The Legacy Effect of the 2005 Kahsmir Earthquake on Post-Seismic Landslide Susceptibility during the Western Monsoon Season in Northern Pakistan

AN EXPLORATORY STUDY ON WHETHER THE TIME BETWEEN AN EARTHQUAKE AND AN UPCOMING RAINFALL SEASON CAN BE UTILIZED TO INFORM EARLY ACTION TO REDUCE THE IMPACTS OF LANDSLIDES

Laura Louise Hielkema
s2370506
Enschede, the Netherlands
September, 2021

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Spatial Engineering.

Supervisors:
Prof. Dr. M.K. van Aalst
Prof. Dr. M. van der Meijde

Chair:
Prof. Dr. V.G. Jetten

External Examiner:
Dr. M. Shafique

ITC Faculty of Geo-Information Science and Earth Observation
University of Twente

DISCLAIMER

Abstract

Landslides are frequently occurring hazards and pose a significant threat to people and property. Especially in mountainous regions, landslide risk is prominent. Both earthquakes and rainfall are important triggers of landslides. This research combines these two triggers by assessing whether dynamic landslide susceptibility analysis based on a combination of post-seismic and rainfall-induced increases in landslide susceptibility in northern Pakistan can inform early action to reduce the impact of those landslides. More specifically, this study examines whether it would have been possible to utilise the time between the 2005 Kashmir earthquake and the first western monsoon season after the earthquake to have predicted where the landslides occurred. Therefore, this study explores a new approach to landslide risk reduction by assessing the possibility of using earthquake parameters to predict the spatial variability of landslides in hopes of reducing the impacts of landslides.

This is done using Bayesian versions of a Generalised Additive Model (GAM) to estimate landslide susceptibility before, during and after the 2005 Kashmir earthquake. To do so, a pre-seismic, a co-seismic and a post-seismic landslide inventory were developed using ASTER satellite images. These were used to assess whether including ground motion parameters in the GAMs show elevated landslide susceptibility in areas and could potentially be applied for early action in the study area to reduce landslide risk during upcoming western monsoon seasons. Contrary to regular landslide predictive modelling, no separate training and validation data sets were used to assess the predictive capacity of the landslide models. Instead, the explanatory models are based on the same parameters per slope unit. For the post-seismic landslide models, the only difference is the inclusion or exclusion of the ground motion parameters. This was done to merely examine the possible benefits of earthquake information on the identification of locations prone to post-seismic landslides.

Three post-seismic GAMs were conducted, one including earthquake ground motion parameters, one excluding them and one with merely the earthquake parameters. The last one was done to evaluate to what extent the earthquake parameters could accurately predict land sliding. For each of the GAMs, the influence of the landslide controlling parameters on landslide occurrence was analysed and compared to assess whether the effects of the earthquake parameters could potentially inform early action to reduce landslide risk. Surprisingly, the fixed and random effects of the earthquake parameters in the GAMs showed no significant influence on landslide occurrence. The post-seismic GAMs including and excluding the earthquake parameters were very similar, and both were similarly accurate. The resulting susceptibility maps for the study area showed only minor differences in susceptibility. Because of minor susceptibility differences, this approach has not been adequate to develop effective early action strategies. Therefore, the focus should be on other landslide risk reduction strategies to reduce the impacts of landslide risk in the study area and elsewhere.

Keywords: 2005 Kashmir earthquake, Post-seismic landslides, Earthquake legacy effect, Western monsoon season, Generalized Additive Model (GAMs), Landslide susceptibility

Acknowledgements

Throughout the thesis process, I have been assisted by professionals who helped and supported me on a topic I had limited knowledge of prior to starting my thesis. Therefore, I had to face a steep learning curve. It has been a challenging but educational year, and I am grateful to have had the supervision and support.

First and foremost, I would like to extend my gratitude to my supervisors, Maarten van Aalst, and Mark van der Meijde. They have been incredibly supportive with weekly meetings to help me get to this end goal. Additionally, I want to thank Muhammad Aufaristama for helping me with the landslide inventories after my initial methodology did not work out. Also, Luigi Lombardo for his additional support with the landslide models. Last but not least, Saad Khan and Muhammed Shafique for providing me with much-needed data and inspiration.

Last year would have been an academically challenging year, even without the global pandemic. Now, most of this year was spent at home. In light of this, I also want to thank my family and friends who provided additional support and have made me realize how important such a comprehensive support system is.

Contents

List of Abbreviations

AUC	Area under the Curve
BSI	Barren Soil Index
CMT	Centroid Moment Tensor
DEM	Digital Elevation Model
$DREF$	Disaster Relief Emergency Fund
EAP	Early Action Protocol
$EWEA$	Early Warning Early Action
EWS	Early Warning System
FBA	Forecast-based Action
FBF	Forecast-based Financing
FPR	False Positive Rate
GAM	Generalized Additive Model
IDW	Inverse Distance Weighting
$IFRC$	International Federation of Red Cross and Red Crescent Societies
$NDVI$	Normalized Difference Vegetation Index
$NDWI$	Normalized Difference Water Index
PGA	Peak Ground Acceleration
PGD	Peak Ground Displacement
PGV	Peak Ground Velocity
RBR	Relativized Burn Ratio
ROC	Receiver Operating Characteristic (curve)
TPR	True Positive Rate
VIF	Variance Inflation Factor

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Natural disasters can have severe impacts on both individuals and communities. Between 2005 and 2015, over 700 thousand lives were lost, disasters injured over 1.4 million people, and in total, over 1.5 billion affected people were reported. Moreover, over 1.3 trillion US dollars in economic losses were estimated (United Nations of Disaster Risk Reduction, 2015, hereafter UNDRR). On top of that, the severity and frequency of these impacts are likely to increase in the future. Contributing factors to the expected increase in impacts are climate change, rapid urbanization, technological hazards, public health emergencies and conflict (International Federation of Red Cross and Red Crescent Societies, 2020, hereafter IFRC). Furthermore, in recent decades, a striking number of countries have suffered from consecutive disasters. Consecutive disasters are events whose impacts overlap both spatially and temporally. The risk of consecutive disasters will increase due to growing exposure, the interconnectedness of human society, and the increased frequency and intensity of non-tectonic hazards (de Ruiter et al., 2020).

1.1.1 Early Warning & Early Action

Over the years, it has become apparent that taking action before a disaster has occurred is generally far more effective than merely acting after the fact. By doing so, it can reduce the impact of disasters by saving more lives, livelihoods, and further reducing damages (Asia Regional Resilience to a Changing ClimateMet Office, International Federation of Red Cross and Red Cresent SocietiesUKaidAnticipation HubRisk Informed Early Action Partnership, 2020). Considering that disaster forecasts have become increasingly reliable over the years, it is not surprising that many humanitarian aid organizations have opted to implement Early Warning and Early Action (EWEA). EWEA is an efficient way to reduce the impacts of disasters. Typically, early action is implemented between an early warning trigger and when the disaster occurs (IFRC, nda). Early warning triggers are pre-established thresholds that are believed to trigger a disaster. EWEA is also referred to as Forecast-based Action (FbA) and anticipatory action. These terms are used interchangeably in this study, depending on the conducted sources.

1.1.2 Forecast-based Financing

The International Federation of Red Cross and Red Crescent Societies (IFRC) and their local partners refer to early action as FbA. FbA aims to minimize the consequences of disasters by acting on forecasts beforehand. It includes prediction, reduction, and prevention of the impact of disasters. Forecast-based Financing (FbF) is the financing system for such early actions. FbF can be funded by the Disaster Relief Emergency Fund (DREF) (International Federation of the Red Cross and German Red Cross, n.d.).

Overall, the IFRC's FbF system consists of three main parts. As discussed prior, there are the thresholds values or triggers. These are determined based on risk analyses and impact assessments. The second part of FbF is the early actions taken when these threshold values are reached to reduce the impacts of the disasters. Finally, the financing aspect of FbF automatically assigns funds when these trigger values are reached. For specific hazards and locations, these aspects are summarized in Early Action

Protocols (EAPs). An EAP acts as a guide but also presents responsibilities for when a threshold is reached (IFRC, ndb). EAPs have been established for many types of hazards, including droughts, floods and cyclones (IFRC, ndb).

EAPs have not yet been established for landslides, potentially due to lead time constraints and the complex nature of these events. Furthermore, landslides are very localized events, which complicates the implementation of EWEA on a large scale. However, following that EWEA for landslides is currently lacking, there is an opportunity to reduce the impacts of disasters through research focusing on circumvention of these difficulties. The overall aim would be to apply FbF to landslides and ultimately mitigate the impacts of landslides.

1.2 Societal relevance

In 2015, the Sendai Framework for Disaster Risk Reduction 2015-2030 was established. Four priorities were set, namely;

> (i) Understanding disaster risk; (ii) Strengthening disaster risk governance to manage disaster risk; (iii) Investing in disaster reduction for resilience and; (iv) Enhancing disaster preparedness for effective response, and to "Build Back Better" in recovery, rehabilitation, and reconstruction (UNDRR, 2015, p.14).

Thus, research focusing on understanding and minimization of landslide risk could be of great societal benefit. More specifically, an improved comprehension of increased landslide risk in specific periods or locations can be used to inform FbA and trigger FbF to reduce risk. As previously mentioned, despite the great efforts of the IFRC, no EAPs are currently available for landslides.

1.3 Landslides

Landslides are large masses of soil debris or rocks moving down a slope, which can have destructive consequences. They are frequently occurring hazards, especially in mountainous regions. Landslides are either caused by natural phenomena or human activity, and they form a significant threat to people and property (Gorum et al., 2013; Owen et al., 2008). Between 1998 and 2017, an estimated 4.8 million people were affected by landslides worldwide, and over 18,000 people lost their lives (World Health Organization, nd).

1.3.1 Post-seismic and rainfall-induced landslides

Both earthquakes and rainfall are two major mechanisms that often trigger landslides (Crosta, 2004). In the literature, this phenomenon is also referred to as triggering or cascading hazards, where a primary hazard can potentially trigger a secondary hazard (Tsiplakidis and Photis, 2019). Ground shaking induced by earthquakes reduces the shear strength of slopes and subsequently increases the susceptibility of the slopes to landslides. The risk of landslides further increases when consecutive seismic events and rainfall occur (Tang et al., 2016). Moreover, successive earthquakes can cause ground cracks, which can act as conduits for rainfall. Over time, this further weakens the ground (Bhandari, nd). Many factors influence the intensity and distribution of co- and post-seismic landslides. These include the earthquake's magnitude, fault characteristics, depth of the epicentre, amplification patterns, and variation in physical features (Gorum et al., 2013; Shafique, 2020; Shafique et al., 2016). In this research, co-seismic landslides are defined as landslides directly triggered by the earthquake and happen during or shortly after the seismic event. Post-seismic landslides are not directly triggered by earthquakes but can occur up to years after the seismic event.

Following a large earthquake, an increased risk of seismic landslides can last up to years after the initial seismic event (Shafique, 2020). The duration of this increase in landslide risk is affected by many different parameters. These include rainfall intensity, cumulative rainfall, changes in soil properties, self-stabilization processes of slopes, and re-vegetation (Huang and Fan, 2013). Other factors that influence increased landslide risk are the development of regional topography through erosion, tectonics, and incision of valleys (Huang and Fan, 2013). The increased susceptibility of the slopes caused by the seismic event combined with other triggers can eventually lead to a landslide. Rainfall-induced landslides are generally caused during periods of heavy rainfall due to an increase in pore pressure and seepage forces (Anderson and Sitar, 1995; Sidle and Swanston, 1982; Sitar, 1992; Terzaghi, 1950, cited by Wang and Sassa, 2003). Increased pore pressure leads to a reduction of effective stress in the soil and, subsequently, reduced soil shear strength. Eventually, this can lead to slope failure, and thus landslides (Brand, 1981; Brenner and Brand, 1985, cited by Wang and Sassa, 2003).

Landslide risk further increases when triggers are combined. Increased landslide risk following an earthquake further intensifies during periods of heavy rainfall, posing a more significant threat to people and property. Therefore, the focus of this study is on post-seismic landslides that occur during periods of heavy rain. Such landslides are caused by a combination of triggers and are thus complex of nature.

1.4 Research aim

As previously mentioned, increased susceptibility to landslides after seismic activity combined with rainfall results in elevated landslide risk. It is then a combination of triggers that lead to the hazard event. This research aims to utilize the time between an earthquake and a rainfall season to determine elevated landslide risk, with the purpose that early action strategies can be implemented to reduce the impacts of those landslides in the upcoming rainy season. To be more specific, the intent of this study is to assess the earthquake's legacy effect on the occurrence and spatial variability of landslides during rainfall seasons. The emphasis is placed on whether the characteristics of an earthquake can tell us where landslides are likely to occur.

By doing so, this study aims to reduce the wickedness of the problem. Wicked problems or unstructured problems are characterized by complexity and their interconnected nature. Furthermore, stakeholders view these problems differently, and there is no single solution to a wicked problem (Elia and Margherita, 2018). The wicked problem framework states that wicked problems arise when dissensus among stakeholders and uncertain knowledge are combined (Georgiadou, 2018). This study reduces the wickedness by improving the understanding of post-seismic landslides during western monsoon seasons in Northern Pakistan and aims to reduce the impacts and risks by evaluating the options of early action.

1.5 Case study

A past earthquake has been selected to determine whether it is possible to utilize the time between an earthquake and an upcoming rainfall season to implement early action strategies. A historic earthquake was chosen because the landslides can validate the predicted elevated landslide risk in the rainfall season after the earthquake. Thus, by establishing an inventory of the effect of the earthquake, it can be assessed whether spatial variability of landslide occurrence during the rainfall season could have been predicted based on the earthquake parameters. Because of available data and previously conducted research on both the earthquake and post-seismic landslides, this study focuses on the 2005 Kashmir earthquake in northern Pakistan.

1.5.1 2005 Kashmir Earthquake

On the 8th of October 2005, the 2005 Kashmir earthquake struck the north of Pakistan and the Kashmir region with a magnitude of 7.6 Mw. The epicentre of the earthquake was located close to the city of Muzaffarabad at a depth of 26 kilometres (Earthquake Engineering Research Institute, 2005). There were an estimated 87,350 fatalities in Pakistan and 10,300 in India (Hussain et al., 2006). The earthquake injured another 69,000 people (United States Geological Survey, nd). Furthermore, the event triggered thousands of landslides, mainly debris flows and rockfalls (Owen et al., 2008).

This study focuses on post-seismic landslides during the first western monsoon seasons after the 2005 Kashmir earthquake. More specifically, the 2005 Kashmir earthquake's legacy effect on landslide occurrence during the western monsoon season is studied. Furthermore, possible mitigation measures to reduce the impacts and occurrence of post-seismic landslides during the western monsoon seasons will be examined.

1.5.2 Study area

The study area's location is in northern Pakistan, which is part of the total area affected by the 2005 Kashmir Earthquake. It includes the towns Balakot and Muzaffarabad, which were severely affected by the 2005 Kashmir earthquake and subsequent landslides (Shafique, 2020). The study area of this research is similar to the area used by Shafique (2020). Figure 1.1 shows the study area, which covers 253.54 km2 and has an elevation between 631 and 2697 meters above sea level.

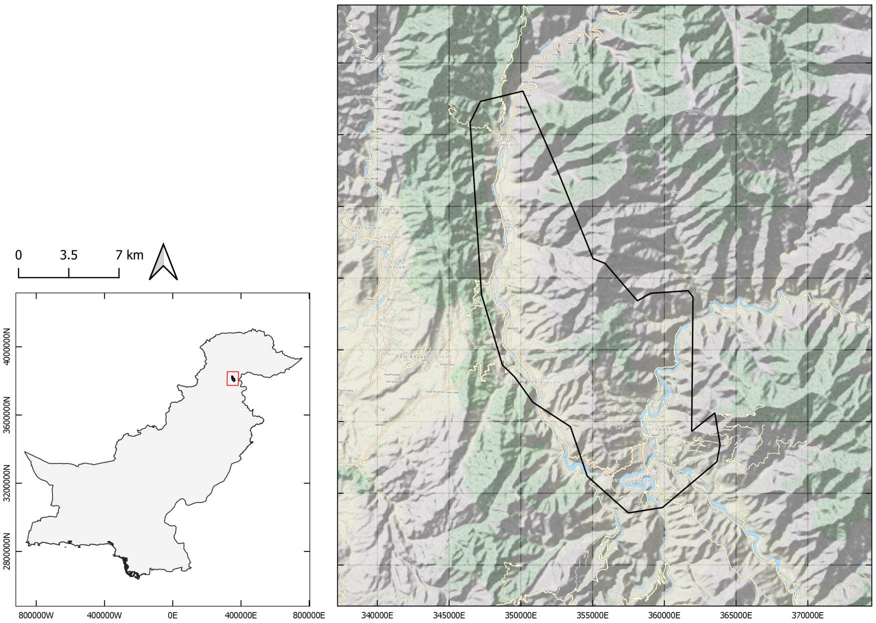


Figure 1.1: Study Area in Northern Pakistan

The study area has been prone to landslides even before the 2005 Kashmir earthquake. However, ever since the earthquake, landslide frequency has drastically increased. The main tectonic structure in the area is the Hazara Kashmir Syntaxis (Sana and Nath, 2016). Several major thrust faults are located in the study area. The fault ruptures in the study area are the Main Boundary Thrust (MBT), the Panjal Fault, the Kaghan Fault and a small part of the Bagh Blind Fault (Khan et al., 2020).

Several geological formations are found in the study area. The most prominent ones are the Muzaffarabad formation, Hazara formation, Murree formation, and Quaternary deposits. A map of the geologic formations in the study area is provided in figure 3.4-d. Table 1.1 provides an overview of the most important geological formations in the study area, as well as the corresponding lithology and the age of the formation. The Murree formation consists of mudstone, sandstone, siltstone, and shale and was formed in the Miocene epoch. The Cambrian Muzaffarabad formation is comprised chiefly of dolomite, quartzite limestone, and sandstone. The Hazara formation consists of slate, limestone, siltstone, and shale. It stems from the pre-Cambrian era, making it the oldest formation present in the study area (Shafique, 2020). According to Owen et al. (2008), after the 2005 Kashmir earthquake, most landslides occurred in the Muzaffarabad formation.

Table 1.1: Most prominent geological formation, including the corresponding lithology and formation age, after (Hussain et al., 2004), (Latif et al., 2008), and (Shafique, 2020).

Geological Formation	**Lithology**	**Age**
Murree Formation	Mudstone, sandstone, siltstone, shale	Miocene
Muzaffarabad Formation	Dolomite, quartzite limestone, sandstone	Cambrian
Quaternary Deposits	Gravel, clay, sand	Quaternary to present
Hazara Formation	Slate, siltstone, limestone, shale	Pre-cambrian

Figure 1.2 shows an example of a large landslide (debris flow) that occurred near Balakot. Debris flows are moving masses of loose mud, rock, soil, water, sand, and air that move down a slope. They are among the most dangerous landslide types (Jakob and Hungr, 2005). An inhabited area is located at the foot of the slope. As a result, people and assets are exposed to further danger if another landslide occurs, illustrating this research's importance.

Figure 1.2: Debris Flow near the town Balakot (Google Earth, 2020)

1.5.3 Western monsoon season

The western monsoonal season generally occurs from late June through September, with the retreating monsoon period until November (Weather & Climate, nd). Based on a 30-year average, Muzaffarabad receives over 1500 mm precipitation a year, of which about a third during the western monsoon season (World Meteorological Organization, nd). Balakot receives over 1700 mm of rainfall a year, of which over 40% falls during the western monsoon season (Weather Atlas, n.d.). During these months, heavy precipitation often results in landslides and flooding (Khattak et al., 2010).

1.6 Research objectives and research questions

This research's main objective is to investigate whether dynamic landslide risk analysis based on earthquake parameters of the 2005 Kashmir earthquake could have informed early action to reduce the impacts of landslides during the first western monsoon season after the earthquake. To reach this objective, the research objectives (RO) and associated research questions (RQ) are as follows:

*R*O1: To develop pre-seismic, co-seismic and post-seismic landslide inventories for the study area and to determine the increased landslide intensity during the western monsoon season after the 2005 Kashmir earthquake.

*R*O2: To develop pre-seismic, co-seismic and post-seismic explanatory hazard models for the study area.

*R*O3: To determine where and to what extent landslide susceptibility increased in the post-seismic inventory compared to pre-earthquake landslide susceptibility.

*R*Q1: What have been the effects of the 2005 Kashmir earthquake on post-seismic landslides in northern Pakistan, specifically during the western monsoon season in the first year after the earthquake?

*R*O4: To assess what parameters affect pre-seismic, co-seismic and post-seismic landslide distribution.

*R*O5: To determine whether the earthquake parameters have a clear influence on landslide distribution in the post-seismic landslide model.

*R*Q2: Compared to pre-seismic landslide distribution, what parameters influence the spatial distribution of post-seismic landslides during monsoon seasons?

*R*O6: To assess the possibilities of risk reduction in the future based on the influence of earthquake parameters on landslide occurrence.

*R*Q4: Is the earthquake legacy effect of the 2005 Kashmir earthquake on landslide risk in combination with exposure information significant enough to allow for early actions in time until the first monsoon season after a possible earthquake?

1.7 Thesis outline

The subsequent chapters are organized as follows. Chapter 2 presents a literature review on landslide hazard, landslide susceptibility assessment, and landslide EWEA. Chapter 3 focuses on the research design, where the applied methods are discussed. The results and discussion are provided in chapter 4, followed by the conclusions in chapter 5. A reflection on the study is given in chapter 6, and the final chapter is the appendix.

Chapter 2

Literature Review

This chapter provides an overview of the available literature and methods that are relevant to this research. To start, section 2.1 discusses landslide hazard and susceptibility assessment methodologies. Section 2.2 provides an overview of the literature that has studied post-seismic and rainfall-induced landslides. Section 2.3 discusses a study that has conducted a seismic hazard analysis of the 2005 Kashmir earthquake. Section 2.4 explores what has already been done by the IFRC in terms of FbA and FbF. Section 2.5 then discusses the challenges with EWEA for landslides, and section 2.6 discusses what Early Warning Systems (EWS) have been established for rainfall-induced landslides.

2.1 Landslide hazard and susceptibility assessment

For a long time, research institutions and governments have focused on landslide hazard, hazard zoning, and risk assessment. Over the years, multiple landslide hazard assessment methodologies have been used. The distinction can be made between quantitative or qualitative methods and direct or indirect methods. Generally, qualitative methods make use of descriptive terms to assess landslide hazard. As a result, they are subjective to the researcher's input. On the other hand, quantitative methods calculate the probabilities of landslide occurrence (Guzzetti et al., 1999). Direct methods are based on the knowledge of the researcher(s) of the geomorphological conditions of the area and its relations to landslide occurrence (Chauhan et al., 2010; Guzzetti et al., 1999; Thiery et al., 2014; Verstappen, 1983). Indirect methods aim to determine landslide-prone areas by assessing mathematical relationships between landslide occurrence (or response variable) and a set of explanatory or predictive variables. In other words, the direct or indirect correlations between a set of predetermined physical factors and slope instability. By use of a statistical function, each explanatory variable is assigned a certain "weight" (Guzzetti et al., 1999; Thiery et al., 2014). Based on this, landslide susceptibility zonation can be done. In landslide susceptibility research, several types of mapping units are commonly used. These are grid-cells, slope units, topographic units, terrain units, and unique-condition units. Which type of mapping unit is best applicable to a study depends on the data and the tools used (Guzzetti et al., 1999).

