Detect Earthquake-Induced Liquefaction: The 2011 Christchurch Earthquake. Remote Sens. 2020, 12, 377. https://doi.org/10.3390/rs12030377.

Reichenbach P, Busca C, Mondini AC, Rossi M.; 2014. The influence of land use change on landslide susceptibility zonation: the Briga catchment test site (Messina, Italy). Environmental Management 54: 1372–1384. https://doi.org/10.1007/s00267-014-0357-0.

Reid, Mark E.; Thomas W. Sisson, Dianne L. Brien; Volcano collapse promoted by hydrothermal alteration and edifice shape, Mount Rainier, Washington. Geology 2001;; 29 (9): 779–782. doi: https://doi.org/10.1130/0091-7613(2001)029<0779:VCPBHA>2.0.CO;2

Sánchez-Sastre LF, Alte da Veiga NMS, Ruiz-Potosme NM, Carrión-Prieto P, Marcos-Robles JL, Navas-Gracia LM, Martín-Ramos P. Assessment of RGB Vegetation Indices to Estimate Chlorophyll Content in Sugar Beet Leaves in the Final Cultivation Stage. AgriEngineering. 2020; 2(1):128-149. https://doi.org/10.3390/agriengineering2010009.

Santangelo, M., Cardinali, M., Rossi, M., Mondini, A.C., Guzzetti, F., 2010. Remote landslide mapping using a laser rangefinder binocular and GPS. Natural Hazards and Earth System Sciences 10, 2539–2546. doi:10.5194/nhess-10-2539-2010.

Savvaidis, P. D. Existing landslide monitoring systems and techniques. In From Stars to Earth and Culture 242–258 (Academia, 2003).

Segoni, S., Pappafico, G., Luti, T. et al. Landslide susceptibility assessment in complex geological settings: sensitivity to geological information and insights on its parameterization. Landslides 17, 2443–2453 (2020). https://doi.org/10.1007/s10346-019-01340-2.

Sidle, R.C., and Chigira, M. (2004). “Landslides and Debris Flows Strike Kyushu, Japan.” EOS, Transactions, American Geophysical Union.

Sidle RC, Bogaard TA. 2016. Dynamic earth system and ecological controls of rainfall-initiated landslides. Earth-Science Reviews 159: 275–291. https://doi.org/10.1016/j.earscirev.2016.05.013.

Sidle RC, Kamai T, Trandafir AC. 2005. Evaluating landslide damage during the 2004 Chuetsu earthquake, Niigata Japan. Eos, Transactions American Geophysical Union 86: 133–136. https://doi.org/10.1029/2005EO130001.Sidle RC, Gomi T, AkasakaM, Koyanagi K. 2018. Ecosystem changes following the 2016 Kumamoto earthquakes in Japan: future perspectives. Ambio 42: 721–734. https://doi.org/10.1007/s13280-017-1005-8.

Shao, Xiaoyi; Chong Xu, Earthquake-induced landslides susceptibility assessment: A review of the state-of-the-art, Natural Hazards Research, Volume 2, Issue 3, 2022, Pages 172-182, ISSN 2666-5921, https://doi.org/10.1016/j.nhres.2022.03.002.

Shimada M., O. Isoguchi, T. Tadono, and K. Isono, "PALSAR Radiometric and Geometric Calibration", IEEE Trans. Geoscience and Remote Sensing, Vol. 47, No. 12, pp.3915-3932, Dec. 2009.

Shinoda, Masahiro; Yoshihisa Miyata, Ushio Kurokawa, Kenichi Kondo; 2019; Regional landslide susceptibility following the 2016 Kumamoto earthquake using back-calculated geomaterial strength parameters; Landslides 16 (2019); © Springer-Verlag GmbH Germany part of Springer Nature 2019.

Shirahama Y, Yoshimi M, Awata Y, Maruyama T, Azuma T, Miyashita Y et al. 2016. Characteristics of the surface ruptures associated with the 2016 Kumamoto earthquake sequence, central Kyushu, Japan. Earth, Planets and Space 68: 1–12. https://doi.org/10.1186/s40623-016-0559-1.

Solari, L. et al. Review of satellite interferometry for landslide detection in Italy. Remote Sens. 12, 1351 (2020).

Song K, Wang F, Dai Z, Iio A, Osaka O, Sakata S. 2019. Geological characteristics of landslides triggered by the 2016 Kumamoto earthquake in Mt. Aso volcano, Japan. Bulletin of Engineering Geology and the Environment 78: 167–176. https://doi.org/10.1007/s10064-017-1097-1.

Sreelakshmi, S.; Vinod Chandra, S. S.; Shaji, E.; 2022, Landslide identification using machine learning techniques: Review, motivation, and future prospects. Earth Sci Inform 15, 2063–2090 (2022). https://doi.org/10.1007/s12145-022-00889-2.

Stanley, T., Kirschbaum, D.B. A heuristic approach to global landslide susceptibility mapping. Nat Hazards 87, 145–164 (2017). https://doi.org/10.1007/s11069-017-2757-y.

Stanley TA, Kirschbaum DB, Benz G, Emberson RA, Amatya PM, Medwedeff W and Clark MK (2021) Data-Driven Landslide Nowcasting at the Global Scale. Front. Earth Sci. 9:640043. doi: 10.3389/feart.2021.640043.

Stefanescu E. R., Bursik M., Cordoba G., Dalbey K., Jones M. D., Patra A. K., Pieri D. C., Pitman E. B. and Sheridan M. F. 2012Digital elevation model uncertainty and hazard analysis using a geophysical flow model, Proc. R. Soc. A.4681543–1563. http://doi.org/10.1098/rspa.2011.0711.

Tamkuan N, Nagai M. Fusion of Multi-Temporal Interferometric Coherence and Optical Image Data for the 2016 Kumamoto Earthquake Damage Assessment. ISPRS International Journal of Geo-Information. 2017; 6(7):188. https://doi.org/10.3390/ijgi6070188.

Tanyas, Hakan; Mauro Rossi, Massimiliano Alvioli, Cees J. van Westen, Ivan Marchesini, A global slope unit-based method for the near real-time prediction of earthquake-induced landslides, Geomorphology, Volume 327, 2019, Pages 126-146, ISSN 0169-555X, https://doi.org/10.1016/j.geomorph.2018.10.022.

Tehrani FS, Santinelli G, Herrera Herrera M (2021) Multi-regional landslide detection using combined unsupervised and supervised machine learning. Geomat Nat Haz Risk 12(1):1015–1038.

Theodoridis, Sergios, and Konstantinos Koutroumbas. Pattern Recognition. San Diego: Academic Press, 1999: 341-342.

Thirugnanam, H. (2023). Deep Learning in Landslide Studies: A Review. In: , et al. Progress in Landslide Research and Technology, Volume 1 Issue 2, 2022. Progress in Landslide Research and Technology. Springer, Cham. https://doi.org/10.1007/978-3-031-18471-0_20

Uemoto, J., Moriyama, T., Nadai, A. et al. Landslide detection based on height and amplitude differences using pre- and post-event airborne X-band SAR data. Nat Hazards 95, 485–503 (2019). https://doi.org/10.1007/s11069-018-3492-8.

Ullo SL, Mohan A, Sebastianelli A, Ahamed SE, Kumar B, Dwivedi, R, Sinha GR (2021) A new mask r-cnn-based method for improved landslide detection. IEEE J Sel Top Appl Earth Observ Remote Sens 14:3799–3810

USGS ShakeMap: https://earthquake.usgs.gov/data/shakemap/. U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center.

Wang, Haojie; Limin Zhang, Kesheng Yin, Hongyu Luo, Jinhui Li, Landslide identification using machine learning, Geoscience Frontiers, Volume 12, Issue 1, 2021, Pages 351-364, ISSN 1674-9871, https://doi.org/10.1016/j.gsf.2020.02.012.

Water and Disaster Management Bureau. 2016. Overview of sediment related disasters triggered by the 2016 Kumamoto earthquake. Available: http://www.mlit.go.jp/river/sabo/jirei/h28dosha/160914_gaiyou_sokuhou.pdf (in Japanese).

WorldClim Database: https://worldclim.org/data/monthlywth.html.

Xie, W.; Li, X.; Jian, W.; Yang, Y.; Liu, H.; Robledo, L.F.; Nie, W. A novel hybrid method for landslide susceptibility mapping-based geodetector and machine learning cluster: A case of Xiaojin county, China. ISPRS Int. J. Geo-Inf. 2021, 10, 93.

Xie, Peng; Haijia Wen, Chaochao Ma, Laurie G. Baise & Jialan Zhang (2018) Application and comparison of logistic regression model and neural network model in earthquake-induced landslides susceptibility mapping at mountainous region, China, Geomatics, Natural Hazards and Risk, 9:1, 501-523, DOI: 10.1080/19475705.2018.1451399.

Xu, Chong; Siyuan Ma; Zhibiao Tan; Chao Xie; Shinji Toda; Xueqiang Huang; Landslides triggered by the 2016 Mj 7.3 Kumamoto, Japan, earthquake; Landslides (2018) 15:551–564; DOI: 10.1007/s10346-017-0929-1. ©Springer-Verlag GmbH Germany part of Springer Nature 2017.

Yamao, M., Sidle, R.C., Gomi, T., and Imaizumi, F. (2016). “Characteristics of Landslides in Unwelded Pyroclastic Flow Deposits Southern Kyushu, Japan.” Natural Hazards and Earth Systems Sciences.

Yamaguchi Y., “Disaster monitoring by fully polarimetric SAR data acquired with ALOS-PALSAR,” Proc. IEEE, vol. 100, no. 10, pp. 2851–2860, Oct. 2012.

Yamazaki, F.; K. Kubo, R. Tanabe and W. Liu, "Damage assessment and 3d modeling by UAV flights after the 2016 Kumamoto, Japan earthquake," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 3182-3185, doi: 10.1109/IGARSS.2017.8127673.

Yamazaki, Fumio; Yuuki Sagawa, and Wen Liu "Extraction of landslides in the 2016 Kumamoto earthquake using multi-temporal Lidar data", Proc. SPIE 10790, Earth Resources and Environmental Remote Sensing/GIS Applications IX, 107900H (9 October 2018); https://doi.org/10.1117/12.2325091.

Yamazaki, F.; Liu, W.; Horie, K. Use of Multi-Temporal LiDAR Data to Extract Collapsed Buildings and to Monitor Their Removal Process after the 2016 Kumamoto Earthquake. Remote Sens. 2022, 14, 5970. https://doi.org/10.3390/rs14235970.

Yu B, Chen F, Muhammad S, Li B, Wang L, Wu M (2017a) A simple but effective landslide detection method based on image saliency. Photogramm Eng Remote Sens 83(5):351–363.

Zhan, Y.; Liu, W.; Maruyama, Y. Damaged Building Extraction Using Modified Mask R-CNN Model Using Post-Event Aerial Images of the 2016 Kumamoto Earthquake. Remote Sens. 2022, 14, 1002. https://doi.org/10.3390/rs14041002.

Zhong, Cheng; Yue Liu, Peng Gao, Wenlong Chen, Hui Li, Yong Hou, Tuohuti Nuremanguli & Haijian Ma (2020) Landslide mapping with remote sensing: challenges and opportunities, International Journal of Remote Sensing, 41:4, 1555-1581, DOI: 10.1080/01431161.2019.1672904.

## 8.2 Chapter 3

Abdelkader, E. Mohammed, On the hybridization of pre-trained deep learning and differential evolution algorithms for semantic crack detection and recognition in ensemble of infrastructures, Smart and Sustain. Built Environ. (2021), https://doi.org/10.1108/SASBE-01-2021-0010.