The most important methods applied in landslide susceptibility fall into five categories. First, geomorphological landslide hazard mapping, which is essentially a qualitative method, based on the researcher's ability to estimate the potential and actual slope failures (Bosi et al., 1985; Godefroy and Humbert, 1983; Hansen et al., 1995; Humbert, 1977; Kienholz et al., 1983, 1984; Seeley and West, 1990; Zimmermann et al., 1986, cited by Guzzetti et al., 1999). A second approach is the heuristic or index-based approach. This method is based on prior knowledge of the causes of landslide occurrence in the studied area. Thus, it depends on the researcher's understanding of the effects of geomorphological processes on the studied area. The factors are weighted based on their expected influence on landslide occurrence. Another methodology category is landslide distribution analysis, where landslide density maps of past and present landslide deposits are compared to predict future landslide distribution. Landslide distribution analysis is an indirect quantitative approach. The fourth method, also an indirect quantitative approach, is the use of geotechnical or physically-based models. It is based on physical laws that affect slope instability. The last landslide susceptibility approach makes use of statistical-based models. This approach relies on analysing functional relationships between landslide distribution in the past and present and a set of instability factors. Various multivariate statistical techniques have been applied in this approach. The most commonly used methods are linear and logistic regression, discriminant analysis and neural networks (Carrara, 1983; Carrara et al., 1991, 1995; Chung et al., 1995; Mark, 1992; Neely and Rice, 1990; Neuland, 1976; Roth, 1983; Van Westen, 1993, 1994; Yin and Yan, 1988, cited by Guzzetti et al., 1999). These types of methods are often intertwined with machine learning, where the majority of the data set is used as training data. The remaining data is used for validation to assess the predictive capacity of the model (Althuwaynee et al., 2012; Pourghasemi et al., 2013; Pradhan et al., 2010; Yang et al., 2019).

2.2 Post-seismic and rainfall-induced landslides

This study focuses on a combination of earthquake and rainfall-induced landslides. Over the years, several studies have been conducted on the effects of earthquakes on landslide activity, also in relation to rainfall. These will be discussed in this section.

Lin et al. (2006) studied the link between the Chi-Chi earthquake and rainfall-induced landslides that occurred after the earthquake. This was done by comparing eight SPOT images from 1996 until 2001 of landslides in the Choushui River watershed. They found that not only the frequency of subsequent landslides increased but also the spatial distribution of landslides had changed compared to before the Chi-Chi earthquake (Lin et al., 2006).

Furthermore, Yang et al. (2018) studied the effect of the 2008 Wenchuan earthquake on landslide activity. This was done by comparing multi-year high spatial resolution and high temporal remote sensing images from before and after the earthquake. According to their results, the Wenchuan earthquake significantly damaged vegetation, which as a result, triggered landslides. Over the years, vegetation in the area has been recovering, and post-seismic landslide activity has decreased accordingly (Yang et al., 2018).

Many researchers have examined both co-seismic and post-seismic landslides caused by the 2005 Kashmir earthquake. Several studies indicated stabilizing trends of slopes in the area (Khan et al., 2013; Khattak et al., 2010; Saba et al., 2010). However, as indicated by Shafique (2020), there were limitations to these studies. They were conducted using either part of the area affected by the earthquake or used a limited set of co- or post-seismic landslides.

Therefore, according to Shafique (2020), these studies might have underestimated landslide risk. Shafique (2020) then analysed pre-earthquake and post-earthquake landslide datasets to estimate the spatial and temporal unfolding of landslides, which were induced by the 2005 Kashmir earthquake. A decline in post-seismic landslides was indicated through visual image interpretation, with an acceleration from 2010 to 2018. However, the rate of decline is slower than for post-seismic landslides of other large earthquakes. Furthermore, he studied the relationship between landslides and rainfall in the area. Mean monthly rainfall data was obtained from the Pakistan Meteorological Department. Data from 2004-2018 was used and compared with the post-seismic landslide inventory. Figure 2.1 shows the landslide area in square kilometres and the mean monthly rainfall data in millimetres. Despite periods with hefty rainfall and monsoon seasons, the landslide areas show a continual decline. This reveals the temporal decay of post-seismic landslides in the area (Shafique, 2020).

Although Shafique (2020) studied the relations of rainfall with post-seismic landslides of the 2005 Kashmir earthquake, this has not been done in detail for monsoon seasons specifically, and it has not been compared to land sliding during monsoon seasons before the 2005 Kashmir earthquake. Furthermore, this study is unique in exploring a new landslide susceptibility approach to detect an apparent spatial earthquake legacy effect on landslide susceptibility.

Because a lot of research has been conducted focusing on post-seismic landslides after the 2005 Kashmir Earthquake, extensive data is available, allowing for further research. Therefore, the 2005 Kashmir earthquake has been selected as a base for this research. Many researchers have studied post-seismic landslides after the 2005 Kashmir Earthquake. However, an analysis of elevated landslide risk caused by the earthquake during the western monsoon season has not yet been conducted. This study is progressive and unique by studying whether the time between the 2005 Kashmir earthquake and the first western monsoon season could have been used to reduce the impacts of those landslides. Gaining insight into the earthquake's legacy effect on the extent and spatial variability of landslide risk could greatly benefit the IFRC for disaster risk reduction.

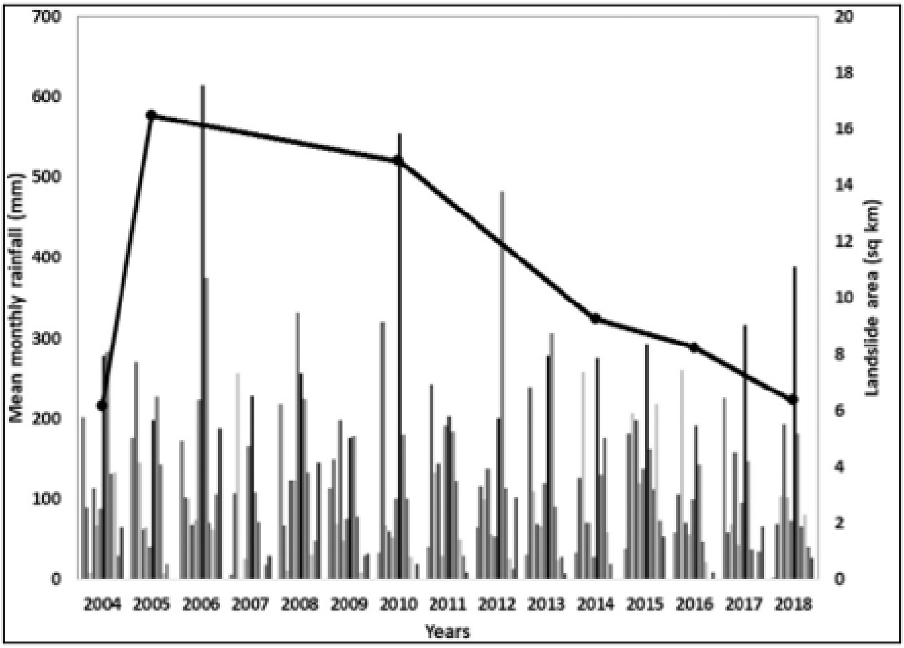


Figure 2.1:

"Comparisons of the temporal mean monthly rainfall with landslide dynamics. The bar graph represents the mean monthly rainfall, and the line represents the landslides area."

(Shafique, 2020)

2.3 Seismic hazard analysis of the 2005 Kashmir earthquake

Khan et al. (2020) have conducted a seismic hazard analysis in Muzaffarabad, Pakistan, with the moment tensor solution of the Kashmir earthquake. The moment tensor solution is the mathematical representation of the fault rupture. The output is a seismic hazard map for the study area. They found that ridges and slopes facing away from the centroid moment tensor (CMT) generally led to amplification of the seismic response, and de-amplification generally occurs in valleys and on the lower end of slopes that are faced towards the fault. Furthermore, they evaluated the effect of topography on co-seismic landslides. This was done by correlation of the amplification pattern with data on co-seismic landslides. This showed that over half of the landslides occurred on slopes facing away from the CMT, while a little over a quarter of landslides were recorded on slopes facing towards the CMT. As cited in Khan et al. (2020), this is in line with previous research (Ashford et al., 1997; Ashford and Sitar, 1997). As indicated in the article, the seismic hazard map, the methodology, and the results can be used to identify potential landslide zones when combined with other factors, such as local geology and anthropogenic factors (Khan et al., 2020). Furthermore, the model could be combined with rainfall forecasts to predict where the risk of both post-seismic and rainfall-induced landslides will increase.

2.4 Forecast-based Action & Forecast-based Financing

As discussed in the introduction, the IFRC refers to early actions as FbA. These are summarized in EAPs. To apply for funding through FbF, a national society needs to develop an EAP while adhering to a set of guidelines (IFRC, nda). Figure 2.2 provides the process for establishing a new EAP. Starting with risk assessment of the hazard, then identification of forecasts, then establishing threshold levels and selecting the early actions that will be used. Based on these aspects, the national society can develop an EAP. The EAP needs to be approved by the National Society management and the National Technical Committee. Upon approval, the EAP comes into force. Once a danger level is exceeded, funding is released, and the pre-established early actions are enforced (IFRC, nda).

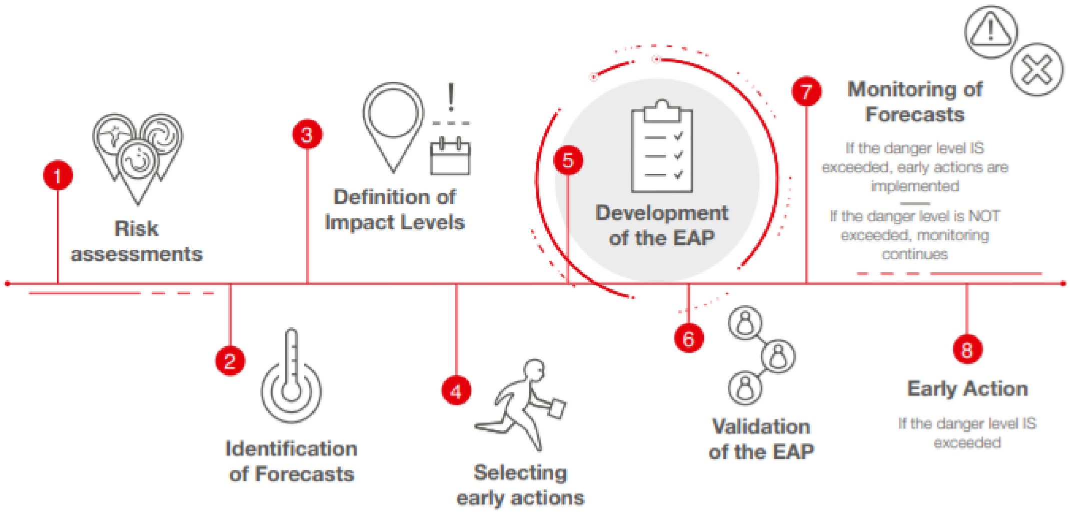


Figure 2.2: Step by step development process for Early Action Protocols (IFRC, nda)

EAPs that are currently in operation focus on several types of hazards, such as floods, droughts, dzuds, heat waves, volcanic ash and cyclones. The funding is also discussed in these EAPs. To further elaborate, some EAPs will briefly be discussed.

The EAP for cyclones in Mozambique first states the actors that have prepared and are responsible for implementation. These include the Mozambique Red Cross, the German Red Cross and the National Institute of Disaster Management. The threshold value at which early actions will be initiated was based on a 30-year analysis focused on historical data and impacts of past cyclones in Mozambique. The danger level was set at expected wind speeds of 120 km/h at landfall. The lead time for cyclones in Mozambique is 72 hours (IFRC, 2019a,b). According to the EAP, the main impacts of cyclones are damages to buildings and infrastructure, crops and agricultural tree losses, losses of critical assets and increases in waterborne diseases. Several early actions were established based on the available time between the trigger and the predicted cyclone, the Mozambican Red Cross ability to act with partners, and the identified impacts. The prioritized impacts are increases in diseases and the destruction of houses and classrooms. First of all, the EAP suggests supplying the communities with basic materials and essential tools to strengthen their rooftops and mud walls against wind and rains. Another measure aims to reduce waterborne diseases by distributing chlorine (Certeza) and buckets to communities before the cyclone hits. By doing so, people have clean water available in the days after the cyclone, which will reduce the number of disease cases (IFRC, 2019a,b).

In 2019, an EAP for cold waves in Peru was approved. Cold waves have a lead time of five days. The impacts that are prioritized in the EAP are morbidity and mortality of livestock and acute respiratory infections. The proposed early actions are the provision of materials for house insulation, temporary shelters for alpacas, veterinary kits, and warm clothes for young children. Furthermore, the focus will be on raising health promotion and disease prevention awareness (IFRC, 2019a).

Another EAP for Peru that was approved in 2019 focuses on flooding. The lead time is ten days. The prioritised impacts are safe water access, damages to assets, infrastructure and livelihoods, hygiene conditions and health risks. Some of the early actions discussed in the EAP include promotion of hygiene, prevention of diseases, provision of water filters and cash grants (IFRC, 2019a).

As previously discussed, currently, no EAPs have been established for landslides. This research could contribute to the EAP development process by providing information for the first four stages of an EAP development, namely risk assessments, identification of forecasts, definition of impact levels and early actions selection. Thus, in the longer run, this study's results could help establish an EAP for landslides in the study area and, consequently, help reduce the impacts of landslides.

2.5 Early warning early action landslides

EWEA is applied less for landslides than many other hazards. Most likely, this is due to the complex nature of landslides and failure mechanisms (Xu et al., 2020). Furthermore, landslides are very localized events, making effective EWEA strategies far more challenging. Because of this, relatively high increases in risk are needed to trigger a warning accurately. Lead time constraints could be another possible reason. In this study, the focus is on a combination of earthquake and rainfall-induced landslides. Major earthquakes have never been accurately predicted in the past, and it is not expected to be possible anytime soon. Only the probability that an earthquake will occur can be calculated (USGS, nd). Considering that earthquake-induced landslides happen during or very shortly after a seismic event, this leaves little to no room for the implementation of early action. For rainfall-induced landslides, some EWS have been applied. Typically, they use rainfall intensity duration thresholds, as well as meteorological modelling to derive rainfall forecasts that could trigger a warning if the threshold is exceeded (Piciullo et al., 2018). However, although also a concern with other hazard types, defining these thresholds is an issue. Thresholds that are too high result in lead times too short for emergency plans or could even miss the event. On the other side, lower thresholds values could lead to false warnings (Coughlan de Perez et al., 2015; Intrieri et al., 2013; Nadim and Intrieri, 2011).

2.6 Early Warning Systems for rainfall-induced landslides

Even though prediction of landslides is challenging, several studies have developed Early Warning Systems (EWS) for rainfall-induced landslides. Some of them are discussed in this section.

Xu et al. (2020) observed over a hundred landslides in Western China. They found that landslides typically undergo an evolution period from deformation until slope failure, in three clear phases. These creep slope failure phases have been referred to as initial, constant and accelerated deformation (Saito, 1961; Xu et al., 2011, cited by Xu et al., 2020). In the study, they developed an EWS for landslides, as well as a warning model. This is done using the rate of deformation and the improved tangent angle as early warning indicators (Xu et al., 2020). A four-level warning system is used based on the creep slope failure phases. Where initial deformation leads to no warning, constant deformation equals attention. When initial acceleration is reached, a caution warning is given, a vigilance warning is issued when a slope is in the medium-term acceleration phase and when eminent sliding is reached, an alarm warning is given (Xu et al., 2020). By the time the study was published in December 2020, the EWS had successfully predicted landslides 11 times (Xu et al., 2020).

Another study focusing on landslide EWS was conducted by Greco et al. (2013). They created a stochastic real-time landslide predictor for rainfall-induced landslides. They have combined the FlaIR model with a point rainfall stochastic model. The FLaIR model uses a slope mobility function that links landslide occurrence with antecedent rainfall characteristics. When a storm occurs, the expected value of the mobility function is based on empirical evaluation of only the historical storms that are characteristically similar to the occurring one (Greco et al., 2013). The predictor for rainfall-induced landslides was calibrated by using almost 48 years of hourly rainfall data collected by the rain gauge of Lanzo, located near the Pessinetto slope in Northern Italy. During the 48 year observation period, six earth flows had occurred. After calibration and validation, the model provided reliable predictions of the slope mobility function up until a lead time of six hours. It was then tested as part of an EWS for earth flows at the slope of Pessinetto. Two threshold values were determined, one for alert and one as an alarming level. They found that it is possible to gain some lead time hours for risk mitigation procedures if the threshold levels were accurately set (Greco et al., 2013).

Segoni et al. (2014) also focused on an EWS for rainfall-induced landslides. In this case for Tuscany, with a focus on a larger area of 23.000 km2. The EWS uses intensity duration rainfall thresholds, real-time rainfall data, as well as rainfall forecasts. It was produced in a WebGIS, where it is possible to focus on real-time rainfall data and forecasts at different lead times, with a maximum lead time of 48 hours. Furthermore, it is possible to focus on the entire area or zoom in on smaller areas. Because Tuscany has high variability in physical features, 25 subdivided alert zones were used. Each of these alert zones has a different threshold to account for the differences in physical characteristics. Each alert zone is monitored separately by making use of 332 rain gauges. As a result, warnings can be given for each of the alert zones separately. Thresholds may vary in time depending on the rainfall paths provided by the rain gauges (Segoni et al., 2014).

Even though these studies have been successful on local and regional scales, they are typically far more challenging to apply to larger areas. The risk of false alarms is much higher. Furthermore, they merely account for rainfall-induced landslides. This study aims to utilize the time between the 2005 Kashmir earthquake and the first western monsoon season after the earthquake to determine which early actions can be used to reduce the impacts of landslides during the rainy season. This approach has not yet been explored in FbF literature and thus could provide an exciting opportunity for landslide risk reduction by increasing the lead time for early action. Besides the study area, it could help reduce impacts of landslides during wet seasons after major earthquakes in other places.

Chapter 3

Research Design

The research design chapter outlines the methods and data used to achieve the research objectives. First, section 3.1 focuses on the establishment of the pre-seismic, co-seismic and post-seismic landslide inventories. Section 3.2 discusses the landslide controlling parameters that have been included in this study and explains the reasons to include each parameter. Section 3.3 outlines the methods applied to obtain the data matrices. Section 3.4 focuses on the landslide explanatory models, in this case, Generalized Additive Models (GAMs). Lastly, section 3.5 described the applied methodology for the landslide susceptibility maps.

3.1 Landslide inventories

To compare landslide activity before and after the 2005 Kashmir earthquake, several landslide inventories were established. A pre-seismic inventory for 2005 based on ASTER satellite imagery taken on the 9th of September 2005, a co-seismic inventory for 2005 based on ASTER satellite imagery taken on the 27th of October 2005, and a post-seismic inventory for 2007 taken on the 24th of April 2007. The pre-seismic and co-seismic inventories are based on satellite images taken before and right after the 2005 Kashmir earthquake on the 8th of October. For the post-seismic inventory, it was essential to include the monsoon seasons of 2006, which lasted from June through September, with the retreating monsoon period until November (Weather & Climate, nd). Suitable satellite imagery of right after the monsoon season was not available, and thus an image was selected from 2007. The satellite image used was taken before the start of the monsoon season of that year to make sure that the inventory only includes the monsoon season of 2006.

An overview of the landslide identification criteria can be found in table 3.1. The landslides for the pre-seismic inventory were extracted based on Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), cloud bands and slope threshold values. The landslides for the co- and post-seismic inventories were based on change detection using satellite imagery from the previous year. In the co-seismic and post-seismic inventory, an area was categorized as a landslide if the Barren Soil Index (BSI) increased by at least 20%, the NDVI with a negative change of at least 35% and if the Relativized Burn Ratio (RBR) showed an increase of 20%. Furthermore, the identified area was only classified as a landslide if it consisted of at least 5 connected pixels.

The co-seismic inventory was compared to the inventory of Sato et al. (2007) to test the landslide criteria that were applied. Sato et al. (2007) developed a landslide inventory for 2005 using SPOT-5 images. They mainly were identical, so it was assumed that the criteria used were sufficiently accurate to conduct the three inventories. The assumption was made because there was no opportunity to validate the results in the study area. It should be noted that the derived inventories are notably different from the inventories of Shafique (2020). Shafique (2020) established landslide inventories using ASTER and SPOT satellite images for 2004, 2005, 2010, 2014, 2016 and 2018. Thus, in theory, the 2004 and 2005 inventories could be used to validate this study's pre-seismic and co-seismic inventories. Shafique (2020) used a combination of visual interpretation of satellite images and field observations to derive the inventories, thus using a different method. Furthermore, the landslide polygons seem to be included manually instead of through image classification, reducing the preciseness of the inventories. Due to the differences in methods, the inventories cannot be precisely compared to the inventories of this study. Nevertheless, the spatial distribution of landslides in the 2004 and 2005 inventories is similar to this study's inventories. However, landslide count and size are very different, but this is expected as the inventories in this study were derived through change detection. In contrast, the inventories of Shafique (2020) reflect the aggregated number of landslides. Moreover, in this study, the landslides smaller than 900 m^2 are not included. In light of the major differences in the construction of the inventories, the decision was made to compare the co-seismic inventory to the inventory established by Sato et al. (2007), as it is a more reliable comparison.

Table 3.1: Landslide identification criteria for the pre-, co-, and post-seismic inventories. Pre-processing features for the pre-seismic inventory and change criteria for establishing co- and post-seismic inventories.

Pre-seismic inventory			**Co-, and post-seimic inventories**		
Pre-Processing Feature for Masking			**Landslide Criteria**		
Type	**Feature**	**Rule**	**Type**	**Feature**	**Rule**
Spectral	Normalized Difference Snow Index (NDSI)	Mask Snow	**Spectral and Temporal**	Barren Soil Index (BSI) Difference	> +20%
Spectral and Vector	Normalized Difference Water Index (NDWI)	Mask Water	**Spectral and Temporal**	Normalized Difference Vegetation Index (NDVI) Difference	> -35%
Spectral	Cloud Band	Mask Cloud	**Spectral and Temporal**	Relativized Burn Ratio (RBR) Difference	> +20%
Spatial	Slope	Mask < 10 degrees	**Spatial**	Size	> 5 connected pixels

The Digital Elevation Model (DEM) used in this research has a resolution of 30m. Therefore, only landslides larger than 900 m^2 have been selected. Furthermore, their impact is also significantly larger than smaller landslides. However, if polygons covering an area of fewer than 900 m^2 are located within a 15-meter proximity of nearby landslides, they were merged to the nearest landslide polygon and treated as a single landslide. They were kept in the data set because it is likely that they occurred at the same time. Note that this was merely done for the landslides smaller than 900 m^2, landslides within 15-meter proximity of other landslides that are larger than 900 m^2 were not merged and kept as separate landslides.

The polygons of the landslide inventories were transferred to points by selecting the point with the highest elevation. This was done to select the most likely initiation point of the landslide. The nearest DEM value was used if there was no point DEM value within the landslide polygon. It is important to note that in some cases a landslide polygon represents two separate landslides that appear to be one. However, part of this problem was eliminated after comparing the landslide polygons of the inventories with the ones from the previous year(s).The spatial distribution and landslide size of the three landslide inventories are displayed in figure 3.1, and the corresponding information on landslide count, size and total landslide area is provided in table 3.2.

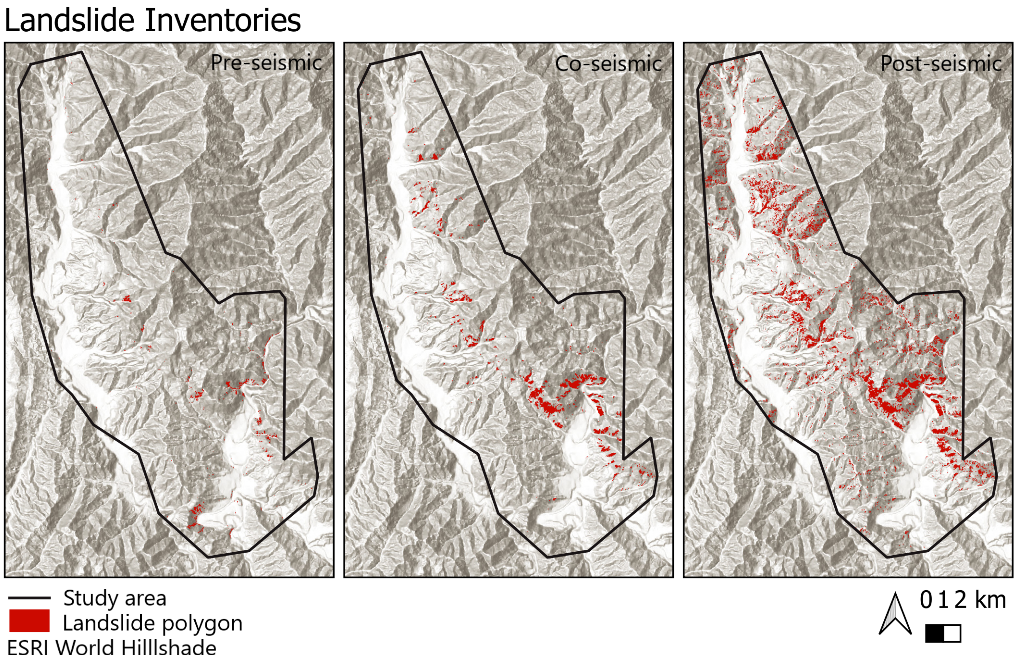


Figure 3.1: Maps of landslides in the pre-seismic, co-seismic and post-seismic landslide inventories

Table 3.2: Landslide count, total landslide area and average landslide size for the pre-seismic, co-seismic and post-seismic inventories

Year	**Pre-Seismic**	**Co-Seismic**	**Post-Seismic**
Number of landslides	242	411	2922
Total landslide area	1.36 km2	4.93 km2	13.86 km2
Average landslide size	5619.31 m2	11998.94 m2	4745.00 m2

3.2 Landslide controlling parameters

After developing the landslide inventories, a subsequent step was to determine the parameters that influence landslide activity. In total, 14 landslide controlling factors were used in this study that were selected based on a literature review of prior landslide studies. These parameters are elevation, slope, slope aspect, CMT angle, slope curvature, distance to fault, geology, land cover, road network, drainages, and mean monthly rainfall. The Peak Ground Velocity (PGV), Peak Ground Acceleration (PGA), and the Peak Ground Displacement (PGD) of the 2005 Kashmir earthquake, developed by Khan et al. (2020), were also used. Figures 3.3, 3.4, and 3.5 display the maps of each of these parameters. The mean monthly rainfall data from the Pakistan Meteorological Department that was used by Shafique (2020) is used to compare the rainfall records with the pre-, co- and post-seismic landslide variation.

The parameters were chosen for several reasons. The geology of the affected area determines landslide factors, including material and magnitude. The elevation, slope, slope curvature, and slope aspect affect the frequency of landslides as well as their spatial attribution (Gorum et al., 2011; Kamp et al., 2008; Korup, 2010, cited by Shafique et al., 2016). Steep slopes and land cover are considered major factors of landslides initiation (Basharat et al., 2016). The road network and drainages are also considered in this study as they could affect slope stability due to possible undercutting action (Basharat et al., 2016). Besides slope aspect, the CMT direction or angle was also included. CMT angle was chosen because Khan et al. (2020) found that ridges and slopes facing away from the centroid moment tensor generally led to amplification of the seismic response, and de- amplification generally occurs in valleys and on the lower end of slopes that are faced towards the fault. Furthermore, the study showed that more than half of the landslides occurred on slopes facing away from the CMT, while a little over a quarter of landslides were recorded on slopes facing towards the CMT. As cited in Khan et al. (2020) this is in line with previous research Ashford et al. (1997); Ashford and Sitar (1997). The CMT angle was categorized into three categories; 1. Slopes facing towards fault rupture, 2. Slopes facing away from fault rupture, and 3. Slopes facing in other directions.

The DEM of the study area was used to visualize the elevation. The DEM of the study area was then used to derive the slope, slope aspect, and slope curvature in QGIS. The geology of the study area was derived from the geology map of Shafique (2020), which was based on prior maps created by Hussain et al. (2004), and Latif et al. (2008). The geology information was georeferenced in QGIS and covers most of the study area. Figures 3.3, 3.4, and 3.5 show maps of each of the parameters for the study area. The Copernicus Dynamic Global Land Cover Layer was used to determine the study area's land cover types (Buchhorn et al., 2019). The classified land cover data is from 2015. Although 2015 is roughly a decade later than the dates of the landslide inventories, it was determined that the land cover has hardly changed through google earth observations of these years. Thus the data will be accurate enough for this study. The road network was obtained through extraction of OpenStreetMap and checked with Google Earth for any changes over the years. The PGV, PGA, PGD of the 2005 Kashmir earthquake were developed by Khan et al. (2020), and cut to the extent of the study area in QGIS. The fault rupture was derived by Khan et al. (2020) and was used to calculate the distance to a fault rupture. The maps of the different parameters are displayed below.

Mean monthly rainfall data from the Pakistan Meteorological Department was used to reflect on the rainfall factor for landslide occurrence. The rainfall data in millimetres for 2004, 2005, and 2006 is displayed in table 3.3. The rainfall data is for the study area used by Shafique (2020). However, since the area is very similar to the study area of this study, it was decided that the data will be sufficiently accurate to use. To justify, a study area comparison is shown in figure 3.2, which shows that the study area is almost identical. Initially, identical study areas were selected. Due to some data constraints, slight alterations were made to the study area to accommodate the available data.

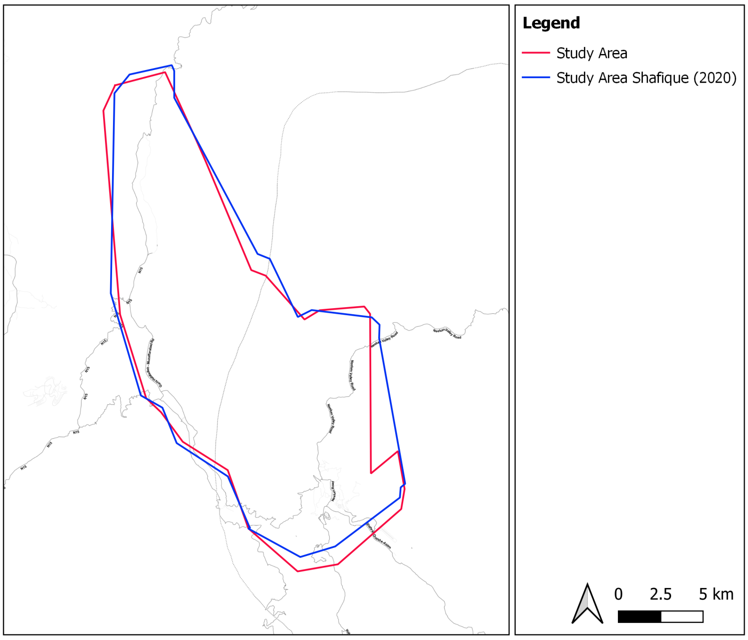


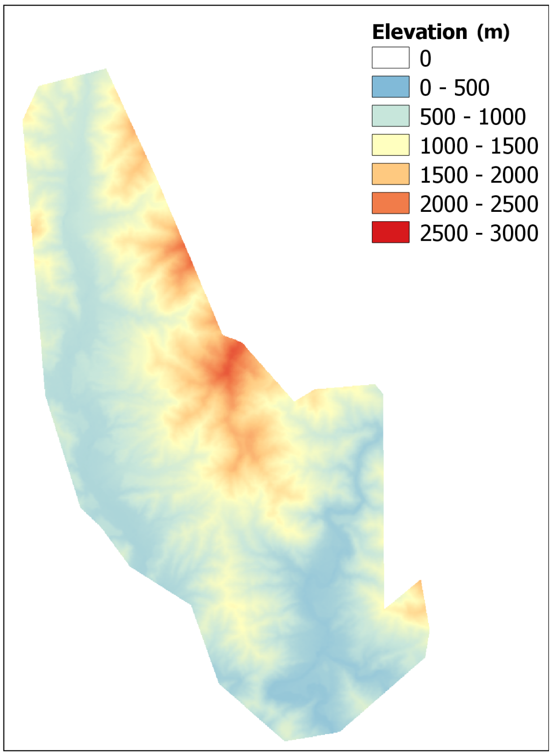
Figure 3.2: Study area of this study and study area used by Shafique (2020) showing the overlap to justify the use of rainfall data from the Pakistan Meteorological Department.

The pre-seismic inventory contains detectable landslides also from before 2004, and thus the rainfall of 2004 is not a completely accurate representation. However, because the rainfall parameter for the pre-seismic inventory is somewhat of an average of multiple years, it is thus considered a constant. Furthermore, it is assumed that the rainfall parameter remains constant throughout the inventories. The spatial variability of the rainfall is therefore not considered in this research. This was done because this study aims to assess the earthquake's legacy effect on landslide occurrence. Therefore, although many of the landslides are triggered by seismic activity and rainfall, the focus is on the earthquake parameters. In other words, the focus is on identifying landslide-prone slopes that are likely to fail when heavy rainfall occurs. Including spatial variability in rainfall would overcomplicate the study and deviate from focusing on the earthquake's legacy effect. Therefore, to effectively assess, the rainfall parameter is assumed to be constant throughout.

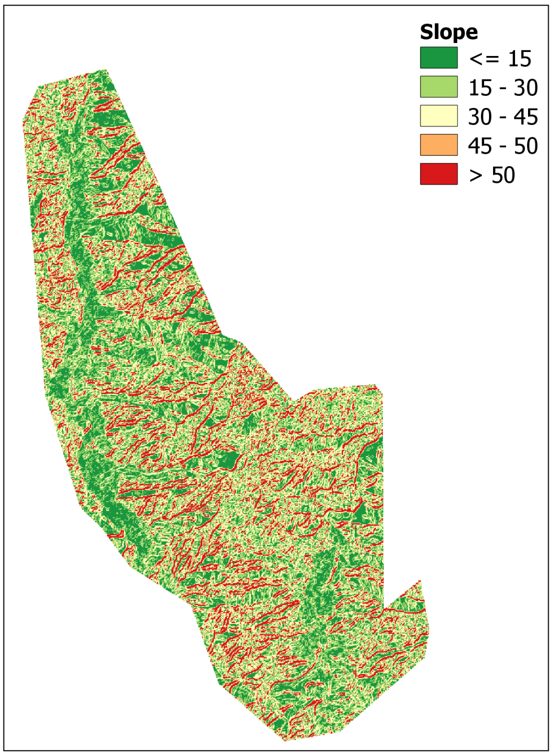
Table 3.3 shows the monthly rainfall data. The number of asterisks indicates the month the satellite image was taken to derive each of the inventories. One asterisk (*) is used for the pre-seismic inventory, two asterisks (**) for the co-seismic inventory and the post-seismic inventory is indicated by three asterisks (***). As shown in table 3.3, the mean monthly rainfall of the pre-seismic inventory is 131.60 mm and 143.02 in the post-seismic inventory, which further indicates the rainfall remained rather constant. The 30-day rainfall average of the co-seismic inventory is a lot lower, with merely 66.70 mm. However, satellite images of the pre-seismic and co-seismic inventories are only 48 days apart. Therefore, the 30-day average is not a valid representation.

Table 3.3: Mean Monthly Rainfall in mm Shafique (2020) (Pakistan Meteorological Department). The number of asterisks (*) indicate the date of satellite imagery that was used for the pre- (*), co- (**), and post-seismic (***) inventories. Based on the dates of the satellite imagery, the 30 day rainfall average for each of the inventories was calculated.

	2004	**2005**	**2006**	**2007**
January	201.1	175.1	171.9	6.2
February	89.2	270.1	101.6	107.1
March	8.8	145.6	100.1	256.5
April	112.6	61.9	68.1	63***
May	67	63.9	74	125.8
June	88.1	40.6	223	165.4
July	278.1	197.4	613.8	227.9
August	282.3	226.4	374.9	108.5
September	132.3	143*	70.8	72
October	132.8	7.6**	62.3	0
November	29.7	19.1	105.3	18.5
December	65	0	187.4	29
Total	1487	1350.7	2153.2	1179.9
Monthly Average	123.9	112.6	179.4	98.3
Date Image	Up until Image date		30 day average	
* 01/01/2004 - 09/09/2005	2710.9		131.6	
** 09/09/2005 - 27/10/2005	106.7		66.7	
*** 27/10/2005 - 24/04/2007	2593.5		143	

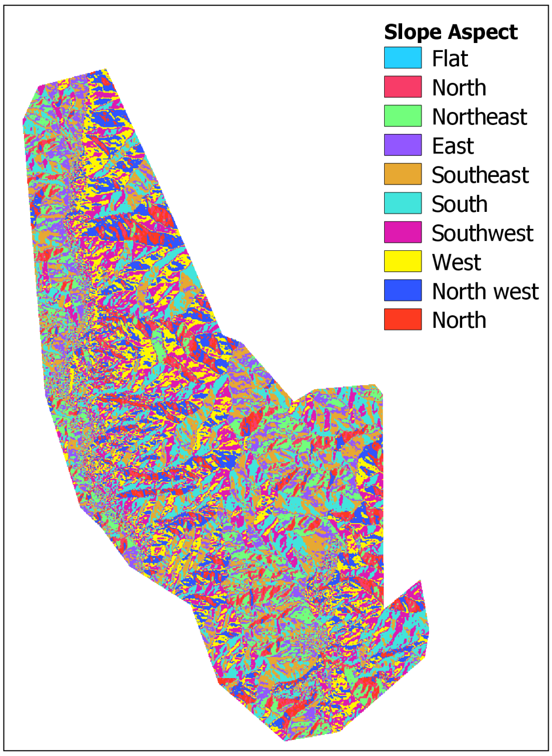


(a) Elevation (m)



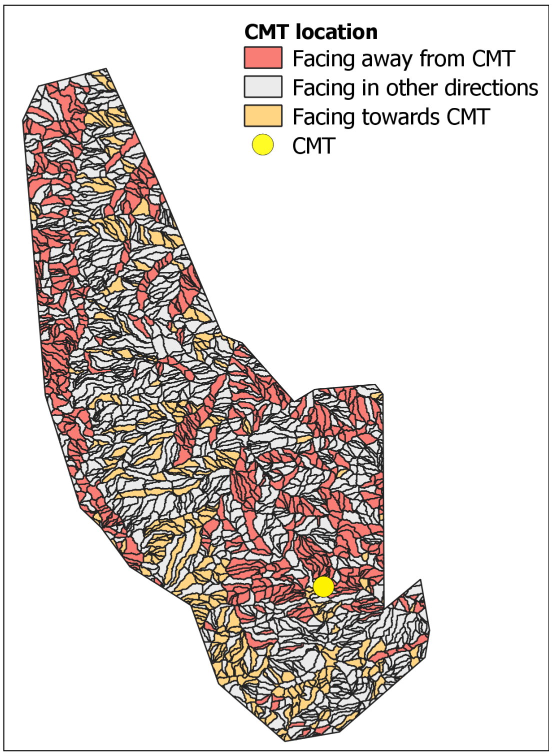
(b) Slope

(c) Slope curvature

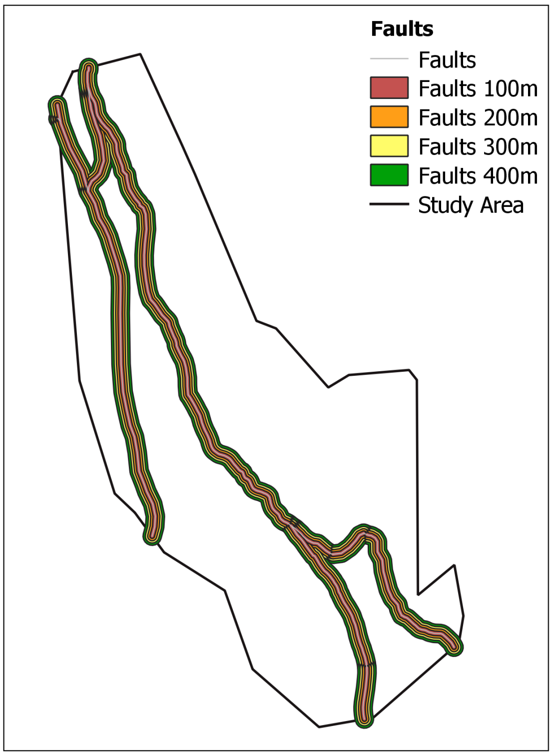


(d) Slope aspect

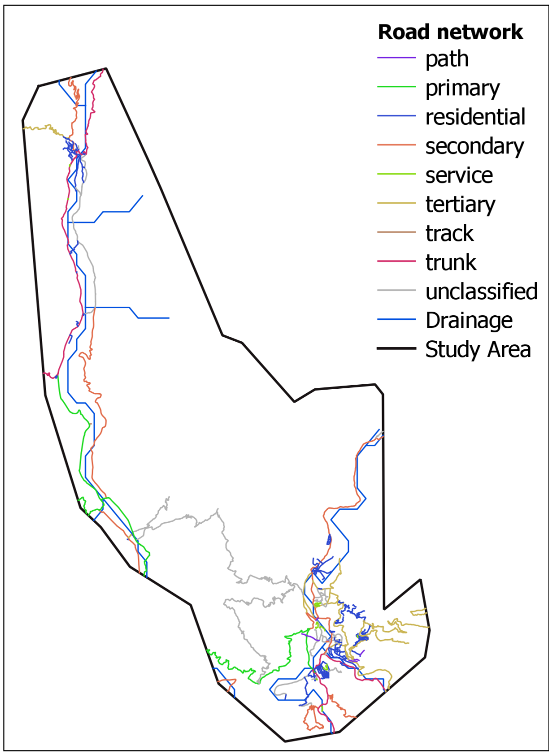
Figure 3.3: Maps of the elevation, slope, slope curvature and slope aspect parameters



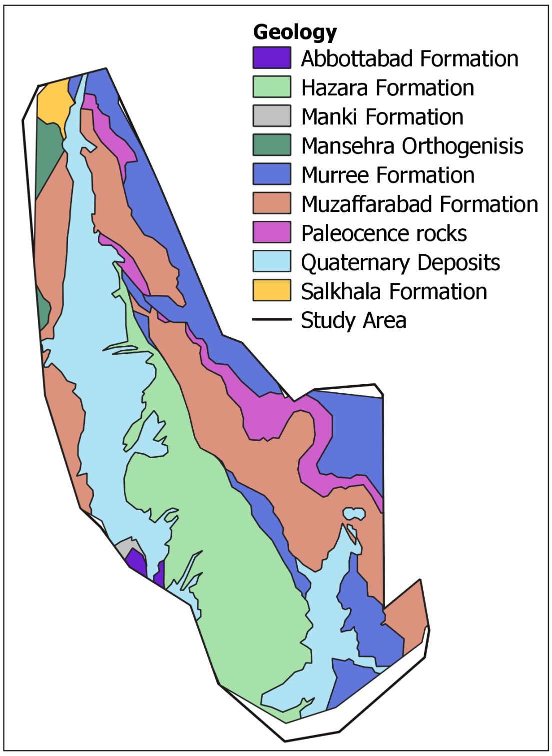
(a) Centroid Moment Tensor direction



(b) Faults

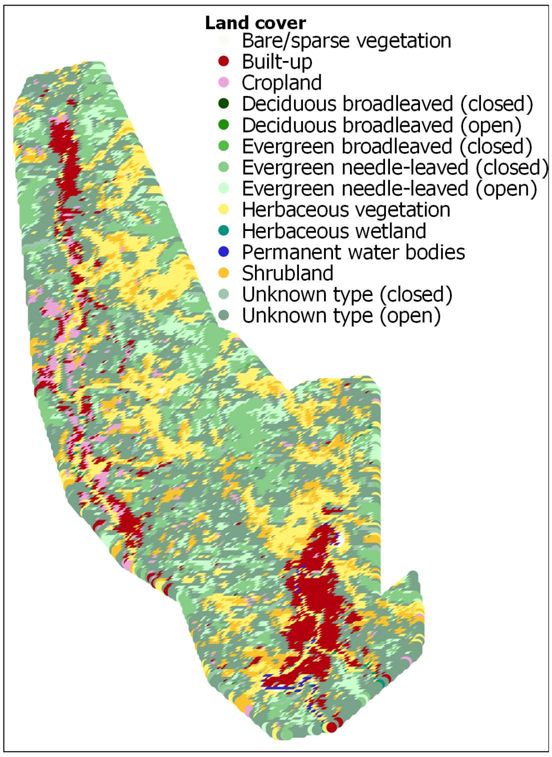


(c) Road network and drainage

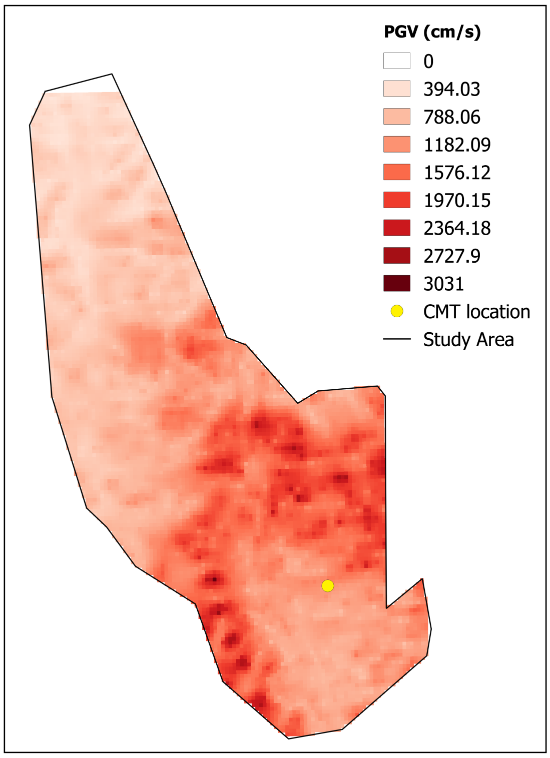


(d) Geology

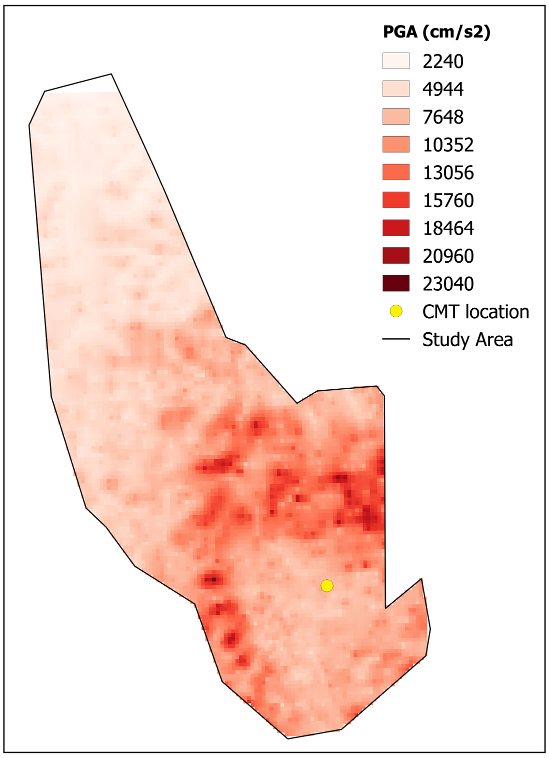
Figure 3.4: Maps of the CMT direction, faults, road network and drainage and geology parameters



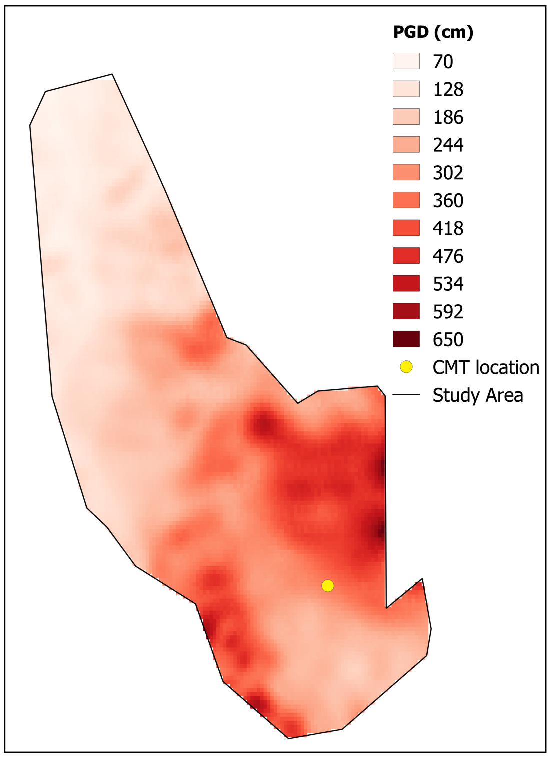
(a) Land cover



(b) Peak Ground Velocity (cm)



(c) Peak Ground Acceleration (cm/s^2)



(d) Peak Ground Displacement (cm)

Figure 3.5: Maps of the land cover, PGV, PGA and PGD parameters

3.3 Data matrix

After collecting the data on each parameter, a data matrix was created to incorporate landslide occurrence and the landslide parameters in the explanatory models. To obtain the data matrix, the study area was divided into mapping units, in this case, slope units. For landslide occurrence, the most likely initiation points of the landslides were then used to count the number of landslides within each slope unit. Subsequently, the landslide occurrence was determined, using a presence-absence structure, categorised as 'landslide' or 'no landslide'. The programme QGIS was used to establish the data matrix. The slope units were used to link all information of the different parameters, as well as landslide occurrence for 2005 (pre-seismic), 2005 (co-seismic) and 2007 (post-seismic). This was done in a variety of ways. The mean of each slope unit's elevation, slope, slope aspect, and slope curvature was used. The ground motion parameters PGV, PGA and PGD, have a resolution of 270 meters. The PGD was already obtained in a raster format. PGV and PGA were interpolated using Inverse Distance Weighting (IDW) interpolation, with the same cell size as the PGD. Due to the limited precision of the data, in many cases, only one value for PGV, PGA and PGD were located in a slope unit. Therefore, only the mean value of the ground motion parameters PGV, PGA and PGD was used. The percentage of land cover and geology type that cover each slope unit was calculated. Distance to fault, road network and drainage was done using buffers. For fault, buffers of 100, 200, 300 and 400 meters were calculated, and for the road network and drainage, a 50-meter buffer zone was used. Similar to geology and land cover, the percentage of each slope unit that falls within these buffers was calculated. Figure 3.6 provides an overview of the steps taken to derive the data matrix, including the input data used for the parameters and the processing steps for the data matrix. For reproducibility purposes, the exact QGIS processing steps are discussed in appendix 1.