Allstadt, K.E., and Thompson, E.M., 2021, Inventory of liquefaction features triggered by the 7 January 2020 M6.4 Puerto Rico earthquake: U.S. Geological Survey data release, https://doi.org/10.5066/P9HZRXI9.

Atzori S, Tolomei C, Antonioli A, Boncori JPM, Bannister S, Trasatti E, Pasquali P, Salvi S. The 2010-2011 Canterbury, New Zealand, seismic sequence: multiple source analysis from InSAR data and modeling. J Geophys Res 2012;117: B08305.

Baik, H.; Son, Y.-S.; Kim, K.-E. Detection of Liquefaction Phenomena from the 2017 Pohang (Korea) Earthquake Using Remote Sensing Data. Remote Sens. 2019, 11, 2184. https://doi.org/10.3390/rs11182184.

Bezdek, James C. Pattern Recognition with Fuzzy Objective Function Algorithms. Boston, MA: Springer US, 1981. https://doi.org/10.1007/978-1-4757-0450-1.

Bi C, Fu B, Chen J, Zhao Y, Yang L, Duan Y, et al. Machine learning based fast multi-layer liquefaction disaster assessment. World Wide Web 2019; 22(5): 1935–50.

Boardman, J. W., & Kruse, F. A., 1994, Automated Spectral Analysis: A Geological Example Using AVIRIS Data, North Grapevine Mountains, Nevada: in Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing, Environmental Research Institute of Michigan, Ann Arbor, MI, pp. I-407 - I-418.

Brandenberg SJ, Zimmaro P, Stewart JP, et al. Next-generation liquefaction database. Earthquake Spectra. 2020;36(2):939-959. https://doi.org/10.1177/8755293020902477.

Brown LJ, Weeber JH. 1992. Geology of the Christchurch Urban area. Scale 1:25 000. Institute of Geological and Nuclear Sciences geological map 1. 1 sheet + 104 p.

Brown LJ. 2001. Canterbury. In: Rosen MA, White PA, editors. Groundwaters of New Zealand. Wellington: New Zealand Hydrological Society Inc; p. 441–459.

Burger, Wilhelm; Mark J. Burge (2010). Principles of Digital Image Processing Core Algorithms. Springer Science & Business Media. pp. 110–111. ISBN 978-1-84800-195-4.

Canterbury Geotechnical Database. 2013. https://canterburygeotechnicaldatabase.projectorbit.com.

Chen, Runnan; Yuexin Ma, Lingjie Liu, Nenglun Chen, Zhiming Cui, Guodong Wei, Wenping Wang, Semi-supervised anatomical landmark detection via shape-regulated self-training, Neurocomputing, Volume 471, 2022, Pages 335-345, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2021.10.109.

CGD, 2013; Canterbury Geotechnical Database Liquefaction Interpreted from Aerial Photography, Map Layer CGD0200 - 11 Feb 2013 ⟨https://canterburygeotechnicaldatabase.projectorbit.com⟩.

Chleborad, A. F., Schuster, R. L. (1990). Ground failure associated with the Puget Sound region earthquakes of April 13, 1949, and April 29, 1965, U.S. Geological Survey Open File Report 90-687.

Cox SC, Rutter HK, Sims A, Manga M, Weir JJ, Ezzy T, White PA, Horton TW, Scott D. 2012. Hydrological effects of the Mw 7.1 Darfield (Canterbury) earthquake, 4 September

2010, New Zealand. New Zeal J. Geol. Geophy. 55:231–247. https://doi.org/10.1080/00288306.2012.680474.

Cubrinovski M, Bray JD, Taylor M, Giorgini S, Bradley BA, Wotherspoon L, Zupan J. 2011. Soil liquefaction effects in the central business district during the February 2011 Christchurch Earthquake. Seismol Res Lett. 82:893–904. https://doi.org/10.1785/gssrl.82.6.893.

Ding, C., and H. Peng. "Minimum redundancy feature selection from microarray gene expression data." Journal of Bioinformatics and Computational Biology. Vol. 3, Number 2, 2005, pp. 185–205.

Darbellay, G. A., and I. Vajda. "Estimation of the information by an adaptive partitioning of the observation space." IEEE Transactions on Information Theory. Vol. 45, Number 4, 1999, pp. 1315–1321.

Dumitru, C. O.; S. Cui, D. Faur and M. Datcu, "Data Analytics for Rapid Mapping: Case Study of a Flooding Event in Germany and the Tsunami in Japan Using Very High-Resolution SAR Images," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, no. 1, pp. 114-129, Jan. 2015, https://doi.org/10.1109/JSTARS.2014.2320777.

ElGharbawi, Tamer; Masayuki Tamura; Estimating deformation due to soil liquefaction in Urayasu city, Japan using permanent scatterers, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 109, 2015, Pages 152-164, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2015.09.002.

Fairchild, Mark D. (2005). "Color and Image Appearance Models". Color Appearance Models. John Wiley and Sons. p. 340. ISBN 0-470-01216-1.

Feng, Z., Zhou, Q., Gu, Q., Tan, X., Cheng, G., Lu, X., ... & Ma, L. (2022). Dmt: Dynamic mutual training for semi-supervised learning. Pattern Recognition, 130, 108777.

Fisher, R. A. "The Use of Multiple Measurements in Taxonomic Problems." Annals of Eugenics, Vol. 7, pp. 179–188, 1936. Available at https://digital.library.adelaide.edu.au/dspace/handle/2440/15227.

Fowles, G. R. Introduction to Modern Optics. New York: Dover, 1989.

Geyin M, Maurer BW, Bradley BA, Green RA, van Ballegooy S. CPT-based liquefaction case histories compiled from three earthquakes in Canterbury, New Zealand. Earthquake Spectra. 2021;37(4):2920-2945. https://doi.org/10.1177/8755293021996367.

Geyin M, Maurer B, Christofferson K. An AI driven, mechanistically grounded geospatial liquefaction model for rapid response and scenario planning. Soil Dyn Earthq Eng 2022; 159:107348.

Ghosh S, Huyck CK, Greene M, Gill SP, Bevington J, Svekla W, DesRoches R, Eguchi RT. Crowdsourcing for rapid damage assessment: the global earth observation catastrophe assessment network (GEO-CAN). Earthq Spectra 2011;27(S1):S179–98.

Gnädinger, F.; Schmidhalter, U. Digital Counts of Maize Plants by Unmanned Aerial Vehicles (UAVs). Remote Sens. 2017, 9, 544. https://doi.org/10.3390/rs9060544.

Gonzalez, R. C., R. E. Woods, and S. L. Eddins. Digital Image Processing Using MATLAB. New Jersey, Prentice Hall, 2003, Chapter 11.

Green RA, Cubrinovski M, Cox B, et al. Select Liquefaction Case Histories from the 2010–2011 Canterbury Earthquake Sequence. Earthquake Spectra. 2014;30(1):131-153. https://doi.org/10.1193/030713EQS066M.

Guo, Y., T. Hastie, and R. Tibshirani. "Regularized linear discriminant analysis and its application in microarrays." Biostatistics, Vol. 8, No. 1, pp. 86–100, 2007.

Hacıefendioğlu, K., Başağa, H.B. & Demir, G. Automatic detection of earthquake-induced ground failure effects through Faster R-CNN deep learning-based object detection using satellite images. Nat Hazards 105, 383–403 (2021). https://doi.org/10.1007/s11069-020-04315-y.

Haralick, R.M., K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification", IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, 1973, pp. 610-621.

Hamada M, Towhata I, Yasuda S, Isoyama R. Study on permanent ground displacement induced by seismic liquefaction. Comput Geotech 1987; 4(4):197–220.

Hastie, T., R. Tibshirani, and J. Friedman. The Elements of Statistical Learning, second edition. New York: Springer, 2008.

Ishitsuka K, Tsuji T, Matsuoka T. Detection and mapping of soil liquefaction in the 2011 Tohoku earthquake using SAR interferometry. Earth Planets Space 2012; 64:1267–76.

Jain, Anil K., Fundamentals of Digital Image Processing, Englewood Cliffs, NJ, Prentice Hall, 1989, pp. 150–153.

Gradient Weight, MATLAB Toolbox.

Jain, Anil K., and Farshid Farrokhnia. "Unsupervised Texture Segmentation Using Gabor Filters." Pattern Recognition 24, no. 12 (January 1991): 1167–86. https://doi.org/10.1016/0031-3203(91)90143-S.

Jolliffe, I. T. Principal Component Analysis. 2nd ed., Springer, 2002.

Kajihara, Kazuhiro; Hiroki Okuda, Takashi Kiyota, Kazuo Konagai, Mapping of liquefaction risk on road network based on relationship between liquefaction potential and liquefaction-induced road subsidence, Soils and Foundations, Volume 60, Issue 5, 2020, Pages 1202-1214, ISSN 0038-0806, https://doi.org/10.1016/j.sandf.2020.07.007.

Kang, Henry R. (1999). Digital Color Halftoning. SPIE Press. p. 1. ISBN 0-8194-3318-7.

Khorrami, R.A., Naeimi, Z., Wing, M., Ahani, H. and Bateni, S.M. (2022) A New Multistep Approach to Identify Leaf-Off Poplar Plantations Using Airborne Imagery. Journal of Geographic Information System, 14, 634-651. https://doi.org/10.4236/jgis.2022.146036.

Kumar, Chandan; Snehamoy Chatterjee, Thomas Oommen, Arindam Guha & Abhijeet Mukherjee (2021): Multi-sensor datasets-based optimal integration of spectral, textural, and morphological characteristics of rocks for lithological classification using machine learning models, Geocarto International, https://doi.org/10.1080/10106049.2021.1920632.

Lu, C.-H.; Ni, C.-F.; Chang, C.-P.; Yen, J.-Y.; Chuang, R.Y. Coherence Difference Analysis of Sentinel-1 SAR Interferogram to Identify Earthquake-Induced Disasters in Urban Areas. Remote Sens. 2018, 10, 1318. https://doi.org/10.3390/rs10081318

Mallat Stéphane, CHAPTER 12 - Sparsity in Redundant Dictionaries, Editor(s): Mallat Stéphane, A Wavelet Tour of Signal Processing (Third Edition), Academic Press, 2009, Pages 611-698, ISBN 9780123743701, https://doi.org/10.1016/B978-0-12-374370-1.00016-1.

Meyer, Y. Wavelets and Operators. Translated by D. H. Salinger. Cambridge, UK: Cambridge University Press, 1995.

Villemure M, Tom Wilson, D Bristow, Charlotte Brown, https://www.researchgate.net/publication/284714549_Liquefaction_ejecta_clean-up_in_Christchurch_during_the_2010-2011_earthquake_sequence_2012_NZSEE_Conference.

Morgenroth, J., Hughes, M.W. & Cubrinovski, M. Object-based image analysis for mapping earthquake-induced liquefaction ejecta in Christchurch, New Zealand. Nat Hazards 82, 763–775 (2016). https://doi.org/10.1007/s11069-016-2217-0.

New Zealand Aerial Mapping (NZAM), 2011.

New Zealand Geotechnical Database. Liquefaction and Lateral Spreading Observations. Map Layer NZGD0300 - 11 Feb 2013. Available online: https://www.nzgd.org.nz/Map (accessed on 22 January 2020).

Nhat-Duc, Hoang; Tran Van-Duc, Comparison of histogram-based gradient boosting classification machine, random Forest, and deep convolutional neural network for

pavement raveling severity classification, Automation in Construction, Volume 148, 2023, 104767, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2023.104767.

Njock PGA, Shen S-L, Zhou A, Lyu H-M. Evaluation of soil liquefaction using AI technology incorporating a coupled ENN/t-SNE model. Soil Dyn Earthq Eng., 2020; 130:105988.

Oludare, Victor; Landry Kezebou, Karen Panetta, Sos Agaian, "Semisupervised learning for improved post-disaster damage assessment from satellite imagery," Proc. SPIE 11734, Multimodal Image Exploitation and Learning 2021, 117340O (12 April 2021); https://doi.org/10.1117/12.2586232.

Oommen, T.; Baise, L.G.; Gens, R.; Prakash, A.; Gupta, R.P. Documenting earthquake-induced liquefaction using satellite remote sensing image transformations. Environ. Eng. Geosci. 2013, 19, 303–318.

Orense, Rolando P.; Nathan A. Hickman, Brian T. Hill & Michael J. Pender (2014) Spatial evaluation of liquefaction potential in Christchurch following the 2010/2011 Canterbury earthquakes, International Journal of Geotechnical Engineering, 8:4, 420-425, https://doi.org/10.1179/1939787913Y.0000000028.

Ponti, D. J., J. L. Blair, C. M. Rosa, K. Thomas, A. J. Pickering, S. Akciz, S. Angster, J.-P. Avouac, J. Bachhuber, S. Bacon, et al. (2020). Documentation of surface fault rupture and ground deformation features produced by the Ridgecrest M 6.4 and M 7.1 earthquake sequence of July 4 and 5, 2019, Seismol. Res. Lett., https://doi.org/10.1785/0220190322.

Ramakrishnan D, Mohanty KK, Natak SR, Vinu Chandran R. Mapping the liquefaction induced soil moisture changes using remote sensing technique: an attempt to map the earthquake induced liquefaction around Bhuj, Gujarat, India. Geotech Geol Eng 2006;24(6):1581–602.

Rashidian V, Baise LG, Koch M. Using High Resolution Optical Imagery to Detect Earthquake-Induced Liquefaction: The 2011 Christchurch Earthquake. Remote Sensing. 2020; 12(3):377. https://doi.org/10.3390/rs12030377.

Rathje, Ellen M.; Kevin Franke, Remote sensing for geotechnical earthquake reconnaissance, Soil Dynamics and Earthquake Engineering, Volume 91, 2016, Pages 304-316, ISSN 0267-7261, https://doi.org/10.1016/j.soildyn.2016.09.016.

Rollins K, Ledezma C, Montalva G. (Eds). Geotechnical aspects of April 1, 2014, M8.2 Iquique, Chile Earthquake, a report of the NSF- Sponsored GEER Association Team. GEER Assoc. Report No. GEER-038, 2014. ⟨http://www.geerassociation.org/GEER_Post%20EQ%20Reports/Iquique_Chile_2014⟩.

Saraf, A. K.; A. Sinvhal , H. Sinvhal , P. Ghosh & B. Sarma (2002); Satellite data reveals 26 January 2001 Kutch earthquake-induced ground changes and appearance of water

bodies, International Journal of Remote Sensing, 23:9, 1749-1756, https://doi.org/10.1080/01431160110107824.

Sengar, S. S.; Anil Kumar, Sanjay K. Ghosh, Hans Raj Wason, and Partha S. Roy; 2012-a "Liquefaction identification using class-based sensor independent approach based on single pixel classification after 2001 Bhuj, India earthquake," Journal of Applied Remote Sensing 6(1), 063531 (21 May 2012). https://doi.org/10.1117/1.JRS.6.063531.

Sengar, S. S., Kumar, A., Ghosh, S. K., and Wason, H. R., 2012-b: SOFT COMPUTING APPROACH FOR LIQUEFACTION IDENTIFICATION USING LANDSAT-7 TEMPORAL INDICES DATA, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XXXIX-B8, 61–64, https://doi.org/10.5194/isprsarchives-XXXIX-B8-61-2012.

Sharifi-Mood, Mahyar; Daniel T. Gillins, Kevin W. Franke, Jasmyn N. Harper, Steven J. Bartlett, Michael J. Olsen, Probabilistic liquefaction-induced lateral spread hazard mapping and its application to Utah County, Utah, Engineering Geology, Volume 237, 2018, Pages 76-91, ISSN 0013-7952, https://doi.org/10.1016/j.enggeo.2018.01.015.

Smith, A. R. “Color Gamut Transform Pairs”. SIGGRAPH 78 Conference Proceedings. 1978, pp. 12–19.

Stewart, J. P. (Editor), S. J. Brandenberg, P. Wang, C. C. Nweke, K. S. Hudson, S. Mazzoni, Y. Bozorgnia, K. W. Hudnut, C. A. Davis, S. K. Ahdi, et al. (2019). Preliminary report on engineering and geological effects of the July 2019 Ridgecrest earthquake sequence, Geotechnical Extreme Events Reconnaissance Association, Rept. GEER-064, https://doi.org/10.18118/G6H66K.

Townsend, D.; JM Lee, DT Strong, R Jongens, B Smith Lyttle, S Ashraf, B Rosser, N Perrin, K Lyttle, M Cubrinovski, ML Taylor, MW Hughes, T Wilson, P Almond, M Jacka, I McCahon & S Christensen (2016) Mapping surface liquefaction caused by the September 2010 and February 2011 Canterbury earthquakes: a digital dataset, New Zealand Journal of Geology and Geophysics, 59:4, 496-513, https://doi.org/10.1080/00288306.2016.1182929.

van Ballegooy S, Malan P, Lacrosse V, Jacka ME, Cubrinovski M, Bray JD, O’Rourke TD, Crawford SA, Cowan H. 2014. Assessment of liquefaction-induced land damage for residential Christchurch. Earthquake Spectra. 30:31–55.

Wu, D.; X. Luo, G. Wang, M. Shang, Y. Yuan, H. Yan, A highly accurate framework for self-labeled semi-supervised classification in industrial applications, IEEE Transactions on Industrial Informatics 14 (3) (2017) 909–920.

Xu, S., Dimasaka, J., Wald, D. J., & Noh, H. Y. (2022). Bayesian Updating of Seismic Ground Failure Estimates via Causal Graphical Models and Satellite Imagery. arXiv preprint arXiv:2204.07653.

Zhang, W.; F. Ghahari, P. Arduino, E. Taciroglu, A deep learning approach for rapid detection of soil liquefaction using time–frequency images, Soil Dynamics and Earthquake Engineering, Volume 166, 2023, 107788, ISSN 0267-7261, https://doi.org/10.1016/j.soildyn.2023.107788.

Zhu, J., Daley, D., Baise, L.G., Thompson, E.M., Wald, D.J., and Knudsen, K.L., 2015, A geospatial liquefaction model for rapid response and loss estimation: Earthquake Spectra, v. 31, no. 3, p. 1813–1837.

Zhu, J.; L. G. Baise and M. Koch, "Mapping earthquake induced liquefaction surface effects from the 2011 Tohoku earthquake using satellite imagery," 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 2016, pp. 2328-2331, https://doi.org/10.1109/IGARSS.2016.7729601.

Zhu, J., Baise, L. G., Thompson, E. M., 2017, An Updated Geospatial Liquefaction Model for Global Application, Bulletin of the Seismological Society of America, 107, p 1365-1385, https://doi.org/0.1785/0120160198.

Zimmaro, Paolo; Chukwuebuka C. Nweke, Janis L. Hernandez, Kenneth S. Hudson, Martin B. Hudson, Sean K. Ahdi, Matthew L. Boggs, Craig A. Davis, Christine A. Goulet, Scott J. Brandenberg, Kenneth W. Hudnut, Jonathan P. Stewart; Liquefaction and Related Ground Failure from July 2019 Ridgecrest Earthquake Sequence. Bulletin of the Seismological Society of America 2020; 110 (4): 1549–1566. https://doi.org/10.1785/0120200025.

## 8.3 Chapter 4

Aimaiti, Y.; Liu, W.; Yamazaki, F.; Maruyama, Y. Earthquake-Induced Landslide Mapping for the 2018 Hokkaido Eastern Iburi Earthquake Using PALSAR-2 Data. Remote Sens. 2019, 11, 2351. https://doi.org/10.3390/rs11202351.

Ayalew, L., Kasahara, M., Yamagishi, H., 2011. The spatial correlation between earthquakes and landslides in Hokkaido (Japan), a GIS-based analysis of the past and the future. Landslides https://doi.org/10.1007/s10346-011-0262-z.

Bai, Y.; Adriano, B.; Mas, E.; Koshimura, S. Machine learning based building damage mapping from the ALOS-2/PALSAR-2 SAR imagery: Case study of 2016 Kumamoto earthquake. J. Disaster Res. 2017, 12, 646–655.

Barbu, Tudor, (2013) "Variational Image Denoising Approach with Diffusion Porous Media Flow", Abstract and Applied Analysis, vol. 2013, Article ID 856876, 8 pages, 2013. https://doi.org/10.1155/2013/856876.

Bhuyan, Kushanav; Sansar Raj Meena, Lorenzo Nava, Cees J. van Westen , Mario Floris, Filippo Catani, (2022), Mapping landslides through a temporal lens: An insight towards multi-temporal landslide mapping using the U-Net deep learning model. https://doi.org/10.31223/X5DM0B.

Bhuyan, K., Tanyaş, H., Nava, L. et al. (2023) Generating multi-temporal landslide inventories through a general deep transfer learning strategy using HR EO data. Sci Rep 13, 162 (2023). https://doi.org/10.1038/s41598-022-27352-y.

Bishop, C. M., (2006) Pattern Recognition and Machine Learning. Springer, New York, NY, 2006.

Boehm, V., Leong, W. J., Mahesh, R. B., Prapas, I., Nemni, E., Kalaitzis, F., ... & Ramos-Pollan, R. (2022). Deep Learning for Rapid Landslide Detection using Synthetic Aperture Radar (SAR) Datacubes. arXiv preprint arXiv:2211.02869.

Casagli, N., Intrieri, E., Tofani, V. et al. Landslide detection, monitoring and prediction with remote-sensing techniques. Nat Rev Earth Environ 4, 51–64 (2023). https://doi.org/10.1038/s43017-022-00373-x.

Chen C-W, Chen H, Wei L-W, Lin G-W, Iida T, Yamada R. (2017-a). Evaluating the susceptibility of landslide landforms in Japan using slope stability analysis: a case study of the 2016 Kumamoto earthquake. Landslides 14: 1793–1801. https://doi.org/10.1007/s10346-017-0872-1.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017-b). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.

Chen, L. C.; Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, (2018) "Encoderdecoder with atrous separable convolution for semantic image segmentation," in Proc. Eur. Conf. Comput. Vis., 2018, pp. 801–818.

Comert R (2021) Investigation of the Effect of the Dataset Size and Type in the Earthquake-Triggered Landslides Mapping: A Case Study for the 2018 Hokkaido Iburu Landslides. Front. Earth Sci. 9:633665. https://doi.org/10.3389/feart.2021.633665.

Cruden DM, Varnes DJ (1996) Landslides: investigation and mitigation. Chapter 3: Landslide Types and Processes. Transportation Research Board Special Report, 247.

Cui, Y., Bao, P., Xu, C. et al. (2021) Landslides triggered by the 6 September 2018 Mw 6.6 Hokkaido, Japan: an updated inventory and retrospective hazard assessment. Earth Sci Inform 14, 247–258 (2021). https://doi.org/10.1007/s12145-020-00544-8.