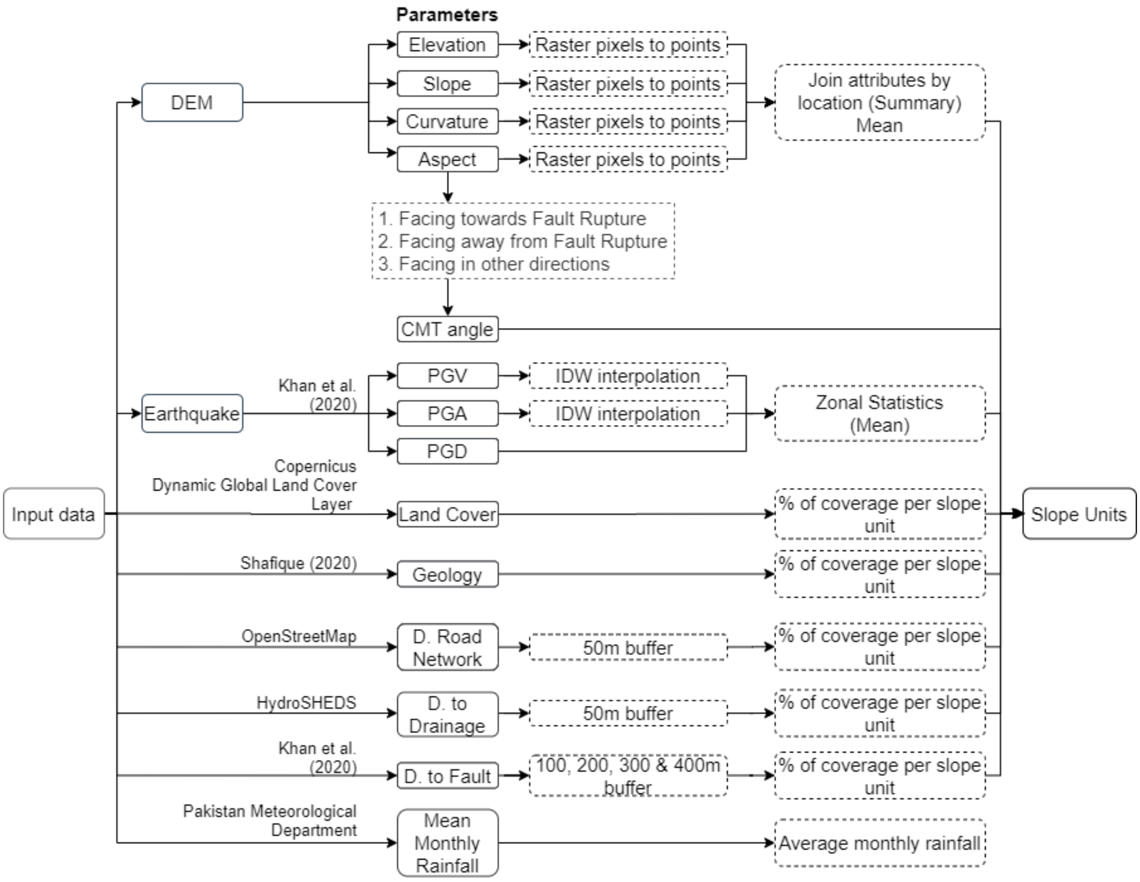


Figure 3.6: Flowchart of the methodology for the data matrix, indicating the processing steps to derive data on the parameters for each of the slope units.

3.4 Generalized Additive Models

To derive the explanatory hazard models, a Bayesian version of the Generalized Additive Models (GAMs) (3.1) has been used. GAMs are regression models that are more capable of deriving non-linear fittings than traditional statistical models. Non-linear explanatory variables can be fitted using various functions to reveal their complex relationships. In other words, a binomial GAM can model both linear (fixed) and non-linear (random) effects. The random-effects can include categorical, ordinal and continuous variables with discrete or ordinal classes. These can be modelled as independent and identically distributed effects (Lombardo et al., 2020). Moreover, GAMs are flexible, as they can be applied to various distribution types (Ma et al., 2020). The GAMs were first introduced by Hastie and Tibshirani (1990). The advantages of GAMs are that explanatory variables can be added by function and their capacity to deal with non-linear relationships between numerous response and explanatory variables.

In total, five explanatory GAMs have been conducted. One pre-seismic model, based on the pre-seismic inventory and all parameters, excluding the earthquake parameters. A co-seismic model, based on the co-seismic inventory and all the parameters. Three post-seismic models, based on the post-seismic inventory. One of the post-seismic models includes all parameters. One post-seismic GAM uses the parameters without the earthquake inputs and one with just the earthquake parameters. To clarify, table 3.4 provides an overview of the three data matrices used for the GAMs. The pre-seismic GAM uses data matrix 1, the co-seismic GAM uses data matrix 2, and for the post-seismic GAMs, all three data matrices are applied.

Table 3.4: An overview of the parameters in the three data matrices. The pre-seismic GAM uses data matrix 1, the co-seismic GAM uses data matrix 2, and the three post-seismic GAMs use data matrices 1, 2 and 3.

Pre-seismic & Post-seismic		**Co-seismic & Post-seismic**		**Post-seismic**	
Data matrix 1		**Data matrix 2**		**Data matrix 3**	
Elevation	D. to drainage	Elevation	D. to drainage		
Slope	D. to roads	Slope	D. to roads		
Curvature		Curvature	PGV*		PGV*
Aspect		Aspect	PGA*		PGA*
Land cover		Land cover	PGD*		PGD*
Geology		Geology	CMT angle*		CMT angle*

*Earthquake parameter

It is important to note that a common approach for landslide susceptibility mapping is to split the data set in both training and validation data to examine the model's predictive capacity. However, contrary to many other studies, the entire dataset is used to train and validate the models in this study. This method has been chosen because the post-seismic models aim to assess the potential effects of the earthquake parameters on landslide susceptibility. The aspect that has been changed is either inclusion or exclusion of the earthquake parameters. The traditional training and validation datasets are typically used to train a model to predict landslides, whereas the GAMs in this study are explanatory models. The validation data in prediction models is used to determine the performance of the model. But, in this study, the primary purpose is to compare the models. Therefore, all other elements are kept constant. Before modelling, slope units with incomplete data were removed from the data matrix.

Furthermore, all remaining covariates (parameters) were rescaled using mean zero-unit variance rescaling. Mean zero-unit variance rescaling is done by subtracting each covariate by its mean and subsequently divided by its standard deviation. The models are set up using a binary presence-absence variable. In other words, each slope unit is categorized in either landside(s), or no landslide(s), which corresponds to a Bernoulli probability distribution.

The Integrated Nested Laplace Approximation (INLA) package was used to derive the explanatory GAMs in the software environment R. The INLA package in R was used to approximate Bayesian inference in Latent Gaussian Models (R-INLA Project, nd). Equation 3.1 provides the equation used for the Bayesian version of the GAM used in this research.

$$\eta(P) = \beta o + \sum_{j=1}^{J} \beta j Z j(s) + f_{geology} + f_{landcover} + f_{roads} + f_{drainage} + f_{cmt} + f_{faults} \tag{3.1}$$

Where:

$$\begin{aligned}
\eta &= \text{Logit link} \\
P &= \text{Landslide susceptibility} \\
\beta_j &= \text{Estimated regression coefficients for each of the covariates } (z_j) \\
z_j &= \text{Covariate} \\
f_{geology} &= \text{Categorical properties} \\
f_{landcover} &= \\
f_{roads} &= \\
f_{drainage} &= \\
f_{cmt} &= \\
f_{faults} &=
\end{aligned}$$

Moreover, the INLA package allows the use of hidden data, in this case, landslide occurrence, which allows for prediction and validation of the model's capacity. By suppressing the variable landslide occurrence, Receiver Operating Curves (ROC) and the corresponding Area Under the Curve (AUC) could be derived. The ROC curves and corresponding AUCs were used to determine the performance and accuracy of each of the GAMs could be determined. However, as explained priorly, in this case, the validation of the model's performance is done on the same data. The ROC curves were calculated to compare the overall performance of the models with varying data matrices. Still, the primary focus is on the relative effect of the covariates within each of the five explanatory GAMs.

3.4.1 Assessment of the landslide controlling parameters to include in the models

After establishing the data matrix, it was essential to determine whether all 13 parameters needed to be included in the explanatory GAMs. As discussed priorly, the parameter rainfall will not be included in the GAMs, because it is assumed to remain constant throughout the studied timespan. The parameters were reevaluated to avoid inter-factor dependencies in the dataset.

To detect inter-factor dependencies, intercorrelations were calculated. Strong intercorrelations in the dataset can be problematic as the variables overly affect the model's output. Different methods have been applied to find these inter-factor dependencies, all of which were done using a 95% confidence interval. For most combinations of variables, the Spearman correlation coefficients were calculated. Spearman's correlation was used because not all data is normally distributed. Spearman's correlation is a rank-based correlation method for continuous variables. It is non-parametric and therefore does not require a normally distributed dataset. It is also more robust to outliers. The association between CMT angle (categorical variable) and the other continuous variables is calculated using Eta squared.

Besides intercorrelations in the dataset, detecting multicollinearity in the data was essential. Multicollinearity is a linear relationship between one or more variables. Although similar, it is not the same as correlation. Correlation is a linear relationship between two variables. Multicollinearity can also exist between a single variable and a linear combination of other variables (Alin, 2010). In a regression model between a response variable and explanatory variables, multicollinearity can become an issue as the coefficients are hard to interpret (Alin, 2010). In this case, landslide occurrence is the response variable, and the various parameters are the explanatory variables. In GAMs, multicollinearity could lead to unstable results (Ma et al., 2020). However, merely looking at correlation is insufficient to determine multicollinearity, as it could still exist when all correlations are low. To determine multicollinearity between variables and the linear combination of other variables in the data set, the Variance Inflation Factor (VIF) has been calculated. The VIF measures how much the interaction with other independent variables influences the variance of an independent variable. In other words, it is a measure of multicollinearity. VIF is calculated for each independent variable separately.

3.5 Landslide susceptibility maps

Besides a comparison of the relative effects of the covariates in each of the five GAMs, it is important to assess the spatial variability and intensity of landslide susceptibility among the five GAMs. This was done to determine whether the earthquake's legacy effect of the 2005 Kashmir earthquake affects the spatial variability of landslide susceptibility in the study area. And ultimately determine whether the elevated landslide susceptibility is significant enough to apply to early action. Therefore, four landslide susceptibility maps were created. A pre-seismic, a co-seismic map and two post-seismic susceptibility maps. The post-seismic GAM, which includes only the earthquake parameters, was excluded from this part of the analysis because it was merely meant to evaluate the influence of the earthquake parameters on landslide occurrence.

Apart from the landslide susceptibility, the susceptibility uncertainty was also determined to reflect on the certainty of the landslide susceptibility in each slope unit. Based on the GAMs of each inventory, the mean of the probabilities is used to define the susceptibilities of each slope unit. The susceptibility uncertainty is the difference between the 97.5 percentile and the 2.5 percentile of the same probability. Maps were made of the landslide susceptibility and susceptibility uncertainty to reflect on the changes in landslide susceptibility in the pre-seismic, co-seismic and post-seismic models.

Chapter 4

Results & Discussions

This chapter presents the results and discussion of the conducted study. Section 4.1 provides the reasoning behind the exclusion of some of the parameters in the GAMs. Where subsection 4.1.1 discusses the correlations and associations of the parameters in the data matrix. Subsection 4.1.2 examines the multicollinearity of the parameters and the reasoning behind altering data matrices. Section 4.2 provides the fixed and random effects of the five GAMs and provides an analysis of the parameters that influence landslide occurrence in the GAMs. Furthermore, the effect of the earthquake parameters on landslide occurrence is analyzed. Section 4.3 sheds light on the accuracy of each of the GAMs, followed by section 4.4, which provides the susceptibility maps of the pre-seismic, co-seismic and the two post-seismic GAMs. Subsection 4.4.1 then reflects on the differences between these susceptibility maps. The final section, section 4.5, discusses the applicability to FbF.

4.1 Generalized Additive Models preparation

As described in chapter 3, before conducting the GAMs, the data was re-evaluated to avoid inter-factor dependencies in the data. First, intercorrelations were calculated. This was done to avoid over-representation of data in the GAMs. The correlation and association matrix is displayed in appendix 3, and the important intercorrelations are described in section 4.1.1. Besides intercorrelations in the data, VIFs of the parameters were calculated to determine whether any multicollinearity could be detected in the data. The multicollinearity is described in section 4.1.2.

4.1.1 Correlations and associations parameters

Although somewhat arbitrary limits, the correlations in this study have been categorized as follows. The correlations of (-)0.0 – (-)0.19 are seen as neglectable correlations, (-)0.20 – (-)0.39 are low correlations, (-)0.40 – (-)0.59 are considered moderate, (-)0.60 – (-)0.79 as moderately high, and (-)0.8 – (-)1 as a very strong correlation (Sinscov and Campbell, 2002). In the appendix, the different strengths of the correlations are indicated using colours. The moderate and moderately high correlations are not directly relevant and are therefore not discussed in this section, but explanations for the correlations are provided in the appendix. In this section, only the strong correlations will be discussed that need to be assessed before conducting the GAMs. Strong intercorrelations in the dataset are problematic when conducting a GAM because they can overly influence the model's outcome.

There are two groups of parameters that need to be reassessed considering their intercorrelations. Very strong intercorrelations are found between the fault buffers. The ground motion parameters PGV, PGA, and PGD also show very strong intercorrelations. Table 4.1 shows the intercorrelations of the fault buffers, and table 4.2 provides the very strong correlations found among the ground motion parameters.

Table 4.1: Intercorrelation table, showing the Spearmann's correlation coefficients of the fault buffers of 100, 200, 300 and 400 meters.

	Fault 100m	Fault 200m	Fault 300m	Fault 400m
Fault 100m	NA	0.919	0.865	0.825
Fault 200m	0.919	NA	0.949	0.912
Fault 300m	0.865	0.949	NA	0.965
Fault 400m	0.825	0.912	0.965	NA

All fault buffers show very high correlations among them. Fault distance 100-meters and fault distance 200-meters correlate with a value of 0.919. This means that almost all slope units containing a fault 100-meter buffer also include a 200-meter fault buffer. The correlation of faults 200-meters and 300-meters is 0.949. The 300 and 400-meters fault buffers are correlated with a value of 0.965. 100 and 300-meters fault distances are correlated with a value of 0.865. Fault distances 100 and 400-meters are correlated with a value of 0.825, and fault distances 200 and 400-meters show a correlation of 0.912. The very strong correlations can easily be explained by the locations of these buffers. The fault distance parameters are adjacent buffers, and thus it is highly expected that slope units fall within multiple of these buffers. Considering the adjacent features of the fault buffers, the intercorrelations of the fault buffers directly adjacent to each other are expected to be higher, which is reflected in the correlations.

Table 4.2: Intercorrelation table showing the Spearmann's correlations coefficients of the ground motion parameters PGV, PGA and PGD.

	PGV	PGA	PGD
PGV	NA	0.960	0.952
PGA	0.960	NA	0.905
PGD	0.952	0.905	NA

The ground motion parameters are also strongly correlated. PGV and PGA show a correlation of 0.960, PGA and PGD show a correlation of 0.905 and PGV and PGD correlate with a value of 0.952. Thus, the values of ground motion parameters can quite accurately be defined through the value of the other. Because the ground motion parameters reflect the amplification pattern of the earthquake, the velocity, acceleration and displacement are expected to show strong correlations. However, as priorly mentioned, this is problematic when applying a GAM due to the overrepresentation of data.

The very strong correlations of the ground motion parameters and the fault buffers can lead to unreliable results. Therefore, they are cause for concern, and need to be altered before conducting the GAMs.

4.1.2 Multicollinearity

Another important reason to focus on the strong correlations between ground motion parameters and the fault buffers is the risk of multicollinearity in the model. Multicollinearity is a linear relationship between one or more variables. Although similar, it is not the same as correlation. Correlation is a linear relationship between two variables. Multicollinearity can also exist between a single variable and a linear combination of other variables (Alin, 2010). In a regression model between a response variable and explanatory variables, multicollinearity can become an issue as the coefficients are hard to interpret (Alin, 2010). In this case, landslide occurrence is the response variable, and the various parameters are the explanatory variables. In GAMs, multicollinearity could lead to unstable results (Ma et al., 2020). However, merely looking at correlation is insufficient to determine multicollinearity, as it could still exist when all correlations are low. To determine multicollinearity between variables and the linear combination of other variables in the data set, the Variance Inflation Factor (VIF) has been calculated. The VIF measures how much the interaction with other independent variables influences the variance of an independent variable. In other words, it is a measure of multicollinearity. VIF is calculated for each individual independent variable.

The VIFs have been calculated separately for each of the models. As priorly determined, the pre-seismic model does not include the ground motion parameters and fault buffers. The co-seismic model includes all parameters. One post-seismic model includes all parameters, while the other excludes the ground motion parameters and fault buffer data. Table 4.3 displays the VIFs for the pre-seismic, co-seismic and the two post-seismic landslide data sets. Note that VIFs cannot be calculated for categorical variables, and therefore the CMT angle has been left out. Any VIF above 10 indicates serious multicollinearity. VIF values higher than 10 are marked with two asterisks (**), VIF values close to 10 are highlighted with one asterisk (*). Due to limited space, the geology and land cover types have been numbered in the table. The corresponding names are less relevant in this part of the analysis considering the aim is to detect multicollinearity in the data, but the land cover names are provided in appendix 3.

Many land cover types (percentages per slope unit) show extremely high VIF values, which is cause for concern because they indicate multicollinearity in the models. Basically, each land cover type can be accurately predicted based on other parameters, in this case, land cover types. The multicollinearity can be explained by that certain landcover types are almost always present together, and others are hardly ever present in the proximity of each other. For example, closed forest types are not found in built-up areas, and thus if one or more closed forest land cover types are present in a slope unit, it is evident that no built-up will be present. On the contrary, as was already indicated by the correlations, when various closed forest types are present, it is a mixed forest, and other tree species are typically present.

Furthermore, as already indicated by the correlation matrix, PGV, PGA, PGD, and the fault buffer variables show high VIFs. If one ground motion parameter has a relatively high or low value, the other ground motion parameters typically showcase that as well. Moreover, the fault distances also have high VIFs, and thus multicollinearity is present. As already indicated by the correlations in the previous subsection, the buffers are adjacent to each other. Therefore the presence of one fault buffer can imply the presence of the other.

To fix the multicollinearity in the data, some adjusting has been done to the data set. Instead of using percentages for land cover type, the land cover type that covers the majority of the slope unit is selected. Thus, land cover type will be included as a categorical variable in the dataset. The same is done for geology type to assure consistency within the data set. The multicollinearity of the fault buffers is a simple fix by merely using a categorical variable, where slope units will be categorized by the smallest fault buffer that falls within the slope unit. It has also been decided to exclude PGV, and PGA from the model, because the PGV and PGD information is already provided by other variables (mainly PGA). Table 4.3 also provides the new VIFs after adjustment of the data sets. There are no VIFs for fault distance, geology, and land cover since they have been transferred to categorical variables. After adjustment of the data set, no multicollinearity was detected, and thus these data matrices were used for the five explanatory landslide GAMs.

Table 4.3: Variance Inflator Factors (VIFs) of the continuous variables in the datasets. VIFs indicate the presence or absence of multicollinearity in the pre-seismic, co-seismic and the two post-seismic data matrices. The new VIFs after adjustment of the data sets are also provided. The VIF values greater than 10 are indicated with two asterisks (**), one asterisk (*) indicates VIF values close to 10.

	VIF pre-seismic	**New VIF**	**VIF co-seismic**	**New VIF**	**VIF post-seismic excl EQ**	**New VIF**	**VIF post-seismic incl EQ**	**New VIF**
SUArea	1.07	1.01	1.08	1.01	1.07	1.01	1.08	1.01
DEM	2.12	1.01	2.33	1.01	2.12	1.33	2.33	1.01
Slope	1.02	1.00	1.03	1.01	1.02	1.01	1.03	1.01
Aspect	1.15	1.01	1.17	1.01	1.15	1.01	1.17	1.01
Curvature	1.08	1.02	1.09	1.02	1.08	1.02	1.09	1.02

PGV			19.82**				19.82**	
PGA			10.14**	1.00			10.14**	1.00
PGD			9.71*				9.71*	
Geo 1	3.76		3.90		3.76		3.90	
Geo 2	1.05		1.06		1.05		1.06	
Geo 3	1.27		1.30		1.27		1.30	
Geo 4	4.29		4.40		4.29		4.40	
Geo 5	3.61		3.80		3.61		3.80	
Geo 6	1.90		2.02		1.90		2.02	
Geo 7	4.27		4.40		4.27		4.40	
Geo 8	1.15		1.18		1.15		1.18	
Geo 9	1.147		1.16		l.147		1.16	
LC 1	769.15**		771.21**		*769.15		*771.21	
LC 2	149731.67**		150157.86**		149731.67**		150157.86**	
LC 3	38932.54**		39041.79**		38932.54**		39041.79**	
LC 4	6.37		6.39		6.37		6.39	
LC 5	19.96**		20.01**		*19.96		*20.01	
LC 6	12.84**		12.87**		12.84**		12.87**	
LC 7	212221.42**		212818.69**		212221.42**		212818.69**	
LC 8	93886.61**		94151.50**		93886.61**		94151.50**	
LC 9	192712.33**		193255.57**		192712.33**		193255.57**	
LC 10	1998.72**		2004.44**		1998.72**		2004.44**	
LC 11	1869.59**		1875.11**		1869.59**		1875.11**	
LC 12	83592.56**		83827.68**		83592.56**		83827.68**	
LC 13	22640.20**		22704.35**		22640.20**		22704.35**	
LC 14	296528.68**		297366.06**		296528.68**		297366.06**	
Roads 50m	1.70	1.28	l.731	1.30	1.70	1.28	1.73	1.30
Drainage 50m	1.31	1.14	1.34	1.15	1.31	1.14	1.34	1.15
Faults 100m			25.65**				25.65**	
Faults 200m			109.35**				109.35**	
Faults 300m			134.41**				134.41**	
Faults 400m			40.07**				40.07**	

**VIF value over 10

*VIF value close to 10

4.2 Fixed and random effects Generalized Additive Models

Table 4.4 shows the fixed (linear) and random (non-linear) effects of each model. These represent the relation between the explanatory variables (landslide controlling parameters) and the dependent variable (landslide occurrence). Therefore, the fixed and random effects indicate the explanatory variable's strength and direction on landslide occurrence. Thus, the effects indicate the relative influence of each parameter on landslide occurrence. By comparing the fixed and random effects, it can be assessed which parameters influence landslide occurrence in the pre-seismic co-seismic and post-seismic GAMs. Ultimately, the influence of the earthquake parameters can be evaluated to determine a potential earthquake legacy effect on landslide occurrence.