Deng, J.; W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. (2009) ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009. (www.image-net.org)

Dou, J., K. T. Chang, S. Chen, A. Yunus, J. K. Liu, H. Xia, and Z. Zhu. 2015. "Automatic Case-Based Reasoning Approach for Landslide Detection: Integration of Object-Oriented Image Analysis and a Genetic Algorithm." Remote Sensing 7 (4): 4318–4342. doi:10.3390/rs70404318.

Dou J, Yunus AP, Bui DT, Merghadi A, Sahana M, Zhu Z, Chen C-W, Han Z, Pham BT (2020-a) Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan. Landslides 17(3):641–658.

Dou, Jie; Ali P. Yunus, Abdelaziz Merghadi, Ataollah Shirzadi, Hoang Nguyen, Yawar Hussain, Ram Avtar, Yulong Chen, Binh Thai Pham, Hiromitsu Yamagishi, (2020-b) Different sampling strategies for predicting landslide susceptibilities are deemed less consequential with deep learning, Science of The Total Environment, Volume 720, 2020, 137320, ISSN 0048-9697, https://doi.org/10.1016/j.scitotenv.2020.137320.

Fang, Z.; Wang, Y.; Peng, L.; Hong, H. Integration of convolutional neural network and conventional machine learning classifiers for landslide susceptibility mapping. Comput. Geosci. 2020, 139, 104470.

Fang, C.; Fan, X.; Zhong, H.; Lombardo, L.; Tanyas, H.; Wang, X. A Novel Historical Landslide Detection Approach Based on LiDAR and Lightweight Attention U-Net. Remote Sens. 2022, 14, 4357. https://doi.org/10.3390/rs14174357.

Fanos AM, Pradhan B, Mansor S, Yusoff ZM, Abdullah AF (2018), A hybrid model using machine learning methods and GIS for potential rockfall source identification from airborne laser scanning data. Landslides 15(9):1833–1850.

Fu, R.; He, J.; Liu, G.; Li, W.; Mao, J.; He, M.; Lin, Y. Fast Seismic Landslide Detection Based on Improved Mask R-CNN. Remote Sens. 2022, 14, 3928. https://doi.org/10.3390/rs14163928.

GEER Report, 2016; Geotechnical Aspects of the 2016 MW 6.2, MW 6.0, and MW 7.0 Kumamoto Earthquakes GEER Association; geerassociation.org; Kumamoto Earthquakes, Version 1.0, July 2016. Kayen R, Dashti S, Kokusho T, Hazarika H, Franke K, Oettle N, Wham B, Calderon JR, Briggs D, Guillies S, Cheng K, Tanoue Y, Takematsu K, Matsumoto D, Morinaga T, Furuichi H, Kitano Y, Tajiri M, Chaudhary B, Nishimura K, Chu C (2016) Geotechnical aspects of the 2016 Mw 6.2, Mw 6.0, and Mw 7.0 Kumamoto earthquakes. Geotechnical Extreme Events Reconnaissance Association, Version 1.0, July 2016. https://pubs.er.usgs.gov/publication/70185571.

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 Sens. https://doi.org/10.3390/rs11020196.

Ghorbanzadeh, O.; A. Crivellari, P. Ghamisi, H. Shahabi, and T. Blaschke, (2021) “A comprehensive transferability evaluation of U-Net and ResU-Net for landslide detection from Sentinel-2 data (case study areas from Taiwan, China, and Japan),” Scientific Reports, vol. 11, no. 1, pp. 1–20, 2021.

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

GSI (2016-a), Information on the 2016 Kumamoto Earthquakes, Japan, “平成２８年熊本地震に関する情報”. Geospatial Information Authority of Japan (GSI), http://www.gsi.go.jp/BOUSAI/H27-kumamoto-earthquake-index.html.

GSI (2016-b), Aerial imagery for 2016 Kumamoto Earthquakes, Japan, Geospatial Information Authority of Japan (GSI) – in Japanese: https://maps.gsi.go.jp/development/ichiran.html#fukkyukizu.

GSI (2018-a), "The 2018 Hokkaido eastern Iburi earthquake: fault model (preliminary)." Geospatial Information Authority of Japan, GSI 2018. Available: https://www.gsi.go.jp/cais/topic180912-index-e.html.

GSI (2018-b), Geospatial Information Authority of Japan (GSI), “Distribution map of landslides around Atsuma Town caused by the 2018 Hokkaido Eastern Iburi Earthquake,” 2018. https://saigai.gsi.go.jp/3/20180906/iburi hokai_2-zentaizu.pdf. (In Japanese)

GSI (2018-c), Geospatial Information Authority of Japan (GSI). 2018-Hokkaido Eastern Iburi Earthquake. Available online: http://www.gsi.go.jp/BOUSAI/H30-hokkaidoiburi-east-earthquake-index.html#1. (Accessed on 25 Feb. 2023).

Guzzetti, Fausto, 2006, Ph.D. dissertation on Landslide Hazard and Risk Assessment, at the Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität University of Bonn, Bonn, Germany, defended on July 2006.

Hacıefendioğlu, K., Demir, G. & Başağa, H.B. (2021) Landslide detection using visualization techniques for deep convolutional neural network models. Nat Hazards 109, 329–350 (2021). https://doi.org/10.1007/s11069-021-04838-y.

Haddad, R.A., and A.N. Akansu, (1991) "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 39, pp 723-727, March 1991.

He, K.; X. Zhang, S. Ren, and J. Sun, (2016) “Deep residual learning for image recognition,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016, pp. 770–778.

He, Ping; Yangmao Wen, Caijun Xu, Yunguo Chen, Complete three-dimensional near-field surface displacements from imaging geodesy techniques applied to the 2016 Kumamoto earthquake, Remote Sensing of Environment, Volume 232, 2019, 111321, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2019.111321.

Hölbling, D. et al. Comparing manual and semi-automated landslide mapping based on optical satellite images from different sensors. Geosciences 7, 37 (2017).

Huang, L., Xiang, L. yang, (2018-a). Method for meteorological early warning of precipitation induced landslides based on deep neural network. Neural. Process. Lett. https://doi.org/10.1007/s11063-017-9778-0.

Huang, B., Zhao, B., Song, Y., (2018-b). Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. Remote Sens. Environ. 214, 73–86. http://dx.doi.org/10.1016/j.rse.2018.04.050.

Ji, Y.; Sumantyo, J.T.S.; Chua, M.Y.; Waqar, M.M. Earthquake/tsunami damage level mapping of urban areas using full polarimetric SAR data. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2018, 11, 2296–2309.

Jolliffe, I. T. Principal Component Analysis. 2nd ed., Springer, 2002.

Kameda, J., Kamiya, H., Masumoto, H. et al. Fluidized landslides triggered by the liquefaction of subsurface volcanic deposits during the 2018 Iburi–Tobu earthquake, Hokkaido. Sci Rep 9, 13119 (2019). https://doi.org/10.1038/s41598-019-48820-y.

Keefer, D.K., 1984. Landslides Caused by Earthquakes. pp. 406–421.

Kim, Jae Sung and KyoHyouk Kim "Analysis of 2016 Minamiaso landslides using remote sensing and geographic information system," Journal of Applied Remote Sensing 12(3), 036001 (4 July 2018). https://doi.org/10.1117/1.JRS.12.036001.

Kiyota, T., Ikeda, T., Konagai, K. and Shiga, M. (2017). Geotechnical Damage Caused by the 2016 Kumamoto Earthquake, Japan. International Journal of Geoengineering Case histories, http://casehistories.geoengineer.org, Vol.4, Issue 2, p.78-95. https://doi.org/10.4417/IJGCH-04-02-01.

Konishi, Tomohisa; and Yuzo Suga "Landslide detection using polarimetric ALOS-2/PALSAR-2 data: a case study of 2016 Kumamoto earthquake in Japan", Proc. SPIE 10788, Active and Passive Microwave Remote Sensing for Environmental Monitoring II, 107880P (9 October 2018); https://doi.org/10.1117/12.2324030.

Koyanagi, Kenta; Takashi Gomi; Roy C. Sidle; 2020; Characteristics of landslides in forests and grasslands triggered by the 2016 Kumamoto; EARTH SURFACE PROCESSES AND LANDFORMS; Earth Surf. Process. Landforms 45, 893–904 (2020); https://doi.org/10.1002/esp.4781. © 2019 John Wiley & Sons, Ltd.

Kurtz, C., Stumpf, A., Malet, J.-P., Gançarski, P., Puissant, A., Passat, N., 2014. Hierarchical extraction of landslides from multiresolution remotely sensed optical images. ISPRS J. Photogramm. Remote Sens. 87, 122–136. http://dx.doi.org/10.1016/j.isprsjprs.2013.11.003.

Lei, T.; Y. Zhang, Z. Lv, S. Li, S. Liu, and A. K. Nandi, (2019) “Landslide inventory mapping from bitemporal images using deep convolutional neural networks,” IEEE

Geosci. Remote Sens. Lett., vol. 16, no. 6, pp. 982–986, 2019. https://doi.org/10.1109/LGRS.2018.2889307.

Liu W. and F. Yamazaki, "Detection of Landslides Induced by the 2018 Hokkaido Eastern Iburi Earthquake Using Multi-Temporal ALOS-2 imagery," IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 2020, pp. 4096-4099, http://dx.doi.org/10.1109/IGARSS39084.2020.9323252.

Li, Y., Shi, T., Zhang, Y., Chen, W., Wang, Z., Li, H., (2021). Learning deep semantic segmentation network under multiple weakly-supervised constraints for crossdomain remote sensing image semantic segmentation. ISPRS J. Photogramm. Remote Sens. 175, 20–33. http://dx.doi.org/10.1016/j.isprsjprs.2021.02.009.

Li, H., He, Y., Xu, Q. et al. (2022) Detection and segmentation of loess landslides via satellite images: a two-phase framework. Landslides 19, 673–686 (2022). https://doi.org/10.1007/s10346-021-01789-0.

Liu, Wen; Fumio Yamazaki and Yoshihisa Maruyama; 2019; Detection of Earthquake-Induced Landslides during the 2016 Kumamoto Earthquake Using Multitemporal Airborne Lidar Data; Remote Sens. 2019, 11, 2292; https://doi.org/10.3390/rs11192292.

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. 2020; 12(5):894. https://doi.org/10.3390/rs12050894.

Liu Z, Gilbert G, Cepeda JM, Lysdahl AOK, Piciullo L, Hefre H, Lacasse S (2021) Modelling of shallow landslides with machine learning algorithms. Geosci Front 12(1):385–393.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International conference on machine learning (pp. 448-456).

Lu H, Ma L, Fu X, Liu C, Wang Z, Tang M, Li N. (2020) Landslides Information Extraction Using Object-Oriented Image Analysis Paradigm Based on Deep Learning and Transfer Learning. Remote Sensing. 2020; 12(5):752. https://doi.org/10.3390/rs12050752.

Lv, Z.; T. Liu, X. Kong, C. Shi, and J. A. Benediktsson, “Landslide inventory mapping with bitemporal aerial remote sensing images based on the dual-path fully convolutional network,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 13, pp. 4575–4584, 2020.

Martha, T. R., Kerle, N., Van Westen, C. J., Jetten, V. & Kumar, K. V. Object-oriented analysis of multi-temporal panchromatic images for creation of historical landslide inventories. ISPRS J. Photogramm. Remote Sens. 67, 105–119 (2011).

Meena, S.R., Soares, L.P., Grohmann, C.H. et al. Landslide detection in the Himalayas using machine learning algorithms and U-Net. Landslides 19, 1209–1229 (2022). https://doi.org/10.1007/s10346-022-01861-3.

Miura T, Nagai S. (2020) Landslide Detection with Himawari-8 Geostationary Satellite Data: A Case Study of a Torrential Rain Event in Kyushu, Japan. Remote Sensing. 2020; 12(11):1734. https://doi.org/10.3390/rs12111734.

Mohan, A, Singh, AK, Kumar, B, Dwivedi, R. Review on remote sensing methods for landslide detection using machine and deep learning. Trans Emerging Tel Tech. 2021; 32:e3998. https://doi.org/10.1002/ett.3998.

Mondini, A. C. et al. Landslide failures detection and mapping using synthetic aperture radar: past, present and future. Earth Sci. Rev. 216, 103574 (2021).

Moosavi, V., Talebi, A., Shirmohammadi, B., 2014. Producing a landslide inventory map using pixel-based and object-oriented approaches optimized by Taguchi method. Geomorphology 204, 646–656.

Murphy, K. P. (2012) Machine Learning: A Probabilistic Perspective. Cambridge, Massachusetts: The MIT Press, 2012.

Nagi, J., F. Ducatelle, G. A. Di Caro, D. Ciresan, U. Meier, A. Giusti, F. Nagi, J. Schmidhuber, L. M. Gambardella. "Max-Pooling Convolutional Neural Networks for Vision-based Hand Gesture Recognition". IEEE International Conference on Signal and Image Processing Applications (ICSIPA2011), 2011.

Nair, V. and G. E. Hinton. "Rectified linear units improve restricted boltzmann machines." In Proc. 27th International Conference on Machine Learning, 2010.

Nava, L.; Bhuyan, K.; Meena, S.R.; Monserrat, O.; Catani, F. Rapid Mapping of Landslides on SAR Data by Attention U-Net. Remote Sens. 2022, 14, 1449. https://doi.org/10.3390/rs14061449.

Nava, L.; O. Monserrat and F. Catani, (2022-b), "Improving Landslide Detection on SAR Data Through Deep Learning," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 4020405, https://doi.org/10.1109/LGRS.2021.3127073.

NIED, 2016, Distribution map of mass movement by the 2016 Kumamoto earthquake, edited by National Research Institute for Earth Science and Disaster of Japan (in Japanese), National Research Institute for Earth Science and Disaster Resilience (NIED). Retrieved from: http://www.bosai.go.jp/mizu/dosha.html.

Ohki, M., Abe, T., Tadono, T. et al. Landslide detection in mountainous forest areas using polarimetry and interferometric coherence. Earth Planets Space 72, 67 (2020). https://doi.org/10.1186/s40623-020-01191-5.

Osanai, N., Yamada, T., Hayashi, Si. et al. (2019) Characteristics of landslides caused by the 2018 Hokkaido Eastern Iburi Earthquake. Landslides 16, 1517–1528 (2019). https://doi.org/10.1007/s10346-019-01206-7.

Peng, D., Guan, H., Zang, Y., Bruzzone, L., 2022. Full-level domain adaptation for building extraction in very-high-resolution optical remote-sensing images. IEEE Trans. Geosci. Remote Sens. 60, 1–17. http://dx.doi.org/10.1109/TGRS.2021.3093004.

Qi, W.; M. Wei, W. Yang, C. Xu, and C. Ma, (2020) "Automatic mapping of landslides by the ResU-Net" Remote Sensing, vol. 12, no. 15, p. 2487, 2020. https://doi.org/10.3390/rs12152487.

Ronneberger, O.; P. Fischer, and T. Brox, (2015) "U-net: Convolutional networks for biomedical image segmentation," in International Conference of Medical image computing and computer-assisted intervention. Springer, 2015, pp. 234–241.

Selvaraju RR, Cogswell M, Das A et al (2020) Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput vis 128:336–359. https://doi.org/10.1007/s11263-019-01228-7.

Savvaidis, P. D. Existing landslide monitoring systems and techniques. In From Stars to Earth and Culture 242–258 (Academia, 2003).

Soares, L. P.; H. C. Dias, and C. H. Grohmann, (2020) "Landslide segmentation with U-Net: Evaluating different sampling methods and patch sizes," arXiv preprint arXiv:2007.06672, 2020.

Solari, L. et al. Review of satellite interferometry for landslide detection in Italy. Remote Sens. 12, 1351 (2020).

Su, Z.; J. K. Chow, P. S. Tan, J. Wu, Y. K. Ho, and Y.-H. Wang, (2021) "Deep convolutional neural network–based pixel wise landslide inventory mapping", Landslides, vol. 18, no. 4, pp. 1421–1443, 2021. https://doi.org/10.1007/s10346-020-01557-6.

Tehrani FS, Santinelli G, Herrera Herrera M (2021) Multi-regional landslide detection using combined unsupervised and supervised machine learning. Geomat Nat Haz Risk 12(1):1015–1038.

Tong, X.-Y., Xia, G.-S., Lu, Q., Shen, H., Li, S., You, S., Zhang, L., 2020. Land-cover classification with high-resolution remote sensing images using transferable deep models. Remote Sens. Environ. 237, 111322. http://dx.doi.org/10.1016/j.rse.2019.111322.

Uemoto, J., Moriyama, T., Nadai, A. et al. Landslide detection based on height and amplitude differences using pre- and post-event airborne X-band SAR data. Nat Hazards 95, 485–503 (2019). https://doi.org/10.1007/s11069-018-3492-8.

Wang, F., Fan, X., Yunus, A.P. et al. (2019) Coseismic landslides triggered by the 2018 Hokkaido, Japan (Mw 6.6), earthquake: spatial distribution, controlling factors, and possible failure mechanism. Landslides 16, 1551–1566 (2019). https://doi.org/10.1007/s10346-019-01187-7.

Wang, Haojie; Limin Zhang, Kesheng Yin, Hongyu Luo, Jinhui Li, Landslide identification using machine learning, Geoscience Frontiers, Volume 12, Issue 1, 2021, Pages 351-364, ISSN 1674-9871, https://doi.org/10.1016/j.gsf.2020.02.012.

Xu, Chong; Siyuan Ma; Zhibiao Tan; Chao Xie; Shinji Toda; Xueqiang Huang; Landslides triggered by the 2016 Mj 7.3 Kumamoto, Japan, earthquake; Landslides (2018) 15:551–564; https://doi.org/10.1007/s10346-017-0929-1. ©Springer-Verlag GmbH Germany part of Springer Nature 2017.

Xu, Q., Ouyang, C., Jiang, T. et al., (2022) MFFENet and ADANet: a robust deep transfer learning method and its application in high precision and fast cross-scene recognition of earthquake-induced landslides. Landslides 19, 1617–1647 (2022). https://doi.org/10.1007/s10346-022-01847-1.

Yamagishi, H., Yamazaki, F., 2018. Landslides by the 2018 Hokkaido Iburi-Tobu earthquake on September 6. Landslides 15, 2521–2524. https://doi.org/10.1007/s10346-018-1092-z.

Yamazaki, F.; K. Kubo, R. Tanabe and W. Liu, "Damage assessment and 3d modeling by UAV flights after the 2016 Kumamoto, Japan earthquake," 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 2017, pp. 3182-3185, https://doi.org/10.1109/IGARSS.2017.8127673.

Yamazaki, Fumio; Yuuki Sagawa, and Wen Liu "Extraction of landslides in the 2016 Kumamoto earthquake using multi-temporal Lidar data", Proc. SPIE 10790, Earth Resources and Environmental Remote Sensing/GIS Applications IX, 107900H (9 October 2018); https://doi.org/10.1117/12.2325091.

Yang, R.; Zhang, F.; Xia, J.; Wu, C., (2022) Landslide Extraction Using Mask R-CNN with Background Enhancement Method. Remote Sens. 2022, 14, 2206. https://doi.org/10.3390/rs14092206.

Ye, C.; Y. Li, P. Cui, L. Liang, S. Pirasteh, J. Marcato, W. N. Gonc¸alves, and J. Li, “Landslide Detection of Hyperspectral Remote Sensing Data Based on Deep Learning with Constrains,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 12, no. 12, pp. 5047–5060, 2019. https://doi.org/10.1109/JSTARS.2019.2951725.

Yu, B.; F. Chen, C. Xu, L. Wang, and N. Wang, (2021) “Matrix SegNet: A Practical Deep Learning Framework for Landslide Mapping from Images of Different Areas with Different Spatial Resolutions,” Remote Sensing, vol. 13, no. 16, p. 3158, 2021. https://doi.org/10.3390/rs13163158.

Zhang, M.; W. Shi, S. Chen, Z. Zhan, and Z. Shi, (2020-a) “Deep multiple instance learning for landslide mapping”, IEEE Geosci. Remote Sens. Lett., vol. 18, no. 10, pp. 1711–1715, 2020. https://doi.org/10.1109/LGRS.2020.3007183.

Zhang, P.; Xu, C.; Ma, S.; Shao, X.; Tian, Y.; Wen, B., (2020-b) Automatic Extraction of Seismic Landslides in Large Areas with Complex Environments Based on Deep Learning: An Example of the 2018 Iburi Earthquake, Japan. Remote Sens. 2020, 12, 3992. https://doi.org/10.3390/rs12233992.

Zhang, Xiaokang; Weikang Yu, Man-On Pun, Wenzhong Shi, (2023) Cross-domain landslide mapping from large-scale remote sensing images using prototype-guided domain-aware progressive representation learning, ISPRS Journal of Photogrammetry and Remote Sensing, Volume 197, 2023, Pages 1-17, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2023.01.018.

Zhong, Cheng; Yue Liu, Peng Gao, Wenlong Chen, Hui Li, Yong Hou, Tuohuti Nuremanguli & Haijian Ma (2020) Landslide mapping with remote sensing: challenges and opportunities, International Journal of Remote Sensing, 41:4, 1555-1581, https://doi.org/10.1080/01431161.2019.1672904.

## 8.4 Chapter 5

Allstadt, K.E., E.M. Thompson, R.W. Jibson, D.J. Wald, M. Hearne, E.J. Hunter, J. Fee, H. Schovanec, Daniel Slosky, and Kirstie L. Haynie (2021). The USGS ground failure product: Near-real-time estimates of earthquake-triggered landslides and liquefaction. Earthquake Spectra. Volume: 38, Issue: 1, page(s): 5-36. https://doi.org/10.1177/87552930211032685.

Baise, L.G., Akhlaghi, A., Chansky, A., Meyer, M., Moaveni, B., (2021), Updating the Geospatial Liquefaction Database and Model, Final Technical Report, USGS Award #G20AP00029.

Bardet, J. P., Deaton, S., Frost, D., Goel, P., Lettis, W., Moss, R., & Stewart, J. (2001). Initial geotechnical observations of the Bhuj, India, Earthquake of January 26, 2001. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Bardet, J.P, F. Liu (2009) Motions of gently sloping ground during earthquakes. J. Geophys. Res. Earth Surf., 114 (2009), DOI:10.1029/2008JF001107.

Bhattacharya, S., Hyodo, M., Goda, K., Tazoh, T., & Taylor, C. A. (2011). Liquefaction of soil in the Tokyo Bay area from the 2011 Tohoku (Japan) earthquake. Soil Dynamics and Earthquake Engineering, 31(11), 1618-1628.