The values in the table show the posterior mean values of each parameter's fixed and random effects. They also indicate whether there are positive or negative relations to landslide susceptibility. Whether the parameters are significant on a 95% confidence interval and thus affect landslide occurrence in the model is indicated with an asterisk (*). As can be seen in table 4.4, there are many insignificant effects. However, essentially the accuracy of the models is more important than the significance of the posterior means. Furthermore, the insignificant effects still affect the landslide occurrence, but their influence cannot be significantly proven. Considering the purpose is to compare the covariates and to determine which covariates significantly impact landslide occurrence, no parameters have been excluded from the models.

Both the fixed and random effects contain information about the posterior distributions. However, for the categorical covariates (geology, land cover, roads, drainage and CMT angle) and the ordinal variable (fault buffers), this is done for each category separately. They are separate data points instead of a line as for the fixed effects. Therefore, a posterior mean is provided for each class. This means that the random effects are more complex to interpret than the linear effects. Therefore, they are merely compared to each other, focusing on the relative differences. For clarification, it should be noted that the posterior means can be greater than one because they are posterior distribution density means instead of probabilities.

Table 4.4: Means of the posterior marginal densities of the fixed and random effects in the GAMs. The mean fixed effects reflect the influence of the continuous variables on landslide occurrence in the GAMs. The mean random values reflect the effect of the categorical variables and ordinal variable (fault distance) on landslide occurrence. Fixed or random effects that are significant on a 95% confidence interval are highlighted using an asterisk (*).

	Mean fixed and random effects				
	Pre-Seismic	Co-Seismic	Post-Seismic	Post-Seismic	Post-Seismic
	Parameters excl EQ	All parameters	Parameters excl EQ	All parameters	Only EQ parameters
Parameters	**Fixed effects**				
intercept	-3.20*	-2.56*	0.11	0.10	0.14
DEM	-0.74*	0.24*	0.46*	0.45*	NA
Slope	0.15*	0.47	0.02	0.02	NA
Aspect	-0.12	0.04	-0.11	-0.11	NA
Curvature	0.07	0.29*	0.18*	0.18*	NA
SU area	0.36*	0.48*	0.80*	0.80*	NA
PGA	NA	0.41*	NA	-0.05	0.00
	Random effects				
Geology type:					
Abbottabad Formation	-0.11	-0.09	-0.32	0.33	NA
Hazara Formation	0.16	-0.20	-0.17	-0.15	NA
Manki Formation	0.12	-0.03	0.13	0.14	NA

Mansehra Formation	-0.18	0.09	0.23	0.20	NA
Murree Formation	0.52*	0.01	-0.47*	-0.47*	NA
Muzaffarabad Formation	0.30	0.68*	0.37*	0.38*	NA
Paleocence Rocks	-0.05	0.13	0.59*	0.61*	NA
Quaternary Deposits	-0.55*	-0.50*	-0.63*	-0.65*	NA
Salkhala Formation	-0.16	-0.06	0.27	0.27	NA
Land cover type:					
Bare/sparse vegetation	-0.02	0.03	-0.07	-0.07	NA
Built-up	-0.24	0.49	-0.42	-0.42	NA
Cropland	-0.61	-0.81	-0.47	0.48	NA
Deciduous needle-leaved (closed forest)	-0.45	-1.89*	-1.16*	-1.15*	NA
Deciduous needle-leaved (open forest)	-0.38	-1.65*	-0.58*	-0.58*	NA
Herbaceous vegetation	1.65*	2.26*	1.32*	1.32*	NA
Permanent water bodies	-0.06	0.02	-0.17	-0.17	NA
Shrubland	0.77*	2.37*	1.43*	1.43*	NA
Unknown type (closed forest)	-0.44	-0.94	-0.45	-0.43	NA
Unknown type (open forest)	-0.02	0.64*	0.62*	0.62*	NA
Roads:					
No	0.00	0.10	0.03	0.03	NA
Yes	0.00	-0.10	-0.03	-0.03	NA
Drainage:					
No	0.20*	0.05	0.07	0.07	NA
Yes	0.20*	-0.05	-0.07	-0.07	NA
CMT angle:					
Away	NA	0.07	NA	0.09	0.18*
Towards	NA	-0.11	NA	-0.11	-0.21*
Other directions	NA	0.05	NA	0.02	0.03
Faults:					
100m	NA	-0.94*	NA	-0.36*	-0.09
200m	NA	0.46*	NA	0.01	-0.06
300m	NA	0.25	NA	0.09	0.02
400m	NA	0.20	NA	0.12	0.05
>400m	NA	0.05	NA	0.16	0.08

*Significant on a 95% confidence interval

4.2.1 Fixed effects

As priorly mentioned, the fixed or linear effects represent the relation between the explanatory variables and the dependent variable. In this case, the fixed effects reflect the influence of the continuous variables in the data matrix on landslide occurrences in the GAMs. This study aims to determine the impact of the earthquake parameters on landslide susceptibility. However, it will be important to evaluate all fixed effects to assess the overall shift in influence of the parameters across the models.

When looking at elevation, the parameter shows significant fixed values for the GAMs. Landslide occurrence was generally more influenced by elevation before the 2005 Kashmir earthquake, with a fixed effect of -0.74. Thus, indicating a strong negative relation in the pre-seismic model. In the co-seismic, the fixed effect is 0.24. Therefore the influence of elevation on landslide occurrence has become reversed, meaning that compared to pre-earthquake conditions, slopes at higher elevations are now more prone to landslides. This trend continues in the post-seismic GAMs. The fixed effects in the post-seismic GAMs are 0.46 in the post-seismic GAM, excluding the earthquake parameters and 0.45 in the GAM, including the earthquake parameters. Including the earthquake parameters hardly changes the effect of elevation on landslide occurrences. Nevertheless, the 2005 Kashmir earthquake clearly affected slopes at higher elevations and increased their susceptibility to landslides. If it were to be that slopes at lower elevations already failed, then this would be significantly reflected in the co-seismic GAM. Instead, the fixed effects of the GAMs show that the 2005 Kashmir earthquake destabilized slopes at higher elevations and consequently increased landslide occurrence at these altitudes.

Although only significant in the pre-seismic model, the effect of the slope parameter on landslide occurrence slightly changes over the GAMs. In the pre-seismic model, the slope has a significant fixed effect of 0.15, indicating that steeper slopes were slightly more prone to failure than more gradual slopes. The parameter has the most influence in the co-seismic model, with a fixed value of 0.47, indicating that the earthquake affected steeper slopes more. Even though the co-seismic fixed effect is insignificant, it is notable, especially because the fixed values in the post-seismic models are both 0.02. The insignificant fixed effects of 0.02 indicates that the parameter slope does not influence landslide occurrence in the post-seismic GAMs. In many cases, this was because slopes had already failed during or shortly after the earthquake and thus were relatively stable in the post-seismic model.

The fixed effects of the slope aspect are insignificant in each of the GAMs. In the pre-seismic, co-seismic and post-seismic models, the fixed effects are similar, respectively -0.12, -0.11, and -0.11. In the co-seismic model, it has a very low positive effect of 0.04. Furthermore, there is no difference in the mean fixed effects between the post-seismic models including and excluding the earthquake parameters. Considering the fixed effects in the GAMs are low and insignificant, it can be said that slope aspect does not significantly impact landslide occurrence.

In the pre-seismic GAM, slope curvature had an insignificant fixed effect of 0.07. However, this notably changes in the co-seismic and post-seismic GAMs. The curvature of the slope had significantly more influence on landslide occurrence in the co-seismic landslide inventory with a fixed effect of 0.29. This shows that slopes with higher curvature were more affected by the earthquake. After the 2005 Kashmir earthquake, this effect decreased slightly but remained present with positive fixed effects of 0.18. In the post-seismic models, there is no difference when the earthquake parameters are left out or included. However, during and after the earthquake, slope curvature does significantly affect landslide occurrence. Although speculating, a possible explanation could be that due to the curvature of the slopes, the amplification pattern of the earthquake had a more varied impact on the slope, contributing to the destabilization and failure of the curved slopes.

The slope unit area significantly affected landslide occurrence in each of the models. This is expected, considering that a larger area statistically has a greater chance of landslide occurrence. The pre-seismic inventory consists of 242 landslides; there are 411 landslides in the co-seismic inventory and 2922 landslides in the post-seismic inventory. This is also reflected in the fixed effects of each of the models. An increase in the total landslides leads to a higher fixed effect, with 0.36 in the pre-seismic GAM, 0.48 in the co-seismic GAM and 0.80 in the post-seismic models. These numbers are not completely in proportion. However, as previously mentioned, in this study, landslide count and landslide size have not been included in the models. Instead, a presence-absence landslide structure was applied. Furthermore, the highest elevation point of the landslide polygon was selected as the likely initiation point of the landslide. In some cases, this means that the majority of a landslide falls within one slope unit, but the landslide has been appointed to another slope unit in which the likely initiation point is located.

4.2.2 Random effects

Similar to the fixed effects, the random or non-linear effects of the GAMs represent the relation between the explanatory variables and the dependent variable. In this case, the random effects reflect the influence of the categorical variables (geology, land cover, roads, drainage and CMT angle) and the ordinal variable (fault buffers) on landslide occurrences in the GAMs.

In the four GAMs, the random effects of the geology types vary quite a bit. The most notable and influential differences in random effects are discussed. The geology type Muzaffarabad Formation shows a substantial increase in the co-seismic model, from an insignificant effect of 0.30 in the pre-seismic model to a significant effect of 0.68 in the co-seismic model. This is in line with the results of Owen et al. (2008), which found that after the 2005 Kashmir earthquake, most landslides occurred in the Muzaffarabad formation. The significant random effects in the post-seismic models are 0.37 and 0.38. Thus, landsliding became more prominent in the Muzaffarabad formation during and after the 2005 Kashmir earthquake. Besides the Muzaffarabad formation, the random effects of Abbottabad formation, Palaeocene rocks and the Salkhala formation also increased in the co-seismic GAM. However, the corresponding lithology of these formations do have some common features. The Muzaffarabad formation mainly consists of dolomite, quartzite limestone, and sandstone. Quaternary deposits are comprised of gravel, clay, and sand. The corresponding lithology of the Hazara formation is slate, siltstone, limestone, and shale (Table 1.1). Palaeocene rocks are mostly comprised of sandstone, shale, and limestone (Shafique, 2020). The Abbottabad formation is predominantly comprised of dolomite with subordinate quartzite, conglomerate, and siltstone (Qasim et al., 2014). Overall, the earthquake significantly impacted these geology types, causing landslides to be more prominent.

What stands out for the post-seismic model is the major difference in random effects of the Abbotabad formation, respectively -0.32 and 0.33. However, these are insignificant. As shown in the map in figure 3.4-d, the geology type only covers a minor part in the southwest of the study area. There are only 10 slope units out of the 1843 slope units with dominant geology type Abbottabad formation. Out of these 10 slope units, two experienced slope failure in the post-seismic model. These slopes faced away from the CMT angle, fall within the shortest fault buffer of 100 meters, and had relatively low PGA values. Considering, the low negative PGA effect of -0.05, the low random effect of 0.09 of the CMT angle and the larger significant random effect of -0.36 of the shortest fault buffer, the model has classified the geology type Abbottabad formation as a positive effect, because these two slopes did experience slope failure. Logically, this effect would be insignificant considering the small number of slope units.

For land cover types, the changes in random effects will be explained separately for each of the classes. Vegetation typically stabilizes slopes, and as a result, reduces a slope's susceptibility to landslides (Huang and Fan, 2013). Earthquakes can damage vegetation, and thus it is expected to see this reflected in the random effects of the land cover types.

The first landcover type is bare/sparse vegetation. The random effect of bare/sparse vegetation in the pre-seismic GAM is -0.02. The effect is 0.03 in the co-seismic GAM and -0.07 in both the post-seismic models, all of which are insignificant. Considering that a lack of vegetation would make a slope more prone to landslides, a stronger positive effect would be expected. However, only one slope unit has bare/sparse vegetation as the dominant land cover type. For that slope unit, there is no landslide occurrence in either of the three models. Considering there is only one slope unit, this is merely a coincidence. Therefore, the random effects are not a good indication of the effects of bare/sparse vegetation on landslide occurrence.

For land cover type built-up, all random effects are insignificant. The random effect is -0.24 in the pre-seismic model, 0.49 in the co-seismic model and -0.42 in the post-seismic models. Thus, typically landslide occurrence would be less present on built-up land. This is expected, considering that settlements are less often present on slopes, although often near slopes. However, even though an insignificant effect, in the co-seismic model, the random effect is positive. Each of the slope units with land cover type built-up that experienced slope failure in the co-seismic model were either stand-alone built-up areas or located at the edge of built-up areas. Therefore, it makes sense that the random effect was positive. There were only 11 slope units with landslides in the co-seismic GAM, which explains the insufficient evidence to reject the null hypothesis for the random effect. Nevertheless, it does reflect the impact of landslides the 2005 Kashmir earthquake had on the region.

The random effects of cropland are -0.61 in the pre-seismic model, -0.81 in the co-seismic model, -0.47 in the post-seismic model excluding the earthquake parameters and 0.48 in the post-seismic model, including the earthquake parameters. There were no landslides in slope units with major land cover type cropland in the pre-seismic and co-seismic inventory. Thus the negative random effects are expected. On the contrary, there were nine landslides in the post-seismic inventory with dominant land cover type cropland. Eight out of the nine slope units fall within the 100-meter fault buffer, and the average PGA was a little over half of the mean PGA of all the slope units. Five slopes faced in other directions than toward or away from the CMT location. The remaining slopes faced away from the CMT. The low impacts of the earthquake parameters on the slope units explain why the random effect of cropland is positive. Especially the significant negative effect of -0.36 of the 100-meter fault buffer explains the positive influence of cropland on landslide occurrence. However, considering there are only nine failed slopes, an insignificant effect is not surprising.

Deciduous needle-leaved (closed forest) is dense vegetation. Thus adverse random effects are expected. The random effect is -0.45 in the pres-seismic GAM, although insignificant. In the co-seismic GAM, the random effect is significant and strongly negative with a value of -1.89. Thus, considering the dense vegetation, the earthquake did not affect slope units with deciduous needle-leaved closed forest. In the post-seismic GAMs, this effect remains strong and significant with -1.16 in the model excluding the earthquake parameters and -1.15 in the model including the earthquake parameters. The relatively weak random effect in the pre-seismic GAM can simply be explained by the much lower landslide occurrence in that inventory. For decideous needle-leaved (open forest), it is still expected to see negative effects, considering the stabilizing effects of vegetation. However, since open forest is less dense than closed forest, the effects should be less strong. This is exactly what can be observed from the random effects of deciduous needle-leaved (open forest). In the pre-seismic model, the insignificant random effect is -0.38. The random effects are significant in the co-seismic and post-seismic GAMs, with -1.65 in the co-seismic model and -0.58 in both the post-seismic GAMs.

The random effects of herbaceous vegetation are significant and strongly positive throughout the GAMs. In the pre-seismic model the random effect is 1.65, the effect is 2.26 in the co-seismic model and 1.32 in both the post-seismic models. Herbaceous vegetation consists of plants such as grasses, sedges and ferns. Their non-woody stems will not stabilize slopes and will therefore show higher positive random effects throughout the models. Especially in the co-seismic model, the 2005 Kashmir earthquake strongly affected slope units with predominantly herbaceous vegetation.

Slope units that are mostly covered by water bodies should show negative random effects. However, since the dominant land cover type classifies slope units, many will not be covered entirely by permanent water bodies. This is also reflected in the random effects, which are insignificant in all GAMs. The random effect of permanent water bodies is -0.06 in the pre-seismic model, 0.02 in the co-seismic model, and -0.17 in both the post-seismic models.

Shrublands are dominated by low dense shrub vegetation. The random effects of shrubland are significant and positive throughout the GAMs. In the pre-seismic model, the random effect of shrubland is 0.77, in the co-seismic GAM, the effect is 2.37, and in the post-seismic models, the random effects of shrubland are 1.43. Similar to the land cover type herbaceous vegetation, the random effects of shrubland clearly show that shrubland is highly prone to landslides compared to other vegetation types that help stabilize slopes. The 2005 Kashmir earthquake also significantly impacted slopes with shrubland causing large numbers of landslides.

The final two land cover types are unknown type open forest, and unknown type closed forest. Similar to the other closed and open forest types, it is expected that the unknown type closed forest will show stronger negative random effects than the open forest land cover type. As expected, the unknown type closed forest has negative random effects of -0.44 in the pre-seismic model, -0.94 in the co-seismic model, -0.45 in the post-seismic model, excluding and -0.43 in the post-seismic model, including the earthquake parameters. All these random effects are insignificant, but considering they are unknown types of forest, this land cover class most likely includes various tree species, and their exact effects are difficult to determine. The random effects of the unknown type open forest in the pre-seismic GAM is insignificant and -0.02. The random effects are significant and positive in the co-seismic and post-seismic models, with 0.64 in the co-seismic GAM and 0.62 in both the post-seismic models. This means that the open forest is likely not very dense considering the positive random effects. If the open forest were dense, the tree's roots would stabilize the slopes and show negative random effects.

As discussed in chapter 3, the presence or absence of roads and drainage in a slope unit was included due to their possible undercutting actions (Basharat et al., 2016). The random effects of road and drainage were compared to assess whether any undercutting action of roads or drainage was observed in the GAMs.

Whether a road is located in a slope unit or not has little effect on landslide occurrence in any of the models. Only in the co-seismic model, there is a slight negative relationship between roads and landslide occurrence, meaning that the absence of a road increases the susceptibility in the co-seismic GAM. However, the random effects in all the GAMs are almost 0, ranging between -0.10 and 0.10. They are also insignificant.

The presence or absence of drainage in a slope unit does not differ in landslide susceptibility in the pre-seismic model. The significant posterior means of presence and absence are both 0.20. This is possible because the categorical covariates are separate data points. In the co- and post-seismic models, the numbers are close to 0 and insignificant. Therefore, the initial theory that the undercutting action of roads and drainage could lead to landslides has not been observed.

4.2.3 Fixed and random effects of the 2005 Kashmir earthquake parameters

As expected, PGA influences landslide occurrence much more in the co-seismic model compared to the post-seismic model. The significant effect in the co-seismic model is 0.48. This reflects the immediate effect of the 2005 Kashmir earthquake on landslide occurrence, leading to the 411 landslides in the co-seismic inventory. The effect is insignificant and merely -0.05 in the post-seismic model, which indicates that slopes highly affected by PGA had already experienced slope failure during or shortly after the earthquake. However, the fixed effect of 0.48 in the co-seismic model is not extremely strong, instead the effects of other parameters such as the slope curvature, Muzaffarabad Formation, herbaceous vegetation, open forest types and shrubland are significantly stronger. They became more impactful on landslide occurrence due to the effects of the earthquake on these parameters. They are more prone to landsliding anyway, but due to the ground shaking of the 2005 Kashmir earthquake they are significantly more susceptible to landslides. Considering the neglectable fixed effect of PGA in the post-seismic model and the lower effects of many other covariates that were affected by the earthquake, it is clear that no lasting significant effect of PGA on landslide occurrence was observed.

Nevertheless, other earthquake parameters, such as fault distance and CMT angle, still significantly affect slope stability in 2007. Especially slopes located within 100 meters of a fault had a significant effect. Unexpectedly, this was a strong negative relation of -0.36, even though almost around half of the slopes within a 100-meter buffer failed, the other fault distances experienced relatively more landslides in relation to the number of slope units. This explains the negative effect. In the co-seismic model, this was -0.94. The 200-meter fault buffer had a significant fixed effect of 0.46 in the co-seismic GAM. However, this could be explained by the fact that all faults were included in this analysis. The proximity to the CMT location would have likely shown something different. However, in this study the decision was made to incorporate the angle to the CMT location, based on the results of a previous study by Khan et al. (2020).

The only parameter that seems to have a significant effect in the post-seismic model with only the earthquake parameters is the CMT angle. More specifically, slopes facing away and towards the CMT location. In the study conducted by Khan et al. (2020), half of the landslides were reported on slopes facing away from the CMT, and a little over a quarter on slopes facing towards the CMT. The random effect in the GAMs for slopes facing towards the CMT location are negative relations, respectively -0.11, -0.11, -0.21. For slopes facing away from the CMT, these are 0.07, 0.09 and 0.18. The positive random effects of the slopes facing away from the CMT show that most of the landslides occurred on slopes facing away from the CMT location. This is in line with the results of Khan et al. (2020). The negative effects of the slopes facing towards the CMT show that significantly fewer landslides occurred on these slopes. The random effects of slopes facing in other directions are close to zero and insignificant. Thus, the results in this study differ slightly. However, it should be noted that Khan et al. (2020) categorized the slopes slightly different. They used a 60° set of aspect to its angle towards the CMT. This study has used a 45° angle, which makes the results less comparable to the study conducted by Khan et al. (2020) than if the same categorization was used. Nevertheless, the major finding that slopes facing away from the CMT experienced most slope failures is compatible with the results of this study.

4.2.4 Earthquake's legacy effect on landslide occurrence

The five GAMs with varying parameters were compared to see if this approach could identify an apparent earthquake legacy effect on landslide susceptibility or occurrence. As discussed in the previous sections, some evident changes in posterior means of the fixed and random effects were observed. However, these observations were primarily found in non-earthquake parameters, such as slope, elevation, geology and land cover types. To determine a clear earthquake legacy effect on landslide susceptibility, the earthquake parameters need to affect landslide occurrence in the post-seismic GAMs significantly. In the post-seismic model, including the earthquake parameters, only the smallest fault distance had a relatively strong significant influence on slope failure, but this observed random effect is negative. A GAM was conducted with only the earthquake parameters to assess the relative effects of the covariates. Here, landslide occurrence mainly was influenced by the CMT angle, which is in line with previous research (Khan et al., 2020). The results show that the CMT angle is the only earthquake parameter that significantly affects landslide occurrence and distribution. For the other earthquake parameters, no earthquake legacy effect on landslide occurrence was observed. To conclude, after comparing the mean fixed and random effects, no clear influence of the 2005 Kashmir earthquake was observed. Instead, non-earthquake parameters seem to affect landslide occurrence in the model significantly more. Even though, in many cases this is due to the effects of the 2005 Kashmir earthquake on the stability of the slopes.