Baird, Alexander J., Mertcan Geyin, and Brett W. Maurer; On the Relationship Between Geospatial Liquefaction-Model Performance and Quality of Geospatial Data: A Case Study of the 2010-2016 Canterbury Earthquakes.

Boulanger RW, Idriss IM. In: CPT and SPT based liquefaction triggering procedures. Center for Geotechnical Modeling; 2014. Report UCD/CGM-10/2.

Bozzoni, F.; Bonì, R.; Conca, D.; Lai, C.G.; Zuccolo, E.; Meisina, C., 2020, Megazonation of earthquake-induced soil liquefaction hazard in continental Europe. Bull. Earthq. Eng. 2020.

Bozzoni, F.; Bonì, R.; Conca, D.; Meisina, C.; Lai, C.G.; Zuccolo, E. A Geospatial Approach for Mapping the Earthquake-Induced Liquefaction Risk at the European Scale. Geosciences 2021, 11, 32. https://doi.org/10.3390/geosciences11010032.

Box GEP, Cox DR. An analysis of transformations. J Roy Stat Soc B 1964;26(2):211–52.

Brandenberg SJ, Zimmaro P, Stewart JP, Kwak DY, Franke KW, Moss R, Cetin KO, Can G, Ilgac M, Stamatakos J, Weaver T, Kramer SL., 2020, Next-generation liquefaction database. Earthq. Spectra 2020; 36(2):939–59. https://doi.org/10.1177/8755293020902477.

Bray, J. D., Sancio, R., Kammerer, A. M., Merry, S., Rodriguez-Marek, A., Khazai, B., & Dreger, D. (2001). Some Observations of the Geotechnical Aspects of the February 28, 2001, Nisqually Earthquake in Olympia, South Seattle, and Tacoma, Washington. Report sponsored by NSF, PEER Center, UCB, University of Arizona, Washington State University, Shannon and Wilson Inc., and Leighton and Associates.

Bray, J., Cohen-Waeber, J., Dawson, T., Kishida, T., & Sitar, N. (2014). Geotechnical engineering reconnaissance of the August 24, 2014 M6 South Napa earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association. Doi: 10(2.1), 1094-7844.

Breiman L., 2001, Random forest. J. Mach. Learn. 2001; 45:5–32.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. Classification and Regression Trees. Boca Raton, FL: Chapman & Hall, 1984.

Brier GW. Verification of forecasts expressed in terms of probability. Mon Weather Rev 1950; 78(1):1–3. https://doi.org/10.1175/1520-0493(1950)078.

Candia, G., de Pascale, G., Montalva, G., & Ledezma, C. Geotechnical Reconnaissance of the 2015 Illapel Earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Carter, M., & Maurer, B. (2011). Geotechnical Quick Report on the Affected Region of the 23 August 2011 M5. 8 Central Virginia Earthquake near Mineral, Virginia. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Castellanos Abella, E.A.; van Westen, C.J. Generation of a landslide risk index map for Cuba using spatial multi-criteria evaluation. Landslides 2007, 4, 311–325.

Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP, 2002, SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res., 16:321–357.

Chleborad, A. F., & Schuster, R. L. (1990). Ground failure associated with the Puget Sound region earthquakes of April 13, 1949, and April 29, 1965 (No. 90-687). US Geological Survey.

Christianini, N., and J. Shawe-Taylor. An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge, UK: Cambridge University Press, 2000.

Clayton, P., Zalachoris, G., Rathje, E., Bheemasetti, T., Caballero, S., Yu, X., & Bennett, S. (2016). The Geotechnical Aspects of the September 3, 2016 M 5.8 Pawnee, Oklahoma Earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Cornell CA, Luco N. Ground motion intensity measures for structural performance assessment at near-fault sites. In: U.S.-Japan joint workshop and third grantees meeting, U.S.-Japan coop. Res. On urban EQ. Disaster mit. Seattle; 2001. p. 1–8.

Cover T.; P. Hart, "Nearest neighbor pattern classification," in IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, January 1967, https://doi.org/10.1109/TIT.1967.1053964.

Cubrinovski, M., Green, R., Allen, J., Ashford, S., Bowman, E., Bradley, B.A., & Cox, B. (2010). Geotechnical reconnaissance of the 2010 Darfield (New Zealand) earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Cubrinovski, M., Green, R. A., Wotherspoon, L., Allen, J., Bradley, B., Bradshaw, A., & Pender, M. (2011). Geotechnical reconnaissance of the 2011 Christchurch, New Zealand earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Danielson, J.J., and Gesch, D.B., 2011, Global multi-resolution terrain elevation data 2010 (GMTED2010): U.S. Geological Survey Open-File Report 2011–1073, 26 p.

Donchyts, Gennadii, Hessel Winsemius, Jaap Schellekens, Tyler Erickson, Hongkai Gao, Hubert Savenije, and Nick van de Giesen. "Global 30m Height Above the Nearest Drainage (HAND)", Geophysical Research Abstracts, Vol. 18, EGU2016-17445-3, 2016, EGU General Assembly (2016).

Durante MG, Rathje EM. An exploration of the use of machine learning to predict lateral spreading. Earthq. Spectra 2021;1(27). https://doi.org/10.1177/87552930211004613.

Erdik, M., Kamer, Y., Demircioğlu, M., & Şeşetyan, K. (2012). 23 October 2011 Van (Turkey) earthquake. Natural hazards, 64(1), 651-665.

Fan, Y., Li, H., & Miguez-Macho, G., 2013, Global patterns of groundwater table 713 depth. Science, 339(6122), 940-943.

Fernandez ´ A, García S, Galar M, Prati RC. Learning from imbalanced data sets. Springer Nature Switzerland AG; 2019. https://doi.org/10.1007/978-3-319-98074-4.

Fisher, R. A. "The Use of Multiple Measurements in Taxonomic Problems." Annals of Eugenics, Vol. 7, pp. 179–188, 1936. Available at https://digital.library.adelaide.edu.au/dspace/handle/2440/15227.

Freedman D, Diaconis P. On the histogram as a density estimator: L2 theory. Probab. Theor. Relat. Field, 1981; 57(4):453–76. https://doi.org/10.1007/BF01025868.

Friedman JH. Greedy function approximation: a gradient boosting machine. Ann Stat 2001:1189–232.

Geyin M, Baird AJ, Maurer BW. Field assessment of liquefaction prediction models based on geotechnical vs. geospatial data, with lessons for each. Earthq. Spectra 2020; 36(3):1386–411.

Geyin, M.; B.W. Maurer; K. Christofferson; 2022, An AI driven, mechanistically grounded geospatial liquefaction model for rapid response and scenario planning, Soil Dynamics and Earthquake Engineering, Volume 159, 2022, 107348, ISSN 0267-7261, https://doi.org/10.1016/j.soildyn.2022.107348.

Glorot X, Yoshua B. Understanding the difficulty of training deep feedforward neural networks. Proc, thirteenth int, conf, Artif. Intell. Stat., 2010:249–56.

Goldberger, J.; S. Roweis, G. Hinton, and R. Salakhutdinov. Neighbourhood components analysis. In L. K. Saul, Y. Weiss, and L. Bottou, editors, Advances in Neural Information Processing Systems 17, pages 513–520, Cambridge, MA, 2005. MIT Press.

Gómez, J. C., Tavera, H. J., & Orihuela, N. (2005). Soil liquefaction during the Arequipa Mw 8.4, June 23, 2001 earthquake, southern coastal Peru. Engineering Geology, 78(3), 237-255.

Gorsevski, P.V.; Jankowski, P. An optimized solution of multi-criteria evaluation analysis of landslide susceptibility using fuzzy sets and Kalman filter. Comput. Geosci. 2010, 36, 1005–1020.

Goto, K., Sugawara, D., Abe, T., Haraguchi, T., & Fujino, S. (2012). Liquefaction as an important source of the AD 2011 Tohoku-oki tsunami deposits at Sendai Plain, Japan. Geology, 40(10), 887-890.

Green RA, Bommer JJ, 2019, What is the smallest earthquake magnitude that needs to be considered in assessing liquefaction hazard? Earthquake Spectra, 35(3):1441–1464. Russell A. Green, Julian J. Bommer; What is the Smallest Earthquake Magnitude that Needs to be Considered in Assessing Liquefaction Hazard?. Earthquake Spectra 2019;; 35 (3): 1441–1464. doi: https://doi.org/10.1193/032218EQS064M.

Grill, G., Lehner, B., Thieme, M. et al. Mapping the world's free-flowing rivers. Nature 569, 215–221 (2019). https://doi.org/10.1038/s41586-019-1111-9.

Guo, Y., T. Hastie, and R. Tibshirani. "Regularized linear discriminant analysis and its application in microarrays." Biostatistics, Vol. 8, No. 1, pp. 86–100, 2007.

Guyon, I., Elisseeff, A., 2003, An introduction to variable and feature selection, Journal of Machine Learning Research, 3, 1157–1182.

Hastie, T., R. Tibshirani, and J. Friedman (2001). The Elements of Statistical Learning: Data Mining, Inference and Prediction, in Springer Series in Statistics, Springer-Verlag, New York, New York.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. 2nd ed. Springer Series in Statistics. New York, NY: Springer, 2009. https://doi.org/10.1007/978-0-387-84858-7.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., and Jarvis, A., 2005, Very high resolution interpolated climate surfaces for global land areas: International Journal of Climatology, 25(15), 1965–1978.

Hamada, M., Isoyama, R., & Wakamatsu, K. (1996). Liquefaction-induced ground displacement and its related damage to lifeline facilities. Soils and foundations, 36(Special), 81-97.

Hashash, Y., Tiwari, B., Moss, R. E., Asimaki, D., Clahan, K. B., Kieffer, D. S., & Pehlivan, M. (2015). Geotechnical field reconnaissance: Gorkha (Nepal) earthquake of April 25, 2015 and related shaking sequence. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Hauksson, E., Felzer, K., Given, D., Giveon, M., Hough, S., Hutton, K., ... & Yong, A. (2008). preliminary report on the 29 July 2008 mw 5.4 chino Hills, eastern Los Angeles basin, California, earthquake sequence. Seismological Research Letters, 79(6), 855-866.

He H, Bai Y, Garcia EA, Li S, 2008, ADASYN: adaptive synthetic sampling approach for imbalanced learning. In: 2008 international joint conference on neural networks (IJCNN 2008).

Holzer, T. L. (Ed.). (1998). The Loma Prieta, California, earthquake of October 17, 1989 (Vol. 1550). Government Printing Office.

Holzer, T. L., Noce, T. E., Bennett, M. J., Tinsley III, J. C., & Rosenberg, L. I. (2005). Liquefaction at Oceano, California, during the 2003 San Simeon earthquake. Bulletin of the Seismological Society of America, 95(6), 2396-2411.

Huang, Y., & Jiang, X. (2010). Field-observed phenomena of seismic liquefaction and subsidence during the 2008 Wenchuan earthquake in China. Natural Hazards, 54(3), 839-850.

Idriss IM, Boulanger RW. Semi-empirical procedures for evaluating liquefaction potential during earthquakes. Soil Dynam. Earthq. Eng. 2006; 26(2–4):115–30. https://doi.org/10.1016/j.soildyn.2004.11.023.

Ishihara, K., & Koga, Y. (1981). Case studies of liquefaction in the 1964 Niigata earthquake. Soils and foundations, 21(3), 35-52.

James, Gareth • Daniela Witten • Trevor Hastie Robert Tibshirani; An Introduction to Statistical Learning with Applications in R, Springer New York Heidelberg Dordrecht London, 2013, ISBN 978-1-4614-7138-7 (eBook), https://doi.org/10.1007/978-1-4614-7138-7.

Jena, R., B. Pradhan, M. Almazroui, M. Assiri, H-J. Park, Earthquake-induced liquefaction hazard mapping at national-scale in Australia using deep learning techniques, Geoscience Frontiers (2022), doi: https://doi.org/10.1016/j.gsf.2022.101460

JICA (Japanese International Cooperation Agency), 2002. The Study on Earthquake Disaster Mitigation in The Kathmandu Valley, Kingdom of Nepal, vols. 1and 2, Ministry of Home Affairs of His Majesty's Government of Nepal, Kathmandu.

Kayen, R., Thompson, E., Minasian, D., Moss, R. E., Collins, B. D., Sitar, N., ... & Carver, G. (2004). Geotechnical reconnaissance of the 2002 Denali fault, Alaska, earthquake. Earthquake Spectra, 20(3), 639-667.

Kayen, R., Cox, B., Johansson, J., Steele, C., Somerville, P., Konagai, K., Zhao, Y., Tanaka, H. (2008). Geoengineering and Seismological Aspects of the Iwate Miyagi-Nairiku, Japan Earthquake of June 14, 2008. . Geotechnical Extreme Events Reconnaissance (GEER) Association.

Kayen, R., Brandenberg, S. J., Collins, B. D., Dickenson, S., Ashford, S., Kawamata, Y., & Tokimatsu, K. (2009). Geoengineering and seismological aspects of the Niigata-Ken Chuetsu-Oki earthquake of 16 July 2007. Earthquake Spectra, 25(4), 777-802.

Karpouza, M., Chousianitis, K., Bathrellos, G.D. et al. Hazard zonation mapping of earthquake-induced secondary effects using spatial multi-criteria analysis. Nat Hazards 109, 637–669 (2021). https://doi.org/10.1007/s11069-021-04852-0.

Kiyono, J., Fujiwara, T., Hamada, M., Hashimoto, T., Ichii, K., Isoyama, R., & Nozaki, T. (2007). Reconnaissance report on the 2000 Tottori-ken Seibu Earthquake. Doboku Gakkai Ronbunshuu A, 63(2), 374-385.

Knudsen K, Bott J, 2011, Geologic and geomorphic evaluation of liquefaction case histories for rapid hazard mapping. Seismol. Res. Lett., 82:334.

Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of the fourteenth international joint conference on artificial intelligence; 1995. p. 1137–43. San Mateo, CA.

Kotoda K, Wakamatsu K, Midorikawa S. Seismic microzoning on soil liquefaction potential based on geomorphological land classification. Soils Found 1986; 28(2):127–43. https://doi.org/10.3208/sandf1972.28.2_127.

Kramer, Steven L., 1996, Geotechnical Earthquake Engineering, Pearson Publications.

Lehner, B., K. Verdin, and A. Jarvis (2006). HydroSHEDS Technical Documentation, World Wildlife Fund US, Washington, D.C.

Liang L, Figueroa JL, Saada AS. Liquefaction under random loading: unit energy approach. J. Geotech. Eng. 1995; 121(11):776–81. https://doi.org/10.1061/(ASCE)0733-9410(1995)121:11(776).

Lin, A.; L. Wotherspoon; B. Bradley; J. Motha; 2021; Evaluation and modification of geospatial liquefaction models using land damage observational data from the 2010–2011 Canterbury Earthquake Sequence, Engineering Geology, Volume 287, 2021, 106099, ISSN 0013-7952, https://doi.org/10.1016/j.enggeo.2021.106099.

Liu Z, Tesfamariam S., 2012, Prediction of lateral spread displacement: data-driven approaches. Bull Earthq. Eng. 2012; 10(5):1431–54. https://doi.org/10.1007/s10518-012-9366-7.

Luco N, Cornell CA. Structure-specific scalar intensity measures for near-source and ordinary earthquake ground motions. Earthq. Spectra 2007; 23(2):357–92, https://doi.org/10.1193/1.2723158.

Luna, R. (2010). Reconnaissance Report of the May 28, 2009 Honduras Earthquake, M 7.3. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Mackie K, Stojadinovic B. Probabilistic seismic demand model for California highway bridges. J Bridge Eng 2002;6(6):468–81. https://doi.org/10.1061/. (ASCE)1084-0702 (2001) 6:6(468).

Manning, C. D., P. Raghavan, and M. Schütze. Introduction to Information Retrieval, NY: Cambridge University Press, 2008.

Matsuoka M, Wakamatsu K, Hashimoto M, Senna S, Midorikawa S, 2015, Evaluation of liquefaction potential for large areas based on geomorphologic classification. Earthq. Spectra, 31:2375–2395.

Midorikawa S, Wakamatsu K. Intensity of earthquake ground motion at liquefied sites. Soils Found 1988; 28(2):73–84. https://doi.org/10.3208/sandf1972.28.2_73.

Moss RES, Baise LG, Zhu J, Kadkha D. Examining the discrepancy between forecast and observed liquefaction from the 2015 Gorkha, Nepal earthquake. Earthq. Spectra 2017; 33:73–83. https://doi.org/10.1193/120316eqs220m.

Orense RP. Assessment of liquefaction potential based on peak ground motion parameters. Soil Dynam. Earthq. Eng., 2005;25(3):225–40. https://doi.org/10.1016/j.soildyn.2004.10.013.

Miranda, E., Aslani, H., & Blume, J. A. (2014). Brief Report on the September 3, 2000 Yountville/Napa, California Earthquake. Pacific Earthquake Engineering Research (PEER), http://peer.berkeley.edu/publications/yountville-napa_sep-2000.html.

Miura, S., Yagi, K., & Kawamura, S. (1995). Liquefaction damage of sandy and volcanic grounds in the 1993 Hokkaido Nansei-Oki earthquake.

Monaco, P., De Magistris, F. S., Grasso, S., Marchetti, S., Maugeri, M., & Totani, G. (2011). 766 Analysis of the liquefaction phenomena in the village of Vittorito (L'Aquila). Bulletin of Earthquake Engineering, 9(1), 231-261.

Mukunoki, T., Kasama, K., Murakami, S., Ikemi, H., Ishikura, R., Fujikawa, T., & Kitazono, Y. (2016). Reconnaissance report on geotechnical damage caused by an earthquake with JMA seismic intensity 7 twice in 28h, Kumamoto, Japan. Soils and Foundations, 56(6), 947-964.

NASEM (National Academies of Sciences, Engineering, and Medicine), 2016, State of the Art and Practice in the Assessment of Earthquake-Induced Soil Liquefaction and Its Consequences; The National Academies Press: Washington, DC, USA, 2016.

Nikolaou, S., Vera-Grunauer, X., Gilsanz, R., Luque, R., Kishida, T., Diaz-Fanas, G., & Alzamora, D. (2016). GEER-ATC M 7.8 April 16, 2016 Muisne, Ecuador earthquake reconnaissance report. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Olson, S. M., Green, R. A., Lasley, S., Martin, N., Cox, B. R., Rathje, E., & French, J. (2011). Documenting liquefaction and lateral spreading triggered by the 12 January 2010 Haiti earthquake. Earthquake Spectra, 27(S1), S93-S116.

Oommen, T., L. G. Baise, and R. M. Vogel (2011). Sampling bias and class imbalance in maximum likelihood logistic regression, Math. Geosci. 43, 99–120.

Padgett, J.E., et al. (2008) Earthquake Engineering and Structural Dynamics. Wiley InterScience, Hoboken. http://www.interscience.wiley.com.

Papathanassiou, G., Caputo, R., & Rapti-Caputo, D. (2012). Liquefaction phenomena along the paleo-Reno River caused by the May 20, 2012, Emilia (northern Italy) earthquake. Annals of Geophysics, 55(4).

Papathanassiou, G., Ganas, A., & Valkaniotis, S. (2016). Recurrent liquefaction-induced failures triggered by 2014 Cephalonia, Greece earthquakes: Spatial distribution and quantitative analysis of liquefaction potential. Engineering Geology, 200, 18-30.

Pavlides, S., Papathanassiou, G., Valkaniotis, S., Chatzipetros, A., Sboras, S., & Caputo, R. (2013). Rock-falls and liquefaction related phenomena triggered by the June 8, 2008, M. Annals of Geophysics, 56(6), S0682.

Pelletier, J. D., P. D. Broxton, P. Hazenberg, X. Zeng, P. A. Troch PA, G. Y. Niu, Z. Williams, M. A. Brunke, and D. Gochis (2016). A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling, J. Adv. Model. Earth Syst. 8:41–65. https://doi.org/10.1002/2015MS000526.

Pisano, L.; Zumpano, V.; Malek, Z.; Rosskopf, C.M.; Parise, M. Variations in the susceptibility to landslides, as a consequence of land cover changes: A look to the past, and another towards the future. Sci. Total Environ. 2017, 601–602, 1147–1159.

Piya, B. K., 2004. Generation of a geological database for the liquefaction hazard assessment in Kathmandu Valley. Master's thesis, International Institute for Geo-Information Science and Earth Observation, Enschede, Netherlands.

Pokhrel RM, Gilder CEL, Vardanega PJ, De Luca F, De Risi R, Werner MJ, Sextos A. Liquefaction potential for the Kathmandu Valley, Nepal: a sensitivity study. Bull. Earthq. Eng. 2022; 20(1):25–51. https://doi.org/10.1007/S10518-021-01198-7.

Preethaa, S.; Natarajan, Y.; Rathinakumar, A.P.; Lee, D.-E.; Choi, Y.; Park, Y.-J.; Yi, C.-Y. A Stacked Generalization Model to Enhance Prediction of Earthquake-Induced Soil Liquefaction. Sensors 2022, 22, 7292. https://doi.org/10.3390/s22197292.

Rashidian, V.; Baise, L.G., 2020, Regional efficacy of a global geospatial liquefaction model. Eng. Geol. 2020, 272.

Rasmussen CE, Williams CKI. Gaussian processes for machine learning. Cambridge, Massachusetts: MIT Press; 2006.

Rathje, E. M., Kelson, K., Ashford, S. A., Kawamata, Y., Towhata, I., Kokusho, T., & Bardet, J. P. (2006). Geotechnical Aspects of the 2004 Niigata Ken Chuetsu, Japan, Earthquake. Earthquake Spectra, 22(S1), 23-46

Rathje, E. M., Karatas, I., Wright, S. G., & Bachhuber, J. (2004). Coastal failures during the 1999 Kocaeli earthquake in Turkey. Soil Dynamics and Earthquake Engineering, 24(9), 699-712.

Rokach L, Maimon O. Data mining with decision trees: theory and applications. World Scientific Pub Co Inc.; 2008, ISBN 978-9812771711.

Rollins, K., Franke, K., Luna, B. R., Rocco, N., Avila, D., Climent, M. R. (2013). Geotechnical aspects of sept. 5, 2012 m7.6 Samara, Costa rica earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Rollins, K., Ledezma, C., & Montalva, G. (2014). Geotechnical aspects of April 1, 2014, M8. 2 Iquique, Chile Earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Saaty, T.L., 1980, The Analytic Hierarchy Process; McGraw-Hill: New York, NY, USA, 1980.

Seed, H. B., and Idriss, I. M. (1971). ‘‘Simplified procedure for evaluating soil liquefaction potential.’’ J. Geotech. Engrg. Div., ASCE, 97(9), 1249–1273.

Shalev-Shwartz, Shai; Shai Ben-David, Understanding Machine Learning: From Theory to Algorithms, pp. 212 – 218; Cambridge University Press, 2014; https://doi.org/10.1017/CBO9781107298019.019.