4.3 Performance Generalized Additive Models

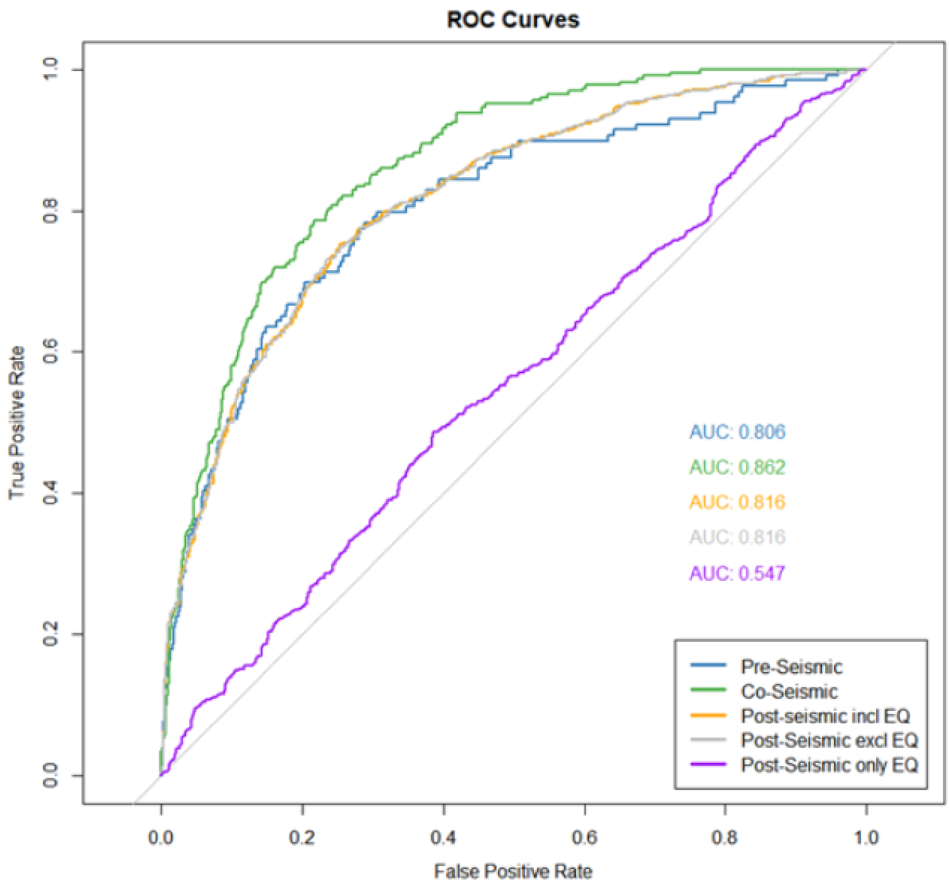


Figure 4.1: ROC curves of the pre-seismic, co-seismic, and the three post-seismic models with corresponding AUC values

The model performances of the GAMs are examined through Receiver Operating Characteristic (ROC) curves. ROC curves are used to show the diagnostic ability of binary classifiers. ROC curves show the trade-off between True Positive Rates (TPR, or sensitivity) and False Positive Rate (FPR, or 1-specificity). Curves closer to the top left of the graph indicate better performances, while curves closer to the centre show less accuracy in landslide prediction. In this case, the model performance has been tested by hiding the models' landslide occurrences or the response variables. In other words, the performance of the GAMs is tested on the dataset that is also used for the training. As previously discussed in chapter 3, this was done because the focus is to assess the potential effects of the earthquake parameters on landslide susceptibility. Accurate landslide prediction is less of a concern. The

ROC curves are displayed in figure 4.1. The performance of the models is summarized by calculating the Area Under the Curve (AUC), which provides the probability that a random positive outcome is ranked higher than a random negative outcome.

The pre-seismic model has the least accuracy, with an AUC of 0.806. The co-seismic model has the best performance with an AUC of 0.862. This is reflected by the strong significant fixed and random effects in the co-seismic model. The fixed and random effects of the post-seismic GAMs remain quite equal between both post-seismic models. Only the Abbattobad Formation, cropland and the 100-meter fault buffers show apparent differences in the random effects. All other non-earthquake parameters remain almost equal after the inclusion of the earthquake parameters. The remaining earthquake parameters have low, insignificant effects. Considering the minor differences in fixed and random effect, the ROC curves unsurprisingly follow almost identical paths. The post-seismic model, including the earthquake parameters and excluding the earthquake parameters, have an AUC of 0.816. Therefore, the ROC curves and corresponding AUC further substantiate the observations drawn from the fixed and random effects that it makes little difference when the earthquake parameters are included in the model to predict post-seismic landslide occurrence. This is reflected by the low AUC of 0.547 of the post-seismic models with only the earthquake parameters. The graph is very close to the centre and thus indicates a low accuracy. This is expected, considering only a few parameters have been used. However, a difference of 0.047 from the centre line further indicates that the earthquake parameters have little effect on the outcomes of the GAMs. To conclude, in the case of the 2005 Kashmir earthquake, including earthquake parameters has little to no impact on the performance and accuracy of the GAMs.

4.4 Susceptibility maps

Susceptibility maps have been made to examine the spatial distribution of the susceptibilities to landslides. No susceptibility map has been made of the model with only the earthquake parameters because it was merely meant to evaluate the effects of the earthquake parameters on the models more thoroughly. The susceptibility maps have been derived by taking the mean of the fitted values per slope unit. The susceptibility uncertainty is the 0.975 and 0.025 quantiles of the fitted values. The susceptibility uncertainties of the maps have been displayed below the susceptibility maps. Not all data was available for the entire study area, and therefore some slope units at the sides of the study area have been left out of the susceptibility analysis. The susceptibility maps and corresponding susceptibility uncertainty maps are displayed in figure 4.2.

The susceptibility maps in figure 4.2 show a slight increase in overall landslide susceptibility in the co-seismic model and a drastic overall increase in susceptibility in the post-seismic models. This is expected, considering the overall stronger fixed and random effects in the co-seismic and post-seismic models. Furthermore, the pre-seismic inventory consists of 242 landslides, the co-seismic inventory of 411 and the post-seismic inventory consists of 2922 landslides. Logically, this is reflected in the susceptibility maps. When comparing it to figure 3.1, the patterns are easily explained. Most of the slope units with high landslide susceptibility correspond with the landslide distribution in the corresponding inventories. To illustrate, an overlay of the susceptibility maps with the landslide inventories can be found in appendix 5.

However, acknowledge that for the dependent variable landslide occurrence, the slope units were categorized as either 'landslide' or 'no landslide', thus not including landslide size and landslide count. Therefore, it should be noted that including landslide count or landslide size would have altered the results of the GAMs and landslide susceptibility maps. Recently, landslide approaches have been introduced that include landslide count instead of merely a landslide presence or absence approach. Thus, focusing on the landslide intensity, which provides more detailed information (Lombardo et al., 2018a).

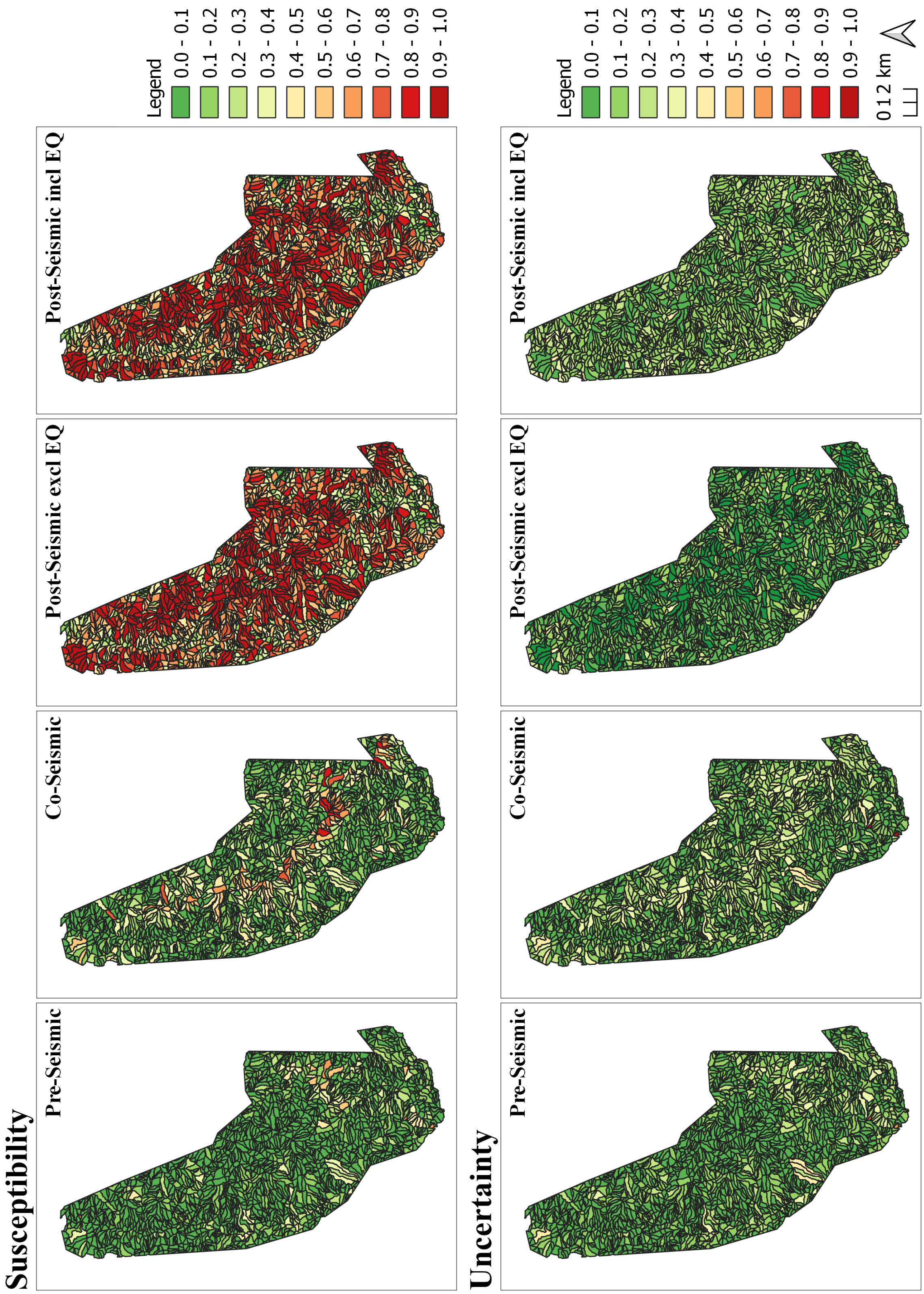


Figure 4.2: Susceptibility & susceptibility uncertainty maps

4.4.1 Susceptibility difference

4.4.1.1 Difference post-seismic and pre-seismic

The difference in susceptibility between the two post-seismic models and the pre-seismic model and the post-seismic model, excluding the earthquake parameters, have been mapped. These are displayed in figure 4.3. As shown in figure 4.3, there are significant susceptibility differences between the post-seismic and pre-seismic models. Even though the 2005 Kashmir earthquake has caused a tremendous increase in post-seismic landslides, the fixed and random effects of the GAMs showed that the earthquake parameters themselves have little effect on the spatial distribution of these landslides. Instead, geomorphological parameters become more important after the seismic activity. Thus, even though the 2005 Kashmir earthquake significantly impacted landslide intensity both during and after the event, this is mostly due to the increased effects of other parameters. A clear spatial effect of the Kashmir earthquake is not observed. Furthermore, slope curvature and slope unit size become more influential. Slope unit size is simply because of the increase in landslide activity.

4.4.1.2 Difference post-seismic models

As shown in the second map in figure 4.3, susceptibility hardly changes after including the earthquake parameters in the post-seismic models. The landslide susceptibility is determined based on the mean of the fitted values. Thus, susceptibility differences between the post-seismic models are logically really minor, as the fixed and random effects in the models are also very similar.

There seems to be a spatial pattern in the susceptibility differences between the models. However, the dark orange and purple slope units merely represent an increase or decrease of 0.1-0.15%, which is far too low to draw any conclusions on the presence of an earthquake legacy's effect on landslide occurrence and spatial variability. Thus, besides a lack of significant changes in the fixed and random effects, no significant difference in the spatial variability of landslide susceptibility after including the earthquake parameters in the GAM is detected.

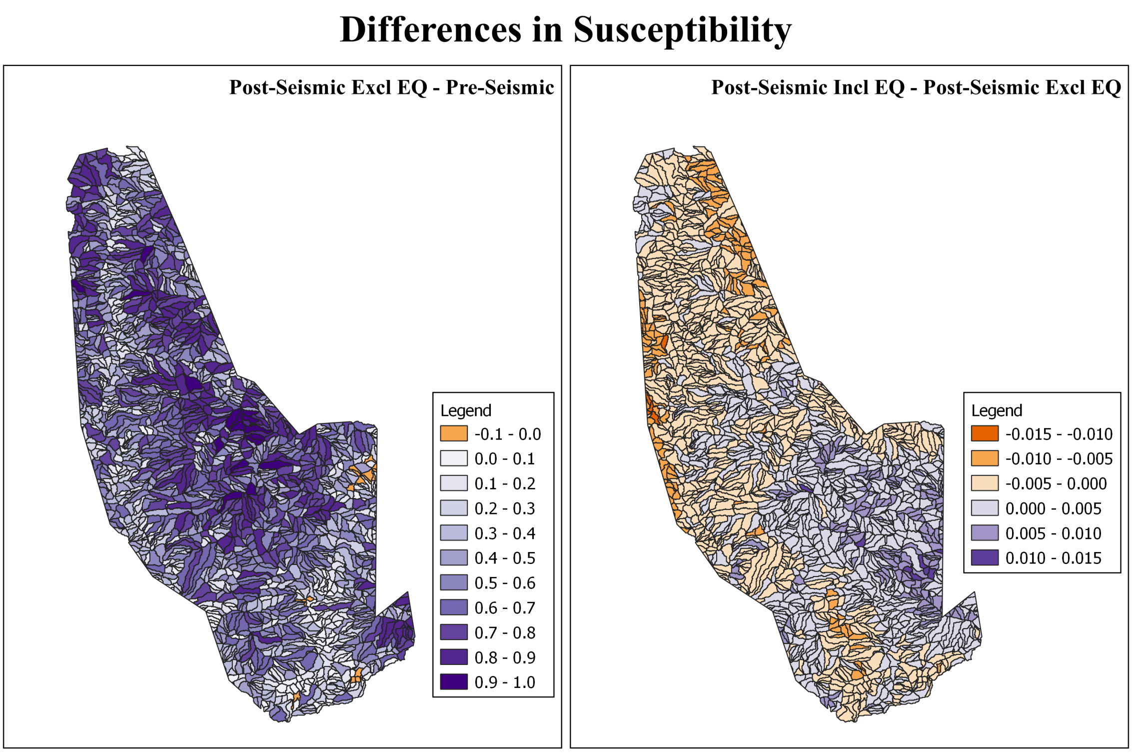


Figure 4.3: Susceptibility differences between the post-seismic models and the pre- and post-seismic model excluding the earthquake parameters

4.5 Applicability of methodology to Forecast based Financing

After comparing the mean fixed and random effects and the spatial variability in landslide susceptibility among the GAMs, this study shows no significant elevated landslide susceptibility when earthquake parameters are included in the GAMs. Between the post-seismic GAMs, including and excluding the earthquake parameters, only slopes facing away from the CMT were significantly more susceptible to landslides after the 2005 Kashmir earthquake. Furthermore, the spatial distribution of landslide susceptibility also stays almost equal after including earthquake parameters. Only a few slopes show elevated landslide susceptibility. However, these susceptibility differences are extremely small. A significantly elevated susceptibility is necessary for successful application to early action. Otherwise, the risk of false and missed landslide alarms is too high.

However, although the earthquake parameters do not show an apparent earthquake legacy effect on landslide occurrence in the GAMs, many other parameters show that the earthquake did significantly affect landslide occurrence in the study area. Parameters such as shrubland, herbaceous vegetation, elevation, slope curvature, and the Muzaffarabad Formation were more susceptible to landslides due to the 2005 Kashmir earthquake. This was concluded from the increases in fixed and random effects of these parameters in the co-seismic and post-seismic models.

Nevertheless, it is impossible to use the earthquake parameters to determine areas at increased risk of landslides, considering the effects of the earthquake parameters themselves do not show significant spatial variance in landslide susceptibility. Unfortunately, this means that the proposed approach to improve lead times for FbF of landslides is not successful. In other words, the time between an earthquake and an upcoming rainfall season cannot be utilized to solve the lead time constraints of landslides. Therefore, it is essential to remain inquisitive about new approaches to overcome the difficulties of FbF and EWEA for landslides. In the meantime, the focus should be on existing EWEA strategies to minimize the impact of landslides.

Chapter 5

Conclusions

This research aimed to investigate whether dynamic landslide risk analysis based on the earthquake parameters of the 2005 Kashmir earthquake could have informed early action to reduce the impacts of landslides during the first western monsoon season after the earthquake. In this chapter, the research results are evaluated to determine whether the main objective has been reached. This will be done according to the six sub-objectives of this research.

To start, the study aimed to develop pre-seismic, co-seismic and post-seismic landslide inventories for the study area and to determine and evaluate the increased intensity of landslides during the first monsoon season after the 2005 Kashmir earthquake. This was done through a comparison of landslide inventories from before, during and after the earthquake. The pre-seismic landslide inventory was extracted based on NDVI, NDWI, cloud bands and slope threshold values. The co- and post-seismic inventories were developed using change detection compared to the previous satellite image. The pre-seismic inventory consists of 242 landslides, with a total landslide area of 1.36 km^2, and an average landslide size of 5619.31 m^2. The co-seismic inventory includes 411 landslides, with a total landslide area of 4.93km^2 and an average landslide size of 11998.94m^2. The post-seismic landslide inventory consists of 2922 landslides, with a total landslide area of 13.86 km^2, and an average landslide area of 4745.00 m^2. According to these inventories, there was a massive increase in landslide occurrence after the 2005 Kashmir earthquake. The effects of the earthquake clearly increased landslide activity. However, the average size of landslides decreased by 15.56% compared to the pre-seismic inventory.

The second sub-objective was to develop pre-seismic, co-seismic and post-seismic explanatory hazard models for the study area based on the 2005 Kashmir earthquake. In total, five explanatory landslide models were conducted; a pre-seismic model, a co-seismic model, and three post-seismic models. A Bayesian version of a GAM was used. The post-seismic models were performed using varying parameters to determine the effects of the 2005 Kashmir earthquake on landslide susceptibility. One post-seismic model includes all parameters, one without the earthquake parameters and one only using the earthquake parameters as input. The response variable 'landslide occurrence' was then concealed from the model to evaluate the accuracy. The ROC curves of the post-seismic GAMs, including and excluding the earthquake parameters, had almost identical ROC curves. The AUCs of the ROC curves were both 0.816. The GAM with only the earthquake parameters showed a low performance and an extremely low accuracy with an AUC of 0.547. To conclude, the inclusion or exclusion of earthquake parameters does not affect the performance and accuracy of the landslide models.

The next step was to determine the extent to which susceptibility to post-seismic landslides had increased during the monsoon season in the first year after the earthquake. Four susceptibility maps were created; one pre-seismic map, one co-seismic and two post-seismic susceptibility maps. The GAM using only the earthquake parameters was not included because this was merely done to more accurately assess the interaction and influence of the earthquake parameters. As expected, there was a drastic increase in susceptibility between the pre-seismic model and the post-seismic landslide models. The spatial susceptibility patterns were in line with the spatial patterns of landslides in the inventories, considering that landslide occurrence was used as the dependent variable and not landslide count or size. Contrary to to initial expectations, including the earthquake parameters in the seismic model had limited effects on the susceptibility in the study area. As expected from the ROC curves, there were only minor differences between the post-seismic susceptibility maps. This was further reflected in the susceptibility difference map.

The fourth sub-objective was to assess what parameters affect pre-seismic, co-seismic and post-seismic landslide distribution. This was done by comparing the posterior mean fixed and random effects of the covariates in the GAMs. Most notable changes were found in the effects of the parameters slope, curvature, geology and the land cover types, CMT angle and the nearest fault distance.

The fixed effects of slope show that in the co-seismic model, the influence of slope on landslide occurrence is significantly more substantial. However, in both the post-seismic models, the fixed values are almost zero. The low fixed effects indicate that steeper slopes stabilized after the 2005 Kashmir earthquake. Similar to the parameter slope, the curvature of the slope had a significantly higher effect on landslide occurrence in the co-seismic model than the pre-seismic one. This indicates that slopes with higher curvature were more affected by the earthquake. After the earthquake, the effect decreased but remained higher than before the earthquake. The geology type Muzaffarabad Formation shows a substantial increase in the co-seismic model. This is in line with a previous study that found that most landslides occurred on slopes within the Muzaffarabad Formation after the 2005 Kashmir earthquake (Owen et al., 2008). After the earthquake, the geology type Palaeocene rocks had a much larger positive effect on landslide occurrence than before. The corresponding lithology of Palaeocene rocks is sandstone, shale, and limestone. This is interesting as some of the other geology types that had a greater impact on landslide occurrence share some similar lithological features. These are the Muzaffarabad formation, the Hazara formation and Quaternary deposits. The Muzaffarabad formation mainly consists of dolomite, quartzite limestone, and sandstone. Quaternary deposits are comprised of gravel, clay, and sand. The corresponding lithology of the Hazara formation is slate, siltstone, limestone, and shale. The main conclusion for land cover types is that dense vegetation with deeper roots has stabilized the slopes and show strong negative random effects in each of the GAMs. The earthquake did not affect these slopes. On the contrary, sparse vegetation and vegetation with non-wooded or shallow roots showed strong positive effects and were highly influenced by the 2005 Kashmir earthquake.

The fifth sub-objective was to determine whether the earthquake parameters clearly influence landslide distribution in the post-seismic landslide model. To determine an earthquake legacy effect on landslide susceptibility, this was a vital step. However, contrary to initial expectations, the earthquake parameters have little influence on landslide occurrence. This is reflected by the fixed and random effects, the ROC curve of the post-seismic GAM with only the earthquake parameters and the post-seismic susceptibility difference map. As already indicated in the previous paragraph, CMT angle and fault distance were the only random effects that were notable. The random effect in the GAMs for slopes facing towards the CMT location are negative relations with a slight negative increase in the post-seismic model. For slopes facing away from the CMT, the opposite is observed. Slopes located within 100 meters of a fault had a significant negative relation of -0.36 in the post-seismic model, including the earthquake parameters. However, the contrary would be expected, that they would be more prone to landslides. This negative relation is also visible in the co-seismic model with -0.94. Overall, the earthquake parameters did not clearly influence landslide susceptibility and therefore no apparent earthquake legacy effect on landslide susceptibility was observed. However, the study did show clear effects on the fixed and random effects of other parameters. Although, no clear earthquake legacy effect was detected from the earthquake parameters, parameters such as land cover types shrubland and herbaceous vegetation, geology type Muzaffarabad formation, and elevation were affected by the 2005 Kashmir earthquake. The fixed and random effects were significantly stronger in the co-seismic and post-seismic GAMs.

The final sub-objective was to assess the possibilities of risk reduction in the future based on the predictive capacity of the models. The study's result shows that it is not beneficial to include earthquake parameters in the case of the 2005 Kashmir earthquake. Even though the effects of several other parameters became stronger in the co- and post-seismic models due to the earthquake's impact, this was not reflected by the earthquake parameters. Thus, using parameters of the 2005 Kashmir earthquake cannot inform early action. In other words, this study shows no evidence that the time between an earthquake and an upcoming rainfall season can be utilized to determine locations where early action strategies can be applied to reduce the impact of landslides. Therefore, including earthquake parameters does not provide further possibilities for landslide risk reduction in the study area.

Chapter 6

Reflection

As previously discussed, including earthquake parameters did not have enough effect in the GAMs for increasing lead times for early action strategies. There are some limitations to the applied methodology that could have partially influenced the results of this study. These will be discussed in this section. However, considering the extremely minor observations of elevated landslide susceptibility after the inclusion of earthquake parameters, it is inevitable that in the case of the 2005 Kashmir earthquake, the proposed new approach would not have been successful in reducing the impacts of landslides.

The reflection is split up into three sections. Section 6.1 reflects on the establishment of the landslide inventories. Section 6.2 provides a discussion on the GAMs, and section 6.3 provides recommendations on further research.

6.1 Landslide inventories

First of all, the establishment of the landslide inventories was done using satellite imagery and change detection. The landslide inventories were established with the most precision possible, considering the possibilities within this study. Furthermore, the landslide inventories were tested to a previous co-seismic inventory established by Sato et al. (2007). No field validation was done, and thus the accuracy of the landslide inventories could have been further improved. It was also surprising that the inventories were notably different than the inventories established by Shafique (2020). However, the landslide polygons were significantly larger and established using a different method, making comparison and validation difficult. Nevertheless, the overall spatial distribution of the landslides in the inventories was similar.

Then, the likely initiation points of the landslides were determined by extracting the highest elevation point for each polygon. This is an estimation, and thus for some landslides, the location of initiation is likely different. In some cases, it could be that polygons in the inventories that represent one landslide could have been two separate landslides in real life. Furthermore, landslides smaller than $900m^2$ that were located within 15-meter proximity of nearby landslides were merged to the nearest landslide polygon and were treated as a single landslide. They were kept in the data set as it is likely they occurred at the same time. However, this cannot be completely validated. The landslide inventories' assumptions and corresponding adjustments were selected carefully but could have led to some inaccuracy in the landslide inventories.

6.2 Generalized Additive Models

To derive the explanatory pre-, co-, and post-seismic landslide models, GAMs were applied. The parameters for the GAMs were carefully selected based on prior landslide models and conducted literature. However, it should be noted that it cannot be undoubtedly stated that the parameters used in this study accurately reflect reality. There will always be a level of uncertainty in landslide modelling. There is a possibility that the data was somewhat oversimplified in the derived GAMs. Furthermore, factors not considered in this research could have affected other parameters or landslide occurrence.

In this study, the decision was made not to split the data into a testing and validation dataset, as the goal was to assess the influence of the available earthquake data on landslide occurrence. As stated previously, this is contradictory to other landslide studies. Thus, the AUC is logically quite high for each model, as both testing and validation are done on the same data. However, this allows for a better comparison between the variables, which was the main objective. Landslide prediction was not the main goal.

Furthermore, a Bayesian version of a GAM was applied in this research. The models are thus set up using a binary presence-absence variable. In this case, landslide occurrence, categorized as either landslide(s) or no landslide(s) for every slope unit. This corresponds to a Bernoulli probability distribution. However, landslide count, as well as landslide size, was not incorporated in the model. It could have been beneficial to use an approach based on landslide counts, which estimates the landslide intensity instead of landslide occurrence, as is proposed by (Lombardo et al., 2018b). This would have provided more information per slope unit than the susceptibility calculated in this research.

As indicated in the results, the Receiver Operating Characteristic (ROC) curves of the GAMs were almost identical for the post-seismic GAMs, including and excluding the earthquake parameters. The corresponding AUC of the two ROC curves was 0.816. To reflect on the decision to include PGA instead of PGV and PGD, the models have also been conducted using the other two ground motion parameters to ensure the correct parameter had been chosen. Including PGV or PGD instead of PGA barely affects the ROC curves. In both cases, the AUC for both models remains 0.816.

6.3 Recommendations

This research was conducted under the assumption that rainfall follows a homogenous pattern. The pre-seismic inventory covers previous monsoon seasons, and the monsoon seasons of 2005 and 2006 are assumed not to be spatially different. There is a possibility that a heterogeneous rainfall pattern has affected the location and occurrence of post-seismic landslides in the study area. This aspect has been excluded from this research. However, studying rainfall patterns and their effect on landslide occurrence would be interesting. This can be done using a similar approach but include the rainfall patterns instead of the earthquake parameters. Thus, merely focusing on rainfall-induced landslides and using rainfall predictions to identify a spatial pattern of landslide risk. However, this does not provide the same lead time opportunity and relies on the accuracy of rainfall forecasts. Nevertheless, further studying options of landslide risk reduction and improvements of EWEA for landslides are vital.

However, using the earthquake's legacy effect on landslide occurrence is going to be challenging. The results of this study have not led to any new information or potential strategy for FbF. In this case, it is not possible to use the time frame between the earthquake and the first western monsoon season to reduce the landslide risk. Considering the low differences in susceptibility, it is not expected that an adjusted method and different choice of landslide controlling parameters would drastically change the results. However, it could be scientifically interesting to examine whether a similar study in a different area would give similar or completely different results. The results of this study, however, do not indicate that this would be the case. The influence of the earthquake parameters needs to show a clear spatial effect on landslide susceptibility for early action strategies to work and be effective. Otherwise, false alarms or missed landslides could have drastic consequences.

As previously discussed, unfortunately, the study's results provide no new opportunities to reduce landslide risk. But, eliminating the possibility of new early action approaches will increase the focus on existing effective strategies or potential new methods. Research should especially focus on combatting lead time constraints and the localized nature of landslides. However, it is often already known which areas are prone to landslides, and we should approach risk by reducing exposure and vulnerability, even before landslide risk drastically increases during and after seismic activity.

Bibliography

Alin, A. (2010). Multicollinearity. *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(3):370–374.

Althuwaynee, O. F., Pradhan, B., and Lee, S. (2012). Application of an evidential belief function model in landslide susceptibility mapping. *Computers & Geosciences*, 44:120–135.

Anderson, S. A. and Sitar, N. (1995). Analysis of rainfall-induced debris flows. *Journal of Geotechnical Engineering*, 121(7):544–552.

Ashford, S., Sitar, N., Lysmer, J., and Deng, N. (1997). Topographic effects on the seismic response of steep slopes. *Bulletin of the seismological society of America*, 87(3):701–709.

Ashford, S. A. and Sitar, N. (1997). Analysis of topographic amplification of inclined shear waves in a steep coastal bluff. *Bulletin of the seismological society of America*, 87(3):692–700.

Asia Regional Resilience to a Changing ClimateMet Office, International Federation of Red Cross and Red Cresent SocietiesUKaidAnticipation HubRisk Informed Early Action Partnership (2020). The future of forecasts: Impact-based forecasting for early action.

Basharat, M., Shah, H. R., and Hameed, N. (2016). Landslide susceptibility mapping using GIS and weighted overlay method : a case study from NW Himalayas , Pakistan. *Arabian Journal of Geosciences*.

Bhandari, R. (n.d.). Landslide hazard mapping methodology – an approach paper. Commisioned by Building Materials Technology Promotion Council, Ministry of Urban Affairs, New Delhi. Unpublished Manuscript.

Bosi, C., Dramis, F., and Gentili, B. (1985). Carte geomorfologiche di dettaglio a indirizzzo applicativo e carte di stabilita su base geomorfologica. *Geologia Applicata e Idrogeologia*, 20:53–62.

Brand, E. W. (1981). No Title. *Proc. 10th ICSMFE*, 3:372–376.

Brenner, R. P. and Brand, E. W. (1985). Field stress path simulation of rain-induced slope failure. *International conference on soil mechanics and foundation engineering*, 11:991–996.

Buchhorn, M., Smets, B., Bertels, L., De Roo, B., Lesiv, M., Tsendbazar, N., Herold, M., and Fritz, S. (2019). Copernicus global land service: Land cover 100m: collection 3: epoch 2019: Globe 2020. Retrieved March, 2020.

Carrara, A. (1983). Multivariate models for landslide hazard evaluation. *Journal of the International Association for Mathematical Geology*, 15(3):403–426.

Carrara, A., Cardinali, M., Detti, R., Guzzetti, F., Pasqui, V., and Reichenbach, P. (1991). Gis techniques and statistical models in evaluating landslide hazard. *Earth surface processes and landforms*, 16(5):427–445.

Carrara, A., Cardinali, M., Guzzetti, F., and Reichenbach, P. (1995). Gis technology in mapping landslide hazard. In *Geographical information systems in assessing natural hazards*, pages 135–175. Springer.

Chauhan, S., Sharma, M., and Arora, M. K. (2010). Landslide susceptibility zonation of the chamoli region, garhwal himalayas, using logistic regression model. *Landslides*, 7(4):411–423.

Chung, C.-J. F., Fabbri, A. G., and Van Westen, C. J. (1995). Multivariate regression analysis for landslide hazard zonation. In *Geographical information systems in assessing natural hazards*, pages 107–133. Springer.

Coughlan de Perez, E., van den Hurk, B., Van Aalst, M., Jongman, B., Klose, T., and Suarez, P. (2015). Forecast-based financing: an approach for catalyzing humanitarian action based on extreme weather and climate forecasts. *Natural Hazards and Earth System Sciences*, 15(4):895–904.

Crosta, G. B. (2004). Introduction to the special issue on rainfall-triggered landslides and debris flows. *Engineering Geology*, 73(3-4):191–192.

de Ruiter, M. C., Couasnon, A., van den Homberg, M. J., Daniell, J. E., Gill, J. C., and Ward, P. J. (2020). Why we can no longer ignore consecutive disasters. *Earth's future*, 8(3):e2019EF001425.

Earthquake Engineering Research Institute (2005). First Report on the Kashmir Earthquake of October 8, 2005.

Elia, G. and Margherita, A. (2018). Can we solve wicked problems? A conceptual framework and a collective intelligence system to support problem analysis and solution design for complex social issues.

Georgiadou, Y. (2018). Cultural theory and policy problem structuring [Powerpoint slides].

Godefroy, P. and Humbert, M. (1983). La cartographie des risques naturels lies aux mouvements de terrain et aux seismes. *Hydrogéologie-Géologie de l'Ingénieur Orléans*, 2:69–90.

Google Earth (2020). Narah, pakistan. 34°32'24.32"n, 73°21'26.77"e. eye alt 2.16 km. Retrieved Octobre 4, 2020.

Gorum, T., Fan, X., van Westen, C. J., Huang, R. Q., Xu, Q., Tang, C., and Wang, G. (2011). Distribution pattern of earthquake-induced landslides triggered by the 12 May 2008 Wenchuan earthquake. *Geomorphology*, 133(3-4):152–167.

Gorum, T., van Westen, C. J., Korup, O., van der Meijde, M., Fan, X., and van der Meer, F. D. (2013). Complex rupture mechanism and topography control symmetry of mass-wasting pattern, 2010 Haiti earthquake. *Geomorphology*, 184:127–138.

Greco, R., Giorgio, g., Capparelli, G., and Versace, P. (2013). Early warning of rainfall-induced landslides based on empirical mobility function predictor. *Engineering geology*, 153:68–79.

Guzzetti, F., Carrara, A., Cardinali, M., and 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.

Hansen, A., Franks, C., Kirk, P., Brimicombe, A., and Tung, F. (1995). Application of gis to hazard assessment, with particular reference to landslides in hong kong. *Geographical Information Systems in assessing natural hazards*, pages 273–298.

Hastie, T. and Tibshirani, R. (1990). Exploring the Nature of Covariate Effects in the Proportional Hazards Model.

Huang, R. and Fan, X. (2013). The landslide story. *Nature Geoscience*, 6(5):325–326.

Humbert, M. (1977). La cartographie zermos-modalites d'etablissement des cartes des zones exposees a des risques lies aux mouvements du sol et du sous-sol. *PUBLICATIONS DU GFGI*, 1(2).

Hussain, A., Mughal, N., Haq, I., and Latif, A. (2004). Geological map of the gari habib ullah area, district mansehra and parts of muzaffarabad district, ajk, geological map series. *Geomatics, Natural Hazards and Risk*.

Hussain, S., Nisar, A., Khazai, B., and Dellow, G. (2006). The Kashmir earthquake of October 8, 2005: impacts in Pakistan. *Earthquake Engineering Research Institute Special Paper*, 8.

IFRC (2019a). Forecast based action by the dref summary 2019. Retrieved July 30, 2021.

IFRC (2019b). Mozambique/africa: To coordinate early actions for preparation and response to cyclones in mozambique. Retrieved July 27, 2021.

IFRC (n.d.a). Early warning early action. Retrieved July 20, 2020.

IFRC (n.d.b). Forecast-based action by the dref. Retrieved July 25, 2020.

International Federation of Red Cross and Red Crescent Societies (2020). Disaster risk management policy - from prevention to response and recovery.

Intrieri, E., Gigli, G., Casagli, N., and Nadim, F. (2013). Brief communication" landslide early warning system: toolbox and general concepts". *Natural Hazards and Earth System Sciences*, 13(1):85–90.

Jakob, M. and Hungr, O. (2005). Debris-flow Hazards and Related Phenomena. *Journal of the Royal Society of Medicine*, 739.

Kamp, U., Growley, B. J., Khattak, G. A., and Owen, L. A. (2008). GIS-based landslide susceptibility mapping for the 2005 Kashmir earthquake region. *Geomorphology*, 101(4):631–642.

Khan, S., Van Der Meijde, M., Van Der Werff, H., and Shafique, M. (2020). The impact of topography on seismic amplification during the 2005 Kashmir earthquake. *Natural Hazards and Earth System Sciences*, 20(2):399–411.

Khan, S. F., Kamp, U., and Owen, L. A. (2013). Geomorphology Documenting fi ve years of landsliding after the 2005 Kashmir earthquake , using repeat photography. *Geomorphology*, 197:45–55.

Khattak, G. A., Owen, L. A., Kamp, U., and Harp, E. L. (2010). Geomorphology Evolution of earthquake-triggered landslides in the Kashmir Himalaya , northern Pakistan. *Geomorphology*, 115(1-2):102–108.

Kienholz, H., Hafner, H., Schneider, G., and Tamrakar, R. (1983). Mountain hazards mapping in nepal's middle mountains maps of land use and geomorphic damages (kathmandu-kakani area). *Mountain Research and Development*, pages 195–220.

Kienholz, H., Schneider, G., Bichsel, M., Grunder, M., and Mool, P. (1984). Mapping of mountain hazards and slope stability. *Mountain Research and Development*, pages 247–266.

Korup, O. (2010). Earthquake-triggered landslides—spatial patterns... - Google Scholar. *Module 1a-Report*.

Latif, A., Afridi, A. G. K., and Majid, A. N. (2008). Geological map of the Balakot quadrangle (Earthquake affected area), District Mansehra, NWFP, Pakistan, Geological Map Series. *Geological Survey of Pakistan, Islamabad-Pakistan*.

Lin, C.-w., Liu, S.-h., Lee, S.-y., and Liu, C.-c. (2006). Impacts of the Chi-Chi earthquake on subsequent rainfall-induced landslides in central Taiwan. *Engineering Geology*, 86(2-3):87–101.

Lombardo, L., Opitz, T., and Huser, R. (2018a). Point process-based modeling of multiple debris flow landslides using inla: an application to the 2009 messina disaster. *Stochastic environmental research and risk assessment*, 32(7):2179–2198.

Lombardo, L., Opitz, T., and Huser, R. (2018b). Point process-based modeling of multiple debris flow landslides using INLA: an application to the 2009 Messina disaster. *Stochastic Environmental Research and Risk Assessment*, 32(7):2179–2198.

Lombardo, L., Tanyas, H., and Nicu, I. C. (2020). Spatial modeling of multi-hazard threat to cultural heritage sites. *Engineering Geology*, 277:105776.

Ma, Y., Ma, B., Jiao, H., Zhang, Y., Xin, J., and Yu, Z. (2020). An analysis of the effects of weather and air pollution on tropospheric ozone using a generalized additive model in Western China: Lanzhou, Gansu. *Atmospheric Environment*, 224:117342.

Mark, R. K. (1992). Map of debris-flow probability, san mateo county, california. 1257-M.

Nadim, F. and Intrieri, E. (2011). Early warning systems for landslides: challenges and new monitoring technologies. In *5th Canadian Conference on Geotechnique and Natural Hazards, Kelowna, BC, Canada*, pages 15–17.

Neely, M. K. and Rice, R. M. (1990). Estimating risk of debris slides after timber harvest in northwestern california. *Bulletin of the Association of Engineering Geologists*, 27(3):281–289.

Neuland, H. (1976). A prediction model of landslips. *Catena*, 3(2):215–230.

Owen, L. A., Kamp, U., Khattak, G. A., Harp, E. L., Keefer, D. K., and Bauer, M. A. (2008). Landslides triggered by the 8 October 2005 Kashmir earthquake. *Geomorphology*.

Piciullo, L., Calvello, M., and Cepeda, J. M. (2018). Territorial early warning systems for rainfall-induced landslides. *Earth-Science Reviews*, 179:228–247.

Pourghasemi, H. R., Jirandeh, A. G., Pradhan, B., Xu, C., and Gokceoglu, C. (2013). Landslide susceptibility mapping using support vector machine and gis at the golestan province, iran. *Journal of Earth System Science*, 122(2):349–369.

Pradhan, B., Oh, H.-J., and Buchroithner, M. (2010). Weights-of-evidence model applied to landslide susceptibility mapping in a tropical hilly area. *Geomatics, Natural Hazards and Risk*, 1(3):199–223.

Qasim, M., Khan, M. A., and Haneef, M. (2014). Stratigraphic characterization of the early cambrian abbottabad formation in the sherwan area, hazara region, n. pakistan: implications for early paleozoic stratigraphic correlation in nw himalayas, pakistan. *Journal of Himalayan Earth Sciences*, 47(1):25.

R-INLA Project (n.d). INLA. Retrieved December 4, 2020.

Roth, R. A. (1983). Factors affecting landslide-susceptibility in san mateo county, california. *Bulletin of the Association of Engineering Geologists*, 20(4):353–372.

Saba, S. B., Meijde, M. V. D., and Werff, H. V. D. (2010). Geomorphology Spatiotemporal landslide detection for the 2005 Kashmir earthquake region. *Geomorphology*, 124(1-2):17–25.

Saito, M. (1961). Failure of soil due to creep. In *Proc. 5th Int. Conf. on SMFE*, pages 315–318.

Sana, H. and Nath, S. K. (2016). In and around the hazara-kashmir syntaxis: A seismotectonic and seismic hazard perspective. *J. Ind. Geophys. Union*, 20:496–505.

Sato, H. P., Hasegawa, H., Fujiwara, S., Tobita, M., Koarai, M., Une, H., and Iwahashi, J. (2007). Interpretation of landslide distribution triggered by the 2005 Northern Pakistan earthquake using SPOT 5 imagery. *Landslides*, 4(2):113–122.

Seeley, M. W. and West, D. O. (1990). Approach to geologic hazard zoning for regional planning, inyo national forest, california and nevada. *Bulletin of the Association of Engineering Geologists*, 27(1):23–35.

Segoni, S., Battistini, A., Rossi, G., Rosi, A., Lagomarsino, D., Catani, F., Moretti, S., and Casagli, N. (2014). an operational landslide early warning system at regional scale based on space–time variable rainfall thresholds. *Nat Hazards Earth Syst Sci Discuss*, 2:6599–6622.

Shafique, M. (2020). Geomorphology Spatial and temporal evolution of co-seismic landslides after the 2005 Kashmir earthquake. *Geomorphology*.

Shafique, M., van der Meijde, M., and Khan, M. A. (2016). Journal of Asian Earth Sciences A review of the 2005 Kashmir earthquake-induced landslides ; from a remote sensing prospective. *JOURNAL OF ASIAN EARTH SCIENCES*, 118:68–80.

Sidle, R. C. and Swanston, D. N. (1982). Analysis of a small debris slide in coastal Alaska. *Canadian Geotechnical Journal*, 19(2):167–174.

Sinscov, T. and Campbell, M. (2002). *Statistics at square one*, pages 111–25. London: Bmj.

Sitar, N. (1992). Conditions leading to the initiation of rainfall-induced debris flows. In *Geotechnical Engineering Division, Specialty Conference on Stability and Performance of Slopes and Embankments-II*, pages 834–839, New York, NY. ASCE.

Tang, C., Van Westen, C. J., Tanyas, H., and Jetten, V. G. (2016). Analysing post-earthquake landslide activity using multi-temporal landslide inventories near the epicentral area of the 2008 Wenchuan earthquake. *Natural Hazards and Earth System Sciences*, 16(12):2641–2655.

Terzaghi, K. (1950). Mechanism of Landslides. In *Application of Geology to Engineering Practice*. Geol. Soc. Am, berkey vol edition.

Thiery, Y., Maquaire, O., and Fressard, M. (2014). Application of expert rules in indirect approaches for landslide susceptibility assessment. *Landslides*, 11:411–424.

Tsiplakidis, J. and Photis, Y. N. (2019). Multihazard Risk Assessment from Qualitative Methods to Bayesian Networks: Reviewing Recent Contributions and Exploring New Perspectives. In *Geospatial Challenges in the 21st Century*, pages 401–429. Springer, Cham.

United Nations of Disaster Risk Reduction (2015). Sendai Framework for Disaster Risk Reduction 2015 - 2030. p.14.

United States Geological Survey (n.d.). Today in Earthquake History.

USGS (n.d.). Can you predict earthquakes? Retrieved August 1, 2021.

Van Westen, C. (1993). Application of geographical information system to landslide hazard zonation.

Van Westen, C. (1994). Gis in landslide hazard zonation: a review, with examples from the andes of colombia. *Mountain environments & geographic information systems*, pages 135–166.

Verstappen, H. T. (1983). Applied geomorphology: geomorphological surveys for environmental development.

Wang, G. and Sassa, K. (2003). Pore-pressure generation and movement of rainfall-induced landslides : effects of grain size and fine-particle content. *Engineering Geology*, 69:109–125.

Weather & Climate (n.d.). Climate and Average Weather in Pakistan.

World Health Organization (n.d). Landslides.

World Meteorological Organization (n.d.). Muzaffarabad. Retrieved September 20, 2020.

Xu, Q., Peng, D., Fan, X., Zhu, X., and He, C. (2020). Presenting some successful cases of regional landslides early warning systems in china. In *Workshop on World Landslide Forum*, pages 279–286. Springer.

Xu, Q., Yuan, Y., Zeng, Y., and Hack, R. (2011). Some new pre-warning criteria for creep slope failure. *Science China Technological Sciences*, 54(1):210–220.

Yang, W., Qi, W., and Zhou, J. (2018). Decreased post-seismic landslides linked to vegetation recovery after the 2008 Wenchuan earthquake. *Ecological Indicators*, 89(February):438–444.

Yang, Y., Yang, J., Xu, C., Xu, C., and Song, C. (2019). Local-scale landslide susceptibility mapping using the b-geosvc model. *Landslides*, 16(7):1301–1312.

Yin, K. and Yan, T. (1988). Statistical prediction models for instability of metamorphosed rocks. In *International symposium on landslides. 5*, pages 1269–1272.

Zimmermann, M., Bichsel, M., and Kienholz, H. (1986). Mountain hazards mapping in the khumbu himal, nepal. *Mountain Research and Development*, pages 29–40.