Singh, R. P., S. Bhoi, and A. K. Sahoo (2002). Changes observed in land and ocean after Gujarat earthquake of 26 January 2001 using IRS data, Int. J. Rem. Sens. 23, 3123–3128.

Snoek, J., H. Larochelle, R. P. Adams. Practical Bayesian Optimization of Machine Learning Algorithms. https://arxiv.org/abs/1206.2944, 2012.

Spacagna R.L., Porchia A., Fabozzi S., Cesarano M., Peronace E. & Romagnoli G. (2022) - Seismic liquefaction assessment in Calabria region in Southern Italy: a geostatistical approach at regional and subregional scale. Ital J. Geosci, 141 (1), 53-68, https://doi.org/10.3301/IJG.2022.04.

Stewart, J. P., Bray, J. D., Seed, R. B., & Sitar, N. (1994). Preliminary report on the principal geotechnical aspects of the January 17, 1994 Northridge earthquake. University of California, Berkeley, Earthquake Engineering Research Center] Report UCB/EERC-94/08. Berkeley: Earthquake Engineering Research Center, University of California.

Stewart, J. P., & Brandenberg, S. J. (2010). Preliminary report on seismological and geotechnical engineering aspects of the April 4, 2010, Mw 7.2 El Mayor-Cucapah (Mexico) earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Sucuoğlu, H. (2000). The 1999 Kocaeli and Düzce-Turkey Earthquakes. Middle East Technical University.

Sugano T, Nozu A, Kohama E, Shimosako KI, Kikuchi Y., 2014, Damage to coastal structures. Soils Found 2014; 54(4):883–901. https://doi.org/10.1016/j.sandf.2014.06.018.

Sun, J., Hutchinson, T., Clahan, K., Menq, F., Lo, E., Chang, W. J., & Ma, K. F. (2016). Geotechnical reconnaissance of the 2016 Mw6. 3 Meinong Earthquake, Taiwan. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Terzaghi, K., Peck, R.B. and Mesri, G. (1996) Soil Mechanics in Engineering Practice. 3rd Edition, John Wiley and Sons, Inc., New York.

Todorovic, Lana; Vitor Silva, 2022, A liquefaction occurrence model for regional analysis, Soil Dynamics and Earthquake Engineering, Volume 161, 2022, 107430, ISSN 0267-7261, https://doi.org/10.1016/j.soildyn.2022.107430.

Taucer, F., Alarcon, J. E., & So, E. (2009). 2007 August 15 magnitude 7.9 earthquake near the coast of Central Peru: analysis and field mission report. Bulletin of Earthquake Engineering, 7(1), 1-70.

Theobald, D. M., Harrison-Atlas, D., Monahan, W. B., & Albano, C. M. (2015). Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. PloS one, 10(12), e0143619.

Thompson, E. M., D. J. Wald, K. Allstadt, and M. Hearne (2016). Combining case history observations with different completeness levels in empirical ground-failure models, Seismol. Res. Lett. 87, no. 2B, 578.

Tohno, I., & Yasuda, S. (1981). Liquefaction of the ground during the 1978 Miyagiken-Oki earthquake. Soils and Foundations, 21(3), 18-34.

Tohno, I., & shamoto, Y. (1985). Liquefaction Damage to the Ground during the 1983 Nihonkai-Chubu (Japan Sea) Earthquake in Akita Prefecture, Tohoku, Japan. Natural disaster science, 7(2), 67-93.

UNDP (United Nations Development Program), 1994. Seismic Hazard Mapping and Risk Assessment for Nepal, His Majesty's Govt. of Nepal, Ministry of Housing and Physical Planning, UNDP/UNCHS Habitat, Subproject NEP/88/054/21.03,1994.

USGS 2015, Executive summary of the M 4.0 - 1km N of Piedmont, California earthquake (2015). United State Geological Survey (USGS). https://earthquake.usgs.gov/earthquakes/eventpage/nc72507396/executive.

USGS 2020, Scientists from the US Geological Survey, Southern California Earthquake Center, and California Division of Mines and Geology. (2000). Preliminary report on the 16 October 1999 M 7.1 Hector Mine, California, earthquake. Seismological Research Letters, 71(1), 11-23.

Van Den Eeckhaut, M., J. Hervás, C. Jaedicke, J. P. Malet, L. Montanarella, and F. Nadim (2012). Statistical modelling of Europe-wide landslide susceptibility using limited landslide inventory data, Landslides 9, No. 3, 357–369.

Vapnik V. The nature of statistical learning theory. New York: Springer; 1995.

Verdugo, R. (2011). Comparing liquefaction phenomena observed during the 2010 Maule, Chile earthquake and 2011 Great East Japan earthquake. In Proceedings of international symposium on engineering lessons learned from the (pp. 1-4).

Wakamatsu, K. (2011). Historic Liquefaction sites in Japan, 745-2008, University of Tokyo Press.

Wald DJ, Allen TI, 2007, Topographic slope as a proxy for seismic site conditions and amplification. Bull. Seismol. Soc. Am., 97:1379–1395.

Wartman, J., Rodriguez-Marek, A., Macari, E. J., Deaton, S., Ramírez-Reynaga, M., Ochoa, C. N., & Ovando-Shelley, E. (2005). Geotechnical aspects of the January 2003 Tecoman, Mexico, earthquake. Earthquake spectra, 21(2), 493-538.

Weiss GM, Provost F. In: The effects of class imbalance and training data size on classifier learning: an empirical study; 2001. https://doi.org/10.7282/t3-vpfwsf95.SN.

Wilson MFJ, O'Connell B, Brown C, Guinan JC, Grehan AJ. Multiscale terrain analysis of multibeam bathymetry data for habitat mapping on the continental slope. Marine Geodesy, vol. 30; 2007. https://doi.org/10.1080/01490410701295962. Issues 1–2.

Xie Y, Ebad SM, Padgett JE, DesRoches R., 2020, The promise of implementing machine learning in earthquake engineering: a state-of-the-art review. Earthq. Spectra 2020; 36(4):1769–801. https://doi.org/10.1177/8755293020919419.

Yang, W., Wang, K., Zuo, W., 2012, Neighborhood Component Feature Selection for High-Dimensional Data, Journal of Computers. 7 (1).

Yasuda, S., Watanabe, H., Yoshida, N., & Mora, S. (1993). Soil Liquefaction During the 1991 Telire-Limón, Costa Rica, Earthquake.

Yasuda S, Towhata I, Ishii I, Sato S, Uchimura T. Liquefaction-induced damage to structures during the 2011 Great East Japan earthquake. J. Jpn. Soc. Civil Eng., 2013; 1:181–93. https://doi.org/10.2208/journalofjsce.1.1_181.

Yilmaz, C.; Silva, V.; Weatherill, G., 2020, Probabilistic framework for regional loss assessment due to earthquake-induced liquefaction including epistemic uncertainty. Soil Dyn. Earthq. Eng. 2020.

Youd TL, Perkins DM, 1978, Mapping liquefaction-induced ground failure potential. J. Geotech. Eng. Div., 104(4):433–446.

Yuan, H., Yang, S. H., Andrus, R. D., & Juang, C. H. (2004). Liquefaction-induced ground failure: a study of the Chi-Chi earthquake cases. Engineering Geology, 71(1), 141-155.

Zhang, J., Wang, Y. An ensemble method to improve prediction of earthquake-induced soil liquefaction: a multi-dataset study. Neural Comput & Applic 33, 1533–1546 (2021). https://doi.org/10.1007/s00521-020-05084-2.

Zhu, J.; Daley, D.; Baise, L.; Thompson, E.; Wald, D.; Knudsen, K., 2015, A Geospatial Liquefaction Model for Rapid Response and Loss Estimation. Earthquake Spectra 2015, 31, 1813–1837.

Zhu, J.; Baise, L.; Thompson, E., 2017, An Updated Geospatial Liquefaction Model for Global Application. Bulletin of the Seismological Society of America 2017, 107, 1365–1385.

Zimmaro, P., & Stewart, J. P. (2016). Engineering reconnaissance following the 2016 M 6.0 Central Italy Earthquake. Geotechnical Extreme Events Reconnaissance (GEER) Association.

Zumpano, V.; Pisano, L.; Filice, F.; Ugenti, A.; de Lucia, D.; Wasowski, J.; Santaloia, F.; Lollino, P. Regional-Scale Seismic Liquefaction Susceptibility Mapping via an Empirical Approach Validated by Site-Specific Analyses. Geosciences 2022, 12, 215. https://doi.org/10.3390/geosciences12050215.

## 8.5 Chapter 6

Abdellah, W.R.; Hirohama, C.; Sainoki, A.; Towfeek, A.R.; Ali, M.A.M. Estimating the Optimal Overall Slope Angle of Open-Pit Mines with Probabilistic Analysis. Appl. Sci. 2022, 12, 4746. https://doi.org/10.3390/app12094746.

Ariana, Diana. 2014. Effect of Seepage and Slope Stability on the Geometry Optimization of the Bukit M Air Laya Mine Slope East Block Tanjung Enim Mining Unit. Geological Engineering, Diponegoro University, Semarang.

Bye, A.; Bell, F. Stability assessment and slope design at Sandsloot open pit, South Africa. Int. J. Rock Mech. Min. Sci. 2001, 38, 449–466.

Chrzanowski, A., et al., Integration of the Global Positioning System with geodetic leveling surveys in ground subsidence studies. CISM Journal ACSGC, 1989. 43(4): p. 377-386.

Fukuzono, T. (1985) A new method for predicting the failure time of a slope, in Proceedings Fourth International Conference and field workshop on landslides, Tokyo, Japan, Japan Landslide Society, Tokyo, pp. 145–150.

Hardy, H. R. and Kimble, E.J. 1991. Application of high-frequency AE/MS techniques to rock slope monitoring. Proceedings of the Fifth Conference on Acoustic Emission/Microseismic Activity in Geologic Structures and Materials, Pennsylvania State University, 11–13. June 1991. Hardy, H.R. (ed.). Trans Tech, Clausthal-Zellerfeld, Germany. pp. 457–477.

Johansson J. M. and T. Edeskär, “Effects of External Water-Level Fluctuations on Slope Stability,” Electron. J. Geotech. Eng., vol. 19, no. K, pp. 2437–2463, 2014.

Kane, W.F. (1998): “Time Domain Reflectometry,” KANE GeoTech, Inc. /ourworld.compuserve.com/homepages/wkane/tdr.htm.

Lee, Thomas S. 2002. Slope Stability and Stabilization Methods Second Edition. John Wiley & Sons, Inc. New York.

Lynch, R.A. and Malovichko, D.A. 2006. Seismology and slope stability in open pit mines. Proceedings of the International Symposium on Stability of Rock Slopes in Open Pit Mining and Civil Engineering, Cape Town. Southern Africa Institute of Mining and Metallurgy, Johannesburg. pp. 375–390.

Luo, X.; M. Salvoni; P. Dight; J. Duan; 2018; Micro-seismic events for slope stability analysis — a case study at an open pit mine, Journal of the Southern African Institute of Mining and Metallurgy, AFRIROCK 2017 International Symposium, ISSN 2411-9717, http://dx.doi.org/10.17159/2411-9717/2018/v118n3a2.

Mark, H., et al., Guidelines for open pit slope design. 2009, Australia, New Zealand and South Africa: CSIRO Publishing.

Mendecki, A. J., 1997. Seismic Monitoring in Mines. Chapman and Hall.