Appendix

Appendix 1: Processing steps parameters in the data matrix

Parameters	Methodology
Elevation	The DEM was converted to a point file. It was then joined to the slope units, using 'join attributes by location (summary)', by calculating the mean of the DEM values for each slope unit.
Slope	The DEM layer was used to calculate slope in QGIS and the slope file was subsequently converted to a point file. It was then joined to the slope units, using 'join attributes by location (summary)', by calculating the mean of the slope for each slope unit.
Slope Curvature	Slope curvature was derived from the DEM. The slope curvature file was then converted to a point file. Next, it was joined to the slope unit, using 'join attributes by location (summary)', by calculating the mean of the slope curvature values for each slope unit.
Slope Aspect	The DEM was used to calculate slope aspect. The slope aspect file was converted to a point file. It was then joined to the slope units, using 'join attributes by location (summary)', by calculating the mean of the slope aspect values for each slope unit.
CMT angle	The angle between the slope unit and the CMT location was calculated in QGIS using the bearing of the slope unit and the bearing of the CMT location. These were then characterized into facing towards the CMT, facing away and in other directions based on angles of 45 degrees.
Distance to Fault	Distance to faults was determined by using a 100-, 200-, 300- and 400-meter buffers. The percentage of each slope units that cover each of these buffer zones was then calculated.
Geology	The percentage of all the geology types that cover each slope unit was calculated using the tool' spatial join by location (summary)' in QGIS.
Land Cover	The percentage of all the land cover types that cover each slope unit was calculated using the tool' spatial join by location (summary)' in QGIS.
Distance to Road Network	A 50-meter buffer zone was created for the road network. The percentage of each slope unit that fall within the 50-meter buffer was incorporated in the data matrix, using the 'overlap analysis' tool.
Distance to Drainage	A 50-meter buffer zone was created for drainages. The percentage of each slope unit that fall within the buffer zone was incorporated in the data matrix.
PGV	The PGV file was converted to a raster by applying IDW interpolation, using 12 variable radius and a power of 2. Using the Zonal Statistics tool in QGIS the mean value per slope unit was calculated
PGA	The PGA file was converted to a raster by applying IDW interpolation, using 12 variable radius and a power of 2. Using the Zonal Statistics tool in QGIS the mean PGA per slope unit was calculated
PGD	The PGD file was already in raster format. Using the Zonal Statistics tool in QGIS the mean PGD value per slope unit was calculateds

Appendix 2: Intercorrelation of parameters

To find the inter-factor dependencies in the data matrix, different methods have been applied. All of which were done using a 95% confidence interval. For most combinations of variables, the Spearman correlation coefficients were calculated. Spearman's correlation was used because not all data is normally distributed. The association between CMT angle (categorical variable) and the other continuous variables is calculated using Eta squared. The correlations of (-)0.0 – (-)0.19 are seen as neglectable correlations, (-)0.20 – (-)0.39 are low correlations, (-)0.40 – (-)0.59 are considered moderate, (-)0.60 – (-)0.79 as moderately high, and (-)0.8 – (-)1 as a very strong correlation. These are highlighted in the data matrix. The land cover and geology types are named in the correlation matrix, but their corresponding names are discussed if relevant. The remaining land cover and geology names are provided in appendix 3.

Most correlations and associations are neglectable to low. The moderate and moderately high correlations will be discussed in this section to reflect on the possible causes of the observed correlations and associations. The very strong correlations that are important to consider before conducting a GAM have been discussed in section 4.1.1.

Elevation has a moderate correlation of -0.570 with geology type 7 (paleocence rocks), a moderate correlation of -0.561 with land cover type 2 (built-up), and a moderate correlation of -0.462 with drainage. All moderate correlations are negative, meaning that paleocence rocks, built-up and drainage are more frequently found at higher elevations. For built-up and drainage, this is expected since built-up areas, and rivers and water bodies are typically located at lower elevations, in between mountains. The moderate negative correlation of elevation with paleocence rocks indicates that paleocence rocks in the study area are more frequently found at higher elevations. Furthermore, elevation is moderately highly correlated with roads, with a correlation of -0.596. Similar to the land cover type built-up and drainage, this is expected because roads are usually located at lower elevations and in and around built-up areas.

Several geology and land cover types show moderate correlations. These are geology type 7 (paleocence rocks) and land cover type 2 (built-up) with a correlation of 0.556, indicating that paleocence rocks and built-up are often present in the same slope unit. A correlation of 0.408 for land cover type 1 (bare/sparse vegetation) and type 11 (permanent water bodies) shows that permanent water bodies are often accompanied by bare vegetation. Land cover type 7 (evergreen broadleaved closed forest) and 8 (decideous needle-leaved closed forest) correlate with 0.429, showing that closed forests often have mixed tree species. Land cover types 7 (evergreen broadleaved closed forest) and 9 (deciduous needle-leaved open forest) are negatively correlated with a value of -0.401. Thus evergreen broadleaved closed forest and open deciduous forest are typically not found in proximity to each other within the same slope units. Land cover types 9 (deciduous needle-leaved open forest) and 12 (shrubland) is correlated with a value of 0.406, which indicates that shrubland and deciduous needle-leaved open forest are frequently found within the same slope unit.

Road and drainage proximity parameters are also correlated, with a correlation of 0.425, indicating that roads and drainage are regularly found within the same slope unit. Furthermore, the parameter roads correlate with land cover type 2 (built-up) with a correlation of 0.529 and 0.426 with geology type 7 (paleocence rocks). As described priorly, land cover type 2 and geology type 7 are also moderately correlated.

Moreover, geology type 7 is correlated with the parameters fault proximity of 300 meters and 400 meters, with correlations of respectively 0.402 and 0.431. This could be a coincidence, and the correlations are only moderate and thus no cause for concern. The correlations with both 300 and 400 meters is most likely due to the slope units located on the edge of the 300 and 400 meters.

Parameter	SU area	DEM mean	Slope mean	Aspect mean	Curva mean	CMT angle	PGV mean	PGA mean	PGD mean	Fault 100m	Fault 200m	Fault 300m	Fault 400m	Geology 1	Geology 2	Geology 3	Geology 4	Geology 5	Geology 6
SU area	NA	0.030	0.032	0.026	0.073	0.000	-0.001	-0.001	0.000	0.110	0.101	0.117	0.113	0.139	0.006	-0.011	0.051	-0.023	0.055
DEM mean	0.030	NA	0.253	0.078	0.065	0.000	0.059	-0.019	0.042	-0.253	-0.295	-0.315	-0.331	-0.011	-0.088	0.028	0.303	0.181	0.302
Slope mean	0.032	0.253	NA	-0.366	0.139	0.001	0.140	0.138	0.147	-0.205	-0.239	-0.235	-0.246	0.157	-0.067	0.053	0.076	0.043	0.073
Aspect mean	0.026	0.078	-0.366	NA	0.007	0.019	-0.009	-0.025	-0.024	0.030	0.039	0.042	0.055	0.023	-0.001	-0.156	-0.092	0.096	0.016
Curva mean	0.073	0.065	0.139	0.007	NA	0.000	0.031	0.027	0.005	-0.005	0.004	0.010	0.004	0.021	0.015	-0.013	-0.008	-0.008	-0.009
CMT angle	0.00	0.00	0.00	0.02	0.00	NA	0.001	0.002	0.000	0.00	0.00	0.00	0.00	0.000	0.000	0.001	0.005	0.000	0.001
PGV mean	-0.001	0.059	0.140	-0.009	0.031	0.001	NA	0.960	0.952	-0.250	-0.290	-0.302	-0.318	0.316	-0.014	-0.225	-0.015	0.026	0.039
PGA mean	-0.001	-0.019	0.138	-0.025	0.027	0.002	0.960	NA	0.905	-0.210	-0.246	-0.252	-0.268	0.302	-0.020	-0.218	-0.027	0.008	0.007
PGD mean	0.000	0.042	0.147	-0.024	0.005	0.000	0.952	0.905	NA	-0.264	-0.302	-0.317	-0.334	0.247	-0.017	-0.248	-0.006	0.097	0.085
Fault 100m	0.110	-0.253	-0.205	0.030	-0.005	0.000	-0.250	-0.210	-0.264	NA	0.919	0.865	0.825	0.003	0.026	0.080	-0.129	-0.101	-0.172
Fault 200m	0.101	-0.295	-0.239	0.039	0.004	0.000	-0.290	-0.246	-0.302	0.919	NA	0.949	0.912	-0.009	0.044	0.096	-0.142	-0.124	-0.201
Fault 300m	0.117	-0.315	-0.235	0.042	0.010	0.001	-0.302	-0.252	-0.317	0.865	0.949	NA	0.965	0.011	0.050	0.087	-0.158	-0.132	-0.216
Fault 400m	0.113	-0.331	-0.246	0.055	0.004	0.005	-0.318	-0.268	-0.334	0.825	0.912	0.965	NA	0.016	0.046	0.083	-0.173	-0.143	-0.231
Geology 1	0.139	-0.011	0.157	0.023	0.021	0.000	0.316	0.302	0.247	0.003	-0.009	0.011	0.016	NA	-0.048	-0.096	-0.358	-0.255	-0.201
Geology 2	0.006	-0.088	-0.067	-0.001	0.015	0.001	-0.014	-0.020	-0.017	0.026	0.044	0.050	0.046	-0.048	NA	-0.013	-0.060	-0.050	-0.030
Geology 3	-0.011	0.028	0.053	-0.156	-0.013	0.006	-0.225	-0.218	-0.248	0.080	0.096	0.087	0.083	-0.096	-0.013	NA	0.025	-0.098	-0.060
Geology 4	0.051	0.303	0.076	-0.092	-0.008	0.002	-0.015	-0.027	-0.006	-0.129	-0.142	-0.158	-0.173	-0.358	-0.060	0.025	NA	-0.212	0.055
Geology 5	-0.023	0.181	0.043	0.096	-0.008	0.002	0.026	0.008	0.097	-0.101	-0.124	-0.132	-0.143	-0.255	-0.050	-0.098	-0.212	NA	0.138
Geology 6	0.055	0.302	0.073	0.016	-0.009	0.000	0.039	0.007	0.085	-0.172	-0.201	-0.216	-0.231	-0.201	-0.030	-0.060	0.055	0.138	NA
Geology 7	0.067	-0.570	-0.344	0.142	0.008	0.002	-0.331	-0.309	-0.331	0.331	0.378	0.402	0.431	-0.107	0.069	-0.061	-0.281	-0.270	-0.212
Geology 8	0.070	0.036	0.011	-0.092	-0.001	0.005	-0.185	-0.184	-0.188	0.087	0.116	0.114	0.122	-0.066	-0.009	0.174	-0.061	-0.049	-0.012
Geology 9	-0.037	-0.107	-0.014	-0.037	0.009	0.001	0.034	0.013	0.016	-0.052	-0.042	-0.040	-0.044	-0.036	0.334	-0.016	-0.074	-0.061	-0.037
Land cover 1	0.050	-0.118	-0.038	-0.023	-0.003	0.000	0.047	0.069	0.049	-0.019	-0.006	0.007	0.012	0.002	-0.008	-0.016	-0.023	-0.038	-0.036
Land cover 2	0.072	-0.561	-0.362	0.029	0.013	0.001	-0.239	-0.200	-0.220	0.295	0.344	0.369	0.379	-0.088	0.085	-0.005	-0.148	-0.147	-0.209
Land cover 3	0.067	-0.200	-0.119	-0.087	0.030	0.001	-0.212	-0.201	-0.226	0.128	0.151	0.153	0.155	-0.035	0.017	0.090	-0.093	-0.161	-0.072
Land cover 4	0.028	0.014	-0.018	-0.004	0.021	0.005	0.002	-0.003	0.014	-0.012	-0.014	-0.016	-0.017	0.046	-0.002	-0.004	-0.017	-0.014	-0.009
Land cover 5	0.007	0.012	-0.001	0.021	-0.023	0.001	0.043	0.038	0.044	-0.017	-0.020	-0.022	-0.024	0.016	-0.003	-0.005	0.016	-0.020	-0.012
Land cover 6	0.030	0.054	0.038	-0.028	0.015	0.000	0.003	-0.002	0.006	-0.017	-0.020	-0.022	-0.024	-0.020	-0.003	-0.005	-0.025	0.067	-0.012
Land cover 7	0.048	0.373	-0.155	0.154	-0.071	0.001	0.127	0.084	0.101	-0.138	-0.173	-0.188	-0.204	0.005	-0.022	-0.005	0.135	0.026	0.069
Land cover 8	0.108	0.275	0.008	0.015	-0.007	0.003	0.093	0.066	0.065	-0.115	-0.132	-0.141	-0.156	0.060	0.001	0.000	0.026	0.040	0.123
Land cover 9	0.153	-0.015	0.193	-0.005	0.040	0.004	-0.043	-0.036	0.008	0.034	0.047	0.044	0.047	-0.012	-0.002	-0.110	-0.023	0.119	0.073
Land cover 10	0.010	-0.213	-0.112	0.005	-0.025	0.000	-0.069	-0.046	-0.072	0.058	0.069	0.069	0.067	-0.072	-0.014	-0.028	-0.023	-0.014	-0.044
Land cover 11	0.032	-0.241	-0.100	-0.008	-0.028	0.000	0.014	0.045	0.036	0.012	0.047	0.065	0.078	-0.019	-0.013	-0.025	-0.066	-0.036	-0.056
Land cover 12	0.183	-0.054	0.288	-0.067	0.088	0.000	0.011	0.014	0.040	-0.002	0.006	0.003	0.000	0.116	0.002	-0.034	-0.022	0.065	0.004
Land cover 13	0.151	0.010	-0.135	0.107	-0.174	0.000	0.057	0.050	0.049	-0.008	-0.019	-0.014	-0.012	0.081	0.077	0.037	0.044	-0.057	-0.030
Land cover 14	0.095	-0.040	0.214	-0.071	0.035	0.000	0.059	0.075	0.005	0.059	0.047	0.051	0.049	0.120	-0.004	0.062	-0.018	-0.084	-0.076
Road 50m	0.086	-0.596	-0.212	-0.081	0.008	0.000	-0.168	-0.103	-0.153	0.234	0.274	0.299	0.302	-0.006	0.056	-0.056	-0.178	-0.127	-0.193
Drainage 50m	0.066	-0.462	-0.232	-0.017	-0.031	0.000	-0.182	-0.149	-0.165	0.194	0.209	0.227	0.233	-0.132	0.052	-0.076	-0.067	-0.045	-0.082

	No significant correlation or association (95% confidence interval)
	Neglectable correlation or association (-)0.0 – (-)0.19
	Low correlation or association (-)0.20 – (-)0.39
	Moderate correlation association (-)0.40– (-)0.59
	Moderately high correlation (-)0.60 – (-)0.79
	Very strong correlation (-)0.80 – (-)1.00

Parameter	Geology 7	Geology 8	Geology 9	Land cover 1	Land cover 2	Land cover 3	Land cover 4	Land cover 5	Land cover 6	Land cover 7	Land cover 8	Land cover 9	Land cover 10	Land cover 11	Land cover 12	Land cover 13	Land cover 14	Road 50m	Drainage 50m
SU area	0.067	0.070	-0.037	0.050	0.072	0.067	0.028	0.007	0.030	0.048	0.108	0.153	0.010	0.032	0.183	0.151	0.095	0.086	0.066
DEM mean	-0.570	0.036	-0.107	-0.118	-0.561	-0.200	0.014	0.012	0.054	0.373	0.275	-0.015	-0.213	-0.241	-0.054	0.010	-0.040	-0.596	-0.462
Slope mean	-0.344	0.011	-0.014	-0.038	-0.362	-0.119	-0.018	-0.001	0.038	-0.155	0.008	0.193	-0.112	-0.100	0.288	-0.135	0.214	-0.212	-0.232
Aspect mean	0.142	-0.092	-0.037	-0.023	0.029	-0.087	-0.004	0.021	-0.028	0.154	0.015	-0.005	0.005	-0.008	-0.067	0.107	-0.071	-0.081	-0.017
Curva mean	0.008	-0.001	0.009	-0.003	0.013	0.030	0.021	-0.023	0.015	-0.071	-0.007	0.040	-0.025	-0.028	0.088	-0.174	0.035	0.008	-0.031
CMT angle	0.006	0.002	0.002	0.000	0.002	0.005	0.001	0.000	0.001	0.001	0.005	0.001	0.000	0.001	0.003	0.004	0.000	0.000	0.000
PGV mean	-0.331	-0.185	0.034	0.047	-0.239	-0.212	0.002	0.043	0.003	0.127	0.093	-0.043	-0.069	0.014	0.011	0.057	0.059	-0.168	-0.182
PGA mean	-0.309	-0.184	0.013	0.069	-0.200	-0.201	-0.003	0.038	-0.002	0.084	0.066	-0.036	-0.046	0.045	0.014	0.050	0.075	-0.103	-0.149
PGD mean	-0.331	-0.188	0.016	0.049	-0.220	-0.226	0.014	0.044	0.006	0.101	0.065	0.008	-0.072	0.036	0.040	0.049	0.005	-0.153	-0.165
Fault 100m	0.331	0.087	-0.052	-0.019	0.295	0.128	-0.012	-0.017	-0.017	-0.138	-0.115	0.034	0.058	0.012	-0.002	-0.008	0.059	0.234	0.194
Fault 200m	0.378	0.116	-0.042	-0.006	0.344	0.151	-0.014	-0.020	-0.020	-0.173	-0.132	0.047	0.069	0.047	0.006	-0.019	0.047	0.274	0.209
Fault 300m	0.402	0.114	-0.040	0.007	0.369	0.153	-0.016	-0.022	-0.022	-0.188	-0.141	0.044	0.069	0.065	0.003	-0.014	0.051	0.299	0.227
Fault 400m	0.431	0.122	-0.044	0.012	0.379	0.155	-0.017	-0.024	-0.024	-0.204	-0.156	0.047	0.067	0.078	0.000	-0.012	0.049	0.302	0.233
Geology 1	-0.107	-0.066	-0.036	0.002	-0.088	-0.035	0.046	0.016	-0.020	0.005	0.060	-0.012	-0.072	-0.019	0.116	0.081	0.120	-0.006	-0.132
Geology 2	0.069	-0.009	0.334	-0.008	0.085	0.017	-0.002	-0.003	-0.003	-0.022	0.001	-0.002	-0.014	-0.013	0.002	0.077	-0.004	0.056	0.052
Geology 3	-0.061	0.174	-0.016	-0.016	-0.005	0.090	-0.004	-0.005	-0.005	-0.005	0.000	-0.110	-0.028	-0.025	-0.034	0.037	0.062	-0.056	-0.076
Geology 4	-0.281	-0.061	-0.074	-0.023	-0.148	-0.093	-0.017	0.016	-0.025	0.135	0.026	-0.023	-0.023	-0.066	-0.022	0.044	-0.018	-0.178	-0.067
Geology 5	-0.270	-0.049	-0.061	-0.038	-0.147	-0.161	-0.014	-0.020	0.067	0.026	0.040	0.119	-0.014	-0.036	0.065	-0.057	-0.084	-0.127	-0.045
Geology 6	-0.212	-0.012	-0.037	-0.036	-0.209	-0.072	-0.009	-0.012	-0.012	0.069	0.123	0.073	-0.044	-0.056	0.004	-0.030	-0.076	-0.193	-0.082
Geology 7	NA	0.009	0.028	0.111	0.556	0.290	-0.017	-0.024	-0.024	-0.229	-0.230	0.055	0.133	0.159	-0.011	-0.054	-0.056	0.426	0.357
Geology 8	0.009	NA	-0.011	-0.011	0.039	0.071	-0.003	-0.004	-0.004	-0.032	0.010	-0.051	-0.019	-0.017	-0.023	0.037	0.072	0.039	0.014
Geology 9	0.028	-0.011	NA	-0.010	0.045	0.097	-0.002	-0.003	-0.003	0.010	-0.050	-0.030	0.015	-0.015	0.000	0.014	0.018	0.048	0.054
Land cover 1	0.111	-0.011	-0.010	NA	0.073	-0.042	-0.002	-0.003	-0.003	-0.061	-0.073	0.069	0.016	0.408	0.029	-0.028	-0.073	0.020	0.123
Land cover 2	0.556	0.039	0.045	0.073	NA	0.199	-0.014	-0.019	-0.019	-0.253	-0.241	-0.028	0.112	0.213	-0.063	-0.078	-0.160	0.529	0.344
Land cover 3	0.290	0.071	0.097	-0.042	0.199	NA	-0.010	-0.014	-0.014	-0.217	-0.182	-0.008	0.051	-0.065	0.106	-0.100	0.015	0.149	0.113
Land cover 4	-0.017	-0.003	-0.002	-0.002	-0.014	-0.010	NA	-0.001	-0.001	0.037	0.015	0.012	-0.004	-0.004	-0.018	-0.011	-0.007	0.041	-0.011
Land cover 5	-0.024	-0.004	-0.003	-0.003	-0.019	-0.014	-0.001	NA	-0.001	0.026	0.007	-0.018	-0.006	-0.005	-0.003	0.029	0.034	-0.024	-0.016
Land cover 6	-0.024	-0.004	-0.003	-0.003	-0.019	-0.014	-0.001	-0.001	NA	0.047	0.043	-0.005	-0.006	-0.005	-0.026	-0.015	-0.008	-0.024	-0.016
Land cover 7	-0.229	-0.032	0.010	-0.061	-0.253	-0.217	0.037	0.026	0.047	NA	0.429	-0.401	-0.063	-0.085	-0.383	0.225	-0.237	-0.222	-0.168
Land cover 8	-0.230	0.010	-0.050	-0.073	-0.241	-0.182	0.015	0.007	0.043	0.429	NA	-0.380	-0.027	-0.079	-0.330	0.117	0.000	-0.173	-0.170
Land cover 9	0.055	-0.051	-0.030	0.069	-0.028	-0.008	0.012	-0.018	-0.005	-0.401	-0.380	NA	0.063	0.091	0.406	-0.123	-0.218	0.010	0.182
Land cover 10	0.133	-0.019	0.015	0.016	0.112	0.051	-0.004	-0.006	-0.006	-0.063	-0.027	0.063	NA	0.082	-0.014	-0.019	-0.061	0.205	0.347
Land cover 11	0.159	-0.017	-0.015	0.408	0.213	-0.065	-0.004	-0.005	-0.005	-0.085	-0.079	0.091	0.082	NA	-0.016	-0.048	-0.125	0.199	0.227
Land cover 12	-0.011	-0.023	0.000	0.029	-0.063	0.106	-0.018	-0.003	-0.026	-0.383	-0.330	0.406	-0.014	-0.016	NA	-0.159	-0.091	0.032	0.001
Land cover 13	-0.054	0.037	0.014	-0.028	-0.078	-0.100	-0.011	0.029	-0.015	0.225	0.117	-0.123	-0.019	-0.048	-0.159	NA	0.026	-0.058	-0.055
Land cover 14	-0.056	0.072	0.018	-0.073	-0.160	0.015	-0.007	0.034	-0.008	-0.237	0.000	-0.218	-0.061	-0.125	-0.091	0.026	NA	-0.051	-0.077
Road 50m	0.426	0.039	0.048	0.020	0.529	0.149	0.041	-0.024	-0.024	-0.222	-0.173	0.010	0.205	0.199	0.032	-0.058	-0.051	NA	0.425
Drainage 50m	0.357	0.014	0.054	0.123	0.344	0.113	-0.011	-0.016	-0.016	-0.168	-0.170	0.182	0.347	0.227	0.001	-0.055	-0.077	0.425	NA

	No significant correlation or association (95% confidence interval)
	Neglectable correlation or association (-)0.0 – (-)0.19
	Low correlation or association (-)0.20 – (-)0.39
	Moderate correlation association (-)0.40– (-)0.59
	Moderately high correlation (-)0.60 – (-)0.79
	Very strong correlation (-)0.80 – (-)1.00

Appendix 3: Names geology and land cover types

The names of the geology and land cover types have been left out of table 4.3 ad the table in appendix 2 to safe space. Before the adjustment of the data set, there were 14 land cover types and 9 geology types. The corresponding names are provided in the table below.

Landcover		**Geology**	
1	Bare/sparse vegetation	1	Abbottabad Formation
2	Built-up	2	Hazara formation
3	Cropland	3	Manki Formation
4	Deciduous broadleaved (closed forest)	4	Mansehra Orthogenisis
5	Deciduous broadleaved (open forest)	5	Murree Formation
6	Evergreen broadleaved (closed forest)	6	Muzaffarabad Formation
7	Evergreen broadleaved (closed forest)	7	Paleocence rocks
8	Deciduous needle-leaved (closed forest)	8	Quaternary Deposits
9	Deciduous needle-leaved (open forest)	9	Salkhala Formation
10	Herbaceous vegetation		
11	Permanent water bodies		
12	Shrubland		
13	Unknown type (closed forest)		
14	Unknown type (open forest)		

Appendix 4: Spatial overlay of landslide susceptibility with the landslide inventories

The map shows the landslide susceptibility maps of the pre-, co-, and post-seismic models overlaid with the corresponding landslide inventories. This was done to illustrate the similar spatial variability's of the landslide susceptibilities and the landslide inventories of the same years. Therefore, demonstrating the accuracy of the GAMs